DÉCRYPTAGE DE LA DYNAMIQUE DE SYSTÈMES BIOMOLÉCULAIRES PAR SIMULATIONS NUMÉRIQUES

Francois Dehez

MÉMOIRE DÉPOSÉ EN VUE DE L'OBTENTION DU DIPLÔME DE

HABILITATION À DIRIGER DES RECHERCHES

par

FRANÇOIS DEHEZ

Université de Lorraine et CNRS

Laboratoire de Physique et Chimie Théoriques

DECIPHERING THE DYNAMICS OF BIOMOLECULAR SYSTEMS WITH NUMERICAL SIMULATIONS

DÉCRYPTAGE DE LA DYNAMIQUE DE SYSTÈMES BIOMOLÉCULAIRES PAR SIMULATIONS NUMÉRIQUES

Soutenue le 30 novembre 2021 à Vandœuvre-lès-Nancy

Devant le jury composé de :

Rapporteurs :

- Mme Annick DEJAEGERE (Université de Strasbourg & CNRS)
- Mme Nathalie REUTER (Université de Bergen)
- M. Thomas SIMONSON (Ecole Polythechnique & CNRS)

Examinateurs :

- M. Christophe CHIPOT (Université de Lorraine & CNRS)
- M. Bruno MIROUX (INSERM)
- Mme Eva PEBAY-PEYROULA (Université Grenoble Alpes & CNRS)
- M. Benoît ROUX (Université de Chicago)

Remerciements

Je tiens tout d'abord à remercier les membres du jury qui ont gentiment accepté de prendre de leur temps pour évaluer ce travail. Merci à Annick, Nathalie et Tom pour avoir sans hésiter accepté le fastidieux travail de rapporteur. Merci à Eva, Bruno, Benoît et Christophe de me faire l'amitié de participer à ce jury.

Je souhaite ici adresser un immense merci à mon épouse, Nathalie, et à nos trois merveilleux enfants, Solal, Tybalt et Gaspard qui m'apportent tous les jours l'amour, le soutien et l'énergie qui me permettent de vivre sereinement et pleinement ma vie de chercheur.

Ma carrière scientifique a été émaillée de nombreuses rencontres dont beaucoup se sont transformées en amitiés, l'écriture de ces quelques lignes me rappelle à quel point celles-ci ont été et sont toujours importantes dans la construction de ma vie scientifique en particulier et de ma vie en générale.

Personne ne sera surpris de me voir commencer par faire part de la profonde affection que j'ai pour Christophe avec qui je chemine depuis mes premiers pas dans la recherche. Je mesure tous les jours ce que m'apporte cette amitié emprunte de respect, de bienveillance et de complicité.

Je suis extrêmement reconnaissant envers Claude et Francesco avec qui j'ai beaucoup appris au cours de ma thèse et de mon post-doctorat et qui ont sans aucun doute façonné le socle du chercheur que je suis devenu. Je dois également énormément à ma rencontre avec Eva autour du transporteur mitochondrial de l'ADP/ATP, celle-ci a été particulièrement importante dans le développement de ma trajectoire scientifique et dans les rencontres que j'ai pu faire par la suite.

J'attache énormément d'importance aux collaborations et amitiés avec Bruno, Hugues, Paul, Joanna, Benoît faites de respect et de confiance, je me sens extrêmement privilégié de pouvoir partager tant de choses avec vous. Je remercie également sincèrement les nombreuses personnes avec qui j'ai pu nouer des collaborations et des amitiés qui m'ont tour à tour enrichies : Mounir, Wensheng, Antonio, Elise, Klaus, Javier, János, Hyungjun, Edmund, Manuella, Michel, Winfried, Andreea, Stéphanie, Irina. Je remercie très chaleureusement tous les étudiant(e)s et post-doctorant(e)s avec qui j'ai eu la chance de travailler : Fabien, Eva-Maria, Daniel, Joël, Axel, Jeff, Nicole, Hong, Flo, Zlatomir, Emma, Margaret, Elodie, Charlotte, Cécilia, Hugo. Ils sont pour beaucoup dans les travaux que je présente dans ce mémoire et je les en remercie très sincèrement.

Préambule

Ce mémoire d'habilitation commence par un chapitre dédié à mon curriculum vitae où sont notamment recensées les listes de mes publications, des encadrements et des contrats dont j'ai pu bénéficier. Les chapitres suivants présentent de manière synthétique et non-exhaustive les activités de recherche que j'ai pu développer au cours de ma carrière. Le mémoire se termine par un dernier court chapitre ou j'aborde quelques projets de recherche.

J'espère que le lecteur me pardonnera l'aspect "collection" de ce manuscrit imposé par une procrastination sans aucun doute exagérée dans la préparation de mon habilitation. J'ai choisi dans ce mémoire de présenter de manière synthétique différents sujets de recherche que j'ai développés sous la forme d'un court résumé en français enrichi de publications sur le même thème qui permettront j'espère de se faire une idée générale de mon activité de recherche.

Après une courte notice contextualisant chronologiquement et thématiquement l'ensemble de mes activités de recherche, la présentation de mes travaux est organisée en 3 chapitres :

- Développements méthodologiques
- Étude de protéines membranaires
- Étude des processus de lésions/réparation de l'ADN.

Ce mémoire ne se termine pas la présentation de quelques projets de recherche dont deux font l'objet d'un financement ANR.

TABLES DES MATIERES

Chapitre 1

I. Curriculum Vitae

Chargé de recherche au CNRS, section 13 (depuis octobre 2003),

Né le 13 novembre 1974 à Montluçon. Marié, 3 enfants

Adresse : Laboratoire de Physique et Chimie Théorique.
Unité mixte de recherche CNRS/UdL No 7019, Université de Lorraine B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex

Téléphone : 03-72-74-50-76, **e-mail** : Francois.Dehez@univ-lorraine.fr

Domaine de recherche

Après une thèse de doctorat dédiée au développement des méthodes mixtes mécanique quantique/mécanique moléculaire et à leur application à la simulation de la séparation des isotopes du lithium par des éthers-couronnes, j'ai effectué un stage postdoctoral dans le département de chimie de l'université de Bologne. J'ai utilisé mon savoir-faire dans le domaine des simulations statistiques pour rationaliser le fonctionnement de nanomachines innovantes telles que les interrupteurs (rotaxanes) et les moteurs moléculaires (caténanes). Après mon entrée au CNRS en octobre 2003, j'ai travaillé au sein de l'équipe de Dynamique des Assemblages Membranaires du Laboratoire Structure et Réactivité des Systèmes Moléculaires Complexes (UMR N° 7565). J'exerce maintenant mon activité de recherche dans le Laboratoire de Physique et Chimie Théoriques (UMR 7019) au sein de l'axe Biophysique et Biochimie. J'ai dédié une partie de mon activité de recherche au développement de nouveaux modèles intermoléculaires simples et précis prenant en compte explicitement les effets de polarisation dans les systèmes complexes. Au centre de mes travaux se trouve maintenant la compréhension des phénomènes de transport au travers des membranes biologiques. Une des questions principales à laquelle j'essaie de répondre est comment s'effectue les changements de conformations au sein des protéines responsables de la communication cellulaire ? J'utilise également les approches théoriques pour rationaliser le sens physiologique des structures de protéines membranaires obtenues par cristallographie aux rayons X, par RMN ou par Cryo-microscopie électronique (Cryo-EM) dans des environnements natifs (nano-disques lipidiques) et non-natifs (détergents). Je m'intéresse particulièrement au fonctionnement des protéines de la membrane interne de la mitochondrie comme le transporteur de l'ADP/ATP (AAC), le transporteur UCP2 et le complexe I de la chaine respiratoire ainsi qu'aux canaux ioniques comme les canaux pentamériques activés par des ligands. Enfin une partie de mon activité est dédiée à l'étude des lésions complexes de l'ADN et à leur reconnaissance par des protéines de la machinerie cellulaire de réparation. Mon domaine d'expertise est la dynamique moléculaire et plus spécifiquement les simulations d'évènements rares et les calculs d'énergie libre ainsi que

l'électrophysiologie computationnelle, techniques que je couple aux informations structurales obtenues par les techniques de cristallographie aux rayons X, de RMN et microscopie électronique.

Parcours

2007- : CNRS – Chargé de Recherches Classe Normale

2003-2007 : CNRS - Chargé de Recherches 2ème classe

2002-2003 : Bourse de recherche européenne Marie Curie Training Mobility Research (F. Zerbetto)
Modélisation de nano-switchs et de nano-moteurs moléculaires. Laboratoire Simulazioni e Modelli Molecolari, Université de Bologne, Italie

1999-2002 : Thèse : Contrat de formation par la recherche du CEA (C. Millot)
Modélisation de la séparation isotopique du lithium par des éthers-couronnes. Équipe de Chimie et Biochimie Théoriques, UMR CNRS/UHP N° 7565

1997-1998 : Diplôme d'étude approfondie, Chimie informatique et théorique, Nancy

Publications et Bibliométrie

71 publications (ACL), 1 chapitre de livre, 2 articles en révision, ~30 communications, ~30 conférences et séminaires invités

h-index : 28 ; Citations : 3,250 (source Google Scholar)

II. Responsabilités

Responsabilités Administratives

- **2006-2017** Membre nommé puis élu du conseil du laboratoire SRSMC
- **2007-2012** Membre nommé du comité hygiène et sécurité du laboratoire SRSMC.
- **Depuis 2015.** Co-responsable de la plateforme SMEC (Simulation, Modélisation et Extraction de connaissance) du CPER 2015-2020, programme IT2MP.
- **Depuis 2018** Membre nommé du conseil de laboratoire, membre du conseil scientifique, membre de la commission informatique du LPCT.
- **2018-2020** Responsable de l'axe Biochimie et Biophysique du LPCT (4 chercheurs CNRS et 4 enseignant-chercheurs)
- **Depuis 2020** Membre du directoire du Mésocentre de Calcul Lorrain Explor.
- **Depuis 2020** Directeur adjoint du LPCT

Activités d'enseignement

- **M2 Chimie informatique et théorique** Université Henri Poincaré : Mécanique moléculaire, Structure 3D des protéines. 15 h/an, 2004/2005 ; 2006/2007
- **3ème année** - **ENSIC/ENSAIA/ENSMN** (filière bioinformatique) : Dynamique moléculaire, Méthode Monte Carlo. 24 h/an, 2004/2005 ; 2005/2006 ; 2006/2007
- **2ème année** - **Filières produits/procédés – ENSIC** : Champs de Force, Introduction à la dynamique moléculaire. 14 h/an, 2004/2005 ; 2006/2007
- **2ème année** - **Filières produits/procédés – ENSIC** : Champs de Force, Introduction à la dynamique moléculaire. 11 h/an, 2007/2008 ; 2008/2009 ; 2009/2010 ; 2010/2011 ; 2011/2012 Technologies de l'information et de la communication.
- **2ème année** - **Filière industries alimentaires – ENSAIA** : Mécanique moléculaire, Dynamique moléculaire. 18 h/an, 2005/2006 ; 2006/2007 ; 2007/2008
- **M2 Sciences de la Vie / Doctorat** Université de Bordeaux 2 : Introduction à la modélisation des protéines membranaires. 15 h/an, 2009/2010 ; 2010/2011 ; 2011/2012 ; 2012/2013 ; 2013/2014
- **L1 Biologie** Université de Lorraine : Technologies de l'information et de la communication. 40 h/an, 2009/2010 ; 2010/2011 ; 2011/2012 ; 2013/2014 ; 2014/2015 ; 2015/2016
- **L1 PACES** Université de Lorraine : Évaluation des méthodes d'analyse appliquées aux sciences de la vie et de la santé. 24 h/an, 2010/2011 ; 2011/2012 ; 2012/2013

III. Encadrements

Licences

- Sixiao Peng - L3 Nancy - *Étude de l'interaction de résidus aromatiques avec des nanotubes de carbone* **2007**
- Hadil Al Tabchi - L3 Nancy - *Interaction de la phospholipase A2 avec un modèle de membrane biologique* **2008**
- Elodie Durand – L3 Biotechnologie du médicament, Nancy - *Simulations de dynamique moléculaire de lésions complexes de l'ADN et reconnaissance par les protéines de réparation* **2015**

Masters

- Julien Schmesser – M1 Sciences-Physiques, Nancy – *Simulations de dynamique moléculaire de photo-lésions dans l'ADN* **2015**

- Gilles Touchagues – M2 ENS de Lyon – *Simulations de dynamique moléculaire de lésions tandem oxydatives dans l'ADN et énergies libres de liaisons ADN/photo sensibilisateurs* **2014**
- Cécilia Hognon – M1 Master BioIngénierie et Médicaments, Nancy. *Étude biophysique du complexe protéine Hbb/ADN par dynamique moléculaire* **2017**
- Elodie Gonis – M1 Master Biochimie, Biologie Moléculaire et Régulations cellulaires, Nancy. *Le comportement et le mode de reconnaissance de la protéine HU de Staphylococcus aureus vis-à-vis de l'ADN lésé par exposition aux rayonnements ultraviolets* **2017**
- Athéna Pactor – M1 - Master Sciences du Vivant - Bioingénierie et Médicaments, Nancy. *Étude de l'interaction du peptide RGDC avec les intégrines par dynamique moléculaire* **2018**
- Julien Colin – M1 Master Sciences-Physiques, Nancy. *Simulations de dynamique moléculaire gros-grains de l'ADN* **2019**
- Charlotte Etienne – M2 Master Sciences du Vivant - Biotechnologies - Génie Cellulaire, Nancy. *Effets du cholestérol sur la dynamique fonctionnelle du récepteur pentamérique 5-HT3* **2019**
- Florence Szczepaniak – Stage de 4ème année ENS, Lyon – *Reconnaissance synaptique* **2019-2020**
- Morgane Viloria - M2 Master Sciences du Vivant - Biotechnologies - Ingénierie Moléculaire, Nancy. *Sumoylation* **2021**

Doctorants

- Latifa Chebil (ENSAIA) – co-encadrement **Juillet-Novembre 2006**. *Modélisation de flavonoïdes en milieu organique.*
- Fabien Archambault (Nancy)– co-encadrement (80 %) **Octobre 2007 – Octobre 2010**. *Développement de modèles polarisables pour la simulation de systèmes d'intérêt biologique.*
- Daniel Bonhenry (Nancy) – co-encadrement (50 %) **Octobre 2010 – Octobre 2013.** *Vectorisation de peptides et de fonctionnelles à visées thérapeutiques à travers les membranes cellulaires. Étude par modélisation moléculaire.*
- Zlatomir Todorov (Grenoble IBS) – co-encadrement (50 %) **Octobre 2015 – Octobre 2019**. *Modélisation de canaux ioniques KIR*
- Hong Zhang (Université de Nankai) – co-encadrement (30 %) **Novembre 2018 - Novembre 2019**. *Modélisation de switchs moléculaires en interaction avec l'ADN.*
- Florence Szczepaniak (Nancy) – Cotutelle Université of Chicago (B. Roux) **Depuis octobre 2020** – *Reconnaissance synaptique*
- Margaret Blazhynska (Nancy) – co-encadrement (50 %) **Depuis octobre 2020** – *Fonctionnement des Protéases*

Post-doctorants

- Dr. Eva-Maria Krammer – **Décembre 2008 – Septembre 2010.** *Étude du transporteur mitochondrial de l'ADP/ATP.*
- Dr. Axel Bidon-Chanal – **Juillet 2010 – Juillet 2012**. *Étude du transporteur mitochondrial de l'ADP/ATP.*
- Dr. Jeffrey Comer – **Janvier 2013 – Juillet 2014**. *Étude de la protéine découplante UCP2*
- Dr. Nicole Holzmann – **Janvier 2015 – Décembre 2015.** *Étude de la protéine p7 du virus de l'hépatite C.*
- Dr. Vijayray Ramadoss – **Janvier 2015 – Décembre 2015.** *Développement d'un outil automatique d'alanine scaning.*
- Dr. Joel Montalvo Acosta – **Janvier 2019 - Janvier 2021.** *Modélisation de protéines fluorescentes pour la microscopie super résolution.*
- Dr. Hankuyl Lee – **Juin 2019- Juin 2020.** *Edition génomique. Mécanismes de reconnaissance de séquences spécifiques d'ADN par la protéine TALE.*

IV. Participation à des jurys de thèse, comités de sélection

Jurys de thèse

- Membre du jury de thèse de Latifa Chebil (ENSAIA, Nancy) **2009**
- Membre du jury de thèse de Fabien Archambault (SRSMC, Grenoble) **2010**
- Membre du jury de thèse de Pankaj Panwar (IBS, Grenoble) **2012**
- Membre du jury de thèse de Daniel Bonhenry (SRSMC, Nancy) **2013**
- Membre du jury de thèse de Mathilde Piel (IBPC, Paris) **2019**
- Membre du jury de thèse de Zlatomir Todorov (IBS, Grenoble) **2019**
- Membre du jury de Thèse de Audrey Deyawe (LPCT, Nancy) **2021**

Comités de suivi de Thèse

- Membre du comité de suivi de thèse de Pankaj Panwar (Institut de biologie structurale, Grenoble)
- Membre du comité de suivi de thèse de Aleksandra Woznicka (Institut de biologie structurale, Grenoble)

- Membre du comité de suivi de thèse de Zlatomir Todorov (Institut de biologie structurale, Grenoble)
- Membre du comité de suivi de thèse de Yannis Brun (LPCT, Nancy)
- Membre du comité de suivi de thèse de Hicham Jabraoui (LPCT, Nancy)
- Membre du comité de suivi de thèse de Tom Miclot (LCPT, Nancy / STEBICEF, Palerme)

Comités de séléction

- Membre de la commission de spécialistes Université Henri Poincaré, Nancy, Section 31 au titre de l'année **2008**.
- Membre d'une commission de spécialistes CHAIRE d'excellence Université de Strasbourg / CNRS année **2010**
- Membre d'une commission de spécialistes CHAIRE d'excellence Nancy université/CNRS année **2010**
- Membre d'une commission de spécialistes CHAIRE d'excellence Nancy université/CNRS année **2011**
- Membre d'une commission de spécialistes CHAIRE d'excellence Nancy université/CNRS année **2012**
- Membre d'une commission de sélection Maitre de Conférence, Université de Lille année **2017**

V. Principales participations à contrats

- ANR Programme Blanc **2006-2010**. *Compréhension du transport ADP/ATP dans les membranes mitochondriales, entre expérience et théorie (TRANS-MIT)* (*participant)*
- CPER Modélisations, Informations et Systèmes numériques (MISN) **2007-2013**. *Modélisation interactive de systèmes biologiques complexes en réalité virtuelle immersive (MODIS)* (participant)
- ANR Programme Blanc **2010–2013**. *Mécanismes moléculaires de transport dans les mitochondries* (MIT- 2M) (participant)
- ANR Programme Blanc **2010–2013**. *Outils pour l'électromanipulation de l'intérieur des cellules biologiques (INTCELL)* (participant)
- Projet PICS – France/USA **2009-2011**. *Simulation de systèmes biologiques avec un champ de force polarisable* (participant)
- Projet Egide – France/Slovénie – Proteus **2010-2011**. *Electroporation of Planar Lipid Bilayers; Experimental and Molecular Dynamics Approach* (participant)

- Projet BQR - Postdoc Nancy université **2010-2011**. *Modélisation du transporteur de l'ADP/ATP* (porteur)
- Laboratoire International Associé – CNRS/Université d'Illinois à Urbana- Champaign **2013-2017, 2017-2021, 2021-2025** (participant)
- Programme PHC France-Chine Cai Yuanpei **2014-2015.** *New nanocarriers for anticancer therapies* (participant)
- Programme PEPS Bio-Math-Info du CNRS **2013-2014**. *Outils théoriques pour la description de grandes transitions conformationnelles dans les protéines membranaires* (participant)
- Projet DDRT/CPER/communauté urbaine du Grand Nancy/FEDER **2013-2015**. *Développements méthodologiques innovants pour la simulation haute performance de système biologique complexes* (co-porteur)
- Projet PRACE preparatory access **2014.** *Large scale MD simulations of supermitochodrial complexes* (porteur)
- Projet PRACE 9th regular call **2014-2015** (13 millions CPU hours). *MitoComp: Large-Scale molecular-dynamics simulations of the entire Mitochondrial Complex I* (porteur)
- Projet CPER **2015-2020** – Programme Innovations Technologiques, Modélisation et Médecine Personnalisée. *Simulation, Modélisation, Extraction de Connaissances* (co-porteur)
- ANR Programme Blanc **2019-2023** – *Targeting ClpP proteases: From dynamics and function to allosteric modulation by drugs* (participant)
- Programme PHC STAR France-Corée **2017-2018** – *Genome Editing with TALE: Insights from all-atom simulations* (porteur)
- Programme LUE Mirabelle+, Université de Lorraine **2018-2019** – *Conception d'inhibiteur du transfert de résistances aux agents microbiens : bio-ingénierie assistée par des approches virtuelles et numériques et appliquée à une relaxase d'élément conjugatif intégratif* (co-porteur)
- Projet Fond Européen de Développement Régional (FEDER) **2018-2021** – *Simulation numérique, bioinformatique structurale et microbiologie moléculaire en synergie pour la lutte contre la dissémination des résistances aux antibiotiques* (co-porteur)
- CNRS Modélisation du vivant - AAP 2021 - **2021-2022** Structural insight into bacterial chromatin assembly by cryo-electron microscopy/tomography and molecular modelling (co-porteur)
- ANR Programme Blanc **2021-2025** – *Molecular mechanism of insecticide action at pentameric receptorsargeting ClpP proteases: From dynamics and function to allosteric modulation by drugs* (responsable scientifique)

VI. Principales collaborations externes

- Prof. Javier Luque, Departament de Fisicoquimica, Faculta de Farmàcia, Universitat de Barcelona, Spain.
- Prof. Klaus Schulten, Theoretical and Computational Biophysics Group, University of Illinois at Urbana- Champaign, USA.
- Prof. Eva Pebay-Peyroula, Dr. Hugues Nury, Dr. Michel Vivaudou. Laboratoire des protéines membranaires, Institut de Biologie Structurale, Grenoble.
- Dr. Bruno Miroux, Dr. Manuella Zoonens. Laboratoire de Biologie Physico-Chimique des Protéines Membranaire, Institut de Biologie Physico-Chimique, Paris.
- Dr. Latifa Chebil, Prof. Mohamed Ghoul, Laboratoire d'ingénierie des biomolécules, Nancy Université, Vandœuvre-lès-Nancy.
- Dr. Joanna Timmins, Institut de Biologie Structurale, Équipe Lésions et Réparation de l'ADN, Grenoble
- Prof. Élise Dumont, Laboratoire de Chimie, École Normale Supérieure, Lyon
- Dr. Paul Schanda, Institut de Biologie Structurale - Biomolecular NMR group, Grenoble
- Dr. Winfried Weissenhorn, Institut de Biologie Structurale - Groupe Entrée et bourgeonnement des virus enveloppés, Grenoble
- Dr. Edmund Kunji, Medical Research Council - Mitochondrial Biology Unit, Cambridge, UK
- Dr. Nicole Zitzmann, Dr. Jason Schnell, Biochemistry Department, University of Oxford, UK
- Prof. Wensheng Cai, Nankai University, Tianjin, China
- Dr. Hyungjun Kim, Department EEWS, Korean Advanced Institute of Science and Technology (KAIST), Daejeon, Corée du Sud
- Prof. Benoit Roux, Biochemistry Department, University of Chicago.

VII. Principales mobilités

- **2006** - Programme HPC-europa : Séjour de 1 mois dans le groupe du Prof. Zerbetto (département de Chimie) à l'université de Bologne (Italie).
- **2007, 2008, 2009, 2013** – 4 séjours de 1 semaine dans le groupe du Prof. Schulten à l'université de l'Illinois à d'Urbana-Champaign (Illinois).
- **2008, 2013** – 2 séjours de 1 semaine dans le groupe du Prof. Luque (Faculté de pharmacie) à l'université de Barcelone (Espagne).

- **2010** – 1 séjour de 1 semaine dans le groupe du Prof. Miklavcic (Faculty of electrical engineering) à l'université de Ljubljana (Slovenie).
- **2012** – 1 séjour de 2 semaines dans le groupe du Prof. Cai (Department of Chemistry) à l'université de Nankai à Tianjin (Chine).
- **2014 (x2), 2015, 2016, 2017, 2018** – 6 séjours de 1 semaine dans le groupe du Prof. Cai (Department of Chemistry) à l'université de Nankai à Tianjin (Chine).
- **2015 (x2), 2016, 2017, 2018** – 5 séjours de 1 semaine dans le groupe du Prof. Hyungjun Kim, KAIST, Daejeon, Corée du sud.
- **2019** – 1 séjour de 1 semaine dans le groupe du Prof. Benoit Roux, University of Chicago, Chicago, USA.

VIII. Liste des publications

VIII.1. Revues à Comité de Lecture

1. Acosta, J. M., Kunji, E., Ruprecht, J., Dehez, F.* & Chipot, C.* Structure, substrate binding, symmetry of the mitochondrial ADP/ATP carrier in its matrix-open state. *Biophys. J.*, **in revision** (2021)
2. Zarkadas, E. Pebay-Peyroula, E., Thompson, M., Schoehn, G., Uchański, T., Steyaert, J., Chipot, C., Dehez, F. Baenziger, J. E. & Nury, H. *Neuron*, **in revision** (2021)
3. Fu, H., Chen, H., Blazhynska, M., Goulard Coderc De Lacam, E., Szczepaniak, F., Pavlova, A., Shao, X., Gumbart, J. C., Dehez, F., Roux, B., Cai, W. & Chipot, C. Accurate determination of protein:ligand standard binding free energies from molecular dynamics simulations, *Nature Protocols*, **accepted** (2021)
4. Caillat, C., Guilligay, D., Torralba, J., Friedrich, N., Nieva, J., Trkola, A., Chipot, C., Dehez, F. & Weissenhorn, W. Structure of HIV-1 gp41 with its membrane anchors targeted by neutralizing antibodies. *Elife*, **10**, e65005 (2021)
5. Zarkadas, E., Zhang, H., Cai, W., Effantin, G., Perot, J., Neyton, J., Chipot, C., Schoehn, G., Dehez, F.* & Nury, H.* The binding of palonosetron and other antiemetic drugs to the serotonin 5-HT3 receptor. *Structure*, **28**, 1131-1140 (2020)
6. Gupta, C., Khaniya, U., Chan, C.K., Dehez, F., Shekhar, M., Gunner, M.R., Sazanov, L., Chipot, C. & Singharoy, A. Charge transfer and chemo-mechanical coupling in respiratory complex I. *J. Am. Chem. Soc.*, **142**, 9220-9230 (2020).
7. Zhang, H., Shao, X., Dehez, F., Cai, W. & Chipot, C. Modulation of membrane permeability by carbon dioxide. *J. Comput. Chem.*, *41*, 421– 426 (2020)
8. Marazzi, M., Gattuso, H., Giussani, A. Zhang, H., Navarrete-Miguel, M., Chipot, C., Cai, W., Roca-Sanjuán, D., Dehez, F.* & Monari*, A. Induced Night Vision by Singlet-Oxygen-Mediated Activation of Rhodopsin. *J. Phys. Chem. Lett.* **10**, 7133-7140 (2019)

9. Hognon, C., Garaude, S., Timmins, J., Chipot, C., Dehez, F.* & Monari, A.* Molecular bases of DNA packaging in bacteria revealed by all-atom molecular dynamics simulations: The case of histone-like proteins in borrelia burgdorferi. *J. Phys. Chem. Lett.*, **10**, 7200-7207, (2019)

10. Terenzi, A., Gattuso, H., Spinello, A., Keppler, B. K., Chipot, C., Dehez, F. Barone, G. & Monari, A. Targeting G-quadruplexes with Organic Dyes: Chelerythrine–DNA Binding Elucidated by Combining Molecular Modeling and Optical Spectroscopy. *Antioxidants*, **8**, 472 (2019)

11. Pinto, D., Fenwick, C., Caillat, C., Silacci, C., Guseva, S., Dehez, F. Chipot, C., Barbieri, S., Minola, A. Jarrossay, D., Tomaras, G. D., Shen, X., Riva, A., Tarkowski, M., Schwartz, O., Bruel, T., Dufloo, J., Seaman, M. S., Montefiori, D. C., Lanzavecchia, A., Corti, D., Pantaleo, G. & Weissenhorn, W. Structural Basis for Broad HIV-1 Neutralization by a Novel MPER-Specific Human Broadly Neutralizing. *Cell Host & Microbe*, **26**, 623-637 (2019)

12. Felix, J., Weinhäupl, K., Chipot, C., Dehez, F., Hessel, A., Gauto, D. F., Morlot, C., Abian, O., Gutsche, I., Velazquez-Campoy, A., Schanda, P. & Fraga, H. Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Sci. Adv.* **5** (9), eaaw3818 (2019).

13. Vucko, T., Pétry, N., Dehez, F., Lambert, A., Monari, A., Lakomy, C., Lacolley, P., Regnault, V., Collet, C., Karcher, G., Pellegrini-Moïse, N. & Lamandé-Langle, S. C-glyco"RGD" as αIIbβ3 and αvβ integrin ligands for imaging applications: Synthesis, in vitro evaluation and molecular modeling. *Bioorg. Med. Chem.* **27**, 4101–4109 (2019).

14. Pieri, E., Ledentu, V., Sahlin, M., Dehez, F., Olivucci, M. & Ferré, N. CpHMD-Then-QM/MM Identification of the Amino Acids Responsible for the Anabaena Sensory Rhodopsin pH-Dependent Electronic Absorption Spectrum. *J. Chem. Theory Comput.* **15**, 4535–4546 (2019).

15. Chagot, M.-E., Morais, R. D. S., Dermouche, S., Lefebvre, D., Manival, X., Chipot, C., Dehez, F.* & Quinternet, M.* Binding properties of the quaternary assembly protein SPAG1. *Biochem. J.* **476**, 1679–1694 (2019).

16. Oestringer, B. P., Bolivar, J. H., Claridge, J. K., Almanea, L., Chipot, C., Dehez, F., Holzmann, N., Schnell, J. R. & Zitzmann, N. Hepatitis C virus sequence divergence preserves p7 viroporin structural and dynamic features. *Sci. Rep.* **9**, 1–10 (2019).

17. Sarre, A., Stelter, M., Rollo, F., De Bonis, S., Seck, A., Hognon, C., Ravanat, J.-L., Monari, A., Dehez, F., Moe, E. & Timmins, J. The three Endonuclease III variants of Deinococcus radiodurans possess distinct and complementary DNA repair activities. *DNA Repair* **78**, 45–59 (2019).

18. Girardon, M., Parant, S., Monari, A., Dehez, F., Chipot, C., Rogalska, E., Canilho, N. & Pasc, A. Triggering tautomerization of curcumin by confinement into liposomes. *ChemPhotoChem,* **3**, 1034-1041 (2019).

19. Zhang, H., Fu, H., Shao, X., Dehez, F., Chipot, C. & Cai, W. Changes in Microenvironment Modulate the B- to A-DNA Transition. *J. Chem. Inf. Model.* **59**, 2324–2330 (2019).

20. Lee, H., Dehez, F., Chipot, C., Lim, H.-K. & Kim, H. Enthalpy–Entropy Interplay in π-Stacking Interaction of Benzene Dimer in Water. *J. Chem. Theory Comput.* **15**, 1538–1545 (2019).

21. Rupakheti, C. R., Roux, B., Dehez, F.* & Chipot, C.* Modeling induction phenomena in amino acid cation–π interactions. *Theor. Chem. Acc.* **137**, 174 (2018).

22. Polovinkin, L., Hassaine, G., Perot, J., Neumann, E., Jensen, A. A., Lefebvre, S. N., Corringer, P.-J., Neyton, J., Chipot, C., Dehez, F., Schoehn, G. & Nury, H. Conformational transitions of the serotonin 5-HT 3 receptor. *Nature* **563**, 275–279 (2018).

23. Oestringer, B. P., Bolivar, J. H., Hensen, M., Claridge, J. K., Chipot, C., Dehez, F., Holzmann, N., Zitzmann, N. & Schnell, J. R. Re-evaluating the p7 viroporin structure. *Nature* **562**, E8–E18 (2018).

24. Zhang, H., Fu, H., Shao, X., Chipot, C., Monari, A., Dehez, F.* & Cai, W.* Conformational changes of DNA induced by a trans-azobenzene derivative via non-covalent interactions. *Phys. Chem. Chem. Phys.* **20**, 22645–22651 (2018).

25. Kurauskas, V., Hessel, A., Dehez, F., Chipot, C., Bersch, B. & Schanda, P. Dynamics and interactions of AAC3 in DPC are not functionally relevant. *Nat. Struct. Mol. Biol.* **25**, 745–747 (2018).

26. Chipot, C., Dehez, F., Schnell, J. R., Zitzmann, N., Pebay-Peyroula, E., Catoire, L. J., Miroux, B., Kunji, E. R. S., Veglia, G., Cross, T. A. & Schanda, P. Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies. *Chem. Rev.* **118**, 3559–3607 (2018).

27. Bonhenry, D., Dehez, F. & Tarek, M. Effects of hydration on the protonation state of a lysine analog crossing a phospholipid bilayer – insights from molecular dynamics and free-energy calculations. *Phys. Chem. Chem. Phys.* **20**, 9101–9107 (2018).

28. Kurauskas, V., Hessel, A., Ma, P., Lunetti, P., Weinhäupl, K., Imbert, L., Brutscher, B., King, M. S., Sounier, R., Dolce, V., Kunji, E. R. S., Capobianco, L., Chipot, C., Dehez, F., * Bersch, B.* & Schanda, P.* How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine. *J. Phys. Chem. Lett.* **9**, 933–938 (2018).

29. Zhang, H., Gattuso, H., Dumont, E., Cai, W., Monari, A., Chipot, C. & Dehez, F.* Accurate Estimation of the Standard Binding Free Energy of Netropsin with DNA. *Molecules* **23**, 228 (2018).

30. Dehez, F., Schanda, P., King, M. S., Kunji, E. R. S. & Chipot, C. Mitochondrial ADP/ATP Carrier in Dodecylphosphocholine Binds Cardiolipins with Non-native Affinity. *Biophys. J.* **113**, 2311–2315 (2017).

31. Gattuso, H., Marazzi, M., Dehez, F. & Monari, A. Deciphering the photosensitization mechanisms of hypericin towards biological membranes. *Phys. Chem. Chem. Phys.* **19**, 23187–23193 (2017).

32. Dehez, F.*, Gattuso, H., Bignon, E., Morell, C., Dumont, E.* & Monari, A.* Conformational polymorphism or structural invariance in DNA photoinduced lesions: implications for repair rates. *Nucleic Acids Res.* **45**, 3654–3662 (2017).

33. Pinheiro, S., Soteras, I., Gelpí, J. L., Dehez, F., Chipot, C., Luque, F. J. & Curutchet, C. Structural and energetic study of cation–π–cation interactions in proteins. *Phys. Chem. Chem. Phys.* **19**, 9849–9861 (2017).

34. Pavel, I.-A., Prazeres, S. F., Montalvo, G., García Ruiz, C., Nicolas, V., Celzard, A., Dehez, F., Canabady-Rochelle, L., Canilho, N. & Pasc, A. Effect of Meso vs Macro Size of Hierarchical Porous Silica on the Adsorption and Activity of Immobilized β-Galactosidase. *Langmuir* **33**, 3333–3340 (2017).

35. Gattuso, H., Dumont, E., Chipot, C., Monari, A. & Dehez, F.* Thermodynamics of DNA: sensitizer recognition. Characterizing binding motifs with all-atom simulations. *Phys. Chem. Chem. Phys.* **18**, 33180–33186 (2016).

36. Holzmann, N., Chipot, C., Penin, F. & Dehez, F.* Assessing the physiological relevance of alternate architectures of the p7 protein of hepatitis C virus in different environments. *Bioorg. Med. Chem.* **24**, 4920–4927 (2016).

37. Bignon, E., Gattuso, H., Morell, C., Dehez, F., Georgakilas, A. G., Monari, A. & Dumont, E. Correlation of bistranded clustered abasic DNA lesion processing with structural and dynamic DNA helix distortion. *Nucleic Acids Res.* **44**, 8588–8599 (2016).

38. Gattuso, H., Durand, E., Bignon, E., Morell, C., Georgakilas, A. G., Dumont, E., Chipot, C., Dehez*, F. & Monari*, A. Repair Rate of Clustered Abasic DNA Lesions by Human Endonuclease: Molecular Bases of Sequence Specificity. *J. Phys. Chem. Lett.* **7**, 3760–3765 (2016).

39. Ramadoss, V., Dehez, F. & Chipot, C. AlaScan: A Graphical User Interface for Alanine Scanning Free-Energy Calculations. *J. Chem. Inf. Model.* **56**, 1122–1126 (2016).

40. Dehez, F.*, Delemotte, L., Kramar, P., Miklavčič, D. & Tarek, M.* Evidence of Conducting Hydrophobic Nanopores Across Membranes in Response to an Electric Field. *J. Phys. Chem. C* **118**, 6752–6757 (2014).

41. Marquardt, R., Hénin, J., Dehez, F. & Chipot, C. Dynamiques moléculaires quantiques et classiques. *Actual. Chim.* **382–383**, 56–62 (2014).

42. Comer, J., Dehez, F., Cai, W. & Chipot, C. Water Conduction through a Peptide Nanotube. *J. Phys. Chem. C* **117**, 26797–26803 (2013).

43. Bonhenry, D., Tarek, M.* & Dehez, F.* Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane. *J. Chem. Theory Comput.* **9**, 5675–5684 (2013).

44. Bidon-Chanal, A., Krammer, E.-M., Blot, D., Pebay-Peyroula, E., Chipot, C., Ravaud, S.* & Dehez, F.* How Do Membrane Transporters Sense pH? The Case of the Mitochondrial ADP–ATP Carrier. *J. Phys. Chem. Lett.* **4**, 3787–3791 (2013).

45. Polak, A., Bonhenry, D., Dehez, F., Kramar, P., Miklavčič, D. & Tarek, M. On the Electroporation Thresholds of Lipid Bilayers: Molecular Dynamics Simulation Investigations. *J. Membr. Biol.* **246**, 843–850 (2013).

46. Zoonens, M., Comer, J., Masscheleyn, S., Pebay-Peyroula, E., Chipot, C., Miroux, B.* & Dehez, F.* Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2. *J. Am. Chem. Soc.* **135**, 15174–15182 (2013).

47. Radzimanowski, J., Dehez, F., Round, A., Bidon-Chanal, A., McSweeney, S. & Timmins, J. An 'open' structure of the RecOR complex supports ssDNA binding within the core of the complex. *Nucleic Acids Res.* **41**, 7972–7986 (2013).

48. Mifsud, J., Ravaud, S., Krammer, E.-M., Chipot, C., Kunji, E. R. S.*, Pebay-Peyroula, E. & Dehez, F.* The substrate specificity of the human ADP/ATP carrier AAC1. *Mol. Membr. Biol.* **30**, 160–168 (2013).

49. Liu, P., Dehez, F., Cai, W. & Chipot, C. A Toolkit for the Analysis of Free-Energy Perturbation Calculations. *J. Chem. Theory Comput.* **8**, 2606–2616 (2012).

50. Ravaud, S., Bidon-Chanal, A., Blesneac, I., Machillot, P., Juillan-Binard, C., Dehez, F., Chipot, C. & Pebay-Peyroula, E. Impaired Transport of Nucleotides in a Mitochondrial Carrier Explains Severe Human Genetic Diseases. *ACS Chem. Biol.* **7**, 1164–1169 (2012).

51. Luque, F. J., Dehez, F., Chipot, C. & Orozco, M. Polarization effects in molecular interactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 844–854 (2011).

52. Chebil, L.*, Chipot, C., Archambault, F., Humeau, C., Engasser, J. M., Ghoul, M. & Dehez, F.* Solubilities Inferred from the Combination of Experiment and Simulation. Case Study of Quercetin in a Variety of Solvents. *J. Phys. Chem. B* **114**, 12308–12313 (2010).

53. Krammer, E.-M., Ravaud, S., Dehez, F., Frelet-Barrand, A., Pebay-Peyroula, E. & Chipot, C. High-Chloride Concentrations Abolish the Binding of Adenine Nucleotides in the Mitochondrial ADP/ATP Carrier Family. *Biophys. J.* **97**, L25–L27 (2009).

54. Archambault, F., Chipot, C., Soteras, I., Luque, F. J., Schulten, K. & Dehez, F.* Polarizable Intermolecular Potentials for Water and Benzene Interacting with Halide and Metal Ions. *J. Chem. Theory Comput.* **5**, 3022–3031 (2009).

55. De Oliveira, E. B., Humeau, C., Chebil, L., Maia, E. R., Dehez, F., Maigret, B., Ghoul, M. & Engasser, J.-M. A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by Candida antarctica lipase B. *J. Mol. Catal. B Enzym.* **59**, 96–105 (2009).

56. Dehez, F., Pebay-Peyroula, E. & Chipot, C. Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel. *J. Am. Chem. Soc.* **130**, 12725–12733 (2008).

57. Dehez, F., Archambault, F., Gutiérrez, I. S., Luque, F. J. & Chipot, C. An ab initio strategy for handling induction phenomena in metal ion complexes. *Mol. Phys.* **106**, 1685–1696 (2008).

58. Delemotte, L., Dehez, F., Treptow, W. & Tarek, M. Modeling Membranes under a Transmembrane Potential. *J. Phys. Chem. B* **112**, 5547–5550 (2008).

59. Soteras, I., Curutchet, C., Bidon-Chanal, A., Dehez, F., Ángyán, J. G., Orozco, M., Chipot, C. & Luque, F. J. Derivation of Distributed Models of Atomic Polarizability for Molecular Simulations. *J. Chem. Theory Comput.* **3**, 1901–1913 (2007).

60. Dehez, F., Ángyán, J. G., Gutiérrez, I. S., Luque, F. J., Schulten, K. & Chipot, C. Modeling Induction Phenomena in Intermolecular Interactions with an Ab Initio Force Field. *J. Chem. Theory Comput.* **3**, 1914–1926 (2007).

61. Dehez, F., Tarek, M. & Chipot, C. Energetics of Ion Transport in a Peptide Nanotube. *J. Phys. Chem. B* **111**, 10633–10635 (2007).

62. Chebil, L., Humeau, C., Anthoni, J., Dehez, F., Engasser, J.-M. & Ghoul, M. Solubility of Flavonoids in Organic Solvents. *J. Chem. Eng. Data* **52**, 1552–1556 (2007).

63. Trebbi, B., Dehez, F., Fowler, P. W. & Zerbetto, F. Favorable Entropy of Aromatic Clusters in Thermophilic Proteins. *J. Phys. Chem. B* **109**, 18184–18188 (2005).

64. Dehez, F., Martins-Costa, M. T. C., Rinaldi, D. & Millot, C. Long-range electrostatic interactions in hybrid quantum and molecular mechanical dynamics using a lattice summation approach. *J. Chem. Phys.* **122**, 234503 (2005).

65. Leigh, D. A., Wong, J. K. Y., Dehez, F. & Zerbetto, F. Unidirectional rotation in a mechanically interlocked molecular rotor. *Nature* **424**, 174–179 (2003).

66. Bottari, G., Dehez, F., Leigh, D. A., Nash, P. J., Pérez, E. M., Wong, J. K. Y. & Zerbetto, F. Entropy-Driven Translational Isomerism: A Tristable Molecular Shuttle. *Angew. Chem. Int. Ed.* **42**, 5886–5889 (2003).

67. Ángyán, J. G., Chipot, C., Dehez, F., Hättig, C., Jansen, G. & Millot, C. OPEP: A tool for the optimal partitioning of electric properties. *J. Comput. Chem.* **24**, 997–1008 (2003).

68. Altieri, A., Bottari, G., Dehez, F., Leigh, D. A., Wong, J. K. Y. & Zerbetto, F. Remarkable Positional Discrimination in Bistable Light- and Heat-Switchable Hydrogen-Bonded Molecular Shuttles. *Angew. Chem. Int. Ed.* **42**, 2296–2300 (2003).

69. Millot, C. & Dehez, F. Modelling of the isotopic exchange reaction of lithium ion complexed by crown-ethers in vacuum and in ethanol. *J. Comput. Methods Sci. Eng.* **2**, 451–456 (2002).

70. Chipot, C., Dehez, F., Ángyán, J., Millot, C., Orozco, M. & Luque, F. J. Alternative Approaches for the Calculation of Induction Energies: Characterization, Effectiveness, and Pitfalls. *J. Phys. Chem. A* **105**, 11505–11514 (2001).

71. Dehez, F., Chipot, C., Millot, C. & Ángyán, J. G. Fast and accurate determination of induction energies: reduction of topologically distributed polarizability models. *Chem. Phys. Lett.* **338**, 180–188 (2001).

72. Celebi, N., Ángyán, J. G., Dehez, F., Millot, C. & Chipot, C. Distributed polarizabilities derived from induction energies: A finite perturbation approach. *J. Chem. Phys.* **112**, 2709–2717 (2000).

73. Dehez, F., Soetens, J.-C., Chipot, C., Ángyán, J. G. & Millot, C. Determination of Distributed Polarizabilities from a Statistical Analysis of Induction Energies. *J. Phys. Chem. A* **104**, 1293–1303 (2000).

VIII.2. Chapitre de Livre

74. Dehez, F. Molecular simulation: a virtual microscope in the toolbox of integrated structural biology. in *From Molecules to Living Organisms: An Interplay Between Biology and Physics* (eds. Pebay-Peyroula, E., Nury, H., Parcy, F., Ruigrok, R. W. H., Ziegler, C. & Cugliandolo, L. F.) 413–436 (Oxford University Press, 2016).

VIII.3. Autres Publications

75. Bonhenry, D., Dehez, F., Tarek, M.
Surface d'énergie libre de transfert d'acides aminés au travers de membranes biologiques, Rapport Grands Challenges GENCI IDRIS **2013**

VIII.4. Logiciels

1. **OPEP :** Optimally Partitioned Electric Properties (référence 59)
   Ángyán, J. G., Chipot, C., Dehez , F., Hättig, C., Jansen, G., Millot, C.
   Programme distribué sous licence GPL à l'adresse :
   http://www.edam.uhp-nancy.fr/Opep/
2. **VMD plugins :**
   ParseFEP **(**référence 41) : analyse de calculs FEP (Free-Energy Perturbation)
   AlaScan (référence 31) : Alanine scanning in silico automatisé

Contextualisation chronologique de mes travaux de recherches

J'ai commencé à m'intéresser à la modélisation moléculaire lors de mon année de maîtrise de Sciences-Physiques à Clermont-Ferrand. C'est assez naturellement, après une discussion téléphonique avec Jean-Louis Rivail qui dirigeait alors le Laboratoire de Chimie Théorique à Nancy, que je me suis inscrit en DEA de Chimie Informatique et Théorique. Mon appétence pour les simulations moléculaires classiques, découvertes en maîtrise, m'a conduit à effectuer mon stage sous la direction de Claude Millot. Claude venait de développer un code de dynamique moléculaire capable de traiter des champs de force explicitement polarisables. Nous avons commencé, en collaboration avec Christophe Chipot et János Angyan, à développer des approches permettant de mettre au point des modèles simples et précis de polarisabilités distribuées à partir de calculs ab initio. J'ai poursuivi ce travail au cours de ma thèse avec Claude et également avec Christophe au cours de mes premières années de Chargé de Recherches.

Financée par un contrat de formation par la recherche du Commissariat à l'Energie Atomique, le sujet principal d'étude de ma thèse était la séparation isotopique du lithium par des éthers-couronnes greffés sur silice. J'ai utilisé les calculs quantiques et les calculs hybrides mécanique quantique/mécanique moléculaire (MQ/MM) pour déterminer les facteurs de séparation isotopique entre ^{7}Li et ^{6}Li. J'ai montré que le phénomène d'enrichissement isotopique était non seulement lié au complexant mais également, et de manière très importante, au solvant. J'ai profité de cette thèse pour implémenter une méthode de somme de réseaux pour traiter les interactions électrostatiques à longue distance dans les méthodes MQ/MM.

J'ai effectué par la suite un stage postdoctoral dans le groupe de Francesco Zerbetto au département de Chimie de l'Université de Bologne. Je me suis impliqué sur le sujet passionnant des moteurs et switch moléculaires que j'ai étudié avec les simulations de dynamique moléculaire classiques. Francesco collaborait activement avec le groupe de David Leigh, un chimiste organicien connu pour avoir inventé des méthodes fondamentales de contrôle de la dynamique et de l'enchevêtrement d'assemblages moléculaires synthétiques. Sur la base d'architectures moléculaires imbriquées mécaniquement (rotaxanes et caténanes), le groupe de Dave venait de synthétiser un moteur moléculaire dont nous avons, grâce au calcul, démontré l'unidirectionalité, une première dans le domaine.

Mon entrée au CNRS en 2003 s'est faite sur la base d'un projet biochimique, l'étude de protéines membranaires interfaciales, dans l'équipe de dynamique des Assemblages membranaires qui venait

juste d'être créée à Nancy. Les moyens de calculs massivement parallèles n'étaient pas encore très développés à l'échelle des laboratoires, et je me suis beaucoup investi alors dans le montage d'un cluster composé de nœuds de calculs "sans disque" qui a permis par la suite à l'équipe d'être plus performante dans la mise en œuvre de simulations de dynamique moléculaire de "gros" systèmes.

Après m'être familiarisé avec la formidable complexité du monde biologique en général et de la dynamique des protéines en particulier, j'ai commencé à travailler avec Mounir Tarek sur les membranes en proposant une méthode d'électrophysiologie computationnelle, permettant de générer un potentiel transmembranaire par l'intermédiaire de deux bains de concentrations ioniques différentes en s'affranchissant de la périodicité du système. Dans la lignée de ces travaux, j'ai également commencé à modéliser un ensemble de processus de transport membranaire (canaux ioniques synthétiques, perméation passive, électroporation, ...)

Ma rencontre avec Eva Pebay-Peyroula a clairement marqué un tournant dans ma recherche non seulement sur le plan des systèmes auxquels je m'étais jusque-là intéressés mais aussi sur le plan des collaborations que j'ai pu développer par la suite. Eva travaillait depuis plusieurs années sur le transporteur mitochondrial de l'ADP/ATP, après en avoir résolu la structure dans sa conformation cytoplasmique en 2003. Très vite, le potentiel des simulations est apparu pour étudier cette protéine responsable des échanges en ADP et ATP dans la mitochondrie. En 2008 nous avons publié avec Eva et Christophe notre premier travail commun décrivant notamment l'association spontanée de l'ADP guidée par la signature électrostatique en forme d'entonnoir du transporteur. Ce premier travail a été suivi par plusieurs études décrivant le fonctionnement de ce transporteur en synergie avec les expériences menées à Grenoble, en particulier par Stéphanie Ravaud, Maîtresse de Conférences dans l'équipe d'Eva. Il a également été à l'origine de plusieurs collaborations avec un des spécialistes des transporteurs mitochondriaux, Edmund Kunji à Cambridge, jusqu'à la soumission d'un article cette année sur la dynamique de AAC dans sa conformation ouverte côté matriciel dont la structure a été résolue par le groupe d'Edmund en 2019 (16 ans après le travail d'Eva !).

C'est par l'intermédiaire d'Eva et du GDR "Protéines Membranaires" que j'ai rencontré Bruno Miroux avec qui j'ai commencé très naturellement à travailler sur d'autres transporteurs mitochondriaux, les protéines découplantes de la chaîne respiratoire (UCP). L'idée de départ était, en ayant recours aux modèles par homologie, de modéliser la dynamique d'UCP comme nous l'avions fait pour AAC, mais la structure controversée de UCP2 publiée en 2011 par le Groupe de James Chou à Harvard Medical School nous amené à un tout autre développement. En 2013, nous avons publié un article démontrant que le détergent utilisé pour solubiliser UCP2 avait clairement altéré le repliement de la protéine, produisant une structure sans aucun sens physiologique. Cet

article a marqué le début d'un "combat" sur l'effet délétère du DPC sur les structures des protéines membranaires et sur le sens physiologique des structures résolues par le groupe de Chou par RMN (phospholamban, HCV p7, transporteur ADP/ATP, ...). Cette aventure et son succès doivent beaucoup à la collaboration avec Paul Schanda dont le travail en RMN sur le comportement du transporteur mitochondrial AAC en DPC et la synergie avec nos travaux de simulation ont permis d'établir de manière formelle l'altération de la structure des protéines solubilisées par ce détergent.

Au travers de la collaboration avec Eva, j'ai commencé à fréquenter l'Institut de Biologie Structurale (IBS) à Grenoble ce qui a été le point de départ d'autres interactions. J'ai commencé à travailler en collaboration avec Joanna Timmins sur un complexe impliqué dans la réparation homologue de l'ADN réparation mettant à profit les simulations de dynamique moléculaire pour rationaliser des données de SAXS (Small-Angle X-Ray Scattering) et proposer un modèle structural pour décrire la reconnaissance d'un simple brin d'ADN par le complexe de réparation. Depuis, j'ai étendu ces travaux à d'autres questions liées aux processus d'endommagement et de réparation de l'ADN en collaboration avec Joanna et également avec Antonio Monari, enseignant-chercheur au LPCT.

Lors de mes premières visites dans le groupe d'Eva, j'avais eu la chance et le plaisir d'échanger avec Hugues Nury, alors étudiant en thèse travaillant à la résolution de nouvelles structures de AAC. Quelques années après et de brillants résultats obtenus sur la structure des canaux pentamériques, Hugues a été recruté à l'IBS en tant que Chargé de Recherches au CNRS. J'ai été très vite embarqué dans la thématique de ces récepteurs-canaux activés par des ligands sur lesquels je travaille activement aujourd'hui. Sur la base des nombreuses structures résolues par le groupe de Hugues, j'utilise les simulations pour caractériser les états fonctionnels des différentes conformations capturées par cryo-microscopie électronique. Un des objectifs est de caractériser au niveau moléculaire la dynamique d'activation allostérique de ces canaux et sa modulation par des pharmacophores.

Enfin, mon activité s'inscrit depuis plusieurs années dans Laboratoire International Associé entre le CNRS et l'université d'Urbana-Champaign dirigé par Christophe Chipot. Clairement ce LIA et les initiatives qui l'ont précédé (PICS) sont des éléments particulièrement importants dans le développement et la structuration de mon activité de recherche. Au-delà, des opportunités importantes de collaborations avec de fantastiques groupes d'expérimentateurs, j'ai pu développer des projets méthodologiques notamment avec le groupe de Klaus Schulten au début de l'aventure et maintenant avec le groupe de Benoit Roux à Chicago. Le périmètre scientifique et humain ainsi que la structuration de ce laboratoire "sans mur" sont indéniablement des éléments cruciaux pour la réussite des projets que je développe.

Chapitre 2 - Développements méthodologiques

I. Modèles polarisables

La plupart des champs de force standards utilisés pour modéliser les interactions moléculaires reposent sur l'approximation additive de paires dans laquelle les phénomènes de polarisation sont pris en compte de manière moyenne par l'intermédiaire d'un potentiel effectif.[1] Dans de nombreux systèmes, notamment biomoléculaires, les phénomènes d'induction peuvent toutefois être prépondérants et nécessiter l'utilisation de potentiel intermoléculaires explicitement polarisables, comme c'est le cas par exemple lorsque l'on veut décrire correctement les interactions π-cations.[2,3] Trois grandes familles de méthodes peuvent être citées pour traiter les effets à n-corps de manière explicite en mécanique moléculaire : les polarisabilités distribuées,[4,5] les charges fluctuantes[6,7] et les oscillateurs de drude.[8]

Depuis plusieurs années, une partie de mon activité de recherche est dédiée au développement de potentiels intermoléculaires polarisables efficaces pour modéliser des systèmes d'intérêt biologique. Les travaux réalisés dans le passé ont permis de mettre au point une méthodologie robuste pour dériver des modèles de polarisabilités distribuées, pour des molécules ou des fragments, à partir d'une cartographie de l'énergie d'induction *ab initio*. Dans la pratique, une grille de points est générée autour de la molécule d'intérêt et l'énergie d'interaction de celle-ci avec une charge polarisante localisée sur l'un des vertex est évaluée.[9–12] L'énergie d'induction est alors modélisée par un développement multipolaire de polarisabilités atomiques, tronqué à un ordre donné. L'ensemble des outils développés a été rassemblé dans un logiciel distribué sous licence GPL, le programme OPEP (*Optimal Partitioning of Electric Properties*).[13] Un effort particulier a été réalisé pour inclure ce modèle dans un potentiel intermoléculaire complet avec pour objectif de disposer d'un potentiel précis (à longue et courte distance) tout en gardant une formulation suffisamment compacte permettant de maintenir des temps de calcul raisonnables. La stratégie repose sur la mise au point de modèles simples composés de charges atomiques reproduisant la signature électrostatique et de polarisabilités de flux de charges et dipolaires isotropes décrivant les propriétés de réponse d'une molécule donnée, stratégie que nous avons éprouvée pour modéliser des interactions π-cation par exemple.[14,15] Par ailleurs, pour dériver le modèle d'interaction moléculaire, la théorie SAPT (*Symmetry Adapted Perturbation Theory*)[16] a été utilisée comme référence permettant d'accéder aux diverses contributions à l'énergie intermoléculaire (équation 1) et d'assurer la robustesse de notre formulation empirique.

$$\Delta \mathscr{U}_{\text{tot}}^{\text{SAPT}} = U_{\text{ele}} + U_{\text{ind}} + U_{\text{exc}} + U_{\text{dis}} + U_{\text{exc-ind}} + U_{\text{exc-dis}} + \delta\text{HF} \qquad (1)$$

Je me suis intéressé au cas de l'interaction d'un ion calcium avec différents groupements chimiques représentatifs de l'environnement de cet ion au site de catalyse d'une métallo-enzyme, l'eau, le formaldéhyde et l'ion formate (figure 1).[17]

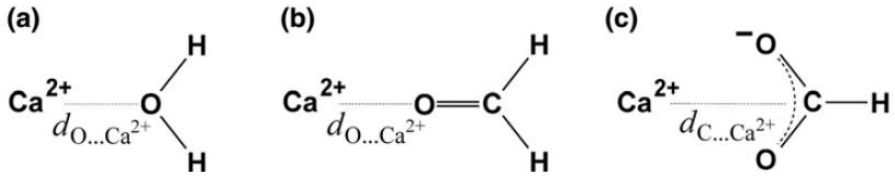


Figure 1. Interaction du cation divalent Ca2+ avec (a) l'eau, (b) le formaldéhyde et (c) l'ion formate. L'approche du cation vers les 3 ligands est considérée le long de l'axe C2.

Nous avons montré qu'un champ de force additif de paires était incapable de reproduire simultanément l'ensemble de ces interactions. En particulier le traitement identique de l'interaction de Ca^{2+} avec l'eau et avec le groupement carbonyle ne permet pas d'assurer la cohésion de la structure du site actif de l'enzyme sur des temps de simulations de l'ordre de quelques nanosecondes. Pour chacune des molécules nous avons dérivé les modèles polarisables tels que décrits précédemment. Nous avons ensuite calculé les énergies d'interaction pour différentes distances intermoléculaires. Une fonction de damping générique a été utilisée pour modéliser la contribution d'échange-induction et ainsi prévenir le problème de polarisation catastrophe rencontré aux courtes distances de séparation. Nous avons par ailleurs employé un potentiel de Lennard-Jones pour décrire les interactions non-électrostatiques. Ces travaux ont montré qu'il était inenvisageable d'injecter directement nos modèles polarisables dans un champ de force additif mais qu'il fallait au contraire dériver l'ensemble des paramètres restants, en particulier ceux relatifs aux interactions de van der Waals. Si cet ajustement est fait de manière adéquate, le potentiel résultant permet de reproduire simultanément toutes les énergies d'interaction moléculaires du calcium avec l'eau, le formaldéhyde et l'ion formate, aux distances d'équilibre (tableau 1) et au-delà.

	$\Delta\mathcal{U}_{ele}$		$\Delta\mathcal{U}^{b}_{ind}$		$\Delta\mathcal{U}^{c}_{vdW}$			$\Delta\mathcal{U}_{tot}$		
	SAPT	MM	SAPT	MM	SAPT	MM	δHF^{d}	$MP2^{e}$	SAPT	MM^{f}
$Ca^{2+}\ldots H_2O$	−51.7	−51.8	−22.1	−23.0	19.8	22.7 (3.6)	−0.8	−53.4	−54.7	**−52.2** (−64.6)
$Ca^{2+}\ldots H_2CO$	−48.4	−45.1	−37.5	−31.3	24.2	15.3 (4.8)	−1.6	−61.2	−63.3	**−61.1** (−71.6)
$Ca^{2+}\ldots HCOO^{-}$	−308.7	−290.0	−62.2	−58.3	65.8	62.8 (1.4)	−0.5	−297.8	−305.6	**−285.6** (−301.6)

Tableau 1. Comparaison des contributions à l'énergie d'interaction au minimum de la courbe quantique obtenues par notre modèle et par un développement SAPT au niveau MP2/6-311++G(d,p). Toutes les valeurs sont en kcal/mol. Les valeurs entre parenthèses sont celles calculées en utilisant directement les paramètres de van der Waals pris dans le champ de force CHARMM27.[18]

Ces modèles prometteurs souffrent toutefois d'une incapacité à reproduire simultanément l'énergie d'interaction aux distances d'équilibre et le comportement de la courbe d'interaction en deçà de ces distances. Pour combler cette lacune, nous avons revisité nos modèles en utilisant comme systèmes d'étude divers cations et anions en interaction avec une molécule d'eau et une molécule de benzène

(cations uniquement) ainsi que le dimère de l'eau (figure 2).[19] Pour la molécule d'eau nous avons développé un modèle électrostatique ou la charge initialement portée par l'oxygène est légèrement décalée vers les hydrogènes le long de l'axe bissecteur de la molécule (modèle 4 points). Ce choix permet d'améliorer très sensiblement la reproduction du potentiel électrostatique de la molécule d'eau ainsi que les énergies d'interaction avec les anions. Par ailleurs nous avons remplacé la fonction de damping générique par la fonction de Tang et Toennies[20] que nous avons paramétrée à partir de la décomposition SAPT des énergies d'interaction.

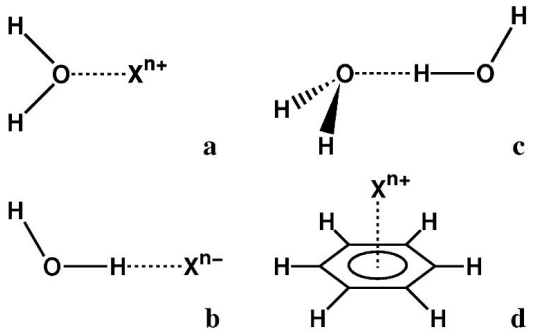


Figure 2. Interaction de l'eau avec un cation le long de l'axe C_2. (b) Interaction de l'eau avec un anion le long de la liaison O-H. (c) Interaction de 2 molécules d'eau formant une liaison hydrogène. (d) Interaction du benzène avec un cation le long de l'axe C_6.

Le reste de l'énergie d'interaction a été modélisé par le potentiel de Buckingham qui à l'inverse de la fonction de Lennard-Jones permet de reproduire non seulement les énergies d'interaction au minimum de la courbe (tableau 2) mais également parfaitement l'ensemble du profil d'interaction quelle que soit la distance considérée (voir figure 3 pour le cas du complexe eau-sodium).

$\Delta\mathscr{U}_{tot,min}$	Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	F^-	Cl^-	Br^-	I^-	H_2O
H_2O	−32.63	−22.91	−16.85	−78.80	−52.98	−22.38	−13.07	−11.14	−9.25	−3.60
	−32.69	**−23.12**	**−16.91**	**−79.85**	**−52.75**	**−22.46**	**−13.00**	**−11.08**	**−9.22**	**−3.60**
benzene	−37.45	−23.41	−17.57	−114.17	−74.25					
	−37.43	**−23.70**	**−17.60**	**−116.41**	**−74.23**					

Tableau 2. Énergies d'interaction (en kcal/mol) pour l'ensemble des systèmes de la figure 2 obtenues par notre modèle et par un calcul quantique de référence (en gras). Les valeurs sont prises au minimum de la courbe quantique. Il est à noter que les distances d'équilibre déterminées par notre modèle coïncident principalement avec les références quantiques.

L'ensemble de ces travaux nous a permis de mettre au point un jeu de règles permettant de dériver des potentiels intermoléculaires complets ayant comme qualités la précision et la compacité.[21]

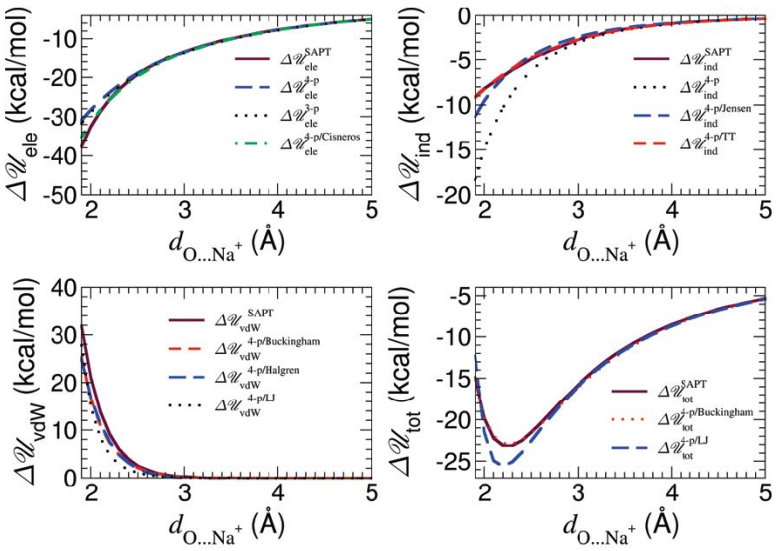


Figure 3. Courbes d'interactions pour le système sodium-eau. La référence SAPT est en trait plein. Trois modèles électrostatiques sont présents : un modèle 3 points, un modèle 4 points avec et sans fonction corrigeant les effets de pénétration. Trois modèles d'induction sont présentés utilisant soit la fonction générique de damping soit la fonction de Tang et Toennies, soit n'utilisant aucun damping. Trois fonctions de van der Waals (lennard-Jones, Halgren, Buckingham) sont utilisées pour ajuster le reste de l'énergie d'interaction auquel a été retranchée la contribution électrostatique du modèle 4 points sans pénétration et la contribution d'induction utilisant la fonction de Tang et Toennies.

Ci-après sont reproduits 2 articles illustrant mes travaux dans le domaine des modèles polarisables :

- Archambault, F., Chipot, C., Soteras, I., Luque, F. J., Schulten, K. & Dehez, F.* Polarizable Intermolecular Potentials for Water and Benzene Interacting with Halide and Metal Ions. *J. Chem. Theory Comput.* **5**, 3022–3031 (2009)
- Rupakheti, C. R., Roux, B., Dehez, F.* & Chipot, C.* Modeling induction phenomena in amino acid cation–π interactions. *Theor. Chem. Acc.* **137**, 174 (2018)

JCTC Journal of Chemical Theory and Computation

Polarizable Intermolecular Potentials for Water and Benzene Interacting with Halide and Metal Ions

Fabien Archambault,[†] Christophe Chipot,*[,†,‡,§] Ignacio Soteras,[¶] F. Javier Luque,[¶] Klaus Schulten,[‡] and François Dehez*[,†]

Équipe de Dynamique des Assemblages Membranaires, UMR 7565, Nancy Université, BP 239, 54506 Vandoeuvre-lès-nancy Cedex, France, Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, and Departament de Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Avgda. Diagonal 643, Barcelona 08028, Spain

Received August 10, 2009

Abstract: A complete derivation of polarizable intermolecular potentials based on high-level, gas-phase quantum-mechanical calculations is proposed. The importance of appreciable accuracy together with inherent simplicity represents a significant endeavor when enhancement of existing force fields for biological systems is sought. Toward this end, symmetry-adapted perturbation theory can provide an expansion of the total interaction energy into physically meaningful, for example, electrostatic, induction and van der Waals terms. Each contribution can be readily compared with its counterpart in classical force fields. Since the complexity of the different intermolecular terms cannot be fully embraced using a minimalist description, it is necessary to resort to polyvalent expressions capable of encapsulating overlooked contributions from the quantum-mechanical expansion. This choice results in consistent force field components that reflect the underlying physical principles of the phenomena. This simplified potential energy function is detailed, and definitive guidelines are drawn. As a proof of concept, the methodology is illustrated through a series of test cases that include the interaction of water and benzene with halide and metal ions. In each case considered, the total energy is reproduced accurately over a range of biologically relevant distances.

1. Introduction

The development of increasingly scalable software and greater access to massively parallel computer architectures allow longer numerical simulations to be performed on continuously more complex molecular assemblies. This, in turn, has brought the investigation of biologically relevant systems over realistic time scales within reach, for example, atomistic simulations of a membrane for hundred of microseconds[1] or folding of small proteins.[2] In the latter examples, the use of a pairwise additive force field, for example, Amber,[3] Charmm,[4] Gromos,[5] or Opls-AA,[6] provides in general a reasonable reproduction of experimental data as long as polarization effects are not dominant.

When, on the contrary, induction effects can no longer be neglected, or vary significantly in the course of the simulation, nonpolarizable potential energy functions rapidly attain the inherent limits of their validity. Two recent studies[7,8] endeavored to describe the folding of a small protein, the human Pin1 WW domain, by means of the Charmm macromolecular force field. Whereas this potential energy function was able to preserve the β-sheet-like native structure over the 200 ns time scale, it failed to fold within 10 μs the

* Corresponding author e-mail: chipot@ks.uiuc.edu (C.C.); Francois.Dehez@edam.uhp-nancy.fr (F.D.).

[†] Nancy Université.

[‡] University of Illinois at Urbana−Champaign.

[¶] Universitat de Barcelona.

[§] On leave from Équipe de Dynamique des Assemblages Membranaires, UMR 7565, Nancy Université, BP 239, 54506 Vandoeuvre-lès-nancy Cedex, France.

10.1021/ct9004189 CCC: $40.75
Published on Web 10/06/2009

protein chain from an extended motif into the native conformation. It was demonstrated that this failure is due to force-field accuracy rather than insufficient sampling—whether or not neglect of polarization phenomena is responsible for protein misfolding remains at this stage a matter of debate, which would require additional investigation to find a definitive answer. Metalloproteins constitute another class of biologically relevant systems where induction effects need to be accounted for. In this case, a metal ion polarizes the environment, which, in turn, alters the binding hierarchy[9] of neighboring chelating agents. In particular,[10,11] when using a standard nonpolarizable potential energy function, the stability of the active site is perturbed on account of the exaggerated interaction of the ion with water, compared to that with either the vicinal carbonyl or carboxylate groups. Ion channels also represent important biological systems, in which small ions permeate through narrow pores, strongly polarizing the walls of the latter[12] and modulate as a result the ionic selectivity.[13] The formation of crystals can also be governed by polarization effects.[14−17] When decomposing the total interaction energy into electrostatic, induction, and van der Waals contributions, the importance of the induction energy suggests that it may play a predominant role in the stability of the crystal and, hence, should not be ignored.

The need for taking into account polarization phenomena explicitly in molecular-dynamics simulations constitutes an ongoing effort that can be traced back to the 1980s.[18] One of the grand challenges for polarizable force fields is the derivation of functions that are suitable for biological simulations. Toward this end, a number of routes have been explored to include polarization effects explicitly. A first route is based on Drude oscillators,[19] which describe electronic induction by the movement of a fictitious mass bonded to polarizable atoms by means of a stiff spring. Another approach, referred to as fluctuating charges,[20−23] consists of changing the charge of the atoms along the simulation. These charges are adjusted according to the principle of partial equalization of orbital electronegativities.[24] A third path, namely, multipole expansions,[21,25−30] uses distributed polarizabilities to model induction energies following a self-consistent procedure to determine the induced dipoles. These methods have been probed in a variety of applications ranging from liquid simulations[31−34] to simulations of DNA.[35] They have been combined with molecular-dynamics simulations of increasing complexity, such as that of polarizable membranes—see for instance ref 36. In the latter reference, Drude oscillators model induction effects and allow a better agreement to be reached between simulation and experiment for the reproduction of the dipole potential that arises at the water−lipid interface.

The main thrust of the present work is to build a polarizable force field that reconciles simplicity and accuracy. The chosen approach relies on the central idea[37] of partitioning the interaction energy into physically meaningful contributions and to determine the associated parameters using quantum-mechanical (QM) methods. A logical roadmap is followed for the construction of a consistent potential energy function.[9,38−43] This consistency imposes a de novo derivation of gas-phase atomic charges. The induction contribution is then derived from ab initio calculations and modified by ad hoc damping functions. Last, van der Waals parameters are determined using an appropriate mathematical expression that guarantees the faithful description of the overall intermolecular interactions at play. Despite the precision of such a description, a number of key elements are still missing for a faithful reproduction of nonbonded interaction energies between molecules over a wide range of distances. In the reproduction of the total quantum-mechanical interaction energy, the rudimentary description needed for large-scale simulations of biological systems is bound to failure if high accuracy is sought, as several terms, like nonmultipolar contributions, are not described explicitly. Choices have, therefore, to be made to obtain a model of sufficient accuracy, while remaining concise. Appropriate functions ought to be determined to include implicitly the missing contributions. In this study, a complete set of definitive guidelines is proposed to derive compact polarizable intermolecular potentials. The approach is illustrated through a series of test cases that include the interaction of water and benzene with halide and metal ions.

2. Theoretical Underpinnings

The QM calculations reported here rely on symmetry-adapted perturbation theory[44,45] (SAPT), which provides a formal, rigorous framework for the expansion of the total interaction energy into meaningful contributions utilized as a reference. In particular, use is made in this work of the SAPT2 expansion, henceforth called SAPT for simplicity, which accounts for correlation effects at the second-order Møller−Plesset[46] (MP2) level of theory. Within SAPT, the interaction energy is decomposed into the individual contributions shown in eq 1. In the latter expression, the first four terms stand for the electrostatic (U_{ele}), induction (U_{ind}), exchange (U_{exc}), and dispersion (U_{dis}) components. In addition, the SAPT expansion embraces two terms corresponding to the coupling between exchange and induction ($U_{exc\text{-}ind}$) and between exchange and dispersion ($U_{exc\text{-}dis}$). Finally, the last term (δHF) accounts for a collection of higher-order induction and exchange−induction terms. Altogether, the contributions are written

$$\Delta\mathscr{U}_{tot}^{SAPT} = U_{ele} + U_{ind} + U_{exc} + U_{dis} + U_{exc\text{-}ind} + U_{exc\text{-}dis} + \delta HF \quad (1)$$

The stated energy components cannot, when taken separately, be studied in a straightforward fashion with a classical decomposition.[47] For a comparison on a one-to-one basis, certain individual contributions need to be summed, as has been done previously by several authors.[9,48−51] It is then possible to cast the description into three potential energy functions, namely, for electrostatic, induction, and van der Waals contributions.

The electrostatic component ($\Delta\mathscr{U}_{ele}^{SAPT}$) corresponds to the first-order contribution to the electrostatic energy. The pure induction term given by SAPT ought to be added to the exchange−induction term, as well as to higher-order induction and exchange−induction terms to encompass the full induction potential ($\Delta\mathscr{U}_{ind}^{SAPT}$). The last contribution, referred

to as the van der Waals potential ($\Delta\mathscr{U}_{\mathrm{vdW}}^{\mathrm{SAPT}}$), consists of the sum of exchange and dispersion terms.

As was put forth previously, a molecular-mechanical expansion can mimic the above-summed QM energies. This includes an electrostatic potential that relies upon atomic charges, an induction term associated with an appropriate damping function, and an ad hoc van der Waals function.[9]

The first order contribution, of electrostatic nature, is modeled by a set of atom-centered point charges and is a rudimentary approach to describe the molecular electrostatic potential. While this description is simple, it has proven[52] to reproduce under most circumstances the target quantity with an appreciable accuracy. Determination of atomic charges is achieved[53] by means of regular grids surrounding the molecule of interest, over which the QM electrostatic potential is mapped. Penetration effects, which represent an important, short-range component of the electrostatic interaction, are evidently absent from the classical, Coulomb potential.[54] The incompleteness of the model is rooted in the computation of the molecular electrostatic potential outside the sphere of convergence.[55] To a large extent, these effects stem from the overlap of electron densities of interacting compounds, which cannot be accounted for by a localized atomic description of the electrostatic potential.[56] Ad hoc functions[57−60] have been shown to mimic the respective short-range interactions. Here, use will only be made of an ansatz inferred from the formulation of Freitag et al. by Cisneros et al.[60] It holds the advantage of being derived explicitly for atomic charges, and compared to other functions, it offers a simpler expression and parametrization. Charges are, therefore, modified as stated in eq 3 in ref 60.

As detailed above, different routes can be chosen for the description of induction phenomena in the classical simulations of biological systems. Here, polarization is modeled by means of distributed polarizabilities[61,62] over subsets of atoms. Previous articles[9,43,63] have demonstrated that the use of charge-flow polarizabilities between chemically bonded atoms supplemented by isotropic dipole polarizabilities on heavy atoms is able to recover the anisotropy of the molecular polarizability with both simplicity and effectiveness. Polarizabilities were determined by means of a fitting procedure[41] that relies on three-dimensional maps of the induction energy evaluated quantum-mechanically. As was asserted previously,[9,43,63] the classical expansion of polarization ought to be damped in order to reproduce the destabilizing effect of exchange-induction effects occurring at short intermolecular distances. In the present work, the Jensen et al.[64] and the Tang and Toennies[65,66] (TT) functions have been chosen. In the former expression, the distances in the interaction tensor are modified by a scalar parameter, a, that appears in eq 16 in ref 64. In the latter expression, truncation of the series is done at the sixth order, as suggested by Millot and Stone,[67] and subsequently rewritten at the third order, as proposed by Meredith and Stone.[68] The final expressions for the induction energy utilized herein are

$$\Delta\mathscr{U}_{\mathrm{ind}}^{\mathrm{MM/TT}} = \frac{1}{2}\sum_{A}\sum_{B\neq A}\Delta Q_t^a T_{tu}^{ab} f_3(\beta;r_{ab}) Q_u^b \qquad (2)$$

$$\Delta Q_t^a = -\sum_{B\neq A}\alpha_{tt'}^{aa'} T_{t'u}^{a'b} f_3(\beta;r_{a'b})(Q_u^b + \Delta Q_u^b) \qquad (3)$$

$$f_3(\beta;r_{ab}) = 1 - \exp(-\beta r_{ab})\sum_{k=0}^{3}\frac{(\beta r_{ab})^k}{k!} \qquad (4)$$

where Q_t^a and ΔQ_t^a are, respectively, the permanent moment and the induced moment of rank t located at site a of molecule A, T_{tu}^{ab} is the element of the electrostatic tensor for the interaction of the multipole moments of ranks t and u, found, respectively, at sites a and b, and $f_3(\beta; r_{ab})$ is the truncated Tang and Toennies damping function, in which β is a scalar parameter. The term $\alpha_{tt'}^{aa'}$ is the atomic polarizability, which describes the change in the multipole moment of rank t at site a resulting from the t'th derivative of the potential created by all other moments at site a'. The above expressions are also detailed in eqs 21 and 22 in ref 50.

The last term of the classical intermolecular potential is the van der Waals contribution, which is modeled in the present work by three alternate formulations depending on the r distance. The first one is the standard 6−12 Lennard-Jones (LJ) expression,[69] namely, $\Delta\mathscr{U}_{\mathrm{vdW}}^{\mathrm{MM/LJ}} = \varepsilon\,[(\sigma/r)^{12} - (\sigma/r)^6]$, where the parameters ε and σ correspond respectively to the depth of the potential well and the distance at which the interparticle potential is zero. The second one, used among others in the Amoeba[70] force field, is the Halgren[71] potential, namely, $\Delta\mathscr{U}_{\mathrm{vdW}}^{\mathrm{MM/Halgren}} = \varepsilon\,[1.07R^*/(r + 0.07R^*)]^7 \times [1.12R^{*7}/(r^7 + 0.12R^{*7}) - 2]$, where R^* is the minimum-energy distance. The third one is the exp-6 Buckingham function,[72] namely, $\Delta\mathscr{U}_{\mathrm{vdW}}^{\mathrm{MM/Buckingham}} = \varepsilon\exp(-A_1 r) - (A_2/r)^6$, where A_1 and A_2 are shape parameters.

3. Computational Details

All molecular geometries were optimized using the Gaussian03 suite of programs[73] at the MP2(Full)/6-311++G(2d,2p) level of approximation. Electronic properties were computed at the MP2(Full)/Sadlej level of approximation, considering that the Sadlej[74−77] basis set provides a very good compromise between the number of Gaussian contractions and accuracy.[78] All individual contributions to the interaction energy were determined with the SAPT2008 program of Jeziorski et al.,[44] interfaced to the ATMOL integral and self-consistent-field package.[79] The QM contributions to the total interaction energy serve as a basis of comparison for the parametrization of the classical model.

The atomic charges of water and benzene were derived employing the Opep package,[42] using 2375 and 8070 grid points, respectively. As shown in Table 1, the molecular electrostatic potential of an isolated water molecule is reproduced with a mean error of about 48% when use is made of atom-centered charges only. As water is one of the most important components in computer simulations of condensed phases, in particular, in biological systems, an appreciable error in the description of its electrostatic potential is not acceptable. An improvement of the rudimentary three-point (3-p) charge model is, therefore, proposed. The simplest way to improve the reproduction of the potential consists in adding a fictitious site along the bisector of the molecule, in the spirit of the TIP4P[80,81] model.

Table 1. Models of Net Atomic Charges and Regenerated Multipole Moments for Water and Benzene at the MP2(Full)/Sadlej//MP2(Full)/6-311++G(2d,2p) Level of Approximation[a]

			molecular multipoles		
	point charges		regenerated		MP2(Full)/Sadlej
water (3-p)	Q_{00}^{O}	–0.672	Q_{10}	–0.749	–0.732
	Q_{00}^{H}	–0.336	Q_{20}	–0.189	–0.231
	rmsd	0.990			
	$\Delta\varepsilon$	48.573			
water (4-p)	Q_{00}^{X}	–1.237	Q_{10}	–0.730	–0.732
	Q_{00}^{O}	0.000	Q_{20}	–0.200	–0.231
	Q_{00}^{H}	–0.6185			
	rmsd	0.231			
	$\Delta\varepsilon$	6.258			
benzene (12-p)	Q_{00}^{C}	–0.124	Q_{20}	–5.553	–5.655
	Q_{00}^{H}	0.124	Q_{40}	120.200	163.700
	rmsd	0.219			
	$\Delta\varepsilon$	20.849			

[a] All multipoles are expressed in atomic units. The root-mean-square deviation (rmsd) between the electrostatic potentials determined quantum-mechanically and regenerated from the point charge models is expressed in 10^{-3} au. The corresponding mean error, $\Delta\varepsilon$, is given in percent.

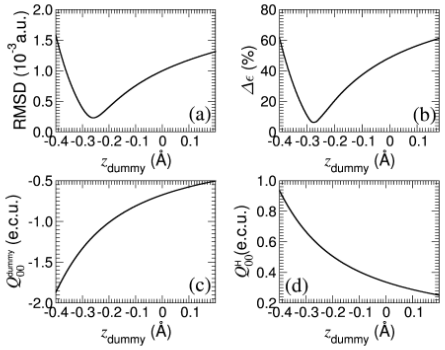


Figure 1. (a) Root-mean-square deviation (RMSD) between the electrostatic potentials determined quantum-mechanically and the point charge model expressed as a function of the position of the dummy atom for water along the C_2 axis. (b) Corresponding mean error ($\Delta\varepsilon$) expressed as a function of the position of the dummy atom for water along the C_2 axis. (c) Value of the atomic charge of the dummy atom expressed as a function of the position of the dummy atom for water along the C_2 axis. (d) Value of the atomic charge of the hydrogen atom expressed as a function of the position of the dummy atom for water along the C_2 axis.

As can be seen in Figure 1, a minimum in the root-mean-square deviation between the QM and the point-charge derived potentials can be found when moving the fictitious site. At this minimum, the additional charge is located 0.276 Å below the oxygen atom (see "Water (4-p)" in Table 1). The resulting enhanced agreement with the target electrostatic potential can be largely explained by the position of the supplementary charge, closer to the center of the molecule. This result is consistent with the nearly isotropic molecular electrostatic potential of water.[54] In the case of benzene, as was shown in a recent article,[82] the use of atom-centered charges is sufficient to model the electrostatic potential with an acceptable accuracy. As can be seen in Table 1, a comparable agreement between the QM potential and that regenerated from point charges is also found here.

Table 2. Models of Distributed Polarizabilities and Regenerated Molecular Polarizabilities of Water and Benzene at the MP2(Full)/Sadlej//MP2(Full)/6-311++G(2d,2p) Level of Approximation[a]

			molecular polarizabilities		
	distributed polarizabilities		regenerated		MP2(Full)/Sadlej
water	$\alpha_{00,00}^{OH}$	–0.808	$\alpha_{10,\,10}$	10.177	9.747
	$\alpha_{1\kappa,1\kappa}^{OO}$	8.1808	$\alpha_{11c,\,11c}$	11.483	10.058
			$\alpha_{11s,\,11s}$	8.180	9.539
	rmsd	0.127			
	$\Delta\varepsilon$	7.000			
benzene	$\alpha_{00,00}^{CC}$	–1.822	$\alpha_{10,\,10}$	47.537	45.121
	$\alpha_{00,00}^{CH}$	–0.280	$\alpha_{11c,\,11c}$	89.089	81.333
	$\alpha_{1\kappa,1\kappa}^{CC}$	7.953	$\alpha_{11s,\,11s}$	89.089	81.333
	rmsd	0.025			
	$\Delta\varepsilon$	2.737			

[a] All polarizabilities are given in atomic units. The root-mean-square deviation (rmsd) between the induction energies determined quantum-mechanically and regenerated from the models of distributed polarizabilities is expressed in 10^{-3} au. The corresponding mean error, $\Delta\varepsilon$, is given in percent.

Table 3. Dipole Polarizability of Ions Computed at the MP2(Full)/Sadlej Level of Approximation

	K^+	F^-	Cl^-	Br^-	I^-
$\alpha_{1\kappa,\,1\kappa}$ (a.u.)	5.174	10.513	30.316	39.074	60.333
$\alpha_{1\kappa,1\kappa}^{exp}$ (a.u.)[a]	5.736	9.313	26.588	35.226	52.704
$\alpha_{1\kappa,1\kappa}^{exp}$ (a.u.)[b]	5.473	10.527	29.760	39.410	60.128

[a] Ref 83. [b] Ref 84.

Penetration of electron clouds results in a short-range contribution, which, for completeness and for attaining the desired accuracy in the reproduction of the intermolecular interaction energy, ought to be included in the electrostatic model. Toward this end, a novel parametrization of the function of Cisneros et al. was probed, wherein the term Ω_{ij} of eq 4 in ref 60 had been modified. Though conceptually simple, the expression proposed to account for penetration effects lacks generality to be of routine practical use for the modeling of large biomolecular assemblies. For this reason, no analytical correction of penetration effects will be introduced explicitly in the intermolecular potential but will instead be incorporated in the van der Waals potential.

The induction energy was mapped employing the Opep package,[42] with grids of 3905 and 3192 points for water and benzene, respectively (see refs 9 and 43 for details). The models of distributed polarizabilities utilized herein rely on a multipole expansion of the molecular polarizability reduced to a combination of zeroth-order charge flow ($\alpha_{00,\,00}$) and first-order isotropic dipole ($\alpha_{1\kappa,\,1\kappa}$) polarizabilities. For clarity, it should be noted that charge-flow polarizabilities are not equivalent to the popular fluctuating charges,[20−23] which are based on a partial equalization of orbital electronegativities and, hence, are very different in spirit. The complete set of parameters fitted simultaneously to quantum-mechanically determined induction energy maps[53] has been shown to reproduce with appreciable accuracy the average molecular polarizability alongside its intrinsic anisotropy. The distributed polarizabilities are reported in Table 2 for water and

benzene and in Table 3 for all the ions considered as polarizable in this work.

Damping parameters are determined with the objective to reproduce the induction energy supplied by SAPT. The first damping function considered is that of Jensen et al., the parameters of which can be found for a number of atom types. For the Tang and Toennies damping function, parameters are given for pair interactions. In some cases, no damping is necessary, due to the importance of the nonmultipolar contribution to the induction energy, as has been documented by Torheyden and Jansen.[49] This contribution to the induction cannot be extracted from the SAPT expansion and, hence, cannot be modeled explicitly in the model proposed here. This missing information can still be added in an appropriately chosen form of the van der Waals potential, compatible with the exponential, repulsive nature of nonmultipolar contributions.

The last contribution to the total energy is the van der Waals potential, which, in the SAPT expansion, is considered as being the sum of exchange and dispersion terms. As outlined previously, both the electrostatic and the induction models adopted here lack a number of contributions that cannot be accounted for to satisfy the stringent criteria of simplicity and tractability. Since the overall model is not sufficiently precise at short distances and when dealing with appreciable nonmultipolar effects, determination of the van der Waals parameters was not carried out on the basis of the exact contribution supplied by the SAPT expansion but using a slightly modified term, which is the total energy ($\Delta\mathscr{U}_{\text{tot}}^{\text{SAPT}}$) minus the electrostatics and induction components of the molecular-mechanical potential. This statement implies that the optimized van der Waals parameters inherently depend on the electrostatic and induction models that have been used here. In other words, modification of either the electrostatic or the polarization term necessarily imposes a requirement that the van der Waals potential be adjusted. Derivation of these parameters was achieved employing a Levenberg-Marquart algorithm.[85] The range of chemically and biologically relevant distances spans from a minimum separation, where the energy is roughly equal to $\Delta\mathscr{U}_{\text{tot, min}}^{\text{SAPT}}$ − 10 kcal/mol, to a quasi-infinite one. The van der Waals contribution was determined on the basis of three alternate functions, that is, Lennard-Jones,[69] Halgren[71] and Buckingham.[72] The exponential nature of the repulsive van der Waals contribution makes the Buckingham exp-6 potential a well-suited candidate to model the van der Waals contribution and short-range interactions, such as nonmultipolar terms.

4. Results and Discussion

4.1. Interaction of Water with Metal Cations. Classical models derived to describe the interaction of cations and anions with water molecules are to a large extent suboptimal. This can be ascribed to three main reasons, recently highlighted in a study covering the interaction of water with a divalent calcium cation.[9] First, a significant error arises on account of the rudimentary representation of the molecular electrostatic potential for water by means of a three-point (3-p) charge model, which misreproduces the target quantity

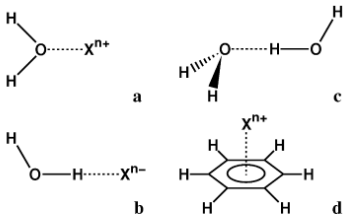


Figure 2. (a) Interaction of a water molecule with a cation along the C_2 axis. (b) Interaction of a water molecule with an anion along the O−H bond. (c) Interaction of a water molecule with another water molecule defining a hydrogen bond. (d) Interaction of the benzene molecule with a cation along the C_6 axis.

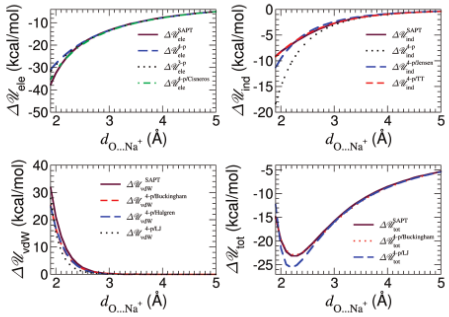


Figure 3. Interaction of a monovalent sodium ion with water. The reference (SAPT) curve is shown as a solid line. Three models for the electrostatic contribution are represented through either no penetration with a three-point (3-p) charge model (dotted line), a four-point (4-p) charge model (long-dashed line), or the function derived by Cisneros et al. (dot-dashed line). Three models for the induction contribution are represented using either no damping (dotted line), the damping function derived by Jensen et al. (long-dashed line), or the damping function derived by Tang and Toennies (TT, short-dashed line). Three models for the van der Waals contribution are fitted to $\Delta\mathscr{U}_{\text{tot}}^{\text{SAPT}} - \Delta\mathscr{U}_{\text{ele}}^{\text{4-p}} - \Delta\mathscr{U}_{\text{ind}}^{\text{4-p/TT}}$ using either a Buckingham (short-dashed line), a Halgren (long dashed line), or a Lennard-Jones (LJ, dotted line) potential. The total energy is computed as the sum of $\Delta\mathscr{U}_{\text{ele}}^{\text{4-p}}$ plus $\Delta\mathscr{U}_{\text{ind}}^{\text{4-p/TT}}$ plus either $\Delta\mathscr{U}_{\text{vdW}}^{\text{4-p/LJ}}$ (short-dashed line) or $\Delta\mathscr{U}_{\text{vdW}}^{\text{4-p/Buckingham}}$ (dotted line).

by up to 50%. Second, the description of the exchange-induction term of the SAPT expansion using the damping function devised by Jensen et al., while constituting a judicious choice, appears to suffer from a number of shortcomings. Last, the most questionable aspect of the proposed polarizable model lies in the use of the 6-12 LJ potential, traditionally utilized in macromolecular force fields. As was shown, however, this potential systematically fails to reproduce in a consistent fashion the exchange and dispersion contributions to the intermolecular energy computed within the SAPT framework. This shortcoming has prompted several authors to turn to more appropriate functions.[6,27,86−88] To address the above issues, the interaction of a series of metal ions with water has been investigated (see Figure 2). For conciseness, only the water−sodium

Table 4. Values of the Jensen et al. Damping Parameter (*a*) for Each Atom Type

	Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	$O(H_2O)$	$X(H_2O)$	$H(H_2O)$
a	0.095	0.085	0.190	0.200	0.105	4.103	4.103	0.358

Table 5. Values of the Damping Parameter (β) in the Tang and Toennies Damping Function[a]

β	Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	F^-
H_2O	2.38	2.00	2.15	2.10	2.10	4.50
benzene	2.10	1.70	1.75	1.85	1.85	

[a] Each value corresponds to a pair of interactions.

interaction will be detailed here (see Figure 3); results for all of the systems are provided in the Supporting Information.

As can be observed in Figure 3, the electrostatic contribution is equally well reproduced using a 4-p or the 3-p charge model. This result is not surprising given that the cation approaches water along the C_2 axis of the latter. Error in the reproduction of the molecular electrostatic potential in the direction of the bisector, on the order of 5% at the equilibrium intermolecular distance, is equivalent for the two models. This, unfortunately, is no longer true when the molecular electrostatic potential is measured quantum-mechanically along the O−H bond. In this event, the error reaches 15% for the 3-p charge model, but only 5% for the 4-p one. The electrostatic term modified by a penetration function is also displayed in Figure 3. This modification suggested by Cisneros et al. improves the accuracy of the model at short distances, even though, for water−cation dimers, the effect of electron-cloud penetration is nearly negligible at distances around the intermolecular equilibrium separation (around 2.2 Å). As underlined in the previous section, this correction to the electrostatic potential will not be introduced explicitly but will, instead, be embedded in the relevant representation of the van der Waals potential, thereby simplifying the model, while remaining physically consistent.

The undamped induction energy of the classical model overestimates the quantity supplied by the SAPT expansion.[9] The damping functions considered here are that of Jensen et al. and that of Tang and Toennies. The latter function is to be favored over the former, as is reflected in the better behavior of the corresponding damped potential. All of the parameters determined for these damping functions are gathered in Tables 4 and 5.

The negligible errors in the reproduction of the electrostatic and induction contributions in water−cation complexes makes the total energy ($\Delta\mathscr{U}_{tot}^{SAPT}$) minus $\Delta\mathscr{U}_{elec}^{4\text{-p}}$ and $\Delta\mathscr{U}_{ind}^{4\text{-p}/TT}$ very similar for the SAPT van der Waals contribution ($\Delta\mathscr{U}_{vdW}^{SAPT}$). Employing a traditional 6-12 LJ function, the repulsive component of the latter results in a systematic failure to match the target van der Waals potential—the associated error is magnified when the induction effects dominate over the Coulomb and dispersion components. Conversely, if use is made of a Halgren or a Buckingham function, reproduction of this potential proves to be extremely accurate. Although depicted in Figure 3, the Halgren function will not be used hereafter, because the exponential part of the exp-6 Buckingham function constitutes a more physically sound framework, capable of absorbing the errors due to penetration effects. Interestingly enough, nonmultipolar contributions to the induction energy appear to be negligible in water−cation interactions and can, thus, be safely omitted. All of the parameters of the aforementioned three van der Waals functions are given in Table 6. These parameters are the raw coefficients for each interaction in the absence of combination rules.

When adding the different terms of the potential energy function, the total interaction energy is reproduced accurately, hence, confirming the applicability of the model (see Tables 7 and 8 for geometrical and energetic properties of the minimum-energy complex). Not too unexpectedly, calculations based on a LJ potential lead to a severely flawed reproduction of the total interaction energy.

4.2. The Water Dimer and Interaction of Water with Anions. In this section, the first four halide ions have been considered to interact with water along one O−H bond of the latter, which corresponds to the most favorable approach of these chemical species. In addition, a water dimer has been examined in the geometry provided in ref 89 (see also Figure 2). The results obtained for these two classes of complexes show similar trends. For conciseness, only the interaction in the water dimer will be detailed here (see Figure 4); results for all of the systems are provided in the Supporting Information.

The importance of the 4-p charge model ought to be emphasized for the reproduction of the electrostatic potential. Even at distances where penetration effects are known to be negligible, the 3-p charge model fails to reproduce the electrostatic contribution, whereas the 4-p charge model describes it successfully. Inclusion of a penetration correction yields an overall agreement between the classical model and the QM reference, irrespective of the intermolecular distance explored. Penetration effects for water−anion systems, as well as for the water dimer are crucial, as they account for up to 50% of the total electrostatic contribution at very short distances. They are, nonetheless, noticeably smaller (approximately 8%) at an intermolecular distance corresponding to the minimum-energy complex. Accordingly, it is reasonable to expect that penetration effects can be transferred in an effective fashion into the van der Waals potential, as discussed previously.

When penetration effects are large, nonmultipolar contributions are anticipated to be equally substantial. In this event, the induction term is burdened by overwhelming nonmultipolar stabilizing effects, which counterbalance the destabilizing effect of the exchange-induction component. These antagonistic phenomena preclude the use of a damping function for the present complexes. As can be seen in Table 5, no damping function is used, except for fluoride, which is the smallest anion with the least nonmultipolar effects. The missing nonmultipolar contribution to the induction energy will, thus, be incorporated into the Buckingham potential.

The conclusions reached in light of the calculations based on the water−cation dimers appear to hold also for the present complexes. Furthermore, the assumption that an exp-6 Buckingham potential, unlike a 6-12 LJ potential, can compensate

Table 6. van der Waals Parameters for H_2O-X Systems[a]

		Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	F^-	Cl^-	Br^-	I^-	H_2O
	ε	19149	66526	120694	94937	94131	5837	11843	12340	17204	14946
(1)	A_1	3.944	4.147	3.842	3.900	3.738	3.598	3.352	3.106	3.132	3.949
	A_2	0.025	0.032	1.418	3.385	2.332	0.019	1.679	2.600	0.034	2.541
	ε	0.041	0.100	0.115	0.433	0.151	0.003	0.021	0.025	0.038	0.041
(2)	R^*	3.412	3.390	3.880	2.848	3.700	5.132	4.113	4.205	4.284	2.864
	ε	468.3	94.27	11.58	1308	7.547	65.52	93.85	323.2	12.20	61468
(3)	σ	1.302	1.715	2.395	1.355	2.405	1.301	1.600	1.566	2.248	0.763

[a] The parameters are obtained using a model of point charges and an induction model of distributed polarizabilities damped by the function of Tang and Toennies. (1) ε and A_n ($n = 1, 2$) correspond to the parameters of the Buckingham function. (2) ε and R^* are the parameters of the Halgren function. (3) ε and σ correspond to the parameters of the Lennard-Jones function. ε is given in kilocalories per mole. All other parameters are in ångströms.

Table 7. Values of the Interaction Distance for All the Minimum-Energy Complexes[a]

$d_{A\ldots B}$	Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	F^-	Cl^-	Br^-	I^-	H_2O
H_2O	1.90	2.22	2.70	1.90	2.30	1.60	2.20	2.40	2.70	2.10
	1.90	**2.25**	**2.66**	**1.90**	**2.30**	**1.60**	**2.20**	**2.40**	**2.70**	**2.10**
benzene	1.90	2.41	2.85	2.02	2.40					
	1.90	**2.41**	**2.90**	**1.92**	**2.40**					

[a] The reference quantum-mechanical values are given in bold. All distances are expressed in ångströms.

Table 8. Values of the Interaction Energy for All the Minimum-Energy Complexes[a]

$\Delta\mathscr{U}_{tot,min}$	Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}	F^-	Cl^-	Br^-	I^-	H_2O
H_2O	–32.63	–22.91	–16.85	–78.80	–52.98	–22.38	–13.07	–11.14	–9.25	–3.60
	–32.69	**–23.12**	**–16.91**	**–79.85**	**–52.75**	**–22.46**	**–13.00**	**–11.08**	**–9.22**	**–3.60**
benzene	–37.45	–23.41	–17.57	–114.17	–74.25					
	–37.43	**–23.70**	**–17.60**	**–116.41**	**–74.23**					

[a] The reference quantum-mechanical values are given in bold. All energies are expressed in kilocalories per mole.

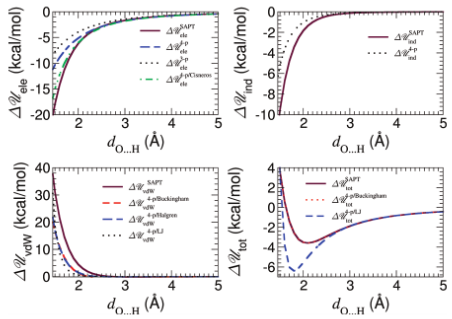


Figure 4. Interaction of two water molecules. The reference (SAPT) curve is shown as a solid line. Three models for the electrostatic contribution are represented through either a three-point (3-p) charge model with no penetration (dotted line), a four-point (4-p) charge model with no penetration (long-dashed line), or the function derived by Cisneros et al. (dot-dashed line). The induction contribution is represented without damping function (dotted line). Three models for the van der Waals contribution are fitted to $\Delta\mathscr{U}_{tot}^{SAPT} - \Delta\mathscr{U}_{ele}^{4\text{-p}} - \Delta\mathscr{U}_{ind}^{4\text{-p}}$ using either a Buckingham (short-dashed line), a Halgren (long dashed line), or a Lennard-Jones (LJ, dotted line) potential. The total energy is computed as the sum of $\Delta\mathscr{U}_{ele}^{4\text{-p}}$ plus $\Delta_{ind}^{4\text{-p}}$ plus either $\Delta\mathscr{U}_{vdW}^{4\text{-p/LJ}}$ (short-dashed line) or $\Delta\mathscr{U}_{vdW}^{4\text{-p/Buckingham}}$ (dotted line).

for the incompleteness of the classical model by embracing its overlooked, sizable contributions proves to be perfectly legitimate and reasonable (see Tables 7 and 8).

In the particular instance of negatively charged compounds, the SAPT expansion can, under certain circumstances, fail to converge.[90] For this reason, the results (not shown here) supplied by an SAPT and an RVS[91] expansion, using the Gamess[92] suite of programs, have been compared. If intramolecular correlation is ignored, the RVS and the SAPT expansions yield the same energies for all the dimers considered here, which provides a safeguard, inasmuch as convergence of the SAPT reference calculation is concerned. Moreover, the recent work of Kim et al.[93] on halide anions interacting with π systems further suggests convergence of the SAPT expansion for such dimers.

4.3. Interaction of Benzene with Metal Cations. The same series of metal cations utilized in the above study of water−cation dimers was considered for the interaction with benzene along its C_6 axis (see Figure 2). This approach corresponds to the most favorable interaction pattern[94] for a π-electron cloud interacting with a positively charged species. For conciseness, only the benzene−sodium interaction will be displayed in Figure 5; results for all systems investigated (viz., Li^+, K^+, Mg^{2+}, and Ca^{2+}) are provided in the Supporting Information.

A twelve-point (12-p) charge model derived from the molecular electrostatic potential tends to overestimate the interaction in cation-π systems, despite the appropriate reproduction of the molecular quadrupole moment of benzene and, more generally, its full molecular electrostatic potential. This overestimation is more prone to be manifested with ions bearing a highly localized charge and, on the contrary, is expected to be less perceptible for a diffuse charge, as in the case of ammonium,[43] let alone guanidinium. The results suggest that the interaction of cations with π systems is not strictly speaking an attractive quadrupole-charge one.[94] A case in point is the benzene−lithium dimer, where the electrostatic energy changes its slope at very short distances,

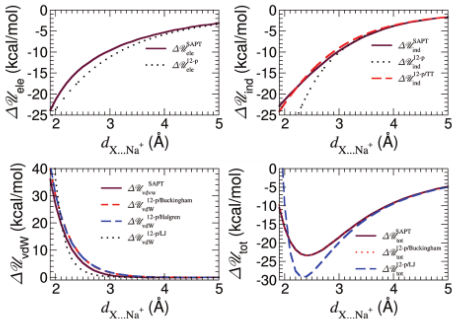


Figure 5. Interaction of a monovalent sodium ion with benzene. The reference (SAPT) curve is shown as a solid line. The electrostatic contribution obtained with a twelve-point (12-p) charge model with no penetration is represented as a dotted line. Two models for the induction contribution are represented using either no damping (dotted line) or the damping function derived by Tang and Toennies (TT, short-dashed line). Three models for the van der Waals contribution are fitted to $\Delta\mathscr{U}_{tot}^{SAPT} - \Delta\mathscr{U}_{ele}^{12\text{-p}} - \Delta\mathscr{U}_{ind}^{12\text{-p/TT}}$ using either a Buckingham (short-dashed line), a Halgren (long dashed line), or a Lennard-Jones (LJ, dotted line) potential. The total energy is computed as the sum of $\Delta\mathscr{U}_{ele}^{12\text{-p}}$ plus $\Delta\mathscr{U}_{ind}^{12\text{-p/TT}}$ plus either $\Delta\mathscr{U}_{vdW}^{12\text{-p/LJ}}$ (short-dashed line) or $\Delta\mathscr{U}_{vdW}^{12\text{-p/Buckingham}}$ (dotted line).

which, in all likelihood, is due to the positively charged lobes of the quadrupole moment interacting with the cation. The resulting destabilizing effect, which largely counterbalances the penetration contribution, is not accounted for in the electrostatic potential and is transferred to a short-distance, repulsive van der Waals term.

As noted above for the cation−water complexes, the undamped induction energy overestimates the values obtained from an SAPT analysis, but inclusion of the Tang and Toennies damping formalism reproduces the SAPT induction profile very nicely. Similarly, the 6-12 LJ potential departs markedly from the SAPT van der Waals contribution, which is always better reproduced when Halgren or Buckingham functions are used. Conclusions drawn on the basis of the water−cation complexes are shown to hold for all other π-cation interactions reported herein (van der Waals parameters are gathered in Table 9, and the geometrical and energetic properties of the minimum-energy complex are given in Tables 7 and 8).

5. Conclusion

The present study addresses the shortcomings of a simplified polarizable potential energy function and how these shortcomings can be overcome to model with an appreciable accuracy intermolecular interactions. It is desirable to have access to a simple polarizable potential energy function that satisfies the criteria of tractability and precision. This necessarily demands that important choices be made, inasmuch as the contributions to be included are concerned. In other words, it is pivotal that all of the underlying physics of the interactions be properly introduced to reflect the various contributions at play.

Table 9. van der Waals Parameters for Benzene−X^{n+} Systems[a]

		Li^+	Na^+	K^+	Mg^{2+}	Ca^{2+}
	ε	3556	8821	38649	103962	61104
(1)	A_1	2.624	2.841	2.924	3.184	2.753
	A_2	2.419	0.067	3.502	4.583	4.661
	ε	0.002	0.016	0.104	0.138	0.157
(2)	R^*	6.331	4.871	4.261	3.471	4.157
	ε	8124	539.9	6.103	33.21	9.573
(3)	σ	2.624	2.841	2.924	1.860	2.753

[a] The parameters are obtained using a model of point charges on atoms with no penetration and an induction model of distributed polarizabilities damped by the function provided by Tang and Toennies. (1) ε and A_n (n = 1, 2) correspond to the parameters of the Buckingham function. (2) ε and R^* are the parameters for the Halgren function. (3) ε and σ correspond to the parameters of the Lennard-Jones function. ε is given in kilocalories per mole. All other parameters are in ångströms.

With this objective in mind, a series of guidelines is proposed. These guidelines can be summarized as follows. The electrostatic contribution to the total energy is described by a truncation of the molecular electrostatic potential at the monopole level, which has proven in many instances to constitute a very reasonable approximation.[52] As indicated by Swart et al.,[95] this description leads to small errors for nonsymmetric systems but can fail in the case of small, symmetric molecules. Such is the case for the water molecule, where fictitious sites are needed to enhance the reproduction of the true molecular electrostatic potential. Accordingly, a TIP4P-like[80,81] potential was proposed for water. In addition to the first-order contribution to the total energy, penetration effects are present at short distances in all interacting systems. This term can be modeled with an exponential function and can be introduced through the exp-6 Buckingham potential. The total interaction energy is further decomposed into an induction component, which, at the classical level, relies on the conjunction of charge-flow polarizabilities between covalently bonded atoms and isotropic dipole polarizabilities determined from QM induction energy maps.

The resulting polarizable models are simple, yet sufficiently accurate for recovering the anisotropy of the molecular polarizability. In most circumstances, the damping function proposed by Tang and Toennies recovers the exchange-induction contribution. This particular function can be expanded up to third order only, which greatly simplifies its calculation. Furthermore, when the nonmultipolar component to the induction energy cannot be neglected, a classical multipole expansion systematically underestimates the polarization part of the interaction energy. Should this situation arise, it is necessary to remove the damping function from the model and encapsulate this stabilizing effect through another suitable function. The remaining terms of the QM expansion that are described explicitly can be incorporated in a single contribution that involves the exp-6 Buckingham potential. The latter is highly appropriate because it encompasses each and every missing term that can be expressed by way of an exponential, repulsive representation.

Extension of the present work to an all-atom description for the van der Waals potential is underway to provide a better three-dimensional description of the short-range

potential. The very encouraging results reported herein suggest that the proposed methodology for deriving polarizable force fields on the basis of a small set of prototypical interacting chemical species can be readily extended to more complex molecular systems.

Acknowledgment. The authors thank János G. Ángyán, Georg Jansen, and Krzysztof Szalewicz for stimulating discussions. The authors are indebted to the Centro de Supercomputación de Catalunya, Barcelona, Spain, and to the Centre Informatique National de l'Enseignement Supérieur, Montpellier, France, for provision of generous amounts of computational time.

Supporting Information Available: Intermolecular energy profiles for water and benzene interacting with halide and metal ions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Yin, Y.; Arkhipov, A.; Schulten, K. *Structure* **2009**, *17*, 882–892.

(2) Lei, H.; Duan, Y. *J. Mol. Biol.* **2007**, *370*, 196–206.

(3) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(4) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(5) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(6) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2003**, *108*, 621–627.

(7) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. *Biophys. J.* **2008**, *94*, 75–77.

(8) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. *Biophys. J.* **2009**, *96*, 3772–3780.

(9) Dehez, F.; Archambault, F.; Soteras, I.; Luque, F. J.; Chipot, C. *Mol. Phys.* **2008**, *106*, 1685–1696.

(10) Shiratori, Y.; Nakagawa, S. *J. Comput. Chem.* **1991**, *12*, 717–730.

(11) Project, E.; Nachliel, E.; Gutman, M. *J. Comput. Chem.* **2008**, *29*, 1163–1169.

(12) Bucher, D.; Guidoni, L.; Maurer, P.; Rothlisberger, U. *J. Chem. Theory Comput.* **2009**, *5*, 2173−2179.

(13) Allen, T.; Andersen, O.; Roux, B. *Biophys. Chem.* **2006**, *124*, 251–267.

(14) Fowler, P. W.; Stone, A. J. *J. Phys. Chem.* **1987**, *91*, 509–511.

(15) Gavezzotti, A. *J. Phys. Chem. B* **2002**, *106*, 4145–4154.

(16) Gavezzotti, A. *J. Phys. Chem. B* **2003**, *107*, 2344–2353.

(17) Welch, G. W. A.; Karamertzanis, P. G.; Misquitta, A. J.; Stone, A. J.; Price, S. L. *J. Chem. Theory Comput.* **2008**, *4*, 522–532.

(18) Barnes, P.; Finney, J. L.; Nicholas, J. D.; Quinn, J. E. *Nature* **1979**, *282*, 459–464.

(19) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

(20) Rappe, A. K.; Goddard, W. A. *J. Phys. Chem.* **1991**, *95*, 3358–3363.

(21) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730–4737.

(22) Patel, S.; Brooks, C. L. *J. Comput. Chem.* **2003**, *25*, 1–16.

(23) Patel, S.; Mackerell, A. D., Jr.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1504–1514.

(24) Sanderson, R. T. In *Chemical Bonds and Bond Energy*; Academic Press: New York, 1976; pp 1−15.

(25) Le Sueur, C. R.; Stone, A. J. *Mol. Phys.* **1994**, *83*, 293–307.

(26) Engkvist, O.; Åstrand, P.-O.; Karlström, G. *Chem. Rev.* **2000**, *100*, 4087–4108.

(27) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497–1506.

(28) Hagberg, D.; Brdarski, S.; Karlström, G. *J. Phys. Chem. B* **2005**, *109*, 4111–4117.

(29) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960–1986.

(30) Darley, M. G.; Popelier, P. L. A. *J. Phys. Chem. A* **2008**, *112*, 12954–12965.

(31) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(32) Masella, M.; Flament, J. P. *J. Chem. Phys.* **1997**, *107*, 9105–9116.

(33) Zhang, Q.; Zhang, X.; Yu, L.; Zhao, D.-X. *J. Mol. Liq.* **2009**, *145*, 58–66.

(34) Zhang, Q.; Zhang, X.; Zhao, D.-X. *J. Mol. Liq.* **2009**, *145*, 67–81.

(35) Babin, V.; Baucom, J.; Darden, T. A.; Sagui, C. *J. Phys. Chem.* **2006**, *110*, 11571–11581.

(36) Harder, E.; MacKerell, A. D., Jr.; Roux, B. *J. Am. Chem. Soc.* **2009**, *131*, 2760–2761.

(37) Claverie, P. In *Intermolecular Interactions: From Diatomics to Biopolymers*; Pullman, B., Ed.; Wiley Interscience: New York, 1978; pp 69−305.

(38) Celebi, N.; Ángyán, J. G.; Dehez, F.; Millot, C.; Chipot, C. *J. Chem. Phys.* **2000**, *112*, 2709–2717.

(39) Dehez, F.; Soetens, J.-C.; Chipot, C.; Ángyán, J. G.; Millot, C. *J. Phys. Chem. A* **2000**, *104*, 1293–1303.

(40) Chipot, C.; Dehez, F.; Ángyán, J. G.; Millot, C.; Orozco, M.; Luque, F. J. *J. Phys. Chem. A* **2001**, *105*, 11505–11514.

(41) Dehez, F.; Chipot, C.; Millot, C.; Ángyán, J. G. *Chem. Phys. Lett.* **2001**, *338*, 180–188.

(42) Ángyán, J. G.; Chipot, C.; Dehez, F.; Hättig, C.; Jansen, G.; Millot, C. *J. Comput. Chem.* **2003**, *24*, 997–1008.

(43) Dehez, F.; Ángyán, J. G.; Soteras, I.; Luque, F. J.; Schulten, K.; Chipot, C. *J. Chem. Theory Comput.* **2007**, *3*, 1914–1926.

(44) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

(45) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorska, M.; Jeziorski, B.; Kucharski, S. A.; Lotrich, V. F.; Misquitta, A. J.; Moszyński, R.; Patkowski, K.; Podeszwa, R.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wheatley, R. J.; Wormer, P. E. S.; Zuchowski, P. S. *SAPT2008*; Department of Physics and Astronomy, University of Delaware: Newark, DE;

Department of Chemistry, University of Warsaw: Warsaw, Poland, 2008.

(46) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(47) Burcl, R.; Chalasiński, G.; Bukowski, R. *J. Chem. Phys.* **1995**, *103*, 1498–1507.

(48) Szalewicz, K.; Patkowski, K.; Jeziorski, B. *Structure Bonding* **2005**, *116*, 43–117.

(49) Torheyden, M.; Jansen, G. *Mol. Phys.* **2006**, *104*, 2101–2138.

(50) Misquitta, A. J.; Stone, A. J. *J. Chem. Theory Comput.* **2008**, *4*, 7–18.

(51) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrasek, J. *J. Chem. Theory Comput.* **2009**, *5*, 982–992.

(52) Chipot, C.; Ángyán, J. G. *New J. Chem.* **2005**, *29*, 411–420.

(53) Ángyán, J. G.; Chipot, C. *Int. J. Quantum Chem.* **1994**, *52*, 17–37.

(54) Vigné-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934–4948.

(55) Stone, A. J. In *Theoretical Models of Chemical Bonding*; Maksić, H., Ed.; Springer-Verlag: Berlin, 1991; Vol. 4, pp 103−131.

(56) Stone, A. J. In *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, U. K., 1996.

(57) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300–7306.

(58) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293–307.

(59) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. *J. Phys. Chem. A* **2003**, *107*, 10353–10359.

(60) Cisneros, G. A.; Na-Im Tholander, S.; Parisel, O.; Darden, T. A.; Elking, D.; Perera, L.; Piquemal, J.-P. *Int. J. Quantum Chem.* **2008**, *108*, 1905–1912.

(61) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.

(62) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047–1064.

(63) Soteras, I.; Curutchet, C.; Bidon-Chanal, A.; Dehez, F.; Ángyán, J. G.; Orozco, M.; Chipot, C.; Luque, F. J. *J. Chem. Theory Comput.* **2007**, *3*, 1901–1913.

(64) Jensen, L.; Åstrand, P.-O.; Osted, A.; Kongsted, J.; Mikkelsen, K. V. *J. Chem. Phys.* **2002**, *116*, 4001–4010.

(65) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726–3741.

(66) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1977**, *66*, 1496–1506.

(67) Millot, C.; Stone, A. J. *Mol. Phys.* **1992**, *77*, 439–462.

(68) Meredith, A. W.; Stone, A. J. *J. Phys. Chem. A* **1998**, *102*, 434–445.

(69) Jones, J. *Proc. R. Soc. London, Ser. A* **1924**, *106*, 441–462.

(70) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(71) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.

(72) Buckingham, R. A.; Hamilton, J.; Massey, H. S. W. *Proc. R. Soc. London, Ser. A* **1941**, *179*, 103–122.

(73) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador1, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(74) Sadlej, A. J. *Theor. Chim. Acta* **1992**, *79*, 123–140.

(75) Sadlej, A. J. *Theor. Chim. Acta* **1992**, *81*, 45–63.

(76) Sadlej, A. J. *Theor. Chim. Acta* **1992**, *81*, 339–354.

(77) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.

(78) Misquitta, A. J.; Stone, A. J.; Price, S. L. *J. Chem. Theory Comput.* **2008**, *4*, 19–32.

(79) Saunders, V. R.; Guest, M. F. *ATMOL*; SERC Daresbury Laboratory: Daresbury, Great Britain, 1976.

(80) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(81) Colonna, F.; Ángyán, J. G.; Tapia, O. *Chem. Phys. Lett.* **1990**, *172*, 55–61.

(82) Dzyabchenko, A. V. *Russ. J. Phys. Chem. A* **2008**, *82*, 758–766.

(83) Coker, H. *J. Phys. Chem.* **1976**, *80*, 2084–2091.

(84) Fajans, K. *J. Phys. Chem.* **1970**, *74*, 3407–3410.

(85) Marquart, D. *SIAM J. Appl. Math.* **1963**, *11*, 431–441.

(86) Gresh, N.; Garmer, D. R. *J. Comput. Chem.* **1995**, *17*, 1481–1495.

(87) Brdarski, S.; Karlström, G. *J. Phys. Chem. A* **1998**, *102*, 8182–8192.

(88) Lago, N. F.; Larranaga, F. H.; Albertí, M. *Eur. Phys. J. D* **2009**, *1*, 75−85.

(89) Mas, E. M.; Szalewicz, K. *J. Chem. Phys.* **1996**, *104*, 7606–7614.

(90) Szalewicz, K. Private communication.

(91) Chen, W.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 14316–14328.

(92) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(93) Kim, D.; Tarakeshwar, P.; Kim, K. S. *J. Phys. Chem. A* **2004**, *108*, 1250–1258.

(94) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303–1324.

(95) Swart, M.; van Duijnen, P. T.; Snijders, J. G. *J. Comput. Chem.* **2000**, *22*, 79–88.

CT9004189

Theoretical Chemistry Accounts (2018) 137:174
https://doi.org/10.1007/s00214-018-2376-z

REGULAR ARTICLE

Modeling induction phenomena in amino acid cation–π interactions

Chetan Raj Rupakheti[1] · Benoît Roux[1,2] · François Dehez[3] · Chris Chipot[3,4,5]

Received: 29 June 2018 / Accepted: 25 September 2018

Abstract
Cation–π interactions are widely recognized as an important class of interactions, notably in biology and supramolecular chemistry, participating in molecular recognition and association phenomena. Numerical simulations relying on additive force fields perform usually poorly in modeling precisely these interactions. It is now well established that accurate reproduction of the interaction energy of a positively charged group bound to the π-electron cloud of an aromatic ring requires an explicit treatment of induction effects by means of polarizable potentials. In this contribution, we compare critically the ability of the CHARMM Drude polarizable force field to describe a series of prototypical cation–π interactions observed in proteins with that of the pairwise additive CHARMM36 force field. Toward this end, potentials of mean force characterizing the binding of amino acid side-chain models, namely ammonium and guanidinium cations, on the one hand, and toluene, para-cresol and 3-methylindole, on the other hand, have been determined within the extended adaptive biasing force framework.

Keywords Recognition and association · Polarization · Free energy · Molecular simulations

1 Introduction

Over thirty years ago, cation–π interactions were recognized as a fourth actor in molecular recognition and association phenomena, in addition to the three other common non-covalent interactions, namely hydrogen bonds, salt bridges and hydrophobic effect [1]. Owing to their short-range nature, it is tempting to regard them as unconventional hydrogen bonds [2], considering that the latter usually form between a donor and a π-electron density, which, in the absence of lone pairs, acts as an acceptor [3]. The seminal work of Kebarle and coworkers brought to light the duality of electrostatic and polarization contributions as the driving force for the binding of a potassium ion to a benzene ring. From a phenomenological standpoint, cation–π interactions arise from the superimposition of two effects of a different nature. The electric field generated by the cation polarizes the π-electron

Published as part of the special collection of articles In Memoriam of János Ángyán.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00214-018-2376-z) contains supplementary material, which is available to authorized users.

✉ François Dehez
francois.dehez@univ-lorraine.fr

✉ Chris Chipot
chipot@ks.uiuc.edu

Chetan Raj Rupakheti
rupakhetic@uchicago.edu

Benoît Roux
chipot@ks.uiuc.edu

[1] Department of Biochemistry and Molecular Biology, Gordon Center for Integrative Science, University of Chicago, Chicago, Illinois 60637–1454, United States

[2] Leadership Computing Facility, Argonne National Laboratory, 9700 Cass Avenue, Building 240, Argonne, Illinois 60439–8643, United States

[3] Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana-Champaign, Unité Mixte de Recherche no 7019, Université de Lorraine, B.P. 70239, 54506 Vandœuvre-lès-Nancy cedex, France

[4] Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, Illinois 61801, United States

[5] Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States

Published online: 15 November 2018

cloud of the aromatic compound, which corresponds to an r^{-2} interaction. In turn, the induced dipole moment of the ring interacts with the polarizing charge through its electrostatic potential, which also corresponds to an r^{-2} interaction. Identification of new non-covalent interactions in simple cation–π complexes was rapidly followed by a host of investigations in biological objects, wherein unusual hydrogen bonds formed between amino and aromatic moieties had long been observed [4–6]. Cation–π interactions have proven strong enough to compete with the hydration of charged, titratable species and promote binding in the hydrophobic pocket of a protein lined with aromatic residues [7–9]. Ubiquitous to a host of components of the cell machinery, they have been proposed to be responsible for the binding of acetylcholine, one of the most extensively studied neurotransmitters, to acetylcholinesterase [10, 11]. Extensive analyses of the protein data bank, or PDB [12], have highlighted a preference for tryptophan to be engaged in an interaction with a cationic amino acid, among which arginine dominates over both lysine and the protonated form of histidine [8]. These analyses have also revealed a pronounced anisotropy of the non-covalent interaction formed by aromatic and cationic side chains, whereby the latter approaches the π-electron cloud in the direction normal to it [13]. In stark contrast to salt bridges, the strength of which is modulated by the dielectric permittivity of the environment, cation–π interactions can sustain exposure to the aqueous surroundings and are, therefore, commonly observed at the surface of proteins [14]. Though cation–π interactions appear to not be pivotal for the stability of proteins [15, 16], they could, nevertheless, participate in their folding [7]. In fact, proximity of a basic, titratable side chain from an aromatic one induces a bias in the local geometry toward the formation of a non-covalent interaction [8].

At the conceptual level, cation–π interactions result from the attraction of a polarizing, positively charged moiety toward the quadrupole engendered by the π-electron cloud of a noteworthily polarizable aromatic ring [7, 9, 17, 18]. Owing to the large polarizability of aromatic species and the strong polarization exerted by the cation, faithful depiction of cation–π interactions falls out of the scope of rudimentary, pairwise additive potential energy functions, which only account for induction phenomena in an average sense [19–21]. The attractive part of the short-range potential describing induction effects, which formally ought to be represented analytically by means of an r^{-4} term [13, 22], is clearly absent from minimalist force fields devoid of an explicit treatment of polarization. Pioneering molecular mechanics-based calculations on prototypical cation–π complexes have emphasized the necessity to incorporate induction phenomena to recover the expected attractive feature of the intermolecular potential, in particular in aqueous environments [23, 24]. One possible route to handle polarization explicitly in numerical simulations, notably in molecular dynamics (MD) simulations, consists in introducing Drude oscillators, formed by auxiliary particles carrying a partial charge and tethered to real atoms by means of a harmonic spring [25, 26]. In response to the local electric field felt by the polarizable sites of the molecule, these auxiliary particles move relative to the real atoms as their respective charge varies, resulting in an overall self-consistent reorganization of the charge distribution induced by the chemical environment. The main advantage of this approach lies in its conceptual simplicity, compared to more sophisticated representations of the induction energy by means of a multipole expansion [27], allowing MD simulations to remain in the framework of a point-charge model, thus limiting the computational overhead incurred in the explicit treatment of polarization [28].

In the present contribution, we probe the ability of two academic force fields to describe the cation–π interaction formed in an aqueous environment between model basic and aromatic amino acid side chains, using high-performance geometric free energy calculations [29]. Specifically, the reversible association of guanidinium and ammonium, on the one hand, with toluene, para-cresol and 3-methylindole, on the other hand, is examined through potential of mean force (PMF) calculations with a potential energy function resting on a two-body additive approximation, and another one accounting for non-additive induction effects by means of classical Drude oscillators [26].

2 Methods

Computational assays In this study, toluene, para-cresol and 3-methylindole serve as proxies for the side chain of phenylalanine, tyrosine and tryptophan, whereas guanidinium and ammonium represent the cationic moiety of arginine and lysine, respectively. Each molecular assembly consisted of an aromatic compound and a cation solvated in a thermalized bath of 1958 water molecules, corresponding to six independent computational assays of dimensions approximately equal to $38 \times 38 \times 38$ Å^3. For the MD simulations performed with a non-additive force field, the ancillary virtual particles and lone pairs of electronegative elements [30] were added to the computational assays using the CharmmGUI server [31].

Molecular dynamics simulations All MD simulations presented herein were carried out with the parallel, scalable program NAMD 2.12 [32]. For the simulations wherein induction phenomena were neglected, the TIP3P model [33] and the CHARMM general force field [34] (CGenFF) were used to describe water and the amino acid side-chain models, respectively. Conversely, for the simulations accounting for mutual polarization, the SWM4-NDP model [35] and the

CHARMM Drude force field [36, 37] were employed for water molecules and the side-chain proxies, respectively. It ought to be mentioned here that the parameters of the non-additive CHARMM potential energy function were taken "out of the box," without any changes. Further refinement could be introduced through the introduction of additional terms or NBFIX corrections [38]. The r–RESPA multiple time-step algorithm [39] was employed to integrate the equations of motion with a time step of 2 and 4 fs for short- and long-range interactions, respectively, in all simulations using a two-body additive force field, and with a time step of 0.5 fs in all simulations involving Drude particles. The SETTLE algorithm was used to constrain the covalent bonds of water molecules to their equilibrium length [40]. Covalent bonds of the side-chain models involving hydrogen atoms were constrained to their equilibrium length by means of the RATTLE algorithm [41]. The temperature and the pressure were maintained at 298 K and 1 atm, respectively, using Langevin dynamics and the Langevin piston method [42]. Long-range electrostatic forces were taken into account by means of the particle mesh Ewald algorithm [43]. A cutoff of 12 Å was utilized to truncate van der Waals and short-range Coulombic interactions. Periodic boundary conditions (PBCs) were applied in the three directions of Cartesian space. Visualization and analyses of all MD trajectories were performed with the VMD program [44].

Free energy calculations The extended adaptive biasing force algorithm [45–47] was employed to determine the PMFs, $w(r)$, that underlie cation–π reversible association, through integration of the average force exerted along the transition coordinate, r. Under these premises, a biasing force is estimated such that, once applied onto the relevant atoms, it yields a Hamiltonian bereft of an average force acting along the transition coordinate, defined here as the Euclidian distance, r, separating the center of mass of the cation, i.e., the nitrogen atom of ammonium, or the carbon atom of guanidinium, from the centroid of the π-electron cloud. The reaction pathway spanning 13 Å, i.e., $2 \leq r \leq 15$ Å, was not stratified into non-overlapping windows and was discretized in bins 0.1 Å wide, wherein samples of the local force acting along r were accumulated. Analyses of the PDB have revealed a marked preference of the cation to approach the aromatic ring in the direction normal to its plane [13, 24]. For this reason, the PMFs were determined enforcing a series of geometric restraints with harmonic potentials centered at $\theta = 90°$ (see Fig. 1). In this event, the cation is able to rotate unhindered about the axis perpendicular to the π–electron cloud. In particular, ammonium can adopt freely a mono-, bi- or tridentate orientation, whereas, for guanidinium, two $-NH_2$ groups approach the aromatic ring. Determination of each PMF required 100 ns of sampling, representing an aggregate time of 600 ns for the present computational investigation.

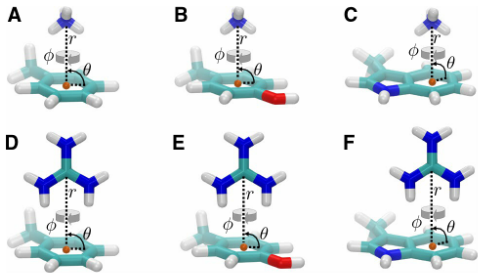


Fig. 1 Reversible association of amino acid side-chain models: ammonium:toluene (**a**), ammonium:para-cresol (**b**), ammonium:3-methylindole (**c**), guanidinium:toluene (**d**), guanidinium:para-cresol (**e**) and guanidinium:3-methylindole (**f**). r denotes the transition coordinate defined as the Euclidian distance between the center of mass of the cation, i.e., the nitrogen atom of ammonium, or the carbon atom of guanidinium, and the centroid of the π–electron cloud. Geometric restraints are enforced on angle θ to ensure perpendicular approach of the cation toward the aromatic ring. The cation can, however, rotate freely about angle ϕ

3 Results

To achieve a fair comparison of the additive CHARMM36 and the non-additive CHARMM Drude force fields, it should be stressed that the latter employed in this study consists of the original version developed by Roux and coworkers for modeling amino acids in the condensed phase [36, 37] and, thus, has not been specifically refined against cation–π reference complexes. Lamoureux and coworkers derived a series of ad hoc Lennard-Jones parameters to model more accurately cation–π interactions in the framework of the Drude force field [37, 48, 49]. Their strategy relies on the addition of an extra Lennard-Jones site located at the centroid of the six-membered ring.

The PMFs evaluated for both the additive CHARMM36 and the non-additive CHARMM Drude force fields are reported in Fig. 2. In all cases, two free energy minima can be observed, namely a contact minimum, where the cation interacts directly with the aromatic moiety, and a solvent-separated minimum, in which the cation–π interaction is mediated by a water molecule. It has been shown previously that occurrence of the second minimum is linked to the restriction of the binding process along the C_6 symmetry axis of the aromatic ring [24]. When cation binding is averaged over all possible orientations, the solvent-separated minimum usually vanishes [24, 49]. Depending on the cation and the π–electron cloud, the first minimum lies between 2.8 and 4 Å, whereas the second minimum emerges between 5.6 and 6.9 Å. For some complexes, explicit inclusion of polarization affects the position of the minima, an observation in line with the findings of Khan et al. [50].

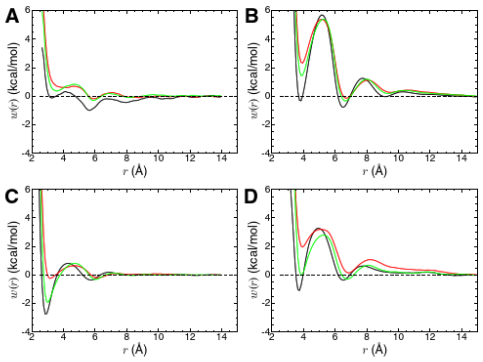


Fig. 2 Potentials of mean force characterizing the reversible association of amino acid side-chain models. Ammonium–π PMFs determined with a two-body pairwise (**a**) and non-additive (**c**) force field. Guanidinium–π PMFs determined with a two-body pairwise (**b**) and non-additive (**d**) force field. Color coding for the aromatic compounds: 3-methylindole (black line), para-cresol (red line) and toluene (green line). The transition coordinate, r, is defined in Fig. 1. All the PMFs obey the classical definition [51], whereby $\lim_{r\to\infty} w(r) := 0$, and thus include the Jacobian term $2r/\beta$ in the gradient of the free energy

Additive CHARMM36 force field For toluene and para-cresol interacting with ammonium and guanidinium, the PMFs obtained with the standard CHARMM36 force field (Fig. 2a, 2b) reveal that, compared to the dissociated state, the contact minimum is predicted to be less stable by 0.3 to 0.6 kcal/mol and 1.4 to 2.3 kcal/mol, for ammonium and guanidinium, respectively. Concerning the solvent-separated minimum, the free energy difference with the dissociated complex is nearly zero. At slight variance with toluene and para-cresol, 3-methylindole described by the pairwise additive force field has a contact minimum corresponding to the same free energy as the dissociated complex, and a solvent-separated minimum favored by 0.8 to 1.0 kcal/mol for both cations. The main difference between the profiles obtained with ammonium or guanidinium is rooted in the free energy barrier separating the two minima, ranging from 0 to 0.5 kcal/mol for the former cation and from 3 to 6 kcal/mol for the latter. Interestingly enough, a similar trend was observed over twenty years ago for ammonium axially constrained along the C_6 symmetry axis of the toluene aromatic ring, using a different academic pairwise additive force field [24].

Non-additive CHARMM Drude force field Resorting to the CHARMM Drude polarizable force field essentially results in lowering the free energy of the contact minima for all complexes, compared to the pairwise additive description, while leaving the depth of the solvent-separated minima globally unchanged (see Fig. 2c, d). As the cation, either ammonium or guanidinium, approaches the π–electron cloud of the aromatic compound, the auxiliary Drude particles attached to the latter move with respect to the real atoms, resulting in an induced dipole ranging at the contact minimum between 0.2 and 0.4 D for toluene and para-cresol, respectively. For para-cresol, explicit description of polarization shifted down the contact minima by 1.0 to 1.5 kcal/mol, compared with the CHARMM36 pairwise additive force field, corresponding to the same free energy as the solvent-separated minima and the dissociated state for ammonium, albeit leaving the situation qualitatively unchanged for guanidinium—i.e., contact > solvent-separated ≈ dissociated state. At variance with the pairwise additive description, the intimate pair formed by toluene and ammonium is predicted by the Drude force field to be more stable than its dissociated state with a free energy difference of 1.9 kcal/mol. For the guanidinium:toluene complex, the contact and the solvent-separated binding motifs are as probable as the dissociated state. The main quantitative and qualitative difference between the additive and the non-additive description is observed for cation:3-methylindole complexes. The use of Drude particles predicts the contact complexes to be the most stable conformation with a free energy of 2.8 and 1.9 kcal/mol for the ammonium and the guanidinium cations, respectively.

When compared to quantum chemical calculations in the gas phase, standard pairwise additive force fields are prone to underestimate grossly the gas-phase interaction energies of cation–π complexes [13, 19, 24, 50]. It should be emphasized here that empirical force fields are tailored to describe the condensed phase and, thus, confronting their relative ability to predict gas-phase quantities, compared to ab initio calculations, is somewhat unfair. One possible way to recover the correct gas-phase cation–π interaction energy in vacuum consists in fine-tuning the Lennard-Jones contribution to the additive force fields either through reparametrization [50], or by introducing an ad hoc function, for instance, by means of a 10–12 or a 4–12 potential [13, 24]. Another strategy toward the accurate description of cation–π systems consists in introducing an explicit polarization term in the potential energy function by means of Drude particles [37, 49, 52], or more elaborate polarizable models relying, for example, on charge flow and dipole polarizabilities [19, 20, 53, 54].

Experiments carried out in an aqueous medium usually predict binding affinities for cation–π complexes much lower than those computed in the gas phase. This apparent discrepancy is now well understood, and the effect of the solvent on the strength of cation–π complexes has been quantified by theoretical calculations [55]. In light of drug-binding experiments, it was shown that cation–π interactions enhance the free energy of association by nearly 3 kcal/mol [9, 56]. CHARMM36 pairwise additive force field is clearly not able to account for such a stabilization free energy. It chiefly

predicts a dissociation for the various cation–π complexes examined here, or slightly favors the solvent-separated minimum by less than 1 kcal/mol for 3-methylindole complexes. Resorting to the CHARMM Drude force field improves the description for 3-methylindole and toluene interacting with ammonium and guanidinium, but leaves the situation essentially unchanged for those complexes involving para-cresol. The effect of explicit polarization further depends on the nature of the cation. When ammonium is the polarizing cation, wherein the charge is rather well localized, the Drude model predicts a binding affinity of 1.9 and 2.8 kcal/mol for toluene and 3-methylindole, respectively. In sharp contrast, for guanidinium, the CHARMM Drude force field leads to a binding affinity of 1.1 kcal/mol for 3-methylindole, but fails to describe for toluene a contact minimum more stable than the dissociated state.

Evidence of cation–π interaction in an aqueous environment is documented in the literature. Lamoureux et al. report a free energy minimum of about −3.5 kcal/mol, inferred from their PMF for the ammonium:toluene complex, using optimized CHARMM Lennard-Jones parameters [49]. Similarly, Chipot et al. report a free energy minimum of −5.5 kcal/mol for the same complex, employing an ad hoc corrective potential in the AMBER force field [24]. Experimental estimates of the association constant for a related cation–π complex have also been reported by Schneider et al. [57], ranging between 2.7 and 3.3 M^{-1}. In stark contrast to the above PMF calculations, wherein the approach of the cation toward the π–electron cloud is geometrically restrained along the normal to the latter, experimental measurements of the binding affinity reflect an orientationally averaged cation–π association. In addition, Dougherty et al. have estimated the interaction energy of methyl ammonium and benzene in water to be on the order of −5.5 kcal/mol, and from their survey of PDB complexes, they have observed that energetically favorable cation–π interactions are rarely completely buried within proteins, but prefer instead to be solvent-exposed [8, 14].

To check whether the interaction with the solvent is the source of the discrepant PMFs for toluene and para-cresol, we computed the corresponding free energy profiles in vacuum. The minimum for para-cresol is about −17.5 kcal/mol, and that for toluene is about −21 kcal/mol. Both minima are found around 2.9 Å in vacuum. It should be noted that the presence of the solvent clearly reduces the cation–π binding by nearly 17 kcal/mol for para-cresol and 19 kcal/mol for toluene. However, we observed a similar trend for the two PMFs (in water and in vacuum), hence suggesting that the discrepant results are rooted primarily in the intermolecular forces in the cation–π complexes.

To identify the differences in the free energy profiles, we also performed an interaction energy calculation in vacuum for the ammonium:para-cresol and ammonium:toluene pairs at their equilibrium distance determined from the PMFs. A 1-ns trajectory formed of 2000 snapshots was generated for each cation–π pair. The average interaction energy was -29.1 ± 1.6 and -34.8 ± 1.6 kcal/mol for the ammonium:para-cresol and ammonium:toluene complex, respectively. The interaction energy follows a similar trend, compared to the free energy, suggesting that the origin of the discrepant results stems from the electrostatic term of the potential energy function.

The partial-charge distribution of toluene and para-cresol atoms reinforces this observation. The partial charge on the carbon atom connected to the hydroxyl group is + 0.297 e.c.u. (electron charge unit) in para-cresol, whereas the same carbon atom in the toluene ring has a partial charge of −0.133 e.c.u. The presence of a positive charge in the para-cresol ring, thus, weakens the molecular quadrupole of the latter, and leads to a reduced interaction with ammonium.

4 Conclusion

From the results of the present contribution, it is clear that introducing explicit polarization by means of a set of Drude particles globally enhances the description of cation–π complexes, compared to a pairwise additive force field. Yet, describing the subtle balance within the intermolecular forces responsible for the recognition and association of cations to aromatic compounds by means of cost-effective polarizable models still remains a daunting challenge. Orabi and Lamoureux were able to recover systematically consistent binding affinities for various cation–π pairs, using an optimized version of the CHARMM Drude force field [49]. Toward this end, they corrected the original polarizable potential energy function, following a strategy akin to that employed for additive force fields [13, 24, 50], namely through fine-tuning of the Lennard-Jones interaction potential by means of ad hoc functionals or parameters handled using the NBFIX facility of the CHARMM36 force field. This strategy has, however, limitations of its own. First, the parametrization strongly depends on the level of theory of the quantum mechanical calculations and on the basis set employed to compute the reference quantities. Orabi and Lamoureux resorted to the standard 6–311++G(*d*,*p*) basis set, which has proven suboptimal in quantifying electric properties, compared to specialized basis sets, e.g., Sadlej, ELP [58, 59]. Moreover, such a reparametrization approach lacks generality, as it requires the optimization of specific pairs of Lennard-Jones parameters for every cation–π complexes, while introducing an additional interaction site at the center of the aromatic ring. This scheme may not be easily extended to anion–π complexes, wherein large penetration and exchange induction effects have proven difficult to capture [20]. Last, cation–π pairing represents one possible

association mode in proteins, but more sophisticated binding motifs, such as cation–π–cation trimers, are also frequently observed. The parameters calibrated for dimers may not necessarily transfer in a straightforward fashion to a trimeric organization. Put together, our data suggest that, while explicit inclusion of induction phenomena definitely improves the description of cation–π interactions, even when in the absence of an ad hoc reparametrization, there is still room for improvement toward reproducing systematically the subtle balance of the intermolecular forces at play.

Acknowledgements The authors gratefully acknowledge fruitful exchanges with Javier Luque (Universidad de Barcelona, Spain) over the years. This work was supported by the National Institutes of Health (through Grant No. NIGMS-2R01 GM-072558) and by the Prof@Lorraine initiative of the University of Lorraine.

References

1. Sunner J, Nishizawa K, Kebarle P (1981) J Phys Chem 85:1814
2. Deakyne CA, Meot-Ner M (Mautner) (1985) J Am Chem Soc 107:474
3. Legon AC, Millen DJ (1987) Acc Chem Res 20:39
4. Perutz MF, Fermi G, Abraham DJ, Poyart C, Bursaux E (1986) J Am Chem Soc 108:1064
5. Burley SK, Petsko GA (1986) FEBS Lett 203:139
6. Levitt M, Perutz MF (1988) J Mol Biol 201:751
7. Dougherty DA (1996) Science 271:163
8. Gallivan JP, Dougherty DA (1999) Proc Natl Acad Sci USA 96:9459
9. Dougherty DA (2013) Acc Chem Res 46:885
10. Dougherty DA, Stauffer DA (1990) Science 250:1558
11. Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, Toker L, Silman I (1991) Science 253:872
12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28:235
13. Minoux H, Chipot C (1999) J Am Chem Soc 121:10366
14. Gallivan JP, Dougherty DA (2000) J Am Chem Soc 122:870
15. Flocco MM, Mowbray SL (1994) J Mol Biol 235:709
16. Mitchell JBO, Nandi CL, McDonald IK, Thornton JM, Price SL (1994) J Mol Biol 239:315
17. Kumpf RA, Dougherty DA (1993) Science 261:1708
18. Mecozzi S, West AP Jr, Dougherty DA (1996) Proc Natl Acad Sci USA 93:10566
19. Dehez F, Soteras I, Luque FJ, Ángyán J, Schulten K, Chipot C (2007) J Chem Theory Comput 3:1914
20. Archambault F, Chipot C, Gutiérrez IS, Luque FJ, Schulten K, Dehez F (2009) J Chem Theor Comput 5:3022
21. Luque FJ, Dehez F, Chipot C, Orozco M (2011) Wires Comput Mol Sci 1:844
22. Stone AJ (1996) The theory of intermolecular forces. Clarendon Press, Oxford
23. Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:4177
24. Chipot C, Maigret B, Pearlman DA, Kollman PA (1996) J Am Chem Soc 118:2998
25. Lamoureux G, Roux B (2003) J Chem Phys 119:3025
26. Lemkul JA, Huang J, Roux B, MacKerell AD (2016) Chem Rev 116:4983
27. Stone AJ (1985) Mol Phys 56:1065
28. Jiang W, Hardy DJ, Phillips JC, Mackerell AD Jr, Schulten K, Roux B (2011) J Phys Chem Lett 2(2):87
29. Chipot C (2014) Wiley Interdiscip Rev Comput Mol Sci 4:71
30. Huang J, Lopes PEM, Roux B, MacKerell AD Jr (2014) J Phys Chem Lett 5:3144
31. Jo S, Kim T, Iyer VG, Im W (2008) J Comput Chem 29:1859
32. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel L, Kalé RD, Schulten K (2005) J Comput Chem 26:1781
33. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926
34. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD Jr (2010) J Comput Chem 31:671
35. Lamoureux G, MacKerell AD Jr, Roux B (2003) J Chem Phys 119:5185
36. Lopes PEM, Lamoureux G, Roux B, Mackerell AD (2007) J Phys Chem B 111:2873
37. Orabi EA, Lamoureux G (2012) J Chem Theory Comput 8:182
38. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102:3586
39. Tuckerman ME, Berne BJ, Martyna GJ (1992) J Phys Chem B 97:1990
40. Miyamoto S, Kollman PA (1992) J Comput Chem 13:952
41. Andersen HC (1983) J Comput Phys 52:24
42. Feller SE, Zhang YH, Pastor RW, Brooks BR (1995) J Chem Phys 103:4613
43. Darden TA, York DM, Pedersen LG (1993) J Chem Phys 98:10089
44. Humphrey W, Dalke A, Schulten K, Molec J (1996) Graphics 14:33
45. Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C (2015) J Phys Chem B 119:1129
46. Fu H, Shao X, Chipot C, Cai W (2016) J Chem Theory Comput 12:3506
47. Lesage A, Lelièvre T, Stoltz G, Hénin J (2017) J Phys Chem B 121:3676
48. Wang L (2012) Annu Rev Biochem 81:615
49. Orabi EA, Lamoureux G (2018) J Phys Chem B 122:2251
50. Khan HM, Grauffel C, Broer R, MacKerell AD, Havenith RWA, Reuter N (2016) J Chem Theory Comput 12:5585
51. Chandler D (1987) Introduction to modern statistical mechanics. Oxford University Press, Oxford
52. Lamoureux G, Orabi EA (2012) Mol Sim 38:704
53. Soteras I, Curutchet C, Bidon-Chanal A, Dehez F, Ángyán J, Orozco M, Chipot C, Luque FJ (2007) J Chem Theory Comput 3:1901
54. Ansorg K, Tafipolsky M, Engels B (2013) J Phys Chem B 117:10093
55. Xu J, Voth GA (2005) Proc Natl Acad Sci U S A 102(19):6795
56. Anderson MA, Ogbay B, Arimoto R, Sha W, Kisselev OG, Cistola DP, Marshall GR (2006) J Am Chem Soc 128:7531
57. Schneider HJ, Schiestel T, Zimmermann P (1992) J Am Chem Soc 114:7698
58. Celebi N, Ángyán JG, Dehez F, Millot C, Chipot C (2000) J Chem Phys 112:2709
59. Chipot C, Dehez F, Ángyán JG, Millot C, Orozco M, Luque FJ (2001) J Phys Chem A 105:11505

II. Calculs d'énergie libre

II.1. Analyse des calculs Free Energy Perturbation (FEP)

L'estimation précise de différences d'énergie libre associées à de nombreux processus chimiques (solvatation, transfert, mutation, ...) reste à bien des égards une des gageures actuelles de la modélisation.[22] L'estimation fiable et précise d'énergies libres nécessite un échantillonnage optimal de l'espace configurationnel, un but difficile à atteindre pour des systèmes moléculaires complexes. L'augmentation de la puissance des ordinateurs et le recours par exemple aux GPU[23–25] permettent de repousser sans cesse les échelles de temps accessibles aux simulations moléculaires sans pour autant parvenir à atteindre les temps caractéristiques de bon nombre de processus biologiques : repliement de protéines, transport, changement conformationnel, ... De nombreuses approches de calculs sont développées depuis plusieurs années pour permettre d'évaluer des énergies libres grâce aux simulations microscopiques, elles peuvent être associées aux classes suivantes : MM/PBSA,[26] perturbation thermodynamique,[27] distribution de probabilités et histogrammes,[28] méthodes hors-équilibre,[29] méthodes hybride Monte-Carlo/dynamique moléculaire,[30] méta-dynamique,[31] force de biais adaptif,[32] ...

Une des stratégies les plus fréquemment employée pour calculer la différence d'énergie libre associée à une mutation alchimique est la perturbation thermodynamique (Free Energy Perturbation). Dans cette approche un système chimique 0 est lentement transformé réversiblement en un système chimique 1 le long d'un paramètre λ. La différence d'énergie libre est calculée comme $\exp(-\beta\Delta A) = \langle \exp(-\beta\Delta U) \rangle_0$ ou ΔU est la différence des énergies potentielles $U(x,\lambda = 1) - U(x,\lambda = 0)$ et $\langle \dots \rangle_0$ est une moyenne configurationnelle représentative de l'état de référence. La même différence d'énergie libre doit être estimée à partir de la transformation inverse de l'état 1 vers l'état 0, on parle alors de transformation bidirectionnelle. La combinaison des données aller-retour peut se faire par la méthode BAR (Bennet Acceptance Ratio)[33] qui permet en pondérant (facteur w) les contributions des états 0 et 1 de minimiser la variance σ^2 de ΔA.

$$\exp(-\beta\hat{\Delta}A) = \frac{\langle w \exp(-\beta\Delta U/2)_0}{\langle w \exp(+\beta\Delta U/2)_1}$$

Un des conditions de convergence du calcul est d'échantillonner à λ donné les mêmes micro-états lors des transformations aller et retour. Cette condition se trouve remplie lorsque le recouvrement des distributions de probabilités, P0(ΔU) et P1(ΔU), représentant respectivement l'état 0 et l'état 1 est maximal. L'analyse détaillée de ce recouvrement n'est pas directe mais est essentielle pour évaluer la stratégie de stratification (le découpage en λ) et d'échantillonnage et ainsi garantir la validité statistique du résultat. Un des autres ingrédients nécessaires à la bonne réalisation d'un calcul d'énergie libre et l'évaluation rigoureuse de l'erreur statistique associée, c'est à dire la variance du calcul d'énergie libre.

Nous avons développé un ensemble d'outils en langage Tcl (figure 4) capable réaliser l'ensemble de ces analyses essentielles à la bonne conduite d'un calcul d'énergie libre. Ceux-ci ont été

rassemblés sous la forme d'un plugin Tcl, ParseFEP,[34] distribué avec le programme VMD[35] permettant ainsi à tous les utilisateurs de la communauté de réaliser simplement ces analyses.

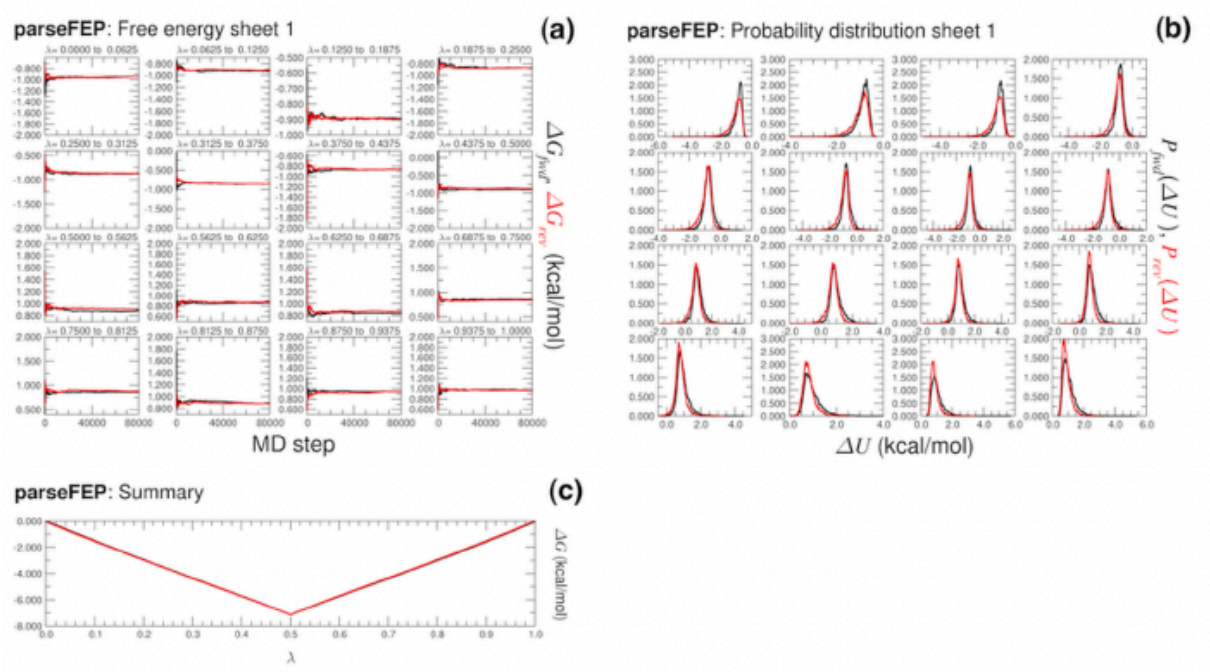


Figure 4. Capture d'une sortie d'écran plugin ParseFEP obtenu dans le cas de la transformation de l'éthane en éthane à partir d'une stratification en 16 fenêtres. L'évolution en fonction du temps de l'énergie libre par fenêtre (a), des histogrammes des probabilités de distributions (b) et de l'énergie libre totale (c).

II.2. In silico Alanine Scaning

L'Alanine Scanning Mutagenesis (ASM)[36,37] est une méthode souvent utilisée pour étudier le rôle spécifique de tel ou tel acide aminé dans le repliement et la stabilité de protéines, pour l'association protéine-ligand ou bien encore protéine-protéine. Plus récemment cette technique a également été employée pour pallier les limitations des études structurales de protéines membranaires, en mettant au point des mutants thermostables de protéines membranaires (en particulier les GPCRs) plus facilement cristallisables.[38]

Les techniques d'Alanine Scaning in Silico représentent une alternative intéressante aux techniques expérimentales même si ces dernières peuvent aujourd'hui se faire à haut débit. En pratique, chaque acide aminé d'une séquence est muté en Alanine, la variation d'énergie libre associée à la mutation permet de caractériser le rôle de l'acide aminé muté dans la formation de complexes ou bien encore dans la stabilité intrinsèque des protéines. L'utilisation courante de ce genre d'approche est limitée par l'effort humain nécessaire à la mise en œuvre de multiples calculs d'énergie libre. Tirant parti des développements faits pour le plugin ParseFep, nous avons développé un plugin pour VMD (figure 5) qui automatise (i) les taches de préparation de setup et des fichiers d'entrée pour chaque mutation et (ii) l'analyse de la convergence et la présentation synthétiques des résultats, minimisation l'intervention manuelle.[39] Cet outil fait appels aux calculs d'énergie libre de type Free Energy Perturbation (FEP), permettant de prédire quantitativement les variations d'énergie libre associées aux mutations en alanine. Grace à une interface graphique conviviale, notre plugin permet

d'identifier rapidement les mutations favorables. Le plugin aide également l'utilisateur final à évaluer la fiabilité du calcul par une inspection visuelle rapide.

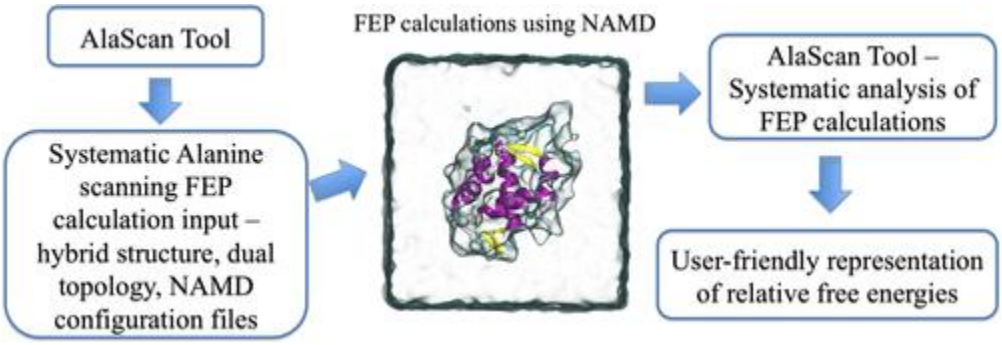


Figure 5. Schéma de principe du plugin AlaScan de VMD

II.3. Solubilité en phase organique

La solubilité est une quantité physico-chimique particulièrement importante dans le contexte de la formulation de nouveaux composés d'intérêt pharmaceutique. Cette propriété permet d'inférer la disponibilité d'une substance pour son absorption *in vivo* ou bien encore comme substrat pour sa transformation chimique. Avec nos partenaires de l'École Nationale Supérieure d'Agronomie de Nancy, nous avons mis au point une stratégie permettant de prédire, à partir de la détermination expérimentale d'une seule solubilité, l'ensemble des solubilités d'un composé dans divers solvants. Celle-ci repose essentiellement sur des calculs de type Free Energy Perturbation (FEP) employés ici pour prédire les énergies libres de solvatation de la quercétine. La quercétine appartient à la famille des flavonoïdes,[40] composés connus pour leurs propriétés anti-oxydantes. Ces derniers intéressent particulièrement nos partenaires qui tentent de transformer ces molécules (pour les rendre plus solubles dans les préparations pharmaceutiques), par biocatalyse enzymatique, en phase organique à l'aide de lipases immobilisées. Dans ce travail nous avons considéré la quercétine dans une série de différents solvants : eau, acétonitrile, acétone, *tert*-amyl alcool (2-methyl-2-butanol) et chloroforme. Pour la quercétine, des jeux de charges atomiques ont été dérivées à partir du potentiel électrostatique créé par la molécule (évalué au niveau quantique) perturbée par les divers solvants. Puis nous avons réalisé des dynamiques moléculaires tout atomes au travers desquelles nous avons calculé les énergies libres de solvatation. Chaque transformation é été faite dans les deux sens (apparition et disparition) ce qui, grâce à la méthode *Bennet Acceptance Ratio* (BAR),[33] a permis d'améliorer la précision statistique des résultats. Par ailleurs pour minimiser l'erreur systématique chaque calcul a été répété 5 fois, les résultats présentés dans le tableau 4 sont une moyenne de ces réalisations.

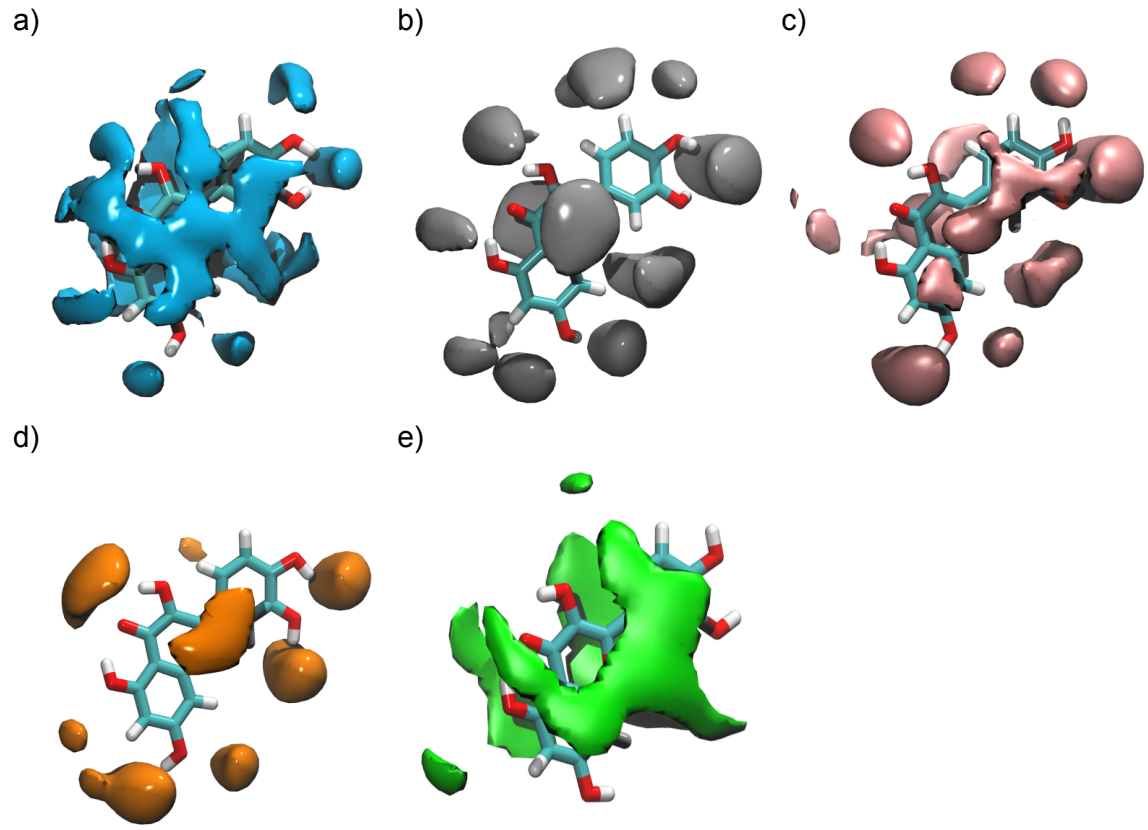


Figure 6. Densité de solvatation autour de la quercétine dans les différents solvants, (a) eau, (b) acétonitrile, (c) acétone, (d) *tert*-amyl alcool et (e) chloroforme.

Les différences relatives entre les énergies libres d'hydratation et de solvatation (ΔΔGeau-solvant) obtenues par la simulation sont tout à fait compatibles avec les mêmes valeurs déduites des mesures de solubilité expérimentales. Il est à noter que pour le chloroforme la solubilité est tellement faible qu'elle n'a pu être mesurée. L'organisation des différents solvants autour de la quercétine, et en particulier des groupements hydroxyles, est tout à fait compatible avec le pouvoir solubilisant de ces solvants (voir figure 6). Dans le cas de l'alcool et de l'acétone, ces groupements sont très bien solvatés, ils le sont en revanche très mal ou pas du tout dans le cas de l'eau et du chloroforme, l'acétonitrile représente lui une situation intermédiaire.

Nos résultats[41] démontrent qu'à partir de calculs théoriques et d'une seule mesure expérimentale de solubilité, il est tout à fait possible de prédire la solubilité d'un composé dans un ensemble d'environnements différents. Nous avons étendu notre étude au cas du dimère de quercétine. Nos simulations montrent que dans l'alcool ou dans l'acétone le dimère se dissocie systématiquement, résultats tout à fait compatibles avec l'important pouvoir solubilisant de ces solvants pour la quercétine. Dans le cas de l'acétonitrile, de l'eau et du chloroforme nous avons pu estimer les énergies libres d'association du dimère par la méthode FEP suivant le schéma de la figure 7. Les résultats montrent qu'il est hasardeux de tenter de prédire la solubilité d'une espèce à partir de ces valeurs, en effet l'association dans le chloroforme apparait moins importante que dans l'eau alors que la solubilité est plus faible. Prédire la contribution du cristal à la solubilité reste une gageure qui ne peut être réduite à l'étude du dimère.[42]

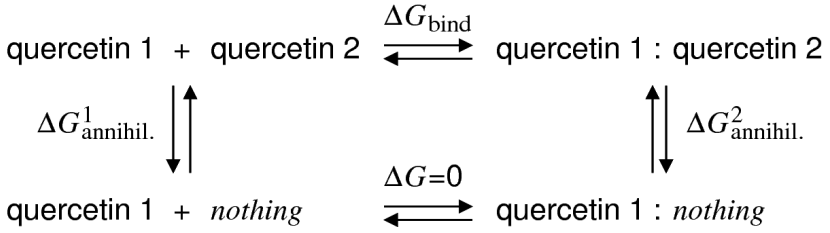


Figure 7. Énergie libre d'association de la quercétine, ΔG_{bind}, mesurée à partir d'un schéma de double annihilation, où l'interaction d'un des solutés avec son environnement et son homologue est progressivement éteinte, *viz.* $\Delta G_{bind} = \Delta G^{1}_{annihil.} - \Delta G^{2}_{annihil.}$.

II.4. Énergie libre d'association ADN-photo sensibilisateur

La compréhension de l'interaction de petites molécules avec les acides nucléiques, et l'ADN, en particulier, a un intérêt biologique considérable. En effet, la sensibilisation de l'ADN par des molécules exogènes ou endogènes peut être liée à l'apparition de lésions et à la modification de l'ADN.[43] Celles-ci peuvent à leur tour évoluer vers la mort cellulaire ou vers des modifications génétiques responsables de cancer.[44] L'interaction de l'ADN avec des molécules sensibilisantes est également largement exploitées pour détruire les cellules cancéreuses en endommagant leur ADN.[45] Malgré leur utilité, l'étude des sensibilisateurs d'ADN reste difficile comme le démontre le manque de structures résolues expérimentalement pour les complexes non-covalents ADN-sensibilisateurs. La variété des motifs d'interaction, association dans le petit sillon, intercalation, insertion rendent difficile la rationalisation de la chimie d'endommagement consécutive à l'association de photo-sensibilisateurs à l'ADN.

Pour prédire de manière quantitative l'énergie libre d'association ADN/molécule, nous avons utilisé un protocole permettant de prendre correctement en compte la forte réorganisation conformationelle de l'ADN lors de la formation du complexe. Nous avons utilisé l'approche FEP en intégrant dans les différentes transformations alchimiques (figure 8) des restreintes positionnelles et orientationnelles garantissant la réversibilité du couplage de l'ADN et de son ligand ainsi que des restreintes conformationnelles sur l'ADN permettant de restreindre l'entropie configurationnelle du système et ainsi faciliter la convergence des calculs d'énergie libre. Nous avons appliqué cette stratégie à un sensibilisateur connu de l'ADN, la benzophénone. En suivant le cycle thermodynamique de la figure 8, nous avons calculé l'énergie libre d'association pour deux motifs possibles d'interaction : association au petit sillon et double insertion, et nous avons pu montrer que l'association au petit sillon était le mode d'interaction le plus probable.[46]

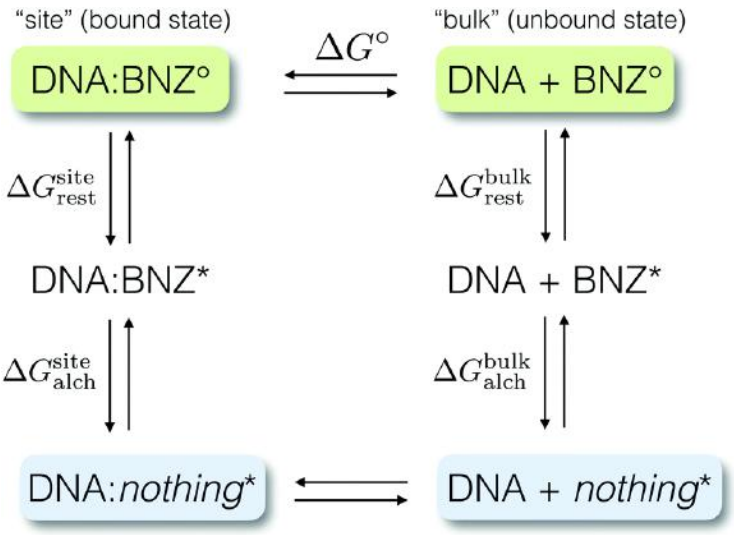


Figure 8. Cycle thermodynamique employé pour déterminer l'énergie libre d'association ADN/benzophénone. * décrit un état restreint. La contribution de chaque restreinte à l'énergie libre est prise en compte explicitement dans l'estimation totale.

La contribution de chaque restreinte à l'énergie libre est prise en compte explicitement dans l'estimation totale.

Ci-après sont reproduits 2 articles illustrant mes travaux dans le domaine des calculs d'énergie libre :

- Chebil, L.*, Chipot, C., Archambault, F., Humeau, C., Engasser, J. M., Ghoul, M. & Dehez, F.* Solubilities Inferred from the Combination of Experiment and Simulation. Case Study of Quercetin in a Variety of Solvents. *J. Phys. Chem. B* **114**, 12308–12313 (2010).
- Gattuso, H., Dumont, E., Chipot, C., Monari, A. & Dehez, F.* Thermodynamics of DNA: sensitizer recognition. Characterizing binding motifs with all-atom simulations. *Phys. Chem. Chem. Phys.* **18**, 33180–33186 (2016).

Solubilities Inferred from the Combination of Experiment and Simulation. Case Study of Quercetin in a Variety of Solvents

Latifa Chebil,*,† Christophe Chipot,‡,§,∥ Fabien Archambault,‡ Catherine Humeau,† Jean Marc Engasser,† Mohamed Ghoul,† and François Dehez*,‡

Laboratoire d'ingénierie des biomolécules (LIBio), Nancy Université, 2 avenue de la forêt de Haye, 54500 Vandœuvre-lès-Nancy, France, Equipe de Dynamique des Assemblages Membranaires, UMR CNRS-UHP No. 7565, Nancy Université, BP 70239, 54506 Vandœuvre-lès-Nancy, France, and Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana-Champaign, 405 North Mathews, Urbana, Illinois 61801

Received: May 19, 2010; Revised Manuscript Received: July 8, 2010

A strategy to infer solubilities from the combination of experiment and all-atom simulations is presented. From a single experimental estimate, the solubility of a substrate can be predicted in various environments from the related free energies of solvation. In the case of quercetin, the methodology was shown to reproduce the experimental solubilities in chloroform, water, acetonitrile, acetone, and *tert*-amyl alcohol within 0.5 log unit. The reliability of the estimates is markedly correlated to the accuracy of the experimental measure and to both the accuracy and precision of the computed free energies of solvation.

Introduction

Solubility is a physicochemical property of paramount importance in the formulation of new pharmaceutical compounds. Precise knowledge of solubility is highly desirable to predict the availability of a chemical species in the context of its absorption in vivo, or of a given substrate in the framework of an organic synthesis. Experimental determination of solubilities, however, constitutes a daunting and time-consuming endeavor when a large number of compounds and/or solvents are involved. In silico approaches, therefore, represent an appealing alternative to predict solubilities. Approximate quantitative structure−property relationship (QSPR) methods[1−3] have been widely employed to estimate this physicochemical quantity. When looking for generality and precision, more rigorous routes have, however, to be explored. The accurate computation of the solubility of a given chemical species requires the knowledge of the free-energy differences associated with its transfer in the solvent and its dissociation from the crystal. Determining in silico the latter quantity, nevertheless, remains somewhat cumbersome. In a series of recent articles, Lindfors and colleagues.[4−7] proposed a predictive model capable of estimating reasonably solubilities in water. For a variety of drugs, they computed, using all-atom Monte Carlo simulations, the free energy of solvation in pure melt at 673.17 K and the free energy associated with the supercooling of the melt from 673.17 to 298 K. Combination of these quantities with hydration free energies allows the solubility of amorphous drugs to be determined in water. In addition, the solubility of the crystal can be inferred from that of the amorphous solid on the basis of experimental melting data, viz., entropy and temperature of melting.

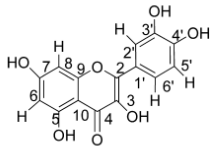


Figure 1. Molecular structure of quercetin.

Evaluation of the crystal contribution represents an equally difficult undertaking, which evidently cannot be performed routinely. Moreover, in the context of chemical synthesis, knowledge of the solubility is important not only in water but also in a variety of solvents that can be used to mediate chemical reactions. Here, a strategy is developed to infer the solubilities of a given compound in different environments by combining theory and experiment. Molecular dynamics simulations and free-energy perturbation (FEP) methodology[8] are employed to predict the relative free energies of solvation in a variety of solvents. It is shown that using these data together with a single experimental solubility measurement allows the full solubility scale to be determined with an appreciable accuracy. The methodology is illustrated in the case of quercetin (3,3′,4′,5,7-pentahydroxy flavone) (see Figure 1), one of the most studied members of the flavonoid family. This compound possesses a strong antioxidant activity, central in preventing the risk of cardiovascular, neurodegenerative, and cancer diseases.[9,10] It is an overall hydrophobic species, which needs to be modified to be incorporated efficiently in a chemical formulation. This alteration can be achieved by means of enzymatic catalysis, employing for instance immobilized *Pseudomonas cepacia* lipase in organic phases,[11] the effectiveness of the process being strongly dependent on the solubility of the substrate.[12]

* To whom correspondence should be addressed. L.C.: tel, +33-383596195; fax, +33-383595778; e-mail, Latifa.chebil@ensaia.inpl-nancy.fr. F.D.: tel, +33-383684098; fax, +33-383684387; e-mail, Francois.Dehez@srsmc.uhp.

† Laboratoire d'ingénierie des biomolécules (LIBio). E-mail: C.H., Catherine.Humeau@ensaia.inpl-nancy.fr; J.M.E., Jean-Marc.Engasser@ensaia.inpl-nancy.fr; M.G., Mohamed.Ghoul@ensaia.inpl-nancy.fr.

‡ Equipe de Dynamique des Assemblages Membranaires. E-mail: C.C, Christophe.Chipot@srsmc.uhp-nancy.fr; F.A., Fabien.Archambault@srsmc.uhp-nancy.fr.

§ University of Illinois at Urbana-Champaign.

∥ On leave from Équipe de dynamique des assemblages membranaires, UMR 7565, Nancy Université, BP 239, 54506 Vandoeuvre-lès-Nancy cedex, France.

10.1021/jp104569k
Published on Web 08/17/2010

Methodology

Molecular Models. The OPLS (optimized potential for liquid simulations) force field developed by Jorgensen and colleagues.[13] was used to describe all the solvents utilized in this work, e.g., water (TIP3P),[14] acetonitrile,[15,16] acetone,[17] and chloroform.[18] *tert*-Amyl alcohol was modeled using the *tert*-butyl alcohol OPLS parameters. For quercetin, all-atom OPLS parameters were used whenever available. For the C2−C1′ bond, a torsional potential was optimized in the spirit of the OPLS force field to reproduce the potential energy surface determined quantum-mechanically at the DFT (B3LYP)[19]/6-31G* level (see Supporting Information). This surface was obtained on the basis of the equilibrium geometry of quercetin, for which the dihedral angle O−C2−C1′−C6′ (φ) was varied by 10° increments. All quantum-chemical calculations were performed using Gaussian03.[20] The resulting torsional potential can be written

$$E(\varphi) = \frac{V_1}{2}[1 + \cos(\varphi)] + \frac{V_2}{2}[1 \cos(2\varphi)] \qquad (1)$$

with $V_1 = -0.40$ kcal/mol and $V_2 = 6.58$ kcal/mol.

The electrostatic fingerprint of quercetin was modeled by means of different sets of atomic charges to account in an average sense for the polarization due to the various environments. In practice, a grid of approximately 15 000 points (distributed spatially between envelopes corresponding to 2 and 4 times the van der Waals radii of the participating atoms) was built around the molecule. The electrostatic potential was evaluated at each point of the grid at the B3LYP/6-31G* level, employing a dielectric continuum representation of the different environments by means of the PCM model of solvation[21] (polarizable continuum model). In a final step, the OPEP[22] package was used to derive the atomic charges from the electrostatic potential maps (see Supporting Information).

Molecular Dynamics (MD) Simulations. For each environment, quercetin, either as a monomer or a dimer, was immersed in a pre-equilibrated cubic box containing approximately 500 explicit molecules of solvent. MD trajectories of 0.5 ns were performed to thermalize each system prior to a production run of 1.5 ns. All MD simulations were carried out using the NAMD package,[23] in the isobaric−isothermal ensemble. The temperature and the pressure were kept constant at 323 K and 1 atm using, respectively, Langevin dynamics and the Langevin piston method.[24] The particle mesh Ewald (PME) algorithm[25] was used to account for long-range electrostatic interactions. The equations of motion were integrated by means of a multiple-time-step algorithm[26] with a time step of 2 and 4 fs for short- and long-range interactions, respectively.

Free Energy and Solubility Calculations. The following strategy was used to compute the solvation free energies. Transfer of quercetin from the low-pressure gaseous phase to the bulk solvent was modeled by means of a double-annihilation[27] of the solute in the two environments, as illustrated in Figure 2.

The reaction pathway connecting the initial and the final states of the annihilation and creation transformations was stratified[28] in a series of 24 strata, or windows, of uneven widths. In each window, 4000 equilibration steps were generated prior to the production of 26 000 steps, over which the ensemble averages were calculated. This sampling strategy represents a total of 1.44 ns for each transformation. In addition to the use of strata of decreasing width toward the end points of the free-energy calculation, which constitutes an artificial and incomplete

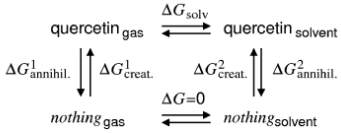


Figure 2. Solvation free energy, ΔG_{solv}, of quercetin measured from a double-annihilation scheme, whereby the interaction of the solute with its environment is progressively turned out, viz. $\Delta G_{solv} = \Delta G^1_{annihil} - \Delta G^2_{annihil}$, or turned on, viz. $\Delta G_{solv} = \Delta G^2_{creat} - \Delta G^1_{creat}$. In practice, neglect of perturbed intramolecular interactions obviates the need to compute the gas-phase contribution, $\Delta G^1_{annihil}$, to the solvation free energy, which is justified by the similarity of the atomic charges determined in the different solvents.

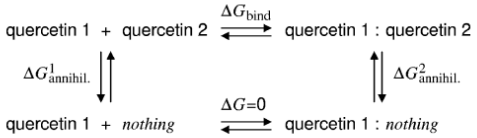


Figure 3. Binding free energy, ΔG_{bind}, of quercetin measured from a double-annihilation scheme, whereby the interaction of the free solute with its environment (left leg) and with its counterpart in the dimer (right leg) is progressively turned out, viz. $\Delta G_{bind} = \Delta G^1_{annihil} - \Delta G^2_{annihil}$. Noteworthily, the left-hand side contribution, $\Delta G^1_{annihil}$, corresponds to the solvation free energy of quercetin, the determination of which is outlined in Figure 2.

remedy to circumvent singularities in the van der Waals potential, a soft-core correction was introduced in the form of a separation-shifted scaling scheme[29] to prevent incoming particles from colliding with the surrounding solvent molecules.

To improve the precision of the estimates, the calculations in each solvent were run bidirectionally, and the free energy was obtained employing the Bennett acceptance ratio (BAR) method,[30] from whence a statistical error was inferred.[31] In general, the statistical error associated with this variance is somewhat smaller than the root mean-square deviation (rmsd) inferred from the free-energy estimates that were measured independently. Furthermore, to enhance the accuracy of the estimates,[28] in relation to finite-sampling biases, each bidirectional measure was conducted five times. The solvation free energies reported in the Results and Discussion correspond to the averages over the five realizations, and their reliability to the root-mean-square deviation (rmsd) thereof. Given the limited cost of the individual free-energy calculations, no specific attempt has been made here to optimize the computational effort against the accuracy of the free-energy estimates. Whereas reducing the number of bidirectional simulations to a lesser, albeit still reasonable number of independent observables, viz., typically three instead of five, would marginally decrease the precision of the estimate, it is strongly recommended, to perform the free-energy calculations bidirectionally to guarantee acceptably low statistical errors.[32]

The binding, or dimerization, free energies of quercetin were also measured in the different solvents, usng the FEP method, based on the thermodynamic cycle shown in Figure 3.

To reach this goal, a double-annihilation scheme was employed, whereby the interaction of the solute with its environment is obliterated in the free and in the bound states. It is worth noting that the first contribution to the binding affinity is equal to the solvation free energy determined previously. The sampling and stratification strategies are identical to that employed for determining solvation free energies. Although it would be desirable to provide BAR estimates, as was done for

TABLE 1: Experimental and Calculated Solubilities (*c*, mol/L) at 323 K of Quercetin in Chloroform, Water, Acetonitrile, Acetone, and *tert*-Amyl Alcohol[a]

	chloroform	water	acetonitrile	acetone	*tert*-amyl alcohol
log c_{exp}	insoluble[b]	−4.52[b]	−2.27[c]	−1.10[c]	−1.17[c]
ΔG_{solv}	−18.44 ± 0.18	−23.52 ± 0.19	−27.52 ± 0.28	−28.99 ± 0.38	−28.88 ± 0.26
log c_{calc}	−8.02	reference	−1.82	−0.81	−0.89
	−8.43	−4.98	reference	−1.27	−1.34
	−8.25	−4.80	−2.09	reference	−1.17
	−8.26	−4.81	−2.10	−1.10	reference
ΔG_{bind}	−3.47 ± 1.17	−6.00 ± 0.93	−1.88 ± 0.87	dissociated	dissociated

[a] c_{calc} is inferred according to eq 2 using successively the experimental solubility determined in each environment as a reference and the free energies of solvation (ΔG_{solv}, kcal/mol) computed using the FEP method. The binding free energies (ΔG_{bind}, kcal/mol) for the dimer of quercetin are also reported. [b] Determined using the same protocol as in Chebil et al.[12] [c] From Chebil et al.[12]

solvation energies, bidirectional simulations impose that restraints be enforced[33] as quercetin is created in the bound state, thereby guaranteeing the formation of the relevant noncovalent interactions. The contribution of these restraints to the binding affinity should in turn be evaluated, which constitute a potential additional source of error. Instead, to improve the accuracy of the estimates, annihilation of the solute was performed in ten independent simulations, from whence an average free energy was inferred, with an rmsd.

The free energy of solvation at a temperature T for any compound can be related to its solubility (c) through

$$c = c_A \exp\left(-\frac{\Delta G_{solv}}{RT}\right) \quad (2)$$

where c_A is the solubility of the pure crystal matter at T. It does not depend on the solvent, thereby allowing the solubilities measured in two different media to be connected via

$$c_2 = c_1 \exp\left(-\frac{\Delta G_{solv2} - \Delta G_{solv1}}{RT}\right) \quad (3)$$

From this equation, it is apparent that the solubility of a given compound can be inferred in any solvent on the basis of the knowledge of a solubility of reference in another solvent, together with the relevant free energies of solvation.

Results and Discussion

In this contribution, the solvation of quercetin is investigated in different media, e.g., water, acetonitrile, acetone, *tert*-amyl alcohol (2-methyl-2-butanol), and chloroform. The experimental solubility of quercetin varies dramatically with these solvents,[12] as can be seen in Table 1. Quercetin is not or is poorly soluble in chloroform, water, and acetonitrile. On the contrary, its solubility is rather high in acetone and *tert*-amyl alcohol, two solvents markedly amphiphilic.

The free energies of solvation (ΔG_{solv}) computed in this work (see Table 1) are consistent with the experimental solubility scale for quercetin, i.e., chloroform < water < acetonitrile < *tert*-amyl alcohol ≈ acetone, with values ranging from ca. −18 to −29 kcal/mol.

The experimental solubility in each media was used successively as a reference to estimate the solubility (c_{calc}) in the remaining solvents. The predicted solubilities reported in Table 1 are all within 0.5 log unit of the corresponding experimental values, irrespective of the reference solvent. Their reliability embraces three contributions (i) the accuracy of the experimental reference solubility, (ii) the relevance of the force field for the description of quercetin-solvent interactions, and (iii) both the precision and the accuracy of the computed solvation free energies. For the first contribution, the larger the solubility, the greater the accuracy. It is, therefore, desirable to have a high reference solubility to guarantee the lowest possible deviation between the predicted and the experimental values. As reported in Table 1, the largest errors for c_{calc} are observed when the low aqueous solubility of quercetin, $(2.98 \pm 0.31) \times 10^{-5}$ mol/L, is used as the reference solubility. Insofar as the force-field contribution is concerned, a reasonable description of electrostatic solute−solvent interactions was achieved by means of distinct sets of atomic charges assumed to account for the polarization of quercetin in the different media in an average sense (see Supporting Information). Finally, the strategy em-

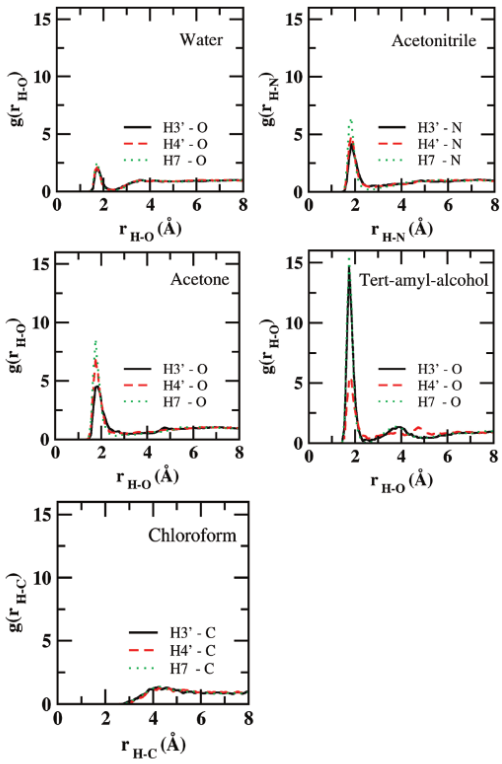


Figure 4. Radial distribution functions for hydroxyl moieties of quercetin and solvents.

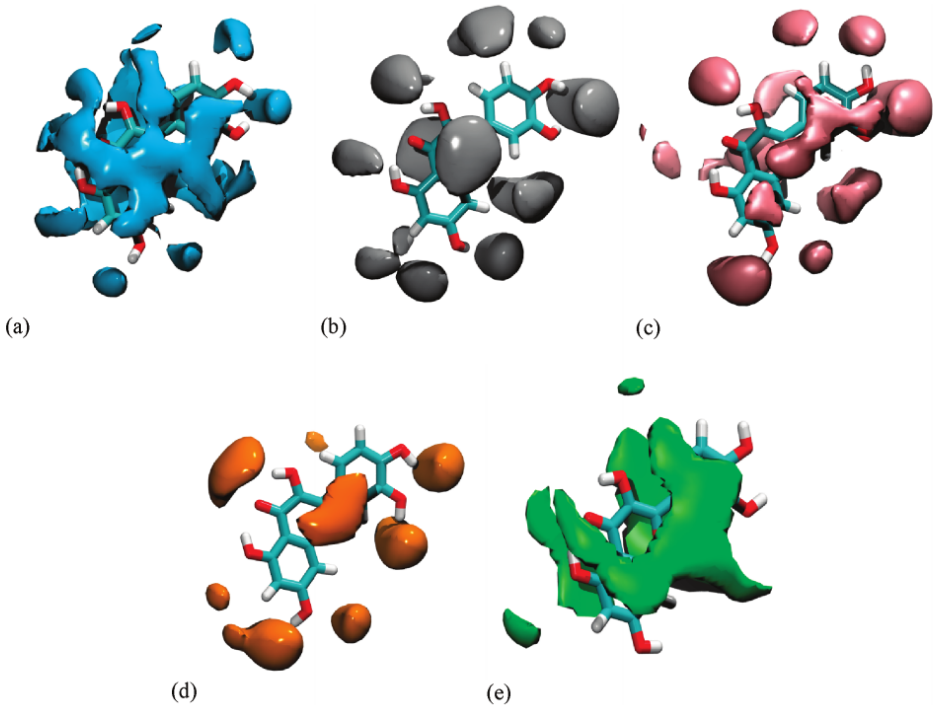


Figure 5. Contour maps of the solvent density around quercetin: (a) water; (b) acetonitrile; (c) acetone; (d) *tert*-amyl alcohol; (e) chloroform.

ployed herein allowed free energies of solvation to be determined with a root-mean-square deviation of less than 0.4 kcal/mol. That the free-energy estimates were obtained within chemical accuracy stems primarily from an appropriate sampling strategy together with the use of the Bennett acceptance ratio (BAR) method.[30] While both the statistical error and the rmsd are less than k_BT, this result is suggestive that a small bias arising from the finite length of the simulations may exist. This bias, which is related to the overall accuracy of the free-energy calculations, is mirrored in the generally appropriate, albeit occasionally imperfect, overlap of the underlying probability distributions for the forward and backward transformations (see Supporting Information).

The procedure utilized here not only provides a reliable prediction of the solubility but also gives access to the atomic detail of the solvation process, which is anticipated to help rationalize the solvent dependence of quercetin solubility. The pair distribution functions (see Figure 4) reveal different solvation patterns for the hydroxyl moieties at positions 3′, 4′, and 7 as a function of the environment. Not too surprisingly, there is no specific organization of chloroform around these hydrophilic groups. Water molecules are only marginally better organized around quercetin, which can be ascribed to the bulky hydrophobic core of the solute. In sharp contrast, amphipathic solvents are capable of interacting concomitantly with both the hydrophobic and the hydrophilic moieties of quercetin, reflected in a steadier organization of the solvent molecules around the solute.

A glance at the probability distributions of Figure 5 suggests, however, that the enhanced solvation of quercetin by acetone and *tert*-amyl alcohol stems from more pronounced hydrophobic and hydrophilic interactions, respectively, compared to the case for acetonitrile.

Interestingly enough, this information provides an additional rational for the optimization of the environment with the goal to increase the solubility of the substrate. In the case of quercetin, a more effective solvent would share the hydrophilic and hydrophobic moieties of *tert*-amyl-alcohol and acetone, respectively. Among other solvents, propan-2-ol meets these criteria. Not too surprisingly, the solubility of quercetin measured in this medium, following the procedure detailed in Chebil et al.,[12] proves to be larger, ∼29 g/L, than the corresponding concentration in *tert*-amyl-alcohol (∼20 g/L) and acetone (∼24 g/L).

To delve further into the solubility of quercetin and the role exerted by the solvent on this process, the dimerization free energy of the solute was measured in the different environments detailed above, as a putative model for the prediction of relative solubilities. As has been shown in the case of cyclodextrins,[34] this route constitutes a rudimentary and certainly incomplete approach toward the estimation of solubilities, although it provides valuable information on the propensity of the corresponding solid to dissociate. Here, simulations at thermodynamic equilibrium in *tert*-amyl alcohol and acetone have led in a systematic fashion to the dissociation of quercetin dimers, consistently with the high solubility in these two solvents. Conversely, a dimerization or binding free energy was measured in water, chloroform, and acetonitrile, following the thermodynamic cycle depicted in Figure 3. At the qualitative level, the nature of the environment yields distinct binding modes (see Supporting Information), hence resulting in dimerization free energies of different magnitudes. Specifically, a robust stacking

motif of quercetin is observed consistently in an aqueous medium. In sharp contrast, the hydrophobic environment of chloroform promotes the formation of dimers loosely bound by a few intermolecular hydrogen bonds and characterized by an appreciable configurational entropy. An intermediate behavior is found in the case of acetonitrile, where a large spectrum of binding motifs is observed, among which are short-lived, stacked, and hydrogen-bonded arrangements. At the quantitative level, the rmsd inferred from the ten independent free-energy calculations is significantly larger than any individual measure of the unidirectional statistical error.[28] To a large extent, this result can be understood by considering the configurational entropy of the dimers, which can hardly be embraced precisely in insufficiently long simulations. Binding free energies have, nonetheless, been estimated with a sufficient accuracy, allowing the different environments to be compared. The absolute value of the free energy of dimerization in acetonitrile is lower than the corresponding estimates in water and chloroform, which is consistent with the higher solubility of quercetin observed in this environment (see Table 1). On the contrary, the dimer model demonstrates to be insufficient to predict the lowest solubility of quercetin in chloroform ($\Delta G_{\text{bind}} = -3.47$ kcal/mol) compared to that in water ($\Delta G_{\text{bind}} = -6.00$ kcal/mol). Clearly, the precise estimation of the solid contribution to the solubility requires that complete models be employed to account for the overall interactions in the crystal, as devised by Lüder et al.,[6] who stated that the free energy required to extract one molecule out of the crystal is certainly different than that necessary to dissociate an isolated dimer.

Conclusion

In silico prediction of solubilities represents a challenging task for computational chemistry, which is in large measure connected with the difficulty to estimate the crystal contribution. Here, demonstration is made that accurate solubilities can be inferred for a wide range of environments at the price of minimal experimental input. The strategy put forth requires the knowledge of a unique experimental solubility together with a set of free energies of solvation. The latter quantities can be computed within chemical accuracy, as illustrated in the case of quercetin in chloroform, water, acetonitrile, acetone, and *tert*-amyl alcohol. In the context of pharmaceutical high-throughput screening, it might be advantageous to consider alternatives to free-energy perturbation, like quantum-mechanical calculations relying on a polarizable continuum description, which are unquestionably faster but also offer a less detailed and somewhat more approximate picture of solvation phenomena.[35] Regardless of the experimental reference, the solubilities were determined within 0.5 log unit. Combined with the current arsenal of tools available to the computational chemist, the present methodology is envisioned to help predict solubilities in a more routine fashion and to constitute a milestone on the road toward the rational design and synthesis of novel molecular compounds.

Supporting Information Available: Atomic charges for quercetin in each solvent. Potential energy surface as a function of φ. Quercetin dimer geometries. Probability distributions underlying the annihilation and creation transformations of quercetin in acetonitrile. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Huuskonen, J.; Livingstone, D. J.; Manallack, D. T. Prediction of drug solubility from molecular structure using a drug-like training set. *SAR QSAR Environ. Res.* **2008**, *19*, 191–212.

(2) Zheng, Y.; Mo, Q.; Liu, Z. The studies of QSPR/QSAR for ionic liquids. *Prog. Chem.* **2009**, *21*, 1772–1781.

(3) Obrezanova, O.; Gola, J. M. R.; Champness, E. J.; Segall, M. D. Automatic QSAR modeling of ADME properties: Blood-brain barrier penetration and aqueous solubility. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 431–440.

(4) Westergren, J.; Lindfors, L.; Höglund, T.; Lüder, K.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility: 1. Free energy of hydration. *J. Phys. Chem. B* **2007**, *111*, 1872–1882.

(5) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility: 2. Free energy of solvation in pure melts. *J. Phys. Chem. B* **2007**, *111*, 1883–1892.

(6) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility. 3. Free energy of solvation in pure amorphous matter. *J. Phys. Chem. B* **2007**, *111*, 7303–7311.

(7) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Persson, R.; Pedersen, M. In Silico prediction of drug solubility: 4. Will simple potentials suffice. *J. Comput. Chem.* **2009**, *30*, 1859–1871.

(8) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(9) Havsteen, B. H. The biochemistry and medical significance of the flavonoids. *Pharmacol. Therapeut.* **2002**, *96*, 67–202.

(10) Heim, K. E.; Tagliaferro, A. R.; Bobilya, D. J. Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. *J. Nutr. Biochem.* **2002**, *13*, 572–584.

(11) Chebil, L.; Anthoni, J.; Humeau, C.; Gerardin, C.; Engasser, J. M.; Ghoul, M. Enzymatic acylation of flavonoids: effect of the nature of substrate, origin of lipase and operating conditions on conversion yield and regioselectivity. *J. Agric. Food Chem.* **2007**, *55*, 9496–9502.

(12) Chebil, L.; Humeau, C.; Anthony, J.; Dehez, F.; Engasser, J. M.; Ghoul, M. Solubility of flavonoids in organic solvents. *J. Chem. Eng. Data* **2007**, *52*, 1552–1556.

(13) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(14) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(15) Grabuleda, X.; Jaime, C.; Kollman, P. A. Molecular dynamics simulation studies of liquid acetonitrile: new six-site model. *J. Comput. Chem.* **2000**, *21*, 901–908.

(16) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field. *J. Comput. Chem.* **2001**, *22*, 1340–1352.

(17) Martin, M. G.; Biddy, M. J. Monte Carlo molecular simulation predictions for the heat of vaporization of acetone and butyramide. *Fluid Phase Equilib.* **2005**, *236*, 53–57.

(18) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. Relative partition coefficients for organic solutes from fluid simulations. *J. Phys. Chem.* **1990**, *94*, 1683–1686.

(19) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, J. T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(21) Tomasi, J.; Persico, M. Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.

(22) Ángyán, J. G.; Chipot, C.; Dehez, F.; Hätting, C.; Jansen, G.; Millot, C. OPEP. A tool for the optimal partitioning of electric properties. *J. Comput. Chem.* **2003**, *24*, 997–1008.

(23) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(24) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

(25) Darden, T. Y., D.; Pedersen, L. Particle mesh Ewald: An N•log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(26) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1990–2001.
(27) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* **1997**, *72*, 1047–1069.
(28) Chipot, C.; Pohorille, A. *Free energy calculations. Theory and applications in chemistry and biology*; Springer Verlag: Berlin, Heidelberg, 2007
(29) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
(30) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
(31) Hahn, A. M.; Then, H. Characteristic of Bennett's acceptance ratio method. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.* **2009**, *80*, 031111.
(32) Pohorille, A.; Jarzynski, C.; Chipot, C. Good practices in free-energy calculations. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(33) Deng, Y.; Roux, B. Computations of standard binding free energies with molecular dynamics simulations. *J. Phys. Chem. B* **2009**, *113*, 2234–2246.
(34) Cai, W.; Sun, T.; Shao, X.; Chipot, C. Can the anomalous aqueous solubility of β-cyclodextrin be explained by its hydration free energy alone. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3236–3243.
(35) Orozco, M.; Luque, F. J. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* **2000**, *100*, 4187–4226.

JP104569K

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 33180

Thermodynamics of DNA: sensitizer recognition. Characterizing binding motifs with all-atom simulations†

Hugo Gattuso,[ab] Elise Dumont,*[c] Christophe Chipot,[abde] Antonio Monari*[ab] and François Dehez*[abe]

Received 2nd September 2016,
Accepted 24th October 2016

DOI: 10.1039/c6cp06078a

www.rsc.org/pccp

We report the investigation of the thermochemical properties of benzophenone interacting with B-DNA studied by all-atom molecular dynamic simulations. In particular, we determine the binding free energy for two competitive binding modes, minor groove binding and double insertion. Our results allow us to quantitatively resolve for the first time the mode of binding of this paradigmatic photosensitizer, indicating a marked preference for minor groove binding. Furthermore, we have settled a protocol allowing for the determination of binding energies in the case of non-covalent interaction with DNA, in particular, tackling the non-trivial problem of the strong reorganization of the DNA imposing extended statistical sampling. Our contribution paves the way to the systematic determination of the thermochemical properties of drugs or pollutants interacting with DNA.

1 Introduction

The comprehension of the interaction of relatively small molecules with nucleic acids, and DNA, in particular, has considerable biological interest. Indeed, the sensitization of DNA by exogenous or endogenous molecules can be related to the insurgence of DNA lesions and modifications. The latter may in turn evolve toward cell death as well as mutagenesis and ultimately carcinogenicity. On the other hand, those kinds of interactions are also widely exploited in cancer therapy. In particular, one could cite the celebrated discovery of the cis-platin drug that is still used in chemotherapy despite its considerable side-effects.[1,2] In the later years, also in order to minimize drug's general toxicity, non-covalent interaction with DNA has been widely exploited.[3,4] DNA non-covalent sensitizers may lead to DNA lesions, and hence apoptosis, by a wide spectrum of mechanisms, such as replication inhibition,[5–9] photosensitization,[10–14] interaction with gene regulatory regions and telomeres.[15–17] Despite their usefulness, the study of DNA sensitizers is somehow complicated by the lack of experimentally resolved structures of the aggregates. Indeed, if in the case of inorganic metal complexes, in particular, Ruthenium based, X-ray crystallographic structures have been resolved,[18] organic aggregates have been characterized to a much lesser extent. Furthermore, the presence of different competitive interacting structures, such as groove binding, intercalation or insertion, has to be taken into account.[19]

A typical case illustrating this difficulty is Benzophenone (BNZ), see Fig. 1. This simple aryl ketone is indeed known to be a paradigmatic, and almost textbook DNA photosensitizer, exerting energy- and electron-transfer, as well as photo-induced H abstraction or Paternó–Büchi reactions.[20] However, the experimental determination of the binding modes with DNA is still missing. The first characterization of two stable interacting modes has been provided only very recently.[21] In particular, it was evidenced that apart from the rather usual minor groove-binding mode, benzophenone presented a novel mode, referred to by the authors double insertion and characterized by the simultaneous expulsion of a full base pair from the Watson and Crick pairing (Fig. 1). Furthermore, it has been shown that the photophysical and photochemical properties of BNZ, and hence its photosensitization capabilities, vary quite consistently between the

[a] *Université de Lorraine – Nancy, Theory-Modeling-Simulation, UMR 7565, Structure et Réactivité des Systèmes Moléculaires Complexes (SRSMC), Vandoeuvre-les-Nancy, France. E-mail: Francois.Dehez@univ-lorraine.fr, Antonio.Monari@univ-lorraine.fr*
[b] *Centre National de la Recherche Scientifique, Theory-Modeling-Simulation, UMR 7565, Structure et Réactivité des Systèmes Moléculaires Complexes (SRSMC), Vandoeuvre-les-Nancy, France*
[c] *Laboratoire de Chimie, UMR 5182, Ecole Normale Supérieure de Lyon, Lyon, France. E-mail: Elise.dumont@ens-lyon.fr*
[d] *Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, USA*
[e] *Laboratoire International Associé, Centre National de la Recherche Scientifique, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

† Electronic supplementary information (ESI) available: Representation of the set of parameters defining the position of BNZ bound in the minor groove. Parameters for each harmonic potential employed to restraint BNZ. Free energy profiles associated with each alchemical transformation. Illustration of the convergence of FEP calculations. See DOI: 10.1039/c6cp06078a

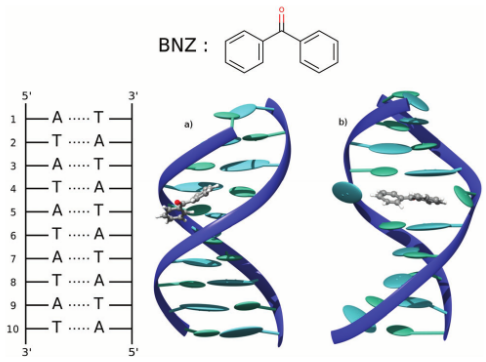


Fig. 1 Cartoon representations of (a) BNZ in the minor groove binding mode (b) BNZ in the double insertion mode.

two modes.[22,23] The presence of a stable double-insertion binding mode has been also independently proposed by Galindo-Murillo *et al.*[24] for organometallic copper complexes. Hence, it will be extremely desirable to achieve a computational methodology able to correctly reproduce the binding free-energy, *i.e.* the thermodynamic properties, of the two configurations that could be used to discriminate between the different DNA/drug binding possibilities. The evaluation of absolute thermodynamics quantities from all-atom simulations remains highly challenging as it requires extensive sampling of the configurational phase space of complex molecular assemblies, an endeavor contrasting markedly with the finite-length of molecular-dynamics simulations. However, thanks to the power of current supercomputers and the development of effective sampling strategies, accurate determination of binding free-energies is now accessible. In recent years, various studies have demonstrated the reliability of free-energy calculations in determining binding affinities for ligand–protein and protein–protein complexes with very low statistical errors. Binding free energies of drugs in DNA were essentially performed through two distinct routes, which, roughly speaking, can be subdivided into two categories—approximate methods, generally relying on an implicit description of the surroundings, and accurate, all-atom methods, based on importance-sampling algorithms[25] to perform alchemical transformations.[26,27] MM/GBSA methods have been widely employed to estimate the free energy of binding of many DNA/drug complexes.[28–30] Conversely, statistical-mechanics-based approaches for alchemical transformations, like free-energy perturbation (FEP)[26,31,32] and thermodynamical integration (TI),[33] were mainly employed to study the thermodynamics of minor groove binders, *e.g.*, distamycin[34] or netropsin.[35] These studies were able to capture the difference of affinities of different molecules, but because of short sampling time, thermodynamics quantities were determined with a low precision and a poor accuracy. It is now commonly admitted that alchemical calculations require extensive sampling of the phase space to ensure a high statistical precision of the free energy estimates. Binding implies a substantial change of configurational entropy that molecular dynamics is not able to capture at thermodynamic equilibrium. To overcome the incomplete sampling due to finite-length simulations, a convenient theoretical framework consists in introducing a set of well-defined geometric restraints, which will reduce the translational, orientational and conformational entropies available to the host–guest complex of interest.[36–38] Numerical schemes aimed at reducing the large variation of configurational entropy characteristics of host–guest association exist in different flavors, tethering the substrate with a variety of restraints of different natures, relevant to the problem at hand.[36,39,40,41] The alchemical transformation, wherein the guest is decoupled reversibly from its environment—either the binding site or the aqueous medium, is carried out with a substrate restrained in its native position, relative orientation and conformation. The loss of configurational entropy due to the geometrical restraints is evaluated subsequently in independent free-energy calculations. Gumbart *et al.*[38] determined the standard binding free energy of a protein–ligand complex, following this alchemical route with a set of adequately chosen restraints, and showed that it coincides within chemical accuracy with the value inferred from an alternate geometric route, relying on the determination of one-dimensional potentials of mean force for the different restrained degrees of freedom of the complex. In this work, we apply the strategy based on alchemical transformations to study the structure and the complexation thermodynamics of BNZ with DNA. Binding affinities of the two putative binding modes proposed by Dumont and Monari,[21] namely minor groove binding and double insertion, were evaluated.

2 Computational details

2.1 DNA/BNZ molecular samples

Minor groove (MinGB) and double insertion (DI) binding modes have been evidenced and characterized by Dumont et Monari.[21] In the present study, each initial conformation of DNA–BNZ complexes was taken from the molecular dynamics simulations performed in the latter reference. In each case, the most representative conformation of a B-DNA double strand decamer composed of alternating thymine-adenine base couples, namely poly(dA-dT)–poly(dA-dT), in complex with BNZ was considered. As depicted in Fig. 1 BNZ was inserted at the level of base pair 5 for both minor groove binding and double insertion. DNA–BNZ complexes and both DNA and BNZ unbound states were embedded in a cubic box of explicit water molecules. For all cases, the edge of the initial box was 60 Å for a total of about 6600 water molecules. 18 potassium cations were added as counter ions to neutralize the global charge of the DNA-containing systems.

2.2 Molecular dynamics simulations

DNA double strand was described using the Amber 99 force field[42] including the bsc0 correction.[43] BNZ was modeled using the generalized amber force field (GAFF)[44] consistently with the strategy used elsewhere.[21] Water molecules were described using the TIP3P model.[45]

All molecular dynamics simulations were performed using the NAMD program,[46] in the *NPT* ensemble. The pressure (1 atm) and the temperature (300 K) were kept constant using, respectively, the Langevin dynamics and the Langevin piston[47] algorithms. Long-range electrostatic interactions were taken into account using the Particle-Mesh-Ewald method.[48] Van der Waals and electrostatic interactions were truncated for distances longer than 9 Å. An integration time step of 2 fs was used for all simulations. Both systems were first minimized for 100 steps using the conjugate gradient algorithm in order to remove close contacts between atoms. DNA heavy atoms were first restrained by means of an harmonic potential. Restraints were slowly released within 10 steps of 100 ps and a trajectory of 40 ns was produced. All figures of DNA interacting with BNZ were rendered using the molecular visualization program Chimera.[49] The structural reorganization of the 10 base-pair double-stranded DNA upon BNZ binding was analyzed using Curves+,[50] a program allowing the comparison of inter- and intra-base pair parameters as well as backbone and groove deformations.

2.3 Alchemical transformations

The standard binding free energy ΔG^0 of BNZ to DNA was evaluated following the alchemical cycle described in Fig. 2. For coupling and decoupling BNZ from the hydrated DNA environment or from the bulk, the position of BNZ with respect to DNA was harmonically restrained by means of a set of spherical coordinates (r, θ, ϕ), whereas its relative orientation was restrained using three Euler angles (Θ, Φ, Ψ) (see Fig. 3) using the collective variable software[51] implemented in NAMD.[46] Furthermore, the double-stranded DNA was harmonically restrained to its bound state conformation thanks to a root mean-square deviation (RMSD) collective variable. The force constant of each harmonic potential has been chosen as low as possible to maintain the system in a conformation identical to that observed in equilibrium simulations (see the ESI,† Table S1). Assignment of the groups of atoms used in the definition of the geometric restraints is arbitrary (see Fig. 3 and the ESI,† Fig. S1), namely P1, P2 and P3 for DNA, and L1, L2 and L3 for BNZ, within the limit that these groups of atoms are not aligned, and their relative position in space fluctuates only marginally. Under these premises, one can show that the binding free energy is an invariant of the geometric restraints. ΔG^0 can be computed as

$$\Delta G^0 = \Delta G^{\text{site}}_{\text{rest.}} + \Delta G^{\text{site}}_{\text{alch.}} - \Delta G^{\text{bulk}}_{\text{rest.}} - \Delta G^{\text{bulk}}_{\text{alch.}} \quad (1)$$

The purpose of introducing geometrical restraints in the alchemical free-energy calculations is twofold. First, appropriately chosen harmonic potentials acting on translational, orientational and conformational degrees of freedom are expected to reduce appreciably the change in configurational entropy associated with the binding process, making sampling of the latter amenable to finite-length, equilibrium simulations.[36,38] Second, by tethering it to the binding site, the substrate is prevented from drifting astray as it is decoupled from its environment, which would otherwise violate thermodynamic micro-reversibility.[39,52,53]

The free energy differences associated with the coupling/uncoupling of BNZ from the hydrated double-strand DNA decamer and bulk water, $\Delta G^{\text{site}}_{\text{alch.}}$ and $\Delta G^{\text{bulk}}_{\text{alch.}}$ respectively, were computed within the Free Energy Perturbation (FEP) framework.[26] The reversible work, $\Delta G^{\text{site}}_{\text{rest.}}$ and $\Delta G^{\text{bulk}}_{\text{rest.}}$, associated with BNZ positional/orientational restraints and with DNA conformational restraint, were evaluated using thermodynamic integration (TI) simulations or analytically.[37] The latter contributions are obtained by summing up the contributions of each individual restraint:

$$\Delta G^{\text{site}}_{\text{rest.}} = \Delta G^{\text{site}}_{\text{RMSD}} + \Delta G^{\text{site}}_{r} + \Delta G^{\text{site}}_{\theta} + \Delta G^{\text{site}}_{\phi} + \Delta G^{\text{site}}_{\Theta} + \Delta G^{\text{site}}_{\Phi} + \Delta G^{\text{site}}_{\Psi} \quad (2)$$

and

$$\Delta G^{\text{bulk}}_{\text{rest.}} = \Delta G^{\text{bulk}}_{\text{RMSD}} + \Delta G^{\text{bulk}}_{r,\theta,\phi} + \Delta G^{\text{bulk}}_{\Theta,\Phi,\Psi} \quad (3)$$

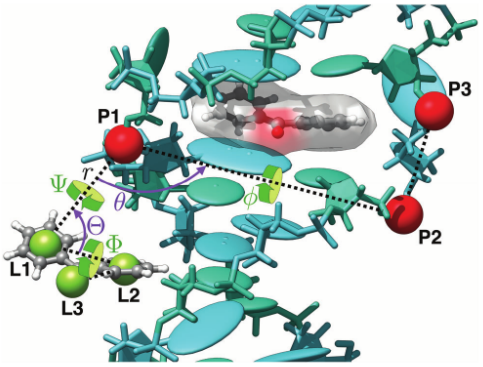


Fig. 3 Representation of the set of parameters defining the position of BNZ in the double insertion mode. For both DNA and BNZ, a triplet of points, P1, P2, P3 (red spheres) and L1, L2, L3 (green spheres), respectively, is defined arbitrarily. The position of BNZ is defined by three spherical coordinates (r(P1–L1), θ(L1–P1–P2), ϕ(L1–P1–P2–P3)), whereas its relative orientation is described by three Euler angles (Θ(P1–L1–L2), Φ(P1–L1–L2–L3), Ψ(P2–P1–L1–L2)). For clarity, inserted BNZ (gray surface) has been split out from the complex.

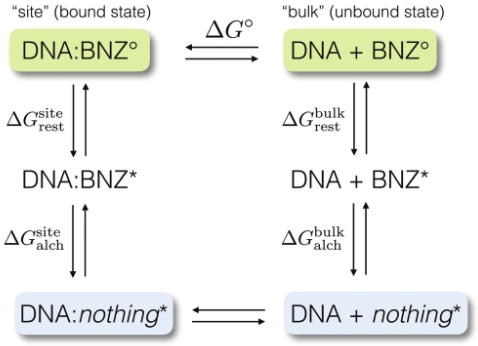


Fig. 2 Alchemical cycle to determine the binding free energy of the DNA/BNZ complex. * refers to a restrained state.

$\Delta G^{\text{bulk}}_{r,\theta,\phi}$ and $\Delta G^{\text{bulk}}_{\Theta,\Phi,\Psi}$ are obtained through numerical integration of analytical expressions (see ESI,† eqn (S1) and (S2)). In practice, the free-energy differences between the different states were evaluated through a series of transformations between non-physical intermediate states (windows) connecting the initial to the final state by means of a coupling parameter λ, varying from 0 to 1. All FEP and TI simulations were performed in both decoupling and coupling directions except for the DNA unbound state system in which the work associated with the RMSD restraint, $\Delta G^{\text{bulk}}_{\text{RMSD}}$, was averaged over a set of 14 independent uncoupling simulations. For FEP simulations, the maximum-likelihood estimator of the free-energy difference was evaluated using data collected in both directions by means of the Bennett Acceptance Ratio (BAR) using the ParseFEP plugin[54] of VMD.[55] Both $\Delta G^{\text{site}}_{\text{alch}}$ and $\Delta G^{\text{bulk}}_{\text{alch}}$ FEP simulations were performed under the same conditions but using different stratification strategies tailored to ensure proper convergence of the free-energy estimates. Each λ window consists of 200 ps of equilibration followed by 800 ps of production. For $\Delta G^{\text{site}}_{\text{alch.}}$, 200 windows were used in the case of the double insertion mode for both coupling and uncoupling simulations (400 ns), whereas 80 windows were employed for minor groove binding (160 ns). $\Delta G^{\text{bulk}}_{\text{alch.}}$ was obtained by stratifying the coupling and uncoupling of BNZ to the bulk in 63 windows (126 ns). Contributions to $\Delta G^{\text{site}}_{\text{rest}}$ (eqn (2)) were determined through TI simulations stratified over 24 windows in both coupling and uncoupling directions (48 ns). For $\Delta G^{\text{bulk}}_{\text{rest.}}$ (eqn (3)), $\Delta G^{\text{bulk}}_{\text{RSMD}}$ was obtained through a series of 14 TI simulations performed only in the uncoupling direction (334 ns), whereas $\Delta G^{\text{bulk}}_{r,\theta,\phi}$ and $\Delta G^{\text{bulk}}_{\Theta,\Phi,\Psi}$ were calculated through numerical integration of analytical expressions.[36,56] In the case of FEP simulations, a soft-core potential was employed in which the interatomic distance, r, used in the Lennard-Jones potential, was shifted $r^2 \rightarrow r^2 + 5 \times (1 - \lambda)$. Furthermore, in order to avoid the so-called "end-point catastrophe", electrostatic interactions of vanishing or appearing particles were linearly coupled to the simulation for λ, varying from 0 to 0.5. At λ values greater than 0.5, electrostatic interactions were fully decoupled from the simulation.

3 Results

3.1 Structural analysis for DNA:BNZ binding modes

In a previous study[21] the locality of the double insertion binding mode deformations on the DNA structure was already described. We give here a more thorough characterization based on specific parameters, namely the opening of the base pairs, Fig. 4, and the ζ dihedral of the DNA backbone, Fig. 5, both compared for base pairs 4, 5 and 6 showing the deviation from their canonical position. First of all, the time evolution of the opening parameter shows good stability in the position of the two ejected bases with an angle of around 120°. The higher intensity in the distribution of this value for base pair 5 compared to 4 and 6 is due to the lack of stabilization of these two nucleobases. In fact while in a canonical environment hydrogen bondings and π-stacking

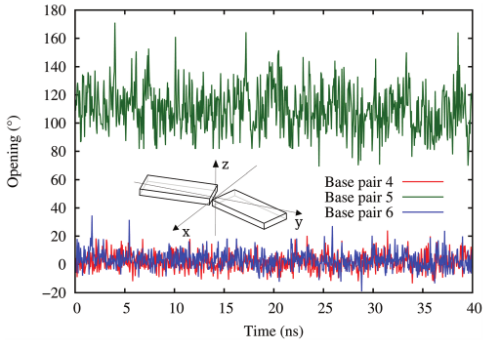


Fig. 4 Variation over time of the opening structural parameter in the double insertion mode.

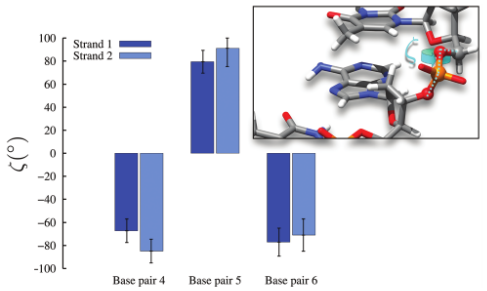


Fig. 5 Average ζ dihedral parameter over the entire molecular dynamic simulation for both strands in the DI binding mode.

interactions stabilize strongly the structure of DNA, in that case the bases are dangling in the water bulk and thus are able to move freely. Moreover, this double ejection can also be clearly described by the movement induced in the ζ dihedral angle of the backbone structure. To follow the ejection of bases, the dihedral flips over and goes from between −60 and −80° to around 90°.

In the case of minor groove binding both local and global DNA structures are much less affected by interaction with BNZ. Only a slight deformation of the backbone width and length can be evidenced during the simulation. Moreover along the 40 ns of molecular dynamic simulation, BNZ slides in the minor groove to reach another equivalent and stable binding position. This sliding phenomenon is not specific to BNZ and has been described previously.[57] The study reported the existence of multiple binding stations, thermodynamically equivalent and separated by free energy barriers of 6 kcal mol^{-1} at most, suggesting that sliding is a likely process.

3.2 Binding free energies

Individual contributions (eqn (1)–(3)) to the free energy of binding of BNZ to DNA are reported in Table 1 for both minor

Table 1 Binding free energy contributions for DI and MinGB binding motifs

	Double insertion		Minor groove binding	
Contributions	Free energies (kcal mol^{-1})	Time (ns)	Free energies (kcal mol^{-1})	Time (ns)
$\Delta G^{\text{site}}_{\text{RMSD}}$	12.8 ± 0.9	48	8.6 ± 0.3	48
$\Delta G^{\text{site}}_{r}$	0.1 ± 0.0	—	0.1 ± 0.0	—
$\Delta G^{\text{site}}_{\Theta}$	0.8 ± 0.0	—	0.9 ± 0.0	—
$\Delta G^{\text{site}}_{\phi}$	1.1 ± 0.0	—	0.9 ± 0.0	—
$\Delta G^{\text{site}}_{\Psi}$	1.0 ± 0.1	—	0.7 ± 0.1	—
$\Delta G^{\text{site}}_{\theta}$	0.2 ± 0.0	—	0.6 ± 0.0	—
$\Delta G^{\text{site}}_{\phi}$	0.6 ± 0.0	—	0.4 ± 0.0	—
$\Delta G^{\text{site}}_{\text{alch}}$	−15.5 ± 0.1	400	−30.0 ± 0.1	160
$\Delta G^{\text{bulk}}_{\text{alch}}$	−35.8 ± 0.0	126	−35.8 ± 0.0	126
$\Delta G^{\text{bulk}}_{\text{RMSD}}$	23.4 ± 1.7	14 × 24	10.3 ± 0.3	48
$\Delta G^{\text{bulk}}_{r,\theta,\phi}$	4.7	—	4.6	—
$\Delta G^{\text{bulk}}_{\Theta,\Phi,\Psi}$	7.2	—	6.6	—
ΔG^{0}	1.6 ± 1.9	910	−3.6 ± 0.4	382

groove and double insertion binding motifs. Systematic errors have been estimated from the hysteresis between backward and forward alchemical transformations with the exception of the work arising from restraining, in the bulk, the conformation adopted by DNA in the double insertion complex ($\Delta G^{\text{bulk}}_{\text{RMSD}}$). Starting from DNA canonical conformation, coupling the latter restraint did not allow the recovery of the target conformation where two bases are expelled out from the double stranded structure. This issue stems essentially from the high level of degeneracy associated with the RMSD collective variable. $\Delta G^{\text{bulk}}_{\text{RMSD}}$ was therefore averaged over a series of 14 independent alchemical transformations, where the RMSD restraint was decoupled.

Free-energy contributions of BNZ positional and orientational restraints are small in the bound state, ∼1 kcal mol^{-1} or less, in line with the very soft harmonic potentials employed in the simulations (see ESI,† Tables S1 and S2). The hysteresis between the corresponding coupling and uncoupling simulations never exceeds at most a few tenth of kcal mol^{-1} (ESI,† Fig. S2). The bulk contributions, $\Delta G^{\text{bulk}}_{r,\theta,\phi,\psi}$ and $\Delta G^{\text{bulk}}_{\Theta,\Psi,\Phi}$, computed analytically, are about one order of magnitude larger than their bound state counterparts. Restraining the conformational space of the double stranded DNA in the bound state contributes by ∼10 kcal mol^{-1} ($\Delta G^{\text{site}}_{\text{RMSD}}$) for both minor groove and double insertion binding motifs. Because the conformation of DNA in the bulk is identical to that observed in the complex where BNZ is bound to the minor groove $\Delta G^{\text{bulk}}_{\text{RMSD}}$ is close to $\Delta G^{\text{site}}_{\text{RMSD}}$. Conversely, maintaining the distorted conformation adopted by DNA when BNZ in doubly inserted requires twice more energy in the unbound state than in the complex. In the case of the double insertion mode, both $\Delta G^{\text{bulk}}_{\text{RMSD}}$ and $\Delta G^{\text{site}}_{\text{RMSD}}$ are affected by a relatively large systematic error compared to the other contributions, *viz.* 1.7 and 0.9 kcal mol^{-1}, respectively (ESI,† Fig. S3).

Free-energy associated with coupling BNZ to the hydrated DNA double strand, $\Delta G^{\text{site}}_{\text{alch}}$, is twice larger for minor groove binding than for double insertion (ESI,† Fig. S4). Backward and forward estimates of $\Delta G^{\text{site}}_{\text{alch}}$ and $\Delta G^{\text{bulk}}_{\text{alch}}$ present a marginal hysteresis, resulting in very small errors (see Table 1 and ESI,† Fig. S5). Convergence of all FEP simulations was checked thanks to the ParseFep plugin[54] of VMD (ESI,† Fig. S6).

4 Discussion

The precise determination of the binding free energy between DNA and a paradigmatic, yet not totally resolved photosensitizer, is of extreme importance to allow the prediction or rationalization of the subsequent chemical and photochemical reactions and hence to assess the possible further DNA damages. Indeed it has been shown that the photochemical pathways strongly depend on the coupling between DNA and the photodrugs, and hence on the structure and the binding mode.[19,22,23,58] In particular, in the case of benzophenone, it has been clearly shown by some of us that hydrogen abstraction from the backbone, a minor photosensitization mechanism, is possible only in minor groove-binding,[23] while triplet–triplet energy transfer is in principle possible from both modes.[22] Furthermore, as shown experimentally by Miranda's group on artificial model systems allowing to clearly dictate the relative position between benzophenone and thymine nucleobases,[59] the structure of the BNZ:thymine aggregates may open the door to the formation of excited state excimer complexes that may ultimately alter the final product distribution also in the case of energy transfer driven phenomena. Hence, the method we provide here allow a preliminary and quantitative discrimination that can orientate the following studies on the photo-reactivity of the drug, as such its importance cannot be underestimated. As a matter of fact it is useful to remind that benzophenone is not only a very useful model system for the study of DNA photosensitization but is also a widespread pollutant, notably present in cigarette smoke. Ketoprophene, a very common antinflammatory drug, also share the core structure, and the photoactive core, with benzophenone hence it can have similar binding modes and subsequently induce similar DNA (photo)sensitization. Our results have allowed for the first time to resolve the benzophenone:DNA binding modes. Specifically we have shown that benzophenone tends to favor minor groove binding that shows a binding free-energy of about −3.6 kcal mol^{-1} (Table 1). It is important to note that the overall statistical error, related to the conformational space sampling, is particularly low and accounts for 0.4 kcal mol^{-1}. There is no experimental binding constant available for BNZ:DNA; however our data are consistent with DNA binding affinities measured for larger organic minor groove binders – *e.g.* −7.7 kcal mol^{-1} for Hoescht 3328 and −8.6 kcal mol^{-1} for Netropsin.[60] Since the possibility of energy transfer from the minor groove binding mode has been clearly evidenced,[22] especially *via* the photophysical population of the benzophenone T_2 excited state[61,62] our results are consistent with the well known photosensitization characteristics of benzophenone. On the other hand the double insertion mode appears much less favorable; indeed not only the binding energy has been found to be slightly positive (1.6 kcal mol^{-1}), but also the

formation of double-insertion is most probably characterized by a much larger kinetic barrier. However, due to the relatively high statistical errors (1.9 kcal mol^{-1}) and difficult sampling of the complex DNA reorganization, double insertion cannot be totally ruled out. We may however certainly affirm that the former is, at most, participating only with a very minor, and even marginal role in the chemistry of benzophenone interacting with DNA.

5 Conclusion

We have reported the thermochemistry, and in particular binding free energy, for the interaction between DNA and the paradigmatic sensitizer benzophenone. The free energy has been obtained using atomistic molecular dynamics and in particular using recently developed methods allowing proper sampling of the conformational space. In particular we have been able to clearly resolve between the two possible interaction structures hypothesized previously, *i.e.* minor groove binding and double insertion. Indeed, the lack of experimentally resolved structures leaves open the possibility of the coexistence of the two modes that happen to have very different photochemical and photophysical behaviors. Our simulation, on the other hand, has clearly pointed toward a strong preference for minor groove binding that should, hence, be considered as the predominant species. From a more methodological point of view we have been able to present a rigorous method to provide thermochemical values of complex DNA/drug aggregates. The great flexibility of B-DNA results in very difficult sampling, which in some cases may lead to relatively large errors. This is also due to the fact that many degrees of freedom, in particular, on the backbone dihedrals, are coupled and entangled in a non-trivial way, while global variables such as RMSD may lead to degeneracy. In the future we plan to extend the present protocol to different interactions between DNA and drugs, in particular, we plan to tackle the non-canonical DNA structures, such as G-quadruplexes, that are extremely significant in controlling cellular senescence and gene promotion.

Acknowledgements

Support from Université de Lorraine and CNRS is gratefully acknowledged. This work was supported by CNRS actions Interdisciplinaires 2016. This work has also been realized in the framework of the COST Action CM1201 ''Biomimetic Radical Chemistry''. The authors wish to thank Mr Gilles Touchagues for his contribution and discussion.

References

1 A.-M. Florea and D. Büsselberg, *Cancers*, 2011, **3**, 1351–1371.
2 E. R. Jamieson and S. J. Lippard, *Chem. Rev.*, 1999, **99**, 2467–2498.
3 K. Gurova, *Future Oncol.*, 2009, **5**, 1685–1704.
4 E. S. Antonarakis and A. Emadi, *Cancer Chemother. Pharmacol.*, 2010, **66**, 1–9.
5 Y. K. Yan, M. Melchart, A. Habtemariam and P. J. Sadler, *Chem. Commun.*, 2005, 4764–4776.
6 G. Suss-Fink, *Dalton Trans.*, 2010, **39**, 1673–1688.
7 Z. Molphy, C. Slator, C. Chatgilialoglu and A. Kellett, *Front. Chem.*, 2015, **3**, 28.
8 M. O'Connor, A. Kellett, M. McCann, G. Rosair, M. McNamara, O. Howe, B. S. Creaven, S. McClean, A. F.-A. Kia, D. O'Shea and M. Devereux, *J. Med. Chem.*, 2012, **55**, 1957–1968.
9 Z. Molphy, A. Prisecaru, C. Slator, N. Barron, M. McCann, J. Colleran, D. Chandran, N. Gathergood and A. Kellett, *Inorg. Chem.*, 2014, **53**, 5392–5404.
10 T. J. Dougherty, C. J. Gomer, B. W. Henderson, G. Jori, D. Kessel, M. Korbelik, J. Moan and Q. Peng, *J. Natl. Cancer Inst.*, 1998, **90**, 889–905.
11 P. Agostinis, K. Berg, K. A. Cengel, T. H. Foster, A. W. Girotti, S. O. Gollnick, S. M. Hahn, M. R. Hamblin, A. Juzeniene, D. Kessel, M. Korbelik, J. Moan, P. Mroz, D. Nowis, J. Piette, B. C. Wilson and J. Golab, *Cancer J. Clin.*, 2011, **61**, 250–281.
12 M. Ethirajan, Y. Chen, P. Joshi and R. K. Pandey, *Chem. Soc. Rev.*, 2011, **40**, 340–362.
13 R. K. Pandey, *J. Porphyrins Phthalocyanines*, 2000, **04**, 368–373.
14 R. R. Allison and C. H. Sibata, *Photodiagn. Photodyn. Ther.*, 2010, **7**, 61–75.
15 S. Neidle and G. Parkinson, *Nat. Rev. Drug Discovery*, 2002, **1**, 383–393.
16 S. Neidle, *FEBS J.*, 2010, **277**, 1118–1125.
17 A. Terenzi, R. Bonsignore, A. Spinello, C. Gentile, A. Martorana, C. Ducani, B. Hogberg, A. M. Almerico, A. Lauria and G. Barone, *RCS Adv.*, 2014, **4**, 33245–33256.
18 B. M. Zeglis, V. C. Pierre and J. K. Barton, *Chem. Commun.*, 2007, 4565–4579.
19 E. Dumont and A. Monari, *Front. Chem.*, 2015, **3**, 43.
20 M. C. Cuquerella, V. Lhiaubet-Vallet, J. Cadet and M. A. Miranda, *Acc. Chem. Res.*, 2012, **45**, 1558–1570.
21 E. Dumont and A. Monari, *J. Phys. Chem. Lett.*, 2013, **4**, 4119–4124.
22 E. Dumont, M. Wibowo, D. Roca-Sanjuan, M. Garavelli, X. Assfeld and A. Monari, *J. Phys. Chem. Lett.*, 2015, **6**, 576–580.
23 M. Marazzi, M. Wibowo, H. Gattuso, E. Dumont, D. Roca-Sanjuan and A. Monari, *Phys. Chem. Chem. Phys.*, 2016, **18**, 7829–7836.
24 R. Galindo-Murillo, J. C. Garcia-Ramos, L. Ruiz-Azuara, T. E. Cheatham and F. Cortes-Guzman, *Nucleic Acids Res.*, 2016, **43**, 5364–5376.
25 T. Lelièvre, G. Stoltz and M. Rousset, *Free energy computations: A mathematical perspective*, World Scientific, 2010.
26 C. Chipot and A. Pohorille, *Free energy calculations*, Springer, 2007.
27 C. Chipot, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 71–89.
28 W. Treesuwan, K. Wittayanarakul, N. G. Anthony, G. Huchet, H. Alniss, S. Hannongbua, A. I. Khalaf, C. J. Suckling, J. A. Parkinson and S. P. Mackay, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10682–10693.

29 E. Dumont and A. Monari, *J. Phys. Chem. B*, 2014, **119**, 410–419.
30 H. Gattuso, T. Duchanois, V. Besancenot, C. Barbieux, X. Assfeld, P. Becuwe, P. C. Gros, S. Grandemange and A. Monari, *Front. Chem.*, 2015, **3**, 67.
31 L. Landau and E. Lifshitz, *Statistical Physics*, Oxford Univ. Press, New York, NY, 1938.
32 R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420–1426.
33 J. G. Kirkwood, *J. Chem. Phys.*, 1935, **3**, 300–313.
34 S. B. Singh, D. E. Wemmer and P. A. Kollman, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 7673–7677.
35 J. Dolenc, C. Oostenbrink, J. Koller and W. F. van Gunsteren, *Nucleic Acids Res.*, 2005, **33**, 725–733.
36 H.-J. Woo and B. Roux, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6825–6830.
37 J. C. Gumbart, B. Roux and C. Chipot, *J. Chem. Theory Comput.*, 2012, **9**, 794–802.
38 J. C. Gumbart, B. Roux and C. Chipot, *J. Chem. Theory Comput.*, 2013, **9**, 3789–3798.
39 S. B. Dixit and C. Chipot, *J. Phys. Chem. A*, 2001, **105**, 9795–9799.
40 D. L. Mobley, A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet and K. A. Dill, *J. Mol. Biol.*, 2007, **371**, 1118–1134.
41 C. Velez-Vega and M. K. Gilson, *J. Comput. Chem.*, 2013, **34**, 2360–2371.
42 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
43 A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
44 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
45 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
46 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
47 S. E. Feller, Y. Zhang, R. W. Pastor and B. R. Brooks, *J. Phys. Chem.*, 1995, **103**, 4613–4621.
48 T. Darden, D. York and L. Pedersen, *J. Phys. Chem.*, 1993, **98**, 10089–10092.
49 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
50 R. Lavery, M. Moakher, J. Maddocks, D. Petkeviciute and K. Zakrzewska, *Nucleic Acids Res.*, 2009, **37**, 5917–5929.
51 G. Fiorin, M. L. Klein and J. Hénin, *Mol. Phys.*, 2013, **111**, 3345–3362.
52 J. Hermans and S. Shankar, *Isr. J. Chem.*, 1986, **27**, 225–227.
53 M. K. Gilson, J. A. Given, B. L. Bush and J. A. McCammon, *Biophys. J.*, 1997, **72**, 1047.
54 P. Liu, F. Dehez, W. Cai and C. Chipot, *J. Chem. Theory Comput.*, 2012, **8**, 2606–2616.
55 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
56 Y. B. Yu, P. L. Privalov and R. S. Hodges, *Biophys. J.*, 2001, **81**, 1632–1642.
57 A. V. Vargiu, P. Ruggerone, A. Magistrato and P. Carloni, *Biophys. J.*, 2008, **94**, 550–561.
58 E. Dumont and A. Monari, *J. Phys. Chem. B*, 2015, **119**, 410–419.
59 P. Miro, V. Lhiaubet-Vallet, M. L. Marin and M. A. Miranda, *Chem. – Eur. J.*, 2015, **21**, 17051–17056.
60 S. A. Shaikh, S. R. Ahmed and B. Jayaram, *Arch. Biochem. Biophys.*, 2004, **429**, 81–99.
61 M. Marazzi, S. Mai, D. Roca-Sanjuán, M. G. Delcey, R. Lindh, L. González and A. Monari, *J. Phys. Chem. Lett.*, 2016, **7**, 622–626.
62 L. Favero, G. Granucci and M. Persico, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10499–10506.

Chapitre 3 - Membranes et Protéines membranaires

Les protéines membranaires sont essentielles pour la communication cellulaire et les processus d'échanges entre compartiments cellulaires ou cellules. Elles jouent donc un rôle clef dans de nombreux processus biologiques. Les transporteurs représentent une classe importante de protéines membranaires responsables du transport d'ions, de petites molécules de toxines ou bien encore de protéines. Ces protéines sont des acteurs incontournables de l'accomplissement des cycles métaboliques et sont également présentes sur le front des agressions externes qu'elles soient biologiques (import de toxines) ou chimiques (import de métaux lourds). Elles jouent également un rôle dans la détoxification cellulaire. Il apparait donc essentiel de comprendre les mécanismes qui régissent le transport transmembranaire. La modélisation moléculaire, et plus précisément la dynamique moléculaire, est dans ce contexte un complément essentiel aux approches expérimentales permettant d'une part de décrire l'ensemble des processus mis en jeu au niveau moléculaire et d'autre part de préciser le sens physiologique des structures tridimensionnelles obtenus in vitro dans des environnements non-natifs potentiellement dénaturants.

I. Transporteur mitochondrial de l'ADP/ATP

Le transport de petites molécules nécessaires dans la grande majorité des cycles métaboliques de la mitochondrie (Figure 9) est essentiel au métabolisme eucaryote. Le passage au travers de la membrane interne de la mitochondrie est sélectivement contrôlé par une famille de protéines membranaires intégrales (la famille des transporteurs mitochondriaux : MCF) chacune transportant un métabolite spécifique.[47] Le passage au travers de la membrane externe est pour sa part assuré par des porines non sélectives capables de transporter des molécules d'au plus 5000 daltons environ.

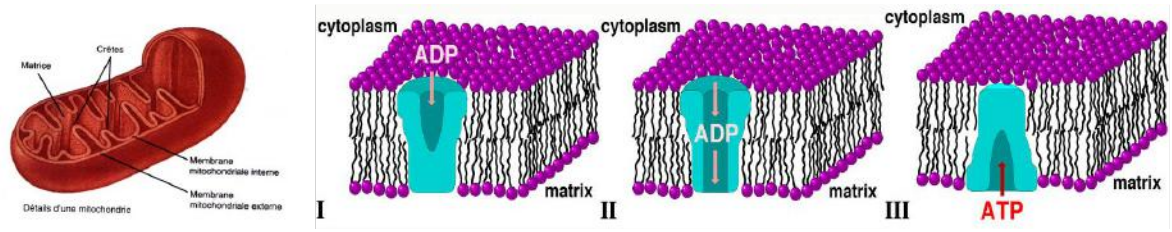


Figure 9. (Gauche) Schéma descriptif de la membrane mitochondriale et (droite) Mécanisme putatif du transport membranaire de l'ADP/ATP au travers d'un transporteur AAC.

Parmi les membres de la famille MCF, AAC joue un rôle physiologique essentiel puisqu'il régit l'économie énergétique des cellules eucaryotes.[48] Il permet d'importer dans la matrice mitochondriale des molécules d'adénosine diphosphate (ADP). Celles- ci sont transformées par

phosphorylation en molécules d'adénosine triphosphate (ATP) qui sont exportées par le même chemin vers le cytoplasme (Figure 9). Ce processus hautement efficace permet d'assurer les besoins de chaque être humain en ATP, à savoir son propre poids chaque 24 heures.

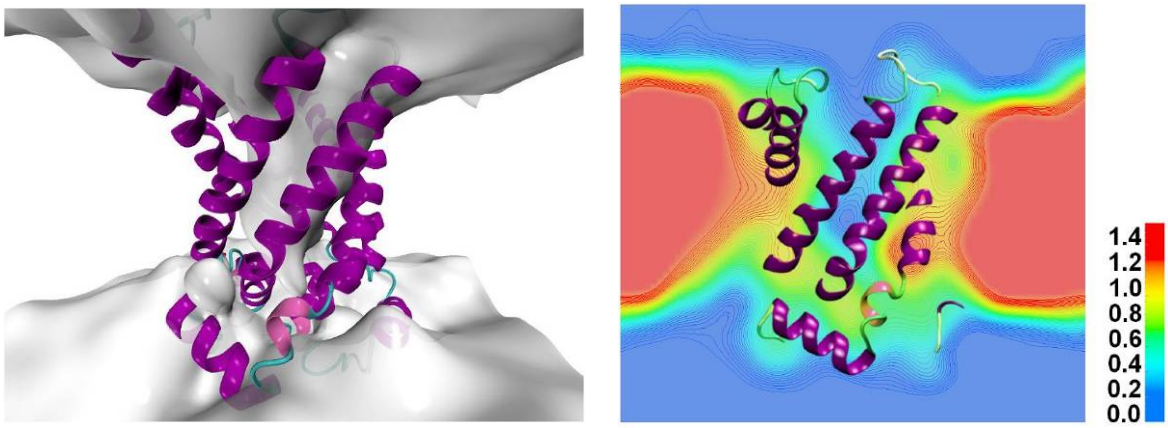


Figure 10. Carte 3D d'isopotentiel électrostatique révélant la présence d'un entonnoir électrostatique capable de guider rapidement les anions vers de fond de la structure (gauche). Plan de coupe du potentiel électrostatique (volts) le long de l'axe longitudinal de l'entonnoir (droite).

La structure cristallographique du transporteur AAC bovin a été résolue à 2.2 A° (PDB:1OKC) par le groupe d'Eva Pebay-Peyroula en présence de l'inhibiteur CATR[49,50] et est restée pendant plusieurs années la seule structure haute résolution d'un MCF, capturée dans son état ouvert côté espace intermembranaire. En 2014, une structure d'un transporteur de levure a été résolue dans le même état conformationnel par le groupe d'Edmund Kunji.[51] Ce même groupe a réussi, plus de 15 ans après la structure 1OKC, à capturer le transporteur dans son état conformationnel ouvert côté matriciel.[52] 1OKC s'organise autour de 6 hélices alpha transmembranaires inclinées par rapport au plan de la membrane, trois d'entre-elles formant un coude au niveau d'un résidu proline. L'arrangement spatial des hélices détermine une large cavité ayant la forme d'un entonnoir de 30 Å de profondeur et de diamètre maximum de 20 Å, largement ouverte sur la face externe de la membrane mitochondriale. La cavité est fermée dans sa partie inférieure du fait de la disposition des hélices coudées. La fermeture est renforcée par la présence de ponts salins et aussi par l'organisation des boucles internes qui forment de courtes hélices enserrant l'extrémité des hélices transmembranaires sur la face interne. Ces interactions impliquent des résidus conservés dans les séquences des transporteurs mitochondriaux de métabolites et pourraient donc être une caractéristique de cette famille.

I.1. Propriétés dynamiques du transporteur mitochondrial AAC dans un environnement lipidique

Nos simulations de dynamique moléculaire (DM) réalisées à partir de la structure cristallographique (code pdb 1OKC) ont permis de caractériser plusieurs propriétés intrinsèques du transporteur AAC. La présence de deux groupes de résidus basiques (arginines et lysines) situes sur les parois de la cavité interne confère à AAC une signature électrostatique très spécifique. L'analyse de cartes de potentiel électrostatique tridimensionnelles révèle la présence d'un entonnoir électrostatique

pouvant potentiellement conduire des espèces chargées négativement vers le fond de la cavité interne de la protéine (Figure 10), i.e. dans la zone d'interconnexion des ponts salins.[53]

Les simulations de dynamique moléculaire ont permis d'observer l'association spontanée d'une molécule d'ADP^{3-} insérée directement dans la protéine libre d'inhibiteur et telle qu'équilibrée dans la membrane. En quelques nanosecondes, le substrat, empruntant le chemin électrostatique en forme d'entonnoir, atteint le fond de la cavité interne de la protéine. Celui-ci converge vers une conformation stable en forte interaction avec les résidus chargés impliqués dans les ponts salins fermant la structure. Par ailleurs la descente de l'ADP^{3-} se fait concomitamment à la réorganisation de ces ponts salins. En particulier, l'interaction de la lysine 32 avec l'acide aspartique 231, rompue dans la protéine sans substrat, se reforme à l'approche de ADP^{3-}. Le site de liaison observé correspond parfaitement au site de liaison suggéré par un ensemble d'expériences biochimiques impliquant les résidus K22, R79, R279 et Y186. La robustesse de ce processus d'association spontanée a été démontrée à la lumière de nombreuses simulations numériques indépendantes et d'une étude concurrente.[54] Une molécule d'ADP^{3-}, positionnée au-dessus de la cavité interne du transporteur, à différentes altitudes et dans diverses orientations, s'associe invariablement à la protéine. De façon tout à fait remarquable, un parcours de plus de 16 Å a pu être observé sur échelle de temps d'environ 40 ns, rendant compte pour la première fois du détail atomique d'un processus d'association spontanée d'un ligand à sa cible. L'analyse minutieuse des trajectoires de dynamique moléculaire rationalise les hypothèses émises à partir de la structure cristallographique quant aux résidus impliqués dans l'association avec l'ADP^{3-} au cours de sa translocation. Par exemple l'observation du glissement du nucléotide le long d'une « échelle » composée de trois tyrosines vient directement illustrer le rôle putatif de ces acides aminés. Ces résultats ont par ailleurs été confortés par des expériences de mutagenèse dirigée.

I.2. Effet inhibiteur des ions chlorures sur l'association de l'ADP au transporteur AAC

L'AAC joue un rôle clé dans les processus métaboliques des cellules, son disfonctionnement peut donc être la source de troubles sévères. La fonction de l'AAC peut être altérée par la mutation de résidus spécifiques ou par l'association d'inhibiteurs puissants. Des expériences menées sur l'AAC de cœur de bœuf ont également montrées que des concentrations importantes en ions magnésiums ou chlorures inhibaient la fonction du transporteur.[55] Mg^{2+} s'associe probablement au substrat qui sous sa forme complexée ne peut plus être transporté. Les ions chlorures en revanche semblent directement affecter les propriétés de AAC.

L'étude des effets d'une forte concentration en chlore sur l'activité de transport d'AAC a été réalisée en combinant de manière synergique des travaux théoriques et expérimentaux. Nos collaborateurs grenoblois, S. Ravaud, E. Pebay-Peyroula (Institut de Biologie Structurale) et A. Frelet-Barrand (Laboratoire de Physiologie Cellulaire Vegetale) ont mesuré l'effet inhibiteur des ions chlorures grâce à des expériences de transport d'ATP radio-marqués au travers d'AAC de plante exprimé dans E.coli. (Figure 11A). D'un point de vue théorique les effets des ions chlorures ont été appréhendés au travers de simulations réalisées pour diverses concentrations en sel : faibles (0.1M et 0.6 M NaCl) et élevées (0.6M NaCl). 30 ns de dynamique moléculaire ont montré que pour apo-

AAC, une concentration élevée en chlore conduit à l'écrantage de résidus chargés essentiels pour la fonction du transporteur.[56]

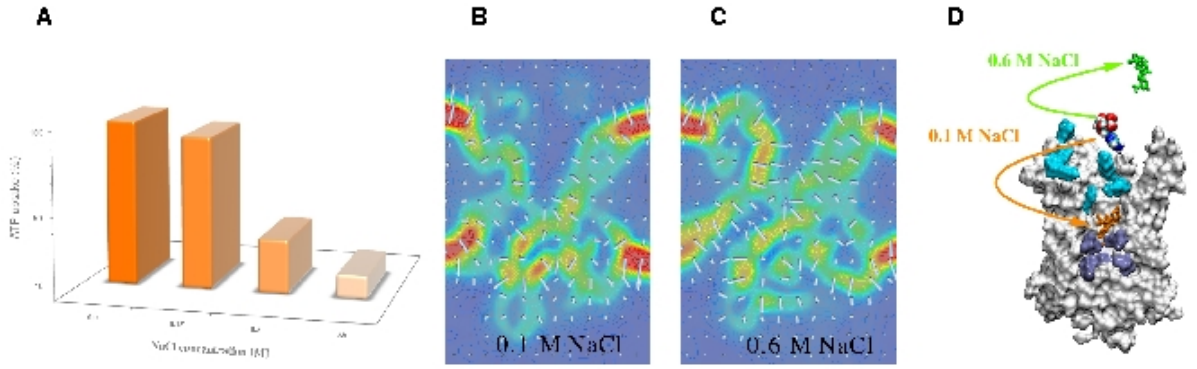


Figure 11. L'activité du transporteur de l'ADP/ATP dépend de la concentration en chlore. (A) Les expériences de transport d'ATP radio-marqués au travers de l'AAC de plante, *Arabidopsis thaliana* exprimé dans des cellules de *E. coli*, montrent que l'import d'ATP est inhibé par une concentration enlevée en chlore. (B, C) La signature électrostatique de *apo*- AAC de cœur de bœuf est affecté par la présence d'ions chlorures comme le montrent les vues en coupe des cartes 3D du potentiel électrostatique pour une faible (0.1M NaCl) et une forte (0.6M NaCl) concentration en sel. (D) Des simulations d'association spontanée du substrat conduisent à liaison de ADP^{3-} à la protéine seulement en l'absence d'une concentration enlevée en ions chlorures.

Ceci a pour conséquence de modifier les propriétés électrostatiques de AAC (figures 11B and 11C). Dans ces conditions, l'entonnoir électrostatique responsable de l'acheminement d'ADP^{3-} au fond de la structure ne pénètre pas la protéine aussi profondément qu'en l'absence de sels. Une série de simulations de dynamique moléculaire d'association protéine-ADP^{3-} a été effectuée pour aller plus loin dans la rationalisation des effets des ions chlorures. Pour l'ensemble de ces tests, la liaison d'ADP^{3-} a été observée seulement en présence de faibles concentrations en chlore (figure 11D). L'analyse de la conservation des résidus spécifiquement ciblé par les ions chlorures dans un ensemble de séquences d'AAC a permis d'étendre nos observations à toute la famille des transporteurs mitochondriaux d'ADP/ATP. L'ensemble de nos résultats montre que la concentration en chlore joue le rôle d'un loquet électrostatique pour la famille des AAC pouvant interdire à l'ADP^{3-} de rejoindre son site de liaison.

I.3. Effets de mutations pathogènes sur les propriétés du transporteur AAC

Plusieurs mutations ponctuelles du transporteur mitochondrial de l'ADP et de l'ATP (AAC) sont associées à des pathologies humaines graves.[57,58] En synergie avec l'équipe de Eva Pebay Peyroula, qui a démontré la relation entre la mutation pathogène et l'activité de transport de l'AAC, nous avons réalisés des simulations de dynamique moléculaire mettant en lumière les mécanismes impliqués pour les six mutations pathogènes connues A90D, L98P, D104G, A114P, A123D et V289M. Pour chacune de ces mutations nous avons réalisé de longues simulations de dynamique moléculaire qui nous ont servi à analyser la dynamique interne des mutants. Nous avons par ailleurs conduit de nombreuses simulations d'association spontanée pour caractériser l'effet de certaines mutations sur l'association de ADP. Plus spécifiquement, nous avons clairement montré que les mutations se traduisent au niveau moléculaire soit par une modification de la topologie du champ

électrique au sein de la protéine membranaire, altérant ainsi l'association spontanée du substrat à la protéine, soit par une diminution de la plasticité de celle-ci, modifiant la capacité de la protéine à changer de conformation (voir figure 12).[59]

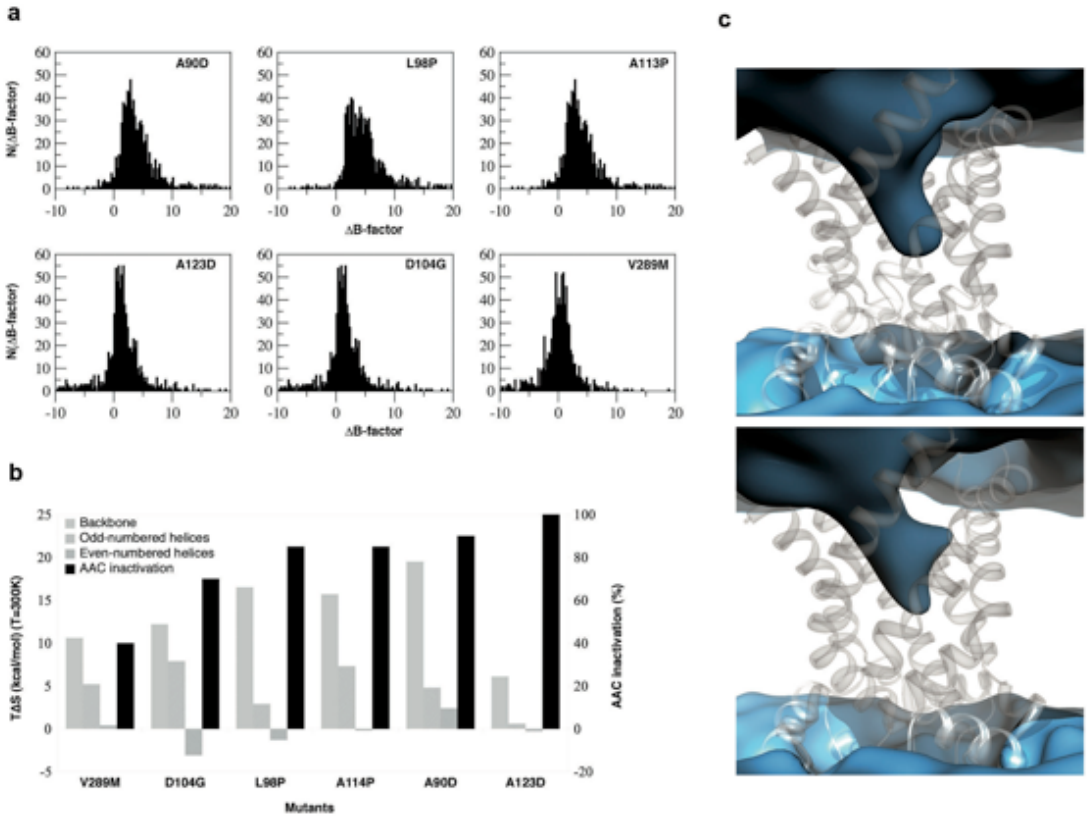


Figure 12. a. Distribution de la différence des facteurs B obtenus pour les atomes du squelette des mutants et de la protéine native (ΔB-facteur e, Å²). b. Différences d'entropies configurationnelles entre la protéine native et chaque mutant mesurées sur les atomes du squelette (décomposes pour les hélices paires et impaires) versus le pourcentage d'inactivation en réponse à la mutation. c. Cartes de potentiel électrostatique pour la protéine native apo (en haut) et le mutant A123D apo (en bas). La signature électrostatique du mutant A123D est fortement altérée.

I.4. Sélectivité de AAC vis-à-vis des substrats ADP^{3-} et ATP^{4-}

Une des caractéristiques spécifiques du transporteur AAC et de ne transporter exclusivement que l'ADP et l'ATP, ni les formes protomères de ces molécules ni aucun substrat analogue tel que GDP, GTP, AMP, GMP ne diffusent au travers de la protéine.[60] Par le biais de simulations de dynamique moléculaire nous avons identifié les bases moléculaires de cette haute sélectivité. Nos résultats, qui viennent appuyer les mesures d'activité de AAC réalisées par nos collègues de l'IBS à Grenoble et au MRC à Cambridge, montrent que l'ensemble des nucléotides est capable de s'associer spontanément à la structure de AAC ouverte coté espace intermembranaire, le groupement phosphate terminal des 6 nucléotides étudiés vient se fixer systématiquement aux mêmes résidus basiques situés dans le fond de la cavité du transporteur (figure 13). Nos simulations montrent en revanche, que à la suite de la fixation du groupement phosphate, seule la partie nucléosidique de l'ADP et l'ATP est capable de s'ancrer spécifiquement au transporteur. Pour l'AMP, il semble que la chaine phosphate soit trop courte pour permettre à l'adénine d'interagir correctement avec les résidus du site de liaison. En raison d'une plus grande polarité par rapport à

l'adénine, le groupement guanine présent dans GMP, GDP et GTP ne parvient pas à établir une interaction stable avec les résidus formant le site reconnaissance de ADP et ATP.[61]

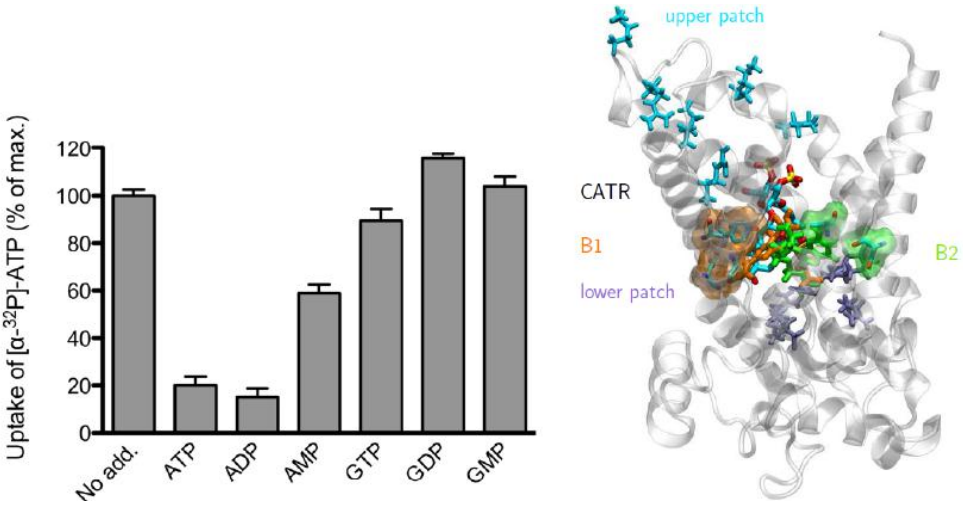


Figure 13. (Gauche) Expériences d'association compétitive d'ATP radio marqué avec des nucléotides analogues, seule la présence de ADP et ATP perturbent significativement l'import de a^{32}P-ATP. (Droite) Vue globale des deux principaux motifs de liaisons de ADP et ATP au fond de la cavité de AAC dans sa conformation ouverte coté espace intermembranaire.

I.5. Modulation de l'activité de transport de AAC sous l'effet du pH

Une autre propriété de AAC est d'avoir une activité qui dépend du pH et qui présente un maximum aux alentours de la valeur du pH dans l'espace intermembranaire (~6.8). Cette dépendance est liée en partie à l'état de protonation du substrat, dont seule la forme déprotonée peut être transportée, mais elle semble également associée à l'état de protonation d'au moins un des résidus de la protéine.

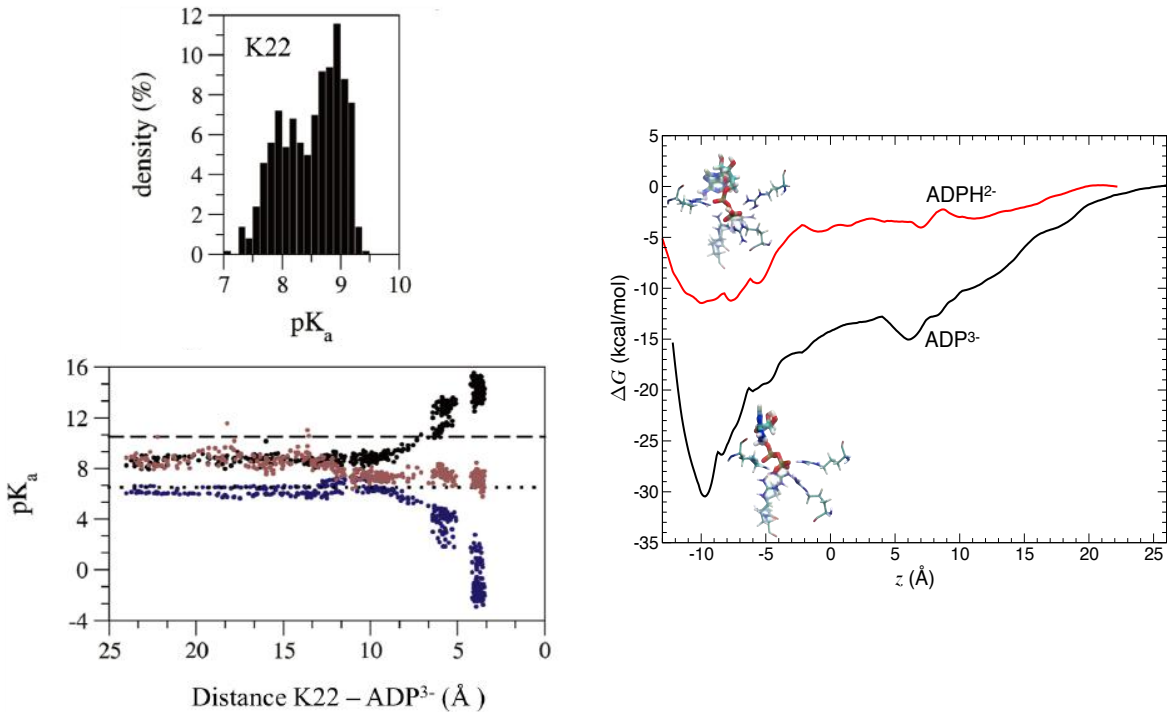


Figure 14. (Gauche) (Haut) Distribution des valeurs de pKa de la lysine 22 pour la protéine en l'absence de substrat. (Bas) Variation du pKa de la lysine 22 (noir) et de l'ADP^{3-} (bleu) au cours de l'association de l'ADP à son site de liaison. (Droite) Profils d'énergie libre association à la complexation par AAC de ADP^{3-} et de ADPH^{2-}

Nos simulations de dynamique moléculaires combinées à des calculs prédictifs de pKa laissent supposer qu'un résidu lysine présent dans la cavité de AAC est impliqué dans la sensibilité au pH. La lysine 22 est le seul acide aminé accessible dont l'environnement induit une réduction importante de son pKa vers des valeurs proches de celle du pH de l'espace intermembranaire (voir figure 14). Des simulations d'association spontanées montrent que l'arrimage de ADP^{3-} s'effectue beaucoup moins facilement lorsque ce résidu est protoné (situation a pH enlevé). Par ailleurs des profils d'énergie libre, associés à l'arrimage du substrat, montrent que l'association de ADPH^{2-} (bas pH) à AAC est beaucoup moins forte que celle de ADP^{3-} (situation à pH physiologique).

Pour démontrer formellement l'implication de la lysine 22 dans la dépendance au pH de l'activité de AAC, nous avons considéré une mutation susceptible de conserver (au moins en partie) l'activité du transporteur tout en modifiant sa réponse à la variation de pH. Des simulations conduites sur le mutant K22R montre que ce dernier est capable de lier l'ADP et que l'arginine en position 22, à l'inverse de la lysine, ne présente pas de déplacement de pKa. Les tests d'activités réalisés par nos collègues de l'IBS sur le mutant K223 en fonction du pH sont venus finalement démontrer la robustesse de notre prédiction (figure 15).[62]

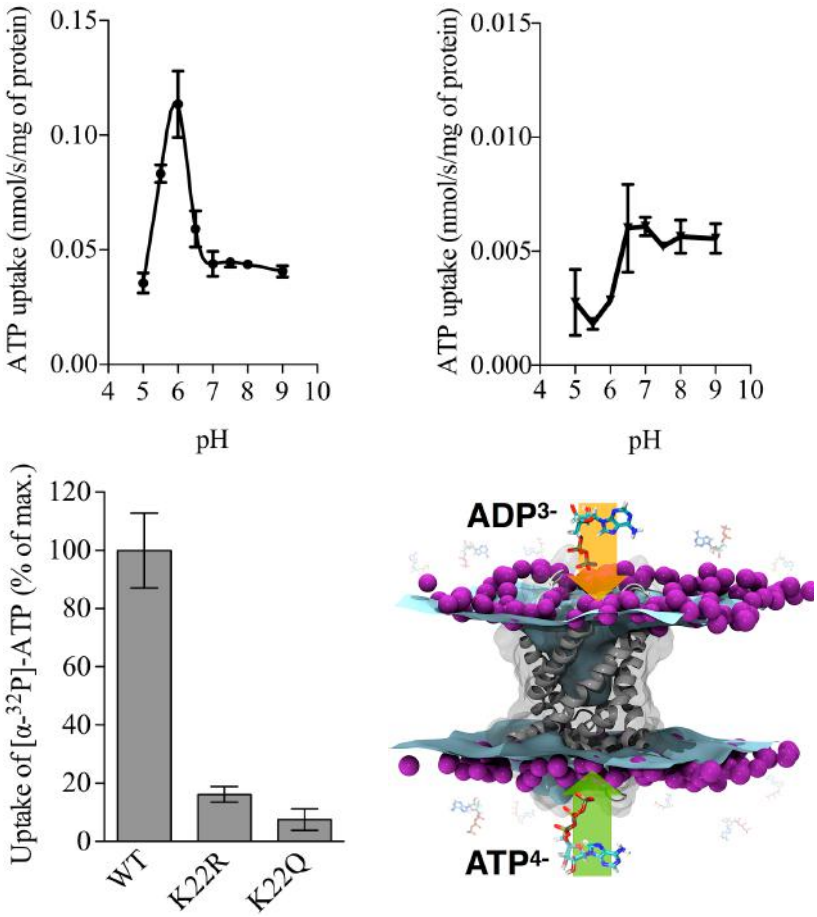


Figure 15. (Haut gauche) Dépendance au pH de l'import de [α-^{32}P]-ATP par le transporteur AAC humain exprimé dans E. coli. (Haut droit) Dépendance au pH de l'import de [α-^{32}P]-ATP par le mutant K22R du transporteur AAC humain exprimé dans E. coli. (Bas gauche) Activité de transport des mutants K22R and K22Q mutants en comparaison de la protéine native. (Bas droit). Illustration schématique de l'échange ADP^{3-}/ATP^{4-} au travers de AAC.

Ci-après sont reproduits 3 articles illustrant mes travaux dédiés au transporteur de l'ADP/ATP:

- Dehez, F., Pebay-Peyroula, E. & Chipot, C. Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel. *J. Am. Chem. Soc.* **130**, 12725–12733 (2008).
- Mifsud, J., Ravaud, S., Krammer, E.-M., Chipot, C., Kunji, E. R. S. *, Pebay-Peyroula, E. & Dehez, F. * The substrate specificity of the human ADP/ATP carrier AAC1. *Mol. Membr. Biol.* **30**, 160–168 (2013).
- Bidon-Chanal, A., Krammer, E.-M., Blot, D., Pebay-Peyroula, E., Chipot, C., Ravaud, S. * & Dehez, F. * How Do Membrane Transporters Sense pH? The Case of the Mitochondrial ADP–ATP Carrier. *J. Phys. Chem. Lett.* **4**, 3787–3791 (2013)

J|A|C|S
ARTICLES
Published on Web 08/26/2008

Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel

François Dehez,[†] Eva Pebay-Peyroula,[‡] and Christophe Chipot*,[†]

Equipe de dynamique des assemblages membranaires, UMR No. 7565 CNRS-UHP, Nancy Université, BP 239, 54506 Vandœuvre-lès-Nancy cedex, France, and Institut de biologie structurale, UMR No. 5075 CEA-CNRS-UJF, 41, rue Jules Horowitz, 38027 Grenoble cedex 1, France

Received May 4, 2008; E-mail: Christophe.Chipot@edam.uhp-nancy.fr

Abstract: The ADP/ATP carrier (AAC) is a membrane protein of paramount importance for the energy-fueling function of the mitochondria, transporting ADP from the intermembrane space to the matrix and ATP in the opposite direction. On the basis of the high-resolution, 2.2-Å structure of the bovine carrier, a total of 0.53 μs of classical molecular dynamics simulations were conducted in a realistic membrane environment to decipher the early events of ADP^{3-} translocation across the inner membrane of the mitochondria. Examination of *apo*-AAC underscores the impermeable nature of the carrier, impeding passive transport of permeants toward the matrix. The electrostatic funnel illuminated from three-dimensional mapping of the electrostatic potential forms a privileged passageway anticipated to drive the diphosphate nucleotide rapidly toward the bottom of the internal cavity. This conjecture is verified in the light of repeated, independent numerical experiments, whereby the permeant is dropped near the mouth of the mitochondrial carrier. Systematic association of ADP^{3-} to the crevice of the AAC, an early event in its transport across the inner membrane, is accompanied by the formation of an intricate network of noncovalent bonds. Simulations relying on the use of an adaptive biasing force reveal for the first time that the proposed binding site corresponds to a minimum of the free energy landscape delineating the translocation of ADP^{3-} in the carrier. The present work paves the way to the design of novel nucleotides and new experiments aimed at unveiling key structural features in the chronology of ADP/ATP transport across the mitochondrial membrane.

Introduction

The mitochondrial ADP/ATP carrier (AAC) is an important constituent of the inner mitochondrial membrane, as it imports ADP and exports ATP toward the cytosol. AAC is related to the respiratory chain, which creates a proton gradient across the inner membrane, further utilized to catalyze the ATP synthesis from ADP and inorganic phosphate via the ATP synthetase. ATP is an essential molecule that provides chemical energy in most cellular reactions—a human being consumes its own weight of ATP per day. To fulfill these requirements, the cell machineries responsible not only for ATP synthesis but also for the transport of the substrates have to be very efficient. ADP^{3-} and ATP^{4-} are imported and exported, respectively, through the inner membrane by a very specific protein, the AAC. The outer mitochondrial membrane can be crossed through porin-like proteins in a somewhat less specific fashion.

Several energy-generating pathways are localized in the matrix of mitochondria (i.e., citric acid cycle and fatty acid oxidation), as well as other pathways (i.e., synthesis and degradation of amino acids). Heat generation is also triggered at the level of the inner mitochondrial membrane by dissipation of the proton gradient.[1] All these processes require the import and export of metabolites or protons. In spite of large differences among these substrates, a common feature emerges: A variety of them are transported by carriers belonging to the same family, viz. the mitochondrial carrier family (MCF)—in the particular instance of protons, translocation is accomplished by uncoupling proteins (UCPs), which are specialized members of the MCF. The latter is characterized by a specific amino acid motif, PxD/ExxK/RxK/R-(20−30 residues)-D/EGxxxxaK/RG, where "a" represents an aromatic residue, repeated three times in the sequence.[2,3] The triplication of about 100 amino acids including this motif most probably stems from an ancestral gene duplication.

Several pathologies involve such carriers.[4] Among all MCF carriers, the AAC is the most abundant in the membrane and represents up to 10% of the proteins that compose the inner membrane of bovine heart mitochondria. For these reasons, it was the first to be characterized.[5,6] In addition, two inhibitors, both acting as strong poisons by blocking nucleotide translocation,[7,8] were extensively utilized as tools for characterizing different conformational states of the carrier. On account of its

[†] Equipe de dynamique des assemblages membranaires, UMR No. 7565 CNRS-UHP, Nancy Université.

[‡] Institut de biologie structurale, UMR No. 5075 CEA-CNRS-UJF.

(1) Ricquier, D.; Casteilla, L.; Bouillaud, F. *FASEB J.* **1991**, *5*, 2237–2242.
(2) Walker, J. E.; Runswick, M. J. *J. Bioenerg. Biomembr.* **1993**, *25*, 435–446.
(3) Jezek, P.; Jezek, J. *FEBS Lett.* **2003**, *534*, 15–25.
(4) Palmieri, F. *Pflug. Arch.* **2004**, *447*, 689–709.
(5) Vignais, P. V. *Biochim. Biophys. Acta* **1976**, *456*, 1–38.
(6) Klingenberg, M. *Arch. Biochem. Biophys.* **1989**, *270*, 1–14.
(7) Vignais, P. V.; Vignais, P. V.; Stanislas, E. *Biochim. Biophys. Acta* **1962**, *60*, 284–300.
(8) Henderson, P. J.; Lardy, H. A. *J. Biol. Chem.* **1970**, *245*, 1319–1326.

10.1021/ja8033087 CCC: $40.75

abundance in natural sources, the structure of the AAC was also the first to be explored by electron microscopy[9] and X-ray crystallography. The AAC could be crystallized in the presence of the inhibitor carboxyatractyloside (CATR),[10] which is known to bind to the carrier from the intermembrane space (IMS), hampering ADP^{3-} import.

The structure was solved to 2.2 Å resolution,[11] revealing for the first time the overall fold of MCF carriers and highlighting the role played by the MCF motif in the folding. Proline residues induce sharp kinks in three of the transmembrane (TM) helices, whereas the basic and acidic amino acids from the first part of the motif link these three helical segments together through salt bridges (see Supporting Information). The three salt bridges are located at the bottom of a large cavity opened toward the IMS, two of them having been already postulated on the basis of revertant studies.[12] The cavity is lined by four patches formed by basic residues, described to be important for the function (for a review of all the mutants and their implications, see ref 13). It contains three aligned tyrosine residues and possesses a slight constriction close to the bottom, consisting of four amino acids that are strictly conserved among ADP/ATP carriers.[14] Even though the position of many amino acids known to alter nucleotide transport could be highlighted, thereby shedding the first light on a putative mechanism, the structure of a single conformation alone does not answer fundamental questions about the access of ADP^{3-} into the cavity, the selective recognition of ADP^{3-} and ATP^{4-}, as well as the structural modifications that allow the transport to occur. Unfortunately, blocking the carrier at key stages of the complete transport process in order to isolate intermediate states remains a daunting task. Inhibitors like CATR or bongkrekic acid (BA) have proven to constitute powerful tools, albeit, considering that they are chemically rather far from true nucleotides, they are admittedly of limited use. Solving structures that participate in the transport mechanism, therefore, require the synthesis of alternate nucleotide mimetics. This appealing strategy, pursued for other membrane proteins, e.g., calcium ATPase,[15,16] has helped pinpoint important intermediate states involved in their function. It is, therefore, becoming clear that, in order to reach the proposed goal, an improved knowledge of the structural and dynamic features of the carrier−nucleotide interaction is necessary.

The existence of a general transport mechanism among all MCF members remains also an open question. In addition, the structure determined by X-ray crystallography is monomeric, without any obvious interface conducive to strong dimerization.[17] Furthermore, the size of the cavity is sufficient to accommodate a nucleotide. On the basis of previous biochemical and functional experiments, it was almost admitted that MCF carriers function as dimers. The three-dimensional structure of the carrier, however, casts doubt on this hypothesis. Recent contributions report that the AAC solubilized in detergents under different conditions is monomeric,[18,19] and that the yeast carrier is able to function as a monomer.[20]

The present research article addresses the first act of the transport mechanism, namely the entrance of ADP^{3-} into the cavity and its subsequent association to putative binding sites. To reach this objective, molecular dynamics (MD) simulations were carried out over realistic time scales, based on the crystallographic structure of the mitochondrial carrier. To pinpoint possible structural modifications of the membrane protein upon transport of the diphosphate nucleotide, the AAC was examined both in its *apo* conformation and with 10 mM ADP and ATP. In addition, free energy calculations relying upon the knowledge of the average force exerted along a unidimensional construct of the reaction coordinate are employed to locate the possible binding sites of ADP^{3-} in the carrier. Inferences drawn from the present set of simulations shed new light on the second stage of the transport mechanism, whereby ADP^{3-} is imported into the matrix.

Methods and Computational Details

Molecular Assays. In this contribution, two distinct assays were considered to model in a realistic environment the mitochondrial carrier in its *apo* form and with 10 mM ADP and ATP. The initial conformation of the AAC corresponds to PDB structure 1OKC, from which the specific inhibitor carboxyatractyloside (CATR) was removed. The carrier was subsequently inserted in a thermalized, fully hydrated palmitoyloleylphosphatidylcholine (POPC) bilayer consisting of 228 lipid units in equilibrium with an aqueous phase formed by 10 553 water molecules. The net excess charge of *apo*-AAC was compensated by 18 chloride counterions. In the second assay featuring 10 mM ADP and ATP, i.e., 10 ADP^{3-} and 10 ATP^{4-}, the initial Cartesian coordinates of the carrier correspond to the end point of the simulation of *apo*-AAC. The the di- and triphosphate nucleotides were placed initially on each side of the lipid bilayer, viz., on the intermembrane space side and on the matrix side, respectively. The additional negative charges of the ADP^{3-} and ATP^{4-} were balanced by 52 sodium counterions. The dimensions of the two assays after appropriate equilibration were ca. $84 \times 75 \times 97$ $Å^3$.

MD Simulation. All simulations were performed using the NAMD simulation package[21] in the isobaric−isothermal ensemble. The pressure and the temperature were fixed at 1 bar and 300 K, respectively, employing the Langevin piston algorithm[22] and softly damped Langevin dynamics. The molecular assays were replicated in the three directions of Cartesian space by means of periodic boundary conditions. The particle-mesh Ewald method[23] was employed to compute electrostatic interactions. The short-range Lennard-Jones potential was smoothly truncated. The *r*-RESPA multiple time-step propagator[24] was used to integrate the equations of motion with a time step of 2 and 4 fs for for short- and long-

(9) Kunji, E. R. S.; Harding, M. *J. Biol. Chem.* **2003**, *278*, 36985–36988.
(10) Dahout-Gonzalez, C.; Brandolin, G.; Pebay-Peyroula, E. *Acta Crystallogr. D: Biol. Crystallogr.* **2003**, *59*, 2353–2355.
(11) Pebay-Peyroula, E.; Dahout-Gonzalez, C.; Kahn, R.; Trézéguet, V.; Lauquin, G. J. M.; Brandolin, G. *Nature* **2003**, *426*, 39–44.
(12) Nelson, D. R.; Felix, C. M.; Swanson, J. M. *J. Mol. Biol.* **1998**, *277*, 285–308.
(13) Nury, H.; Dahout-Gonzalez, C.; Trézéguet, V.; Lauquin, G. J. M.; Brandolin, G.; Pebay-Peyroula, E. *Annu. Rev. Biochem.* **2006**, *75*, 713–741.
(14) Pebay-Peyroula, E.; Brandolin, G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 420–425.
(15) Toyoshima, C.; Nomura, H.; Tsuda, T. *Nature* **2004**, *432*, 361–368.
(16) Jensen, A. M. L.; Sørensen, T. L. M.; Olesen, C.; Møller, J. V.; Nissen, P. *EMBO J.* **2006**, *25*, 2305–2314.
(17) Nury, H.; Dahout-Gonzalez, C.; Trézéguet, V.; Lauquin, G.; Brandolin, G.; Pebay-Peyroula, E. *FEBS Lett.* **2005**, *579*, 6031–6036.
(18) Bamber, L.; Harding, M.; Butler, P. J. G.; Kunji, E. R. S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16224–16229.
(19) Nury, H. Études structurales du transporteur d'ADP/ATP. Ph.D. thesis, Université Joseph Fourier, 2007.
(20) Bamber, L.; Harding, M.; Monné, M.; Slotboom, D. J.; Kunji, E. R. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10830–10834.
(21) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, L.; Kalé, R. D.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
(22) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
(23) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(24) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Phys. Chem. B* **1992**, *97*, 1990–2001.

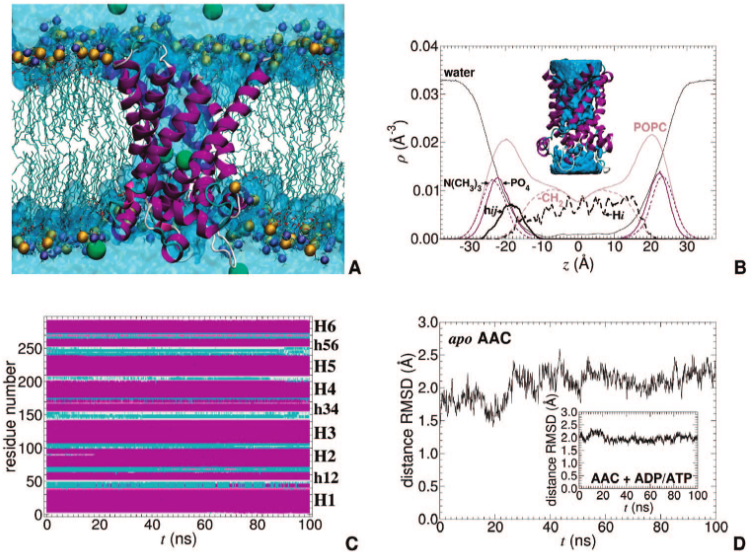


Figure 1. (A) *apo*-AAC in a fully hydrated POPC bilayer after 0.1 μs. TM α-helices are shown as purple ribbons. Chloride counterions are depicted as green van der Waals spheres. A semitransparent, continuous representation is used for water. Lipid units are featured as cyan rods, with their phosphate and choline groups depicted as orange and blue van der Waals spheres, respectively. The image rendering was done with VMD.[41] (B) Number density profiles of the molecular assembly averaged over the entire 0.1-μs trajectory. H_i denotes the subset of TM α-helices, whereas h_{ij} corresponds to the three short α-helical stretches. Inset: Mass-weighted isodensity map of water in *apo*-AAC averaged over the 0.1-μs trajectory and highlighting the absence of conduction between the matrix and the intermembrane space of the mitochondria. (C) Time evolution of the secondary structure of *apo*-AAC. α- and 3_{10}-helices, coils, and turns are highlighted in purple, pink, white, and cyan, respectively. (D) Time evolution of the distance rmsd over the backbone atoms forming *apo*-AAC with respect to the crystallographic structure.[11]

range forces, respectively. Covalent bonds involving a hydrogen atom were constrained to their equilibrium length by means of the Rattle algorithm.[25] The mitochondrial AAC and its environment were described by the all-atom CHARMM27 force field.[26,27] Each assay, i.e., the AAC in its *apo* form or with 10 mM ADP and ATP, was run over a period of 0.1 μs after proper equilibration. To appraise the propensity of the permeant to reach in a systematic fashion a putative binding site located at the bottom of the internal cavity, seven additional, shorter simulations of the second assay were carried out, assigning different sets of initial positions and momenta to an ADP molecule dropped near the mouth of the carrier. These different numerical experiments of duration equal to 17.5, 77, 41, 42, 5.3, 3.6, and 20 ns correspond to a total simulation time of ca. 0.21 μs. Running on the 16 cores of an array of two 2.33-GHz Intel Xeon QC processors communicating via a low-latency Infiniband network, the wall clock time was equal to 0.62 day per nanosecond.

Free Energy Calculations. To investigate the transport of ADP in the cavity of the mitochondrial AAC, a surrogate reaction coordinate was defined as the distance separating the center of mass of the permeant from the centroid of the lower region of helices H1, H3, and H5, i.e., below the hinges formed by the triplet of proline residues P27, P132, and P229, projected onto the z-direction of Cartesian space. Variation of the free energy, $\Delta G(z)$, along z was determined using the ABF method,[28] which relies upon the integration of the average force acting in the direction of z. In the NAMD implementation of ABF,[29] this force is determined within the classical thermodynamic integration formalism.[30] The free energy derivative, $dG(z)/dz$, is estimated locally as the simulation progresses, thereby providing a continuous update of the biasing force. When applied to the system, this bias generates a Hamiltonian exempt of a net average force along z. As a result, all values of the model reaction coordinate are sampled with an equal probability, thus improving markedly the accuracy of the computed free energies. Notably, this method makes no *a priori* assumptions about the final state of the transformation but rather allows the permeant to diffuse freely along the postulated model reaction coordinate. The interval spanned by ADP^{3-}, viz., roughly speaking from the bottom of the cavity to 8 Å above it, was divided into two non-overlapping windows, in which up to 60 ns of MD trajectory were generated, thus representing a simulation time of 0.12 μs for the overall reaction pathway.

Results and Discussion

Structural Features of *apo*-AAC. Analysis of the 0.1-μs MD trajectory is depicted in Figure 1. The number density profiles shed light on the organization of *apo*-AAC in the model membrane. They indicate that the carrier is nicely anchored in the lipid environment, its TM helices, H_i, spanning the width of the hydrophobic core of the bilayer while the short helical stretches, h_{ij}, appear to be essentially buried in the headgroup region. A glimpse at the opposite leaflet reveals that the mouth of *apo*-AAC emerges slightly below the surface of the model membrane. It is noteworthy that peripheral charged residues of the carrier interact strongly with both the choline and the phosphate moieties of the POPC units.

(25) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24–34.
(26) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(27) Feller, S. E.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2000**, *104*, 7510–7515.
(28) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169–9183.
(29) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904–2914.
(30) den Otter, W. K. *J. Chem. Phys.* **2000**, *112*, 7283–7292.

As the simulation of *apo*-AAC proceeds, water molecules from the bulk environment gush into the internal cavity of the membrane protein. Figure 1 illuminates how, within 0.1 μs of MD sampling, the hollow structure of the carrier is completely flooded. It is apparent from the number density profiles that the probability to find water molecules across the protein is nonzero and, hence, evocative of a possible communication between the matrix and the IMS of the mitochondria. Such may not be necessarily true, as the computation of number density profiles implies an averaging in the *x*- and *y*-directions of Cartesian space, normal to the aqueous interface. In reality, a rapid glance at the water isodensity maps of Figure 1 suggests that *apo*-AAC is hermetically close and impervious to the passage of water molecules.

Interestingly enough, flooding of *apo*-AAC is accompanied by the diffusion of chloride counterions toward the wall of the cavity, as highlighted in Figure 1 and in the Supporting Information. On the time scale explored, up to five counterions occupy concomitantly the cavity of the protein, with individual residence times as long as 60 ns. The seven positively charged residues gathered in the lower half of the carrier—viz., R79, R137, R234, R235, R279, K22, and K32—together with those borne by helices H2 and H4, about 10 Å above, in the upper half—viz., K91, K95, K198, and R187—form basic patches toward which the chloride counterions aggregate. Long-lived noncovalent bonds generally correspond to doubly chelated anions interacting with two arginine residues. As will be seen shortly, the asymmetric topological distribution of lysine and arginine amino acids in the cavity is envisioned to enhance the translocation of negatively charged species across the carrier.

Equally remarkable is the fully preserved secondary structure of *apo*-AAC over a period of 0.1 μs. As shown in Figure 1, all TM helices, H_i, and interfacial helical stretches, h_{ij}, remain unaltered on the time scale explored here. Of particular interest, the α-helical content of the carrier remains essentially constant, barring sporadic fraying in TM segment H2 over the first 20 ns, which can be legitimately ascribed to partial reorganization of the protein scaffold upon insertion in the POPC bilayer. Overall random coil content is limited, the small loops connecting the TM segments to the interfacial stretches on the matrix side undergoing rapid transition between coil, γ-turn, and 3_{10}- or α-helical conformations. The marginal distortion of the conformation witnessed on the 0.1-μs time scale is consistent with the moderate distance root-mean-square deviation (rmsd) of ca. 2 Å, measured over backbone atoms with respect to the reference PDB 1OKC structure. This value reflects internal motions of reduced amplitude within the TM scaffold resulting from the solvation of the carrier. On the basis of a 10-fold shorter simulation, Falconi et al. reported a somewhat larger distance rmsd of 3 Å,[31] which can be rationalized by the use of a distinct potential energy function.

While the secondary structure of *apo*-AAC is perfectly conserved in the course of the 0.1-μs MD simulation, the tertiary structure, however, appears to depart somewhat from that of 1OKC, which features the specific inhibitor CATR. A closer look at the reference crystallographic structure reveals a flared aperture toward both the IMS and the matrix, whereas the *apo* conformation of the carrier has a more pronounced cylindrical shape. In line with this observation, Falconi et al. had previously underlined that the volume of the internal cavity increased upon removal of the atractyloside. From a structural perspective, alteration of the tertiary structure of *apo*-AAC stems primarily from an inward motion of the C2 loop on the IMS side and of the N-terminus segment and, to a lesser extent, of the C1 loop and of the less flexible C-terminus segment. A decrease in the rigidity of the mitochondrial carrier resulting from the removal of the inhibitor CATR—which can be viewed as a structural wedge—represents a plausible explanation for the displacement of the C2 loop, albeit possible effects of the lipidic environment should not be ruled out.

Electrostatic Properties of *apo*-AAC. As has been underlined previously, the two basic patches that line the cavity of the mitochondrial carrier are likely to exert an attractive force onto anions present in the vicinity of the protein, e.g., chloride ions. To understand the relationship between the architecture of the AAC and the net electric forces at play, the electrostatic potential has been mapped numerically in the three dimensions of Cartesian space.[32] Figure 2 illuminates how this quantity varies dramatically between the matrix and the mitochondrial IMS. The conical isopotential surface revealed in this figure is suggestive of a privileged passageway anticipated to guide negatively charged permeants toward the bottom of the carrier, where TM helices H1, H3, and H5 interconnect through salt bridges. Notably, the longitudinal axis of this electrostatic funnel is not rectilinear and, hence, not collinear to the normal to the aqueous interface. Yet, the concept of a favored pathway is not contradictory with the impervious nature of the AAC in its closed conformation brought to light by the water iso-density maps. As illustrated in the contour maps of the electrostatic potential, halfway down the AAC, the concentric isopotential surfaces form a constriction region consistent with the topology of the membrane protein.

The isopotential contour map of Figure 2 indicates that any negatively charged species approaching the mouth of the carrier will be rapidly thrusted toward the vestibular space of the latter. In this region, anions can interact favorably with residues K91, K95, K198, and R187 borne by TM helices H2 and H4. To gain further insight into the net electric forces acting on the permeants of the AAC, the electric field has been derived from the three-dimensional maps of the electrostatic potential using the OpenDX open-source visualization package (http://www.opendx.org). In the region of interest, where the electrostatic funnel conducts, the intensity of the electric field ranges from ca. 0.075 to 0.125 V/Å. It is apparent that, below the constriction region, the lower basic patch of the cavity causes the direction of the electric field to be reverted. The diphosphate moiety of ADP^{3-}, initially pointing upward in the vestibular space, is, therefore, expected to reorient as the permeant sinks down in the crevice of the carrier.

This is demonstrated in the second 0.1-μs MD simulation of the AAC with 10 mM ADP and ATP. An ADP^{3-} fragment placed below the constriction region at $t = 0$ undergoes slow reorientation concomitant with its downward diffusion. Figure 2 emphasizes that the diphosphate nucleotide rotates about the central pentose ring, the space swept out by the tilting adenosine- and diphosphate groups being significant. Within ca. 40 ns, the former moiety lands at the bottom of the cavity, roughly parallel to the surface of the lipid bilayer, as the latter associates noncovalently with the positively charged residues of the second basic patch. As will be detailed subsequently, once locked in

(31) Falconi, M.; Chillemi, G.; Di Marino, D.; D'Annessa, B.; Morozzo della Rocca, I.; Palmieri, L.; Desideri, A. *Proteins: Struct. Funct. Bioinf.* **2006**, *65*, 681–691.

(32) Aksimentiev, A.; Schulten, K. *Biophys. J.* **2005**, *88*, 3745–3761.

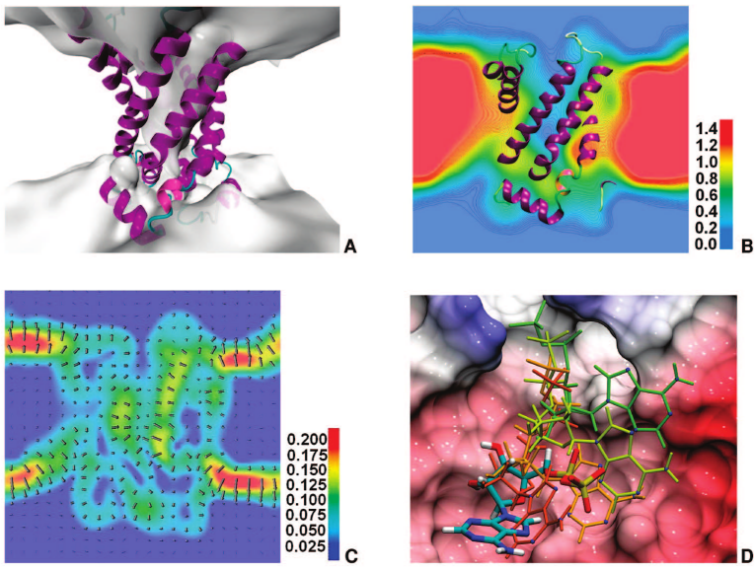


Figure 2. (A) Three-dimensional electrostatic isopotential map of *apo*-AAC revealing a funnel conducive to guide ADP^{3-} toward the bottom of the closed-state conformation. (B) Electrostatic potential contour maps.[32] Cross-sectional view along the longitudinal axis of the funnel. Values given in units of mV. The contour map was generated with OpenDX (http://www.opendx.org). (C) Cross-sectional view of the three-dimensional map of the electric field. The arrows denote the projection of the field in the plane of interest. Values are given in units of mV/Å. (D) Chronological superimposition of ADP^{3-} at $t = 0$, 5, 28, 35, and 70 ns (respectively green, yellow, orange, brown, and colored by element) taken from the 0.1-μs simulation of AAC with 10 mM ADP and ATP. The downward translocation of the permeant in the cavity is accompanied by its tilt, consistent with the local topology of the electric field. The surface of the cavity is colored according to the local value of the electrostatic potential following the coloring scheme of panel B. Orientation of the mitochondrial carrier is identical to that in panels A−C, viz., the longitudinal axis of the protein is collinear with the z-direction of Cartesian space.

place through an intricate network of salt bridges and hydrogen bonds, ADP^{3-} remains frozen in essentially the same orientation and conformation. Legitimately enough, this region of the carrier may be viewed as a binding pocket and the corresponding bound state of the permeant as an intermediary, possibly a recognition pattern, toward its release on the matrix side of the mitochondrial membrane.

Translocation of ADP^{3-} Modifies the Hydrogen-Bond Network of the AAC. One of the structural signatures of the AAC lies in the possibility to form three salt bridges to interconnect helices H1 to H3, H3 to H5, and H5 to H1.[11] Noncovalent interhelical linkage is achieved through the pairs of residues E29:R137, D134:R234, and D231:K32, which are located in the lower half of the carrier. In the *apo* form of the AAC, it is noteworthy that these salt bridges occur only intermittently (see Figure 3). The most glaring example is D231:K32, which, roughly speaking, satisfies the criteria of a salt bridge only in the first third of the 0.1-μs MD trajectory. To a large extent, this result is rooted in the original structure, i.e., PDB 1OKC, which was crystallized with specific inhibitor CATR. When setting up the simulation, the latter was removed from the internal cavity of the carrier. It would seem, however, that thermalization of the molecular assay was insufficient to erase the structural lag by allowing the protein scaffold to relax fully upon removal of the glycoside.

Conversely, the simulation of the AAC with 10 mM ADP and ATP reveals that the initially broken D134:R234 and D231:K32 salt bridges are restored within 35 ns. Yet, whereas the second salt bridge is long-lived, through the remainder of the 0.1-μs simulation, the first is partially disrupted after ca. 15 ns. This result is not completely surprising, given the more localized positive charge borne by the onium moiety of lysine, compared to the guanidinium functional group of arginine. Barring marginal fluctuations, the E29:R137 salt bridge appears to be essentially preserved throughout the MD trajectory.

The time scale characterizing the formation of the three interhelical salt bridges is congruent with that spanned by the simultaneous diffusion and reorientation of ADP^{3-} in the internal cavity of the carrier until locked in place. As the diphosphate nucleotide lands on the bottom of the crevice, it also binds noncovalently positively charged residues of the lower basic patch, viz., K22, R79, and R279. The latter amino acid residues, supplemented by Y186, have been proposed to form the "selectivity filter" of the carrier.[11] It is worth mentioning that site-directed mutagenesis experiments on residues K22 and R79 led to a severely impaired transport of the nucleotide across the mitochondrial membrane.[13]

As highlighted in Figure 3, the resulting salt bridges are established within 40 ns, remaining intact for the rest of the simulation. Additional, albeit weaker contacts are also created with N276 and S227 in the form of hydrogen bonds with the adenosine fragment, as well as with I183, primarily of dispersive nature. The ancillary, nonconventional hydrogen bond formed between oxygen atom O3′ of the pentose ring and the π-electron cloud of Y186 is also worth underlining, consistent with previous observations in model aqueous−π electron interactions.[33]

(33) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Science* **1992**, *257*, 942.

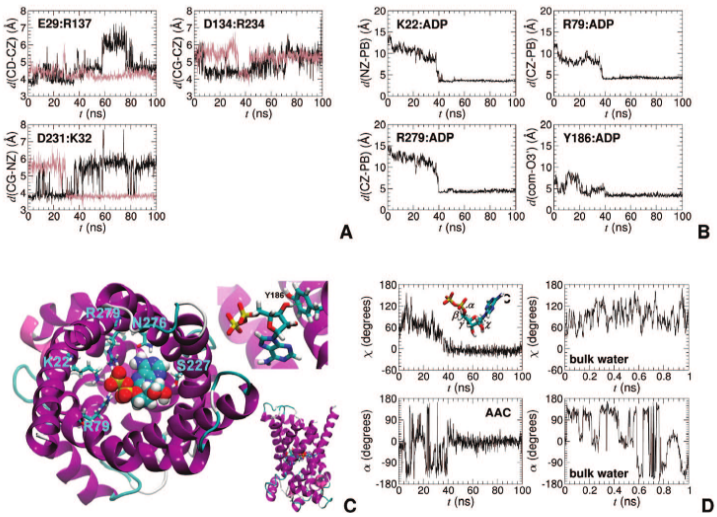


Figure 3. (A) Time evolution of the three salt bridges formed between TM helices H1, H3, and H5 in *apo*-AAC (dark lines) and in the assembly formed by AAC with 10 mM ADP and ATP (light lines). (B) Time evolution of the interaction distances between ADP^{3-} and key residues of AAC. In the case of Y186, "com" stands for the centroid of the aromatic ring. (C) Detail of the noncovalent bonds formed between ADP^{3-} and AAC. The upper inset highlights the pseudo-hydrogen bond established between O3 and the π-electron cloud of Y186. (D) Time evolution of representative torsional angles of ADP^{3-} locked in AAC (left) and in a bulk aqueous environment (right). α and χ stand for the C5 −O5 −P−O3 and O4 −C1 −N9−C8 dihedral angles, respectively. Over the 0.1-μs time scale, the adjacent β and γ torsions remain approximately constant around 180 and −60 , respectively.

The loss of rotational and translational entropy manifested in the binding of ADP^{3-} to the mitochondrial carrier is also accompanied by conformational changes in the permeant. Whereas the latter isomerizes rapidly in bulk water—which can be viewed as a naive model of the vestibular space—in particular through its χ and α torsional angles that vary roughly between 30 and 180 and between −180 and 180 , respectively, it is congealed in the same conformation once locked at the bottom of the crevice. Notably, in the AAC, the diphosphate nucleotide undergoes conformational alterations for approximately the first 40 ns of the complete 0.1-μs trajectory, after which both χ and α remain reasonably close to 0 . Whether in water or in the carrier, the two other dihedral angles, β and γ, plateau around 180 and −60 , respectively, with only marginal deviation from the latter. That the value of χ reached after 40 ns is not sampled in bulk water exemplifies the conformational restraint exerted by the environment on ADP^{3-}, reminiscent of the loss of internal degrees of freedom in protein−ligand association.

Binding of ADP^{3-} Is an Early Event of Its Transport across the Inner Membrane. Simulation of the AAC with 10 mM ADP and ATP suggests that the membrane protein acts as an electrostatic sensor, thrusting the permeant downward along a privileged pathway to reach the bottom of the crevice. Using as a naive reaction coordinate the distance separating the center of mass of ADP^{3-} from the centroid of the lower region of helices H1, H3, and H5 projected onto the *z*-direction of Cartesian space (see Methods section), the permeant was translocated over the entire 0.1-μs trajectory from 17.3 to 10.4 Å, as shown in Figure 4. If the region of the internal cavity to which ADP^{3-} is associated coincides, indeed, with a *true* binding site,[34] a key question, then, remains to be answered: Is this single experiment reproducible, and would the permeant reach the same binding site upon starting from distinct sets of initial positions and momenta? To tackle this issue, seven independent numerical experiments representing a total simulation time of 0.21 μs were carried out, wherein a single ADP^{3-} unit was dropped from different altitudes in the mitochondrial carrier, i.e., different values of the reaction coordinate, and with different orientations.

In the first simulation, the permeant virtually remains at the same altitude over a period of 17.5 ns, viz., approximately 29 Å. Within the first 3.6 ns, the phosphated moiety is anchored to K91, K95, N115, and R187, while the adenosine fragment interacts with Q84, N87, and F88. The subsequent reorientation of ADP^{3-} is accompanied by a large-amplitude motion of adenosine, now located above the diphosphate group and interacting with Y194. In the second trajectory, however, the permeant moves down from 33.8 to 27.6 Å within 77 ns, as highlighted in Figure 4. Over the entire simulation, the diphosphate moiety remains anchored to K91, K95, N115, and R187, while adenosine successively forms π-stacking motifs with Y194 and Y190 either individually or together. Throughout this trajectory, ADP^{3-} appears to be frozen in a highly strained conformation. Similarly coerced *anti* conformations of the permeants, where the phosphated moiety bends toward the purine ring, have been observed previously in related molecular systems.[35]

In the third experiment, translocation of ADP^{3-} proceeds from 33.5 to 21.0 Å over a period of 41 ns. In the first 5.5 ns, its adenosine moiety binds Y194, while the diphosphate fragment interacts with K91. In the remainder of the simulation, the phosphated group interacts sequentially with R187, Y186, and

(34) Robinson, A. J.; Kunji, E. R. S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2617–2622.

(35) Huber, T.; Klingenberg, M.; Beyer, K. *Biochemistry* **1999**, *38*, 762–769.

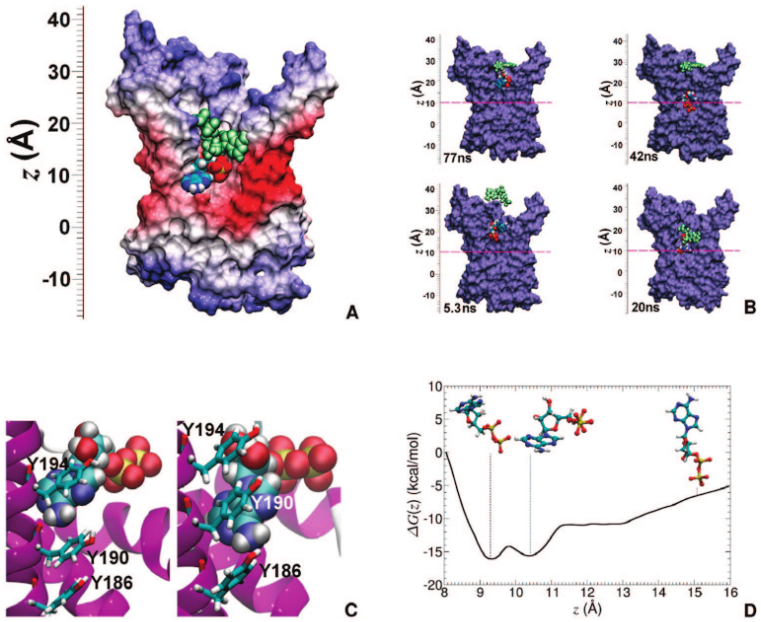


Figure 4. (A) Translocation of ADP^{3-} in the course of the 0.1-μs simulation of AAC in the presence of ADP and ATP. The initial position of ADP^{3-} is featured by pastel green van der Waals spheres. The surface of the carrier is colored according to the local value of the electrostatic potential following the coloring scheme of Figure 2B. (B) Comparison of the trajectories followed by ADP^{3-} in different simulations, starting with distinct initial positions and velocities of the permeant: (i) 77, (ii) 42, (iii) 5.3, and (iv) 20 ns. A dashed pink line characterizes the reference equilibrium position of ADP^{3-} in the binding pocket of AAC after 0.1 μs. (C) Tyrosine ladder formed by residues Y186, Y190, and Y194, along which ADP^{3-} glides, forming sequentially $\pi-\pi$ stacking interactions with the adenosine moiety. (D) Free energy profile delineating the translocation of ADP^{3-} in the cavity of AAC along the z direction of Cartesian space. The order parameter is the distance separating the center of mass of ADP^{3-} from that formed by the lower region of helices H1, H3, and H5, viz., residues 28−37, 133−142, and 230−239, projected onto the z axis. The insets characterize the conformation of ADP^{3-} for different values of the order parameter. ADP^{3-} is shown in the orientation of the carrier depicted in panel A.

Y190 as the nucleotide forms long-lived contacts with both N87 and K91. The most significant distance spanned by the permeant is observed in the fourth simulation depicted in Figure 4. Within 42 ns, ADP^{3-} moves down from 29.3 to 13.1 Å, gliding along Y194, Y190, and Y186.[11] The diphosphate moiety initially interacts with K22, to which residues R79, R234, and R235 are subsequently added. At the same time, the adenosine fragment, located above the phosphated group, binds R279.

The fifth experiment provides a cogent illustration of the role played by the electrostatic funnel to guide the diphosphate nucleotide toward the hypothesized binding site (see Figure 4). A rapid descent of ADP^{3-} from 39.3 to 22.8 Å is witnessed over 5.3 ns. Here again, the adenosine moiety appears to interact successively with the tyrosine ladder formed by residues Y186, Y190, and Y194. The permeant is essentially collinear to the longitudinal axis of the carrier, its diphosphate group pointing downward, ultimately interacting with R187 and K91. In sharp contrast, the sixth simulation reveals that the adenosine moiety is below the phosphated group. Over 3.6 ns, the permeant is translocated from 39.2 to 28.8 Å. The nucleotide interacts preferentially with K91, K95, and N87.

For the seventh and last numerical experiment, the starting position of ADP^{3-} is roughly similar to that chosen in the aforementioned reference, 0.1-μs simulation, the initial set of velocities being different. As shown in Figure 4, within 20 ns, the permeant moves down from 17.3 to 11.5 Å, the adenosine moiety tilting in a sweeping motion comparable to that observed in the longer, 0.1-μs trajectory (see Figure 2), eventually interacting with R79, R234, R235, and D134, while the diphosphate group binds K91.

Put together, the seven independent MD simulations described here illustrate the pronounced sensitivity of the investigated phenomenon to the chosen initial conditions, in particular, over the admittedly limited time scales being explored. They also underscore the complementary roles of electrostatic probe and electrostatic sensor, played respectively by the permeant and the mitochondrial carrier, accounting for the rapid attraction of ADP^{3-} toward the bottom of the crevice. Inappropriate initial orientation and conformation of the permeant can, however, lead to an equally rapidly hampered diffusion across the AAC, ADP^{3-} being generally trapped in the constriction region, where it is congealed in an inert conformation and forms long-lived interactions with its environment. To address the issue of quasi-nonergodicity exemplified in these numerical experiments and, most importantly, reconcile the latter with the free energy landscape that characterizes the transport of the permeant in the mitochondrial carrier, additional simulations were performed using the ABF method.[28,29]

The free energy profile of Figure 4 delineates the translocation of ADP^{3-} in the carrier, roughly speaking from the constriction region to the bottom of the internal cavity, hence covering a range of 8 Å of the order parameter. However naive this order parameter may be—the *true* reaction coordinate being in all probability a multidimensional one—it nonetheless provides a

compelling evidence that the binding site brought to light in the 0.1-μs MD trajectory corresponds to a minimum of the free energy. As has been discussed above, this minimum emerges at ca. 10.4 Å and is characterized by the adenosine moiety of ADP^{3-} lying flat on the bottom of the crevice, while the diphosphate group forms an intricate network of salt bridges with residues K22, R79, and R279.

An interesting feature revealed by the present free energy calculations is the existence of a second minimum, found slightly deeper in the carrier, at ca. 9.3 Å. What distinguishes the two minima of comparable depth and separated by a barrier of less than 2 kcal/mol lies essentially in the orientation of ADP^{3-}. Whereas over the 0.1-μs time scale of the MD trajectory, it remained locked in the same position after 40 ns, ABF allows, through reversible diffusion, its reorientation and isomerization, predominantly about the χ and α dihedral angles. The permeant, initially located at ca. 10.4 Å, rotates about its pentose ring, so that the adenosine fragment, initially beneath it, is now above, pointing toward the vestibule, while the diphosphate group penetrates deeper down in the cavity. This new conformation is conducive to an enhanced induced fit, mirrored in a somewhat lower ADP:AAC interaction energy (see Supporting Information). Contrasting with the reference 0.1-μs MD simulation, the free energy minimum found at 9.3 Å corresponds to a disrupted D231:K32 salt bridge.

A number of previously published data appear to corroborate this result. In particular, residues pertaining to the MCF motif of the yeast carrier have been systematically mutated, and the function of the mutants investigated. Among the mutations of the six residues corresponding to the bovine residues, viz., E29,[36,37] K32,[37] D134,[12,37] R137,[37] D231,[37] and R234,[36,38,39] two main observations were made. First, the residues that have proven to be the most sensitive to inhibitor ATR, CATR, and BA binding and to nucleotide exchange are K32, D134, and D231; i.e., mutations of equivalent residues in yeast lead to almost no binding for ATR or CATR and BA and almost no nucleotide exchange. Second, when mutations are carried out on yeast residues equivalent to K32, D231, and R234, no second-site revertants are detected. The latter suggests that the function cannot be recovered easily and that these residues are likely to play an important role in the mitochondrial function. Put together, these results underscore the importance of the two residues forming salt bridge D231:K32.

In contrast with the latter salt bridge, E29:R137 is only partially preserved through the sole noncovalent interaction of the OE1 oxygen atom with the hydrogen atoms borne by NH2. Compared with the 10.4-Å minimum, the lower position of ADP^{3-} in the carrier promotes a strong interaction of the phosphated moiety with R137.

The region of the mitochondrial carrier that encompasses the two minima is prefaced in both directions by large free energy barriers. Not too surprisingly, the steeper barrier corresponds to the bottom of the crevice, which has been shown to be impervious to the passage of small chemical species, like water molecules (see Figure 1). It may be anticipated that brute-force translocation of ADP^{3-} below 9.3 Å would require a considerable positive work imposed by the structural rearrangement of the TM helices, which evidently is not amenable to classical MD simulations. The barrier arising above 10.4 Å reflects the energetic cost incurred for disrupting the noncovalent bonds formed between the permeant and its binding site. The free energy plateau that follows within the next ca. 2 Å is suggestive of an unhampered translocation and isotropic orientation of the permeant. Beyond 13 Å, the free energy increases again as the diphosphate nucleotide crosses the constriction region of the carrier. It is noteworthy that, in these ABF calculations, the phosphated moiety is systematically predicted to point downward, i.e., in the direction of the lower basic patch, which is congruent with the conclusions drawn from the above independent numerical experiments.

Conclusions

The early events of ADP transport across the mitochondrial AAC[11,13] have been examined based on large-scale MD simulations. Investigation of *apo*-AAC confirms the impervious nature of the carrier, which prevents the exchange of chemical species between the matrix and the IMS of the mitochondria. Simulation in a model membrane over the 0.1-μs time scale underscores the robustness of the α-helical TM scaffold. Three-dimensional mapping of the electrostatic potential brings to light a privileged, funnel-like passageway envisioned to drive rapidly permeants toward the bottom of the membrane protein. Mirroring the distinctive distribution of charged amino acids in the carrier, the revolving directionality of the electric field is suggestive of a possible reorientation of the permeant as the latter descends in the internal cavity.

0.1-μs simulation of AAC with 10 mM ADP and ATP reveals a possible recognition site for ADP^{3-} at the bottom of this cavity. The loss of orientational and conformational entropy of the permeant as it binds the crevice is reminiscent of protein−ligand association. Additional independent numerical experiments, wherein ADP^{3-} is dropped at different altitudes in the mitochondrial carrier, illustrate over the time scales explored the marked sensitivity of the process to the chosen starting positions and momenta. The rationale for probing different starting points stems from our limited knowledge of how the nucleotide approaches the AAC, in particular what is the preferred orientation and conformation of the substrate at a given separation from the mouth of the transporter. Assuming appropriate orientation, the diphosphate nucleotide is thrusted downward, reaching rapidly the hypothesized binding region of the carrier. Conversely, under inadequate initial conditions, ADP^{3-} may remain frozen in a kinetic trap, systematically near the constriction region of the AAC. While the present results do not necessarily imply that the path followed by the permeant is unique, it may be legitimately inferred that, if the transport phenomenon is witnessed on an appreciably short time scale, e.g., 42 ns to cover approximately 16 Å, the corresponding trajectory is thermodynamically relevant. It may further be contended that, given the limited time scale over which it is monitored, translocation of ADP^{3-} is a highly probable event. That the substrate does not reach the postulated binding site, down in the crevice, simply means that sampling is incomplete.

The proposed binding site highlighted in the present classical MD simulations, and identified from structural analyses based on sequence similarities,[34] is shown to correspond to a true minimum of the free energy landscape determined along the longitudinal axis of the membrane protein, employing an

(36) Klingenberg, M.; Nelson, D. R. *Biochim. Biophys. Acta* **1994**, *1187*, 241–244.
(37) Müller, V.; Heidkämper, D.; Nelson, D. R.; Klingenberg, M. *Biochemistry* **1997**, *36*, 16008–16018.
(38) Nelson, D. R.; Lawson, J. E.; Klingenberg, M.; Douglas, M. G. *J. Mol. Biol.* **1993**, *230*, 1159–1170.
(39) Heidkämper, D.; Müller, V.; Nelson, D. R.; Klingenberg, M. *Biochemistry* **1996**, *35*, 16144–16152.

adaptive biasing scheme. The resulting free energy profile features two neighboring minima of comparable depth, which characterize distinct binding modes and orientations of ADP^{3-}, thereby lending support to the hypothesis of a true association site put forth on the basis of the crystallographic structure.[11] The 10 kcal/mol free energy difference between these minima and the constriction region rationalizes the fast downward translocation of the diphosphate nucleotide. In the opposite direction, the steep free energy barrier encountered by ADP^{3-} as it moves farther down the mitochondrial carrier is congruent with the accepted view of an impermeable carrier impeding further progression of the permeant toward the mitochondrial matrix.

The set of simulations reported here emphasizes that the rapid binding of ADP^{3-} to the internal cavity of the AAC constitutes an early event of its transport across the inner membrane, known to span the millisecond time scale.[40] Recognition of the diphosphate nucleotide and association thereof with the mitochondrial carrier is hypothesized to trigger the opening of the latter, an event anticipated to span significantly longer time scales. Opening could occur through the triplet of proline residues P27, P132, and P229, which act as hinges in helices H1, H3, and H5, respectively. Paradoxically, binding of ADP^{3-} to the AAC is accompanied by the formation of salt bridges between these three TM segments, a process envisioned to reinforce the α-helical scaffold rather than disrupt it into a conformation opened toward the matrix.

This apparent paradox is, however, removed when considering the free energy minimum pinpointed at 9.3 Å, employing ABF calculations. Whereas binding of ADP^{3-} to the crevice initially strengthens the bundle of TM helices through the creation of interhelical salt bridges, subsequent reorientation and isomerization of the permeant as it moves deeper down in the carrier ruptures the observed network of noncovalent bonds. This disruption could be viewed as a preamble to the opening of the membrane protein toward the matrix of the mitochondria, albeit an atomic-level description of the complete transport phenomenon, which is likely to embrace both environment-dependent processes and events of stochastic nature, currently remains out of reach. It is, nevertheless, reasonable to believe that, as novel experimental advances are published, theoretical approaches will be used profitably to gain additional insight into the mechanisms that underlie the assisted passage of nucleotides across the mitochondrial membrane.

Such experimental advances imply the design of new blocking agents targeted at trapping intermediate structures that represent potential milestones of the transport process. De novo design of analogues solely in the light of a single crystallographic structure has, however, proven to be a difficult undertaking. In the present work and for the first time, the interaction of the substrate with the mitochondrial carrier is revealed at the atomic level over the initial stage of the translocation pathway. Contacts formed between key residues of the AAC and both the adenosine and the phosphate moieties of ADP^{3-} provide valuable guidance for the rational modification of the substrate or the design of site-directed mutagenesis experiments on the membrane protein.

Acknowledgment. The authors are indebted to Emad Tajkhorshid for insightful discussions. The Centre Informatique National de l'Enseignement Supérieur (CINES), Montpellier, France, and the Centre de Calcul, Recherche et Technologie (CCRT), Arpajon, France, are gratefully acknowledged for generous provision of computational time. This work is funded by the Agence nationale pour la recherche (ANR).

Supporting Information Available: Figures showing the structure of the ADP/ATP mitochondrial carrier, the mass-weighted isodentisy map and residence time of the chloride counterions in *apo*-AAC, interaction energy of ADP^{3-} with the mitochondrial carrier from classical MD and ABF calculations, complete ref 26, and an AVI movie generated from the 42-ns numerical experiment, during which ADP^{3-} was dropped near the mouth of the mitochondrial carrier. This material is available free of charge via the Internet at http://pubs.acs.org.

JA8033087

(40) Gropp, T.; Brustovetsky, N.; Klingenberg, M.; Müller, V.; Fendler, K.; Bamberg, E. *Biophys. J.* **1999**, *77*, 714–726.
(41) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

Molecular Membrane Biology, 2013; 30(2): 160–168

informa
healthcare

The substrate specificity of the human ADP/ATP carrier AAC1

JOHN MIFSUD[1*], STÉPHANIE RAVAUD[2,3,4*], EVA-MARIA KRAMMER[5,6*], CHRIS CHIPOT[5,6], EDMUND R. S. KUNJI[1], EVA PEBAY-PEYROULA[2,3,4] & FRANCOIS DEHEZ[5,6]

[1]*The Medical Research Council, Mitochondrial Biology Unit, Cambridge, UK,* [2]*Université Grenoble 1, IBS, Institut de Biologie Structurale, Grenoble cedex 1, France,* [3]*CEA, IBS, Grenoble,* [4]*CNRS, IBS, Grenoble,* [5]*Université de Lorraine, Vandoeuvre-lès-Nancy Cedex, and* [6]*CNRS, SRSMC, Nancy, France*

(Received 30 August 2012; and in revised form 17 October 2012)

Abstract

The mitochondrial ADP/ATP carrier imports ADP from the cytosol into the mitochondrial matrix for its conversion to ATP by ATP synthase and exports ATP out of the mitochondrion to replenish the eukaryotic cell with chemical energy. Here the substrate specificity of the human mitochondrial ADP/ATP carrier AAC1 was determined by two different approaches. In the first the protein was functionally expressed in *Escherichia coli* membranes as a fusion protein with maltose binding protein and the effect of excess of unlabeled compounds on the uptake of [^{32}P]-ATP was measured. In the second approach the protein was expressed in the cytoplasmic membrane of *Lactococcus lactis*. The uptake of [^{14}C]-ADP in whole cells was measured in the presence of excess of unlabeled compounds and in fused membrane vesicles loaded with unlabeled compounds to demonstrate their transport. A large number of nucleotides were tested, but only ADP and ATP are suitable substrates for human AAC1, demonstrating a very narrow specificity. Next we tried to understand the molecular basis of this specificity by carrying out molecular-dynamics simulations with selected nucleotides, which were placed at the entrance of the central cavity. The binding of the phosphate groups of guanine and adenine nucleotides is similar, yet there is a low probability for the base moiety to be bound, likely to be rooted in the greater polarity of guanine compared to adenine. AMP is unlikely to engage fully with all contact points of the substrate binding site, suggesting that it cannot trigger translocation.

Keywords: *Mitochondrial carrier, nucleotide transport activity, molecular-dynamics simulations*

Introduction

Mitochondrial carriers form a family of transporters that are embedded in the inner mitochondrial membrane. They are responsible for the traffic of a variety of metabolites and small molecules between the cytoplasm and the mitochondrial matrix (Kunji 2004, Palmieri 2004). The substrate specificity of mitochondrial carriers is quite narrow, as only chemically and structurally related compounds are accepted as substrates for import or export. The mitochondrial ADP/ATP carrier (AAC) transports newly synthesized ATP out of the mitochondrion and in return imports ADP. Due to its abundance in heart mitochondria, it is the best-characterized mitochondrial carrier (Klingenberg 2008) and the only one for which a high-resolution structure is available (Pebay-Peyroula et al. 2003). Compared to other transporters, mitochondrial carriers are relatively small, comprising only six transmembrane helices connected through very short loops, delineating a central pathway through which the passage of the substrates occurs. The known AAC1 structure resembles the conformation that is open toward the intermembrane space able to bind ADP (Nury et al. 2006).

The directionality of adenine nucleotide transport is imposed by the concentration gradients of the substrates and by the electrostatic potential of the inner mitochondrial membrane as the ADP/ATP exchange is electrogenic. *In vitro*, it was shown that either ADP or ATP can be transported in both ways (Brustovetsky et al. 1997, 1996). In humans there are

Correspondence: Dr F. Dehez, Université de Lorraine, BP239, 54506 Vandoeuvre-lès-Nancy Cedex, France. Tel: +33 (0)3 8368 4098. E-mail: Francois.Dehez@univ-lorraine.fr; Dr E. Kunji, The Medical Research Council, Mitochondrial Biology Unit, Hills Road, CB2 0XY, Cambridge, UK. Tel: +44 (0) 1223 252850. E-mail: edmund.kunji@mrc-mbu.cam.ac.uk; Dr E. Pebay-Peyroula, Institut de Biologie Structurale, 41, rue Jules Horowitz, 38027 Grenoble cedex 1, France. Tel: +33 (0) 4 3878 3482. E-mail: eva.pebay-peyroula@ibs.fr
*John Mifsud, Stéphanie Ravaud and Eva-Maria Krammer contributed equally to the paper.

ISSN 0968-7688 print/ISSN 1464-5203 online © 2013 Informa UK, Ltd.
DOI: 10.3109/09687688.2012.745175

four different isoforms, which have different expression patterns (Dolce et al. 2005). The substrate specificity of the rat mitochondrial ADP/ATP carrier is largely limited to ADP and ATP (Pfaff and Klingenberg 1968). Surprisingly, the carrier does not transport the related pyrimidine nucleotides GDP and GTP. An unresolved issue is whether AMP can be transported or whether it can interfere with ADP and ATP binding. Residues potentially involved in substrate binding were identified by site-directed mutagenesis and photolabeling (Dalbon et al. 1988, Heidkamper et al. 1996, Muller et al. 1997, Dianoux et al. 2000), but the extent of the substrate binding site was defined only after the structure became available by sequence analyses (Robinson and Kunji 2006, Robinson et al. 2008) and molecular-dynamics simulations (Dehez et al. 2008, Wang and Tajkhorshid 2008). The major contact points of the binding site that interact with the nucleotides consist of positively charged residues that bind the phosphate moieties (contact point I and III) and a hydrophobic pocket that binds the adenine moiety (contact point II) (Figure 1) (Robinson and Kunji 2006, Dehez et al. 2008, Robinson et al. 2008, Wang and Tajkhorshid 2008). The definition of the substrate binding site is not sufficient to explain the narrow substrate selectivity, but it is a crucial step to a deeper understanding of the transport mechanism at the molecular level.

Although the transport selectivity has been explored extensively in mitochondrial membranes (i.e., Duee and Vignais 1968, Pfaff and Klingenberg 1968), the substrate specificity of AAC has not been studied in isolated systems nor has it been defined at the atomic level. Here we have used two different expression systems to analyze the transport of nucleotides by the major human isoform AAC1 (hAAC1). In parallel, the entrance and binding of these nucleotides into the cavity of AAC1 were followed by molecular-dynamics simulations. The transport assays clearly confirm the selectivity for adenine nucleotides comprising 2 or 3 phosphates. In the light of molecular-dynamics simulations performed herein, the molecular basis that determines the selectivity of the carrier in its conformation opened toward the intermembrane space (IMS) can be established. Interestingly, the numerical experiments described in this contribution suggest that both the phosphate and the base moiety ought to be recognized sequentially to ensure binding. Even though the phosphate binding is similar for GDP (GTP) and ADP (ATP), the base recognition is probably less efficient for guanine compared to adenine nucleotides. Our findings confirm experimentally the molecular basis of nucleotide specificity of the ADP/ATP carrier and shed light on the mechanism of selective binding.

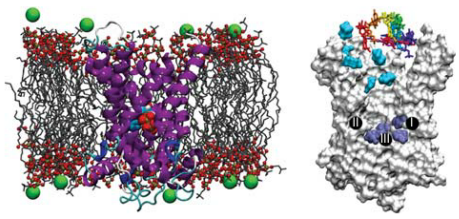


Figure 1. Molecular-dynamics simulations assays. (Left) Typical configuration of the bovine-heart AAC embedded in a fully hydrated POPC bilayer (the chloride counter-ions are depicted as green spheres and water molecules are not represented for clarity) with a nucleotide bound at the bottom of the internal cavity of the carrier. (Right) Starting positions of the nucleotides for the 10 binding assays. Cyan and dark-blue surfaces correspond respectively to nucleotides found in the vicinity of the upper and the lower patches of basic residues. The three major contact points of the binding site that interact with the nucleotides are indicated (I: K22, R79; II: G182, I183; III: R279) (Robinson and Kunji 2006).

Material and methods

Uptake of radioactively labeled ATP by intact Escherichia coli cells expressing hAAC1

The hAAC1 gene was cloned as a fusion to the periplamsic MBP (MBP-hAAC1) as described previously (Ravaud et al. 2012). The recombinant protein was expressed in *Escherichia coli* strain C43 (DE3) (Miroux and Walker 1996). Cells were grown at 37°C in LB (lysogeny broth) medium containing ampicillin (100 mg.l^{-1}) to an A600 of 0.6. Protein expression was induced using 0.1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and took place overnight at 20°C. Cells were harvested by centrifugation for 15 min at 3,000 *g* and 4°C and washed once in 50 mM potassium phosphate buffer pH 7. Cells were sedimented again (3,000 *g*, 15 min, 4°C) and the final pellet was resuspended in 50 mM potassium phosphate buffer pH 7 to a cell density of 0.1 g.ml^{-1} and kept on ice.

The uptake assays were initiated by the addition of 10 µM [α-^{32}P]-labeled ATP (3,000 mCi/mmol; Perkin Elmer) in 30 µl of the IPTG-induced *E. coli* cells. The incubation was performed at 25°C and was terminated after 30 min by the addition of 1 ml ice-cold potassium phosphate buffer followed by rapid filtration through a 0.45 µm filter (Millipore, France) under vacuum. The filters were washed three times with 1 ml ice-cold potassium phosphate buffer and the radioactivity retained on the filters was quantified in 3.5 ml of water in a Multi-Purpose

Scintillation Counter (Beckman Coulter, Fullerton, CA, USA). Competition experiments were carried out with ATP, ADP, AMP-PNP, GDP, and GTP at a final concentration of 100 μM. The nucleotides were added 5 min before the uptake experiment was started and after addition of [α-^{32}P]-labeled ATP, the transport rates were determined at 30 min.

In all experiments, C43 (DE3) *E. coli* cells harboring the corresponding empty vector (pET-20b vector containing only the periplasmic MBP gene) were used as a control.

Uptake of radioactively labeled ADP by intact Lactococcus lactis cells *expressing hAAC1*

Lactococcus lactis strain NZ9000 was transformed with vector PNZ8048 containing the gene for the expression of hAAC1, essentially as described (Kunji et al. 2003). Pre-cultures of *L. lactis* were set up overnight in M17 media to grow at 30°C with no aeration to an optical density of ~ 2.5 (A_{600nm}). The growth in M17 media was initiated at an A_{600nm} of ~ 0.1 from the pre-culture and grown at 30°C until an A_{600nm} of 0.5–0.7 was reached. Induction of protein expression was performed by addition of filtered M17 medium containing nisin A excreted by strain NZ9700 at a dilution of 1:10000, after which growth was continued for a further 3 h.

Cells were collected at 6,000 *g* for 10 min at 4°C and washed once in PIPES buffer (10 mM piperazine-1,4-bis-2-ethanesulfonic acid, 50 mM NaCl, pH 7.0). The final pellet was resuspended to A_{600nm} of 20 and stored on ice. After 30 min pre-incubation at room temperature the transport assays performed at 30°C were started by adding an equal volume of [^{14}C]-ADP radiolabel (2.22 GBq mmol^{-1}, Perkin Elmer) to the cell suspension. When inhibitor studies were performed, the inhibitor was added to the mixture 5 min before the uptake experiment commenced. Samples of 15% of the total volume were taken at appropriate time-points, quenched with ice-cold PIPES buffer and filtered on a vacuum manifold through cellulose nitrate membranes (0.45 μm pore, Whatman). The filters were washed with 3 ml of cold buffer and dissolved in 3 ml of liquid scintillant (Ultima Gold AB, Perkin Elmer), which were counted in a TriCarb liquid scintillant counter (Perkin Elmer).

Uptake of radioactively labeled ADP by fused membrane vesicles of Lactococcus lactis expressing hAAC1

Cells expressing hAAC1 were harvested at 6000 *g* for 10 min at 4°C, washed in PIPES buffer and resuspended in 1/40 of the original volume. Cell lysis was performed by mechanical disruption with one pass at 33,000 pa in a cell disruptor (Constant Cell Disruption Systems, Daventry, UK). Whole cells were removed by a low speed spin at 10,800 *g* for 10 min at 4°C. Membranes were harvested by two ultracentrifugation steps at 138,000 *g* for 40 min at 4°C, washing with PIPES buffer in between. Pellets were resuspended in PIPES buffer to a total protein concentration of 5 mg.ml^{-1} and stored in liquid nitrogen.

E. coli polar lipid extract and egg yolk phosphatidylcholine (both at 20 mg.ml^{-1} in chloroform, Avanti Polar lipids) were mixed in a 3:1 weight ratio and dried by agitation under nitrogen. The lipids were resuspended in a homogenizer in cold PIPES buffer and frozen in liquid nitrogen. Isolated lactococcal membranes were fused with liposomes by freeze-thawing (Driessen et al. 1985). First, membranes (1 mg total protein concentration) were mixed with liposomes (5 mg) in PIPES buffer and a substrate for loading of the fused membranes (5 mM final concentration from 10× stock in PIPES buffer, pH 7). The membrane mixture was frozen in liquid nitrogen and thawed to room temperature seven times before storage in liquid nitrogen to create multilamellar fused membranes. The effect of bovine heart cardiolipin (Avanti Polar Lipids) on transport was investigated by adding different amounts of cardiolipin to the mixed vesicles and liposomes prior to freeze-thawing. Prior to performing transport assays, fused membrane vesicles were extruded 11 times through a polycarbonate nucleopore track etch membrane (1 μm, Whatman) to form single lamellar vesicles of defined size. The extruded membranes were concentrated five-fold by ultracentrifugation at 300,000 *g* for 30 min at 4°C and resuspended in buffer containing 5 mM substrate. External substrate was removed by passing the fused membranes through a Sephadex G-75 gel filtration column (3.5 ml bed volume, GE Healthcare), eluting in a 1 ml volume. After 30 min pre-incubation at room temperature transport assays were started by diluting the loaded vesicles four-fold into buffer containing [^{14}C]-ADP (2.22 GBq mmol^{-1}, Perkin Elmer) at the stated final concentrations. Uptakes, which were performed at 30°C, were stopped by quenching with ice-cold PIPES buffer and filtration on a vacuum manifold through cellulose nitrate membranes (0.45 μm pore, Whatman). The filters were washed with 3 ml of cold buffer and then dissolved in 3 ml of scintillation liquid (Ultima Gold AB, Perkin Elmer). The amount of accumulated radio-label was determined in a TriCarb liquid scintillation counter (Perkin Elmer).

Molecular-dynamics simulations

The spontaneous association of six nucleotide analogues (AMP, ADP, ATP, GMP, GDP and GTP) was investigated by means of molecular-dynamics (MD) simulations employing a methodology described previously (Dehez et al. 2008). The initial molecular assay consisted of the bovine heart AAC in its conformation open toward the cytosol as revealed by X-ray crystallography (PDBID:1OKC) (Pebay-Peyroula et al. 2003). After removal of the inhibitor CATR bound to the crystal structure, AAC was inserted in a thermalized, fully hydrated palmitoyloleylphosphatidyl-choline (POPC) bilayer consisting of 228 lipid unit and approximately 10,000 water molecules (Figure 1). All molecular assemblies were built and analyzed with VMD (Humphrey et al. 1996). The net excess charge of AAC was counterbalanced by means of 17 chloride counter ions. This corresponds to a chloride concentration of about 0.1 M. The complete molecular assays involving the carrier in its hydrated membrane environment together with one of the aforementioned nucleotide were described by the all-atom CHARMM27 (Mackerell et al. 1998, Feller and Mackerell 2000) force field for the protein, the nucleotides, water and ions. A modified united-atom version of the force field was utilized for the lipids (Henin et al. 2008). CMAP corrections (Mackerell et al. 2004) were introduced for the protein. All MD simulations were carried out using the NAMD (Phillips et al. 2005) package, in the isobaric-isothermal ensemble. The temperature and the pressure were kept constant at 300 K and 1 atm using, respectively, Langevin dynamics and the Langevin piston method. The particle mesh Ewald (PME) algorithm was employed to account for long-range electrostatic interactions (Darden et al. 1993). The equations of motion were integrated by means of the resp-A multiple-time step algorithm (Tuckerman et al. 1992) with a time step of 2 and 4 fs for short- and long-range interactions, respectively. Covalent bonds involving hydrogen atoms were constrained to their equilibrium length by means of the Rattle algorithm (Andersen 1983). The molecular assemblies were first equilibrated at 300 K under 1 atm for approximately 100 ns, thereby providing the initial configuration for the subsequent binding assays. Each nucleotide was successively inserted at the mouth of the carrier with 10 randomly chosen initial orientations (Figure 1). Following a short equilibration of the water molecules around the substrate, a series of unbiased MD trajectories up to 100-ns long was generated. Binding of the substrates was characterized by considering how long-lived – specifically over time scales exceeding 15 ns, and steady is the anchoring of the phosphate and the nucleoside moieties at the bottom of the cavity.

Results and discussion

Substrate specificity of the human ADP/ATP carrier

The substrate specificity of the human ADP/ATP carrier hAAC1 was investigated by expressing the carrier in two different bacterial hosts. In the first approach hAAC1 is expressed in *E. coli* membranes in a functional form by fusing the periplasmic maltose binding protein to the N-terminus of the carrier. It has been shown previously that its main transport properties are retained in this heterologous system (Ravaud et al. 2012). [α-^{32}P]ATP uptake experiments revealed an apparent K_m for ATP of 23.7 ± 5 μM and a V_{max} of 14.6 ± 0.6 nmol/min/mg of protein consistent with previously published data (De Marcos Lousa et al. 2002). The measured transport activity could also be completely inhibited by the addition of the specific inhibitors carboxy-atractyloside (CATR) and bongkrekic acid (BA). To study the substrate specificity, transport of [α-^{32}P]ATP was measured in the presence and absence of 100 μM of competing nucleotides to test whether these compounds can compete with the radio-labeled substrate for binding or can inhibit the transport activity. ATP import was reduced by around 80% by competition with excess ATP or ADP and to a much lower extent by competition with other nucleotides (Figure 2). ADP exhibited the strongest inhibition on ATP import. In the presence of AMP, the uptake is reduced by about 40%, while in the presence of GTP, GDP and GMP the uptake level remains essentially unaffected.

In the second approach hAAC1 was expressed in the cytoplasmic membrane of *Lactococcus lactis* (Figure 3A). Uptake into whole cells of *L. lactis* expressing AAC1 was observed by the exchange of intracellular adenine nucleotides for radio-labeled ADP, but not in the control strain (Figure 3B). Complete inhibition of uptake was achieved by the addition of 10 μM CATR or 2.5 μM BA in agreement with published data on human AAC isoforms 1, 2 and 3 expressed in yeast mitochondria (De Marcos Lousa et al. 2002), AAC4 expressed in *E. coli* and refolded (Dolce et al. 2005) and AAC of rat liver mitochondria (Pfaff and Klingenberg 1968). To see if the expression levels and uptake rates could be improved, the N-terminus was truncated, as has been done for other members of the mitochondrial carrier family (Monne et al. 2005). The N-terminus of the hAAC1 was truncated (△1–3 human AAC1) to a length identical to the optimal truncation for the yeast Aac2p △2–19, but no significant improvement in the

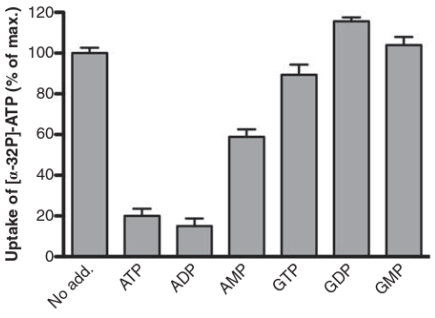


Figure 2. Effect of the nature of the nucleotide on [α-^{32}P]ATP uptake into *E. coli* cells expressing hAAC1. IPTG-induced *E. coli* cells harboring plasmid MBP-hAAC1 were incubated in potassium phosphate buffer containing 10 μM labeled ATP. The uptake assay was performed in the absence or in the presence of 100 μM of non-labeled nucleotides and carried out for 30 min. The uptake was terminated by the addition of cold buffer and the samples were filtered. The rate of transport in the absence of additives (No add.) was set to 100%. The control nucleotide uptake (empty vector) was subtracted. Data are the mean of three independent experiments. The average values with standard deviation are shown.

uptake rate was observed (data not shown). Another factor that was investigated was the requirement of cardiolipin, which is known to improve the activity of bovine AAC1 reconstituted in proteoliposomes (Kramer and Klingenberg 1980). Lactococcal membranes are suitable for the expression of mitochondrial carriers, as they naturally contain a large amount of cardiolipin (32% of total lipid), but the length and saturation of the alkyl chains may differ from mitochondrial cardiolipin. Different additions of bovine cardiolipin to the membrane fusions were tried, and the optimal one corresponded approximately to 6% of total lipid and led to a 2-fold increase in transport rate (data not shown).

The substrate specificity of hAAC1 was determined by two routes. The first strategy was similar to the one used with *E. coli* cells. Radio-labeled ADP uptake in whole cells was monitored in the presence of a 1000-fold excess of a non-labeled compound (Figure 4A). High competition for substrate binding or transport inhibition was observed with ADP and ATP, as observed for hAAC1 expressed in *E. coli* cells. Furthermore, an 85% and 80% reduction in the transport rate was observed with an excess of AMP or ADP-glucose, respectively. NADH reduced the transport rate by 60%, unlike NAD^+, which exhibited no effect. Furthermore, the purine nucleotides, (deoxy-thymidine triphosphate, and deoxy-uridine triphosphate, reduced the rate to 45%, but the corresponding di- and monophosphate nucleotides did not. The competition with the oxy-form of uridine triphosphate only led to a reduction of 30%, similar to the guanosine tri- and di-phosphate and FAD, although the latter effect may be due to interference with liquid scintillation counting because of the yellow colour. The purine cytidine-triphosphate, but not the diphosphates or the deoxy form, reduced the uptake by 40%.

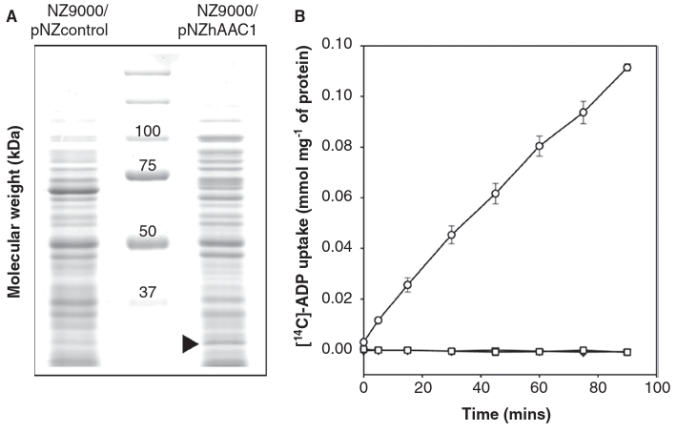


Figure 3. Expression and transport of ADP by the human AAC1 expressed in *Lactococcus lactis*. (A) (A) Coomassie brilliant blue-stained SDS-PAGE gel of the cytoplasmic membrane of the control and hAAC1-expressing strain. The arrowhead indicates the expressed hAAC1. (B) Radio-labeled ADP uptake by whole *Lactococcus* cells in the absence of inhibitors (circles) or in the presence of the specific inhibitors bongkrekic acid (squares) or carboxy-atractyloside (triangles). The results are presented as the average and standard deviation from three independent experiments.

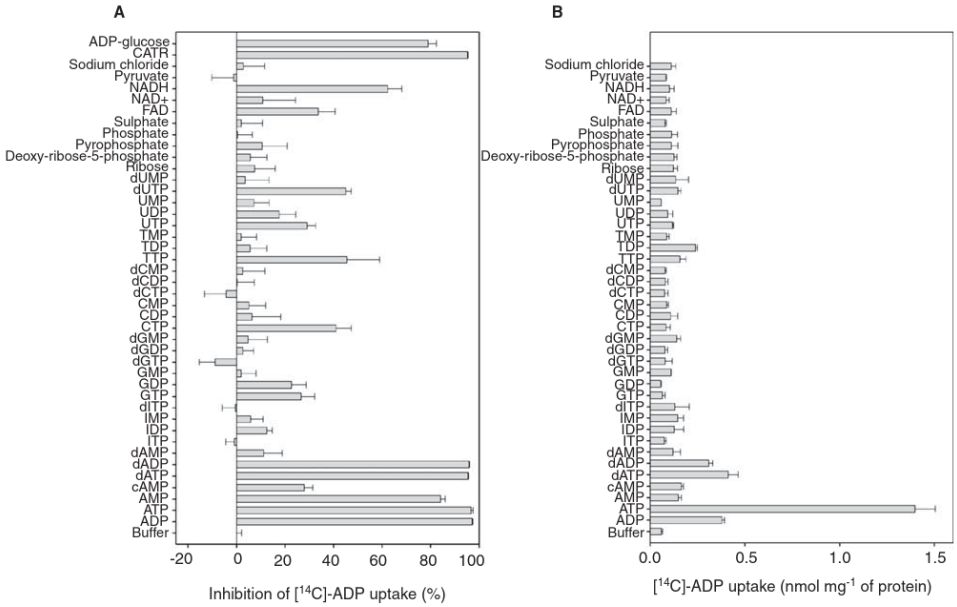


Figure 4. Substrate specificity of the human AAC1 protein. (A) Initial uptake rate of radio-labeled ADP in the presence or absence of a 1000-fold higher concentration of the indicated non-radio-labeled compounds. Uptake was started by the addition of 1.25 μM [^{14}C]-ADP to *Lactococcus* cells at final A_{600nm} of 8 and after 1 h, the uptake was quenched by adding cold buffer and the sample was filtered. (B) Initial uptake rate of radio-labeled ADP into fused membrane vesicles pre-loaded with 5 mM of the indicated compound. Uptake will only occur if the internal compound can be transported out of the vesicle. Substrate exchange was started by the addition of 1.34 μM [^{14}C]-ADP and the initial rate was determined by measuring the accumulated radio-labeled substrate after 10 min. Experiments were carried out in triplicate and the average with standard deviation is shown.

To test whether these compounds are actually transported by the carrier, unlabeled compound was loaded into the fused membrane vesicles and the uptake of radio-labeled ADP was monitored (Figure 4B). Only when the compound is exported from the fused membrane vesicles can uptake of the label occur. Exchange was observed only in the case of ADP, ATP and potentially dADP and dATP in agreement with literature. No transport was observed for the other competitive inhibitors, such as NADH, dUTP, or AMP.

Our results are consistent with previously reported data indicating the strict specificity of ADP/ATP carriers for ADP and ATP (Pfaff and Klingenberg 1968, De Marcos Lousa et al. 2002, Dolce et al. 2005).

Molecular-dynamics simulations

MD simulations have demonstrated the ability of the 1OKC structure to recognize and bind spontaneously ADP (Dehez et al. 2008, Wang and Tajkhorshid 2008). Free-energy calculations have further revealed two binding motifs differing solely by the nucleoside position, the phosphate moiety remaining in close contact with the lower patch of basic residues. Anchoring of the substrate deep inside the carrier cavity, therefore, appears to constitute an early event of the overall ADP/ATP exchange mechanism. The spatial organization of two specific patches of basic residues within the cavity confers to the protein a very specific electrostatic signature consisting of a funnel responsible for conveying ADP down the cavity, towards the bottom of the carrier. However robust and fast, the binding process is very sensitive to any alteration of this signature, possibly induced by either single-point mutations of charged residues, or electrostatic screening in response to a high concentration of small anions such as chloride (Krammer et al. 2009, Ravaud et al. 2012). At the core of the exchange process lies the selectivity of AAC towards ADP and ATP only – neither AMP nor any other nucleotide analogue, e.g., a guanosine phosphate, is transported. In this work, all-atom numerical simulations were employed to gain new insight into the possible role

Table I. Statistical analysis of the association assays. Binding is considered successful when the nucleotide is found inside the cavity, with the phosphate and the nucleoside moieties steadily bound to the bottom of the carrier over time scales in excess of 15 ns.

	Nucleotide inside AAC	Anchoring of the phosphate moiety	Anchoring of the nucleoside moiety
AMP	10	2	1
ADP	10	4	4
ATP	10	2	2
GMP	10	5	1
GDP	10	3	1
GTP	10	1	1

of the 1OKC conformation in the selectivity of the mitochondrial carrier. A series of 10 association experiments were carried out for each of the following nucleotides, AMP, ADP, ATP, GMP, GDP and GTP. In all trajectories, spontaneous binding of the substrate to either the top or the bottom basic patch was monitored. A statistical analysis of the different association assays is reported in Table I and shows that all substrates are recognized by AAC in its conformation open toward the IMS; for all of them at least one stable binding state was observed in agreement with the systematic competition of theses substrates for ADP binding observed in *L. lactis* expressing AAC1 (Figure 4A). At the bottom of the cavity, the binding motif of the phosphate moiety constitutes the common denominator for all nucleotides involving residues K22, R79, R235 and R279 (Figure 5). The situation is somewhat different for those nucleosides that are primarily found at two different positions, B1 and B2 (Figure 5), regardless of the nucleotides. In line with our previous findings (Dehez et al. 2008), B1, which overlaps partially with the binding site of CATR revealed by the crystal structure (Pebay-Peyroula et al. 2003), involves residues Y186, R187, S227 and G224, whereas in B2, the nucleosides adopt a conformation wherein the base moiety interacts with A284 and S21. Sliding and anchoring of the phosphate group appears to be a fast and robust event of the overall association process (occurring on a time scale spanning approximately 10 ns) independent of the nucleoside and the number of terminal phosphate groups. In sharp contrast, reorientation of the sugar and base moieties occurs on a longer time scale and appears to represent the rate-limiting step of the complete binding event. Interestingly enough, stable association of the nucleoside at the bottom of the cavity seems to be easier in the case of ADP and ATP. For the latter two substrates, anchoring of the phosphate moiety at the bottom of the cavity has led systematically to the binding of the nucleoside group to the carrier. It is noteworthy that the polarity of guanine is about three times higher than adenine (Jasien and Fitzgerald 1990) resulting in a weaker interaction with the

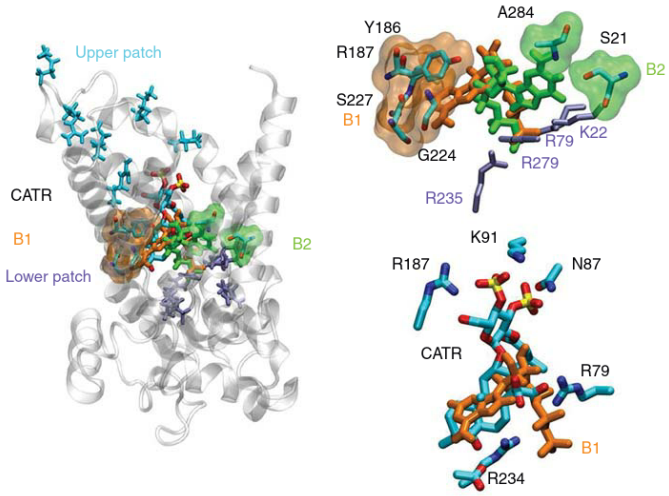


Figure 5. Most populated binding motifs (B1 and B2) revealed by the association-assays. (Left) Global view of the main binding positions of the nucleotides at the bottom of the cavity of AAC near its 1OKC conformation. (Top right) Detailed view of B1 and B2 binding motifs. (Bottom right) Overlap between CATR and B1 binding motifs.

hydrophobic pocket that binds the adenine moiety (contact point II). Furthermore, it might be argued that the lesser polarity of adenine results in a weaker interaction with the local, oriented electric field created by the inner cavity of the transporter and, hence, to a greater orientational entropy, compared to guanine.

Conclusion

In this contribution, the selectivity of the mitochondrial human AAC has been explored employing two expression assays for the carrier, namely *E. coli* and *L. lactis*. Experiments of competitive inhibition reveal a clear preference for the ADP (dADP) and ATP (dATP) tandem, albeit do not rule out the possible interaction with alternate substrates, e.g., AMP, GMP, GDP and GTP, as soon as an appreciable ratio between the concentrations in substrate and inhibitor is involved. These experiments, however, provide only a convoluted picture of the binding specificity and transport selectivity of the latter. Uptake experiments of radio-labeled ADP into fused membrane vesicles preloaded with inhibitors, nevertheless, indicate that, inasmuch as transport is concerned, selectivity only favors ADP and ATP. At the molecular level, numerical simulations indicate that all the nucleotides considered in the computational section of the present study bind to the carrier in a nonspecific fashion through their phosphate moiety, which is consistent with the experiments of competitive inhibition at high concentration ratio. Binding of the base to the carrier appears to be more specific and has been observed systematically for both ADP and ATP once their phosphate group is anchored steadily within the protein. These results suggest that the chemical signature of the nucleobase and the topology of the nucleotide, namely the distance between the base and the terminal phosphate moiety, constitute two stringent criteria likely to modulate the binding of nucleotide analogues to AAC in a conformation akin to the CATR inhibited-state. These observations provide the background for rationalizing the substrate specificity of the carrier, but need to be further complemented to unveil the molecular basis of the transport selectivity of AAC towards ADP and ATP only. In this context, one of the remaining, most challenging tasks consists in isolating alternative AAC conformations and transient binding states that accompany substrate translocation.

Acknowledgements

The collaboration was funded by the European Community's Seventh Framework Programme FP7/2007–2013 under grant agreement no. HEALTH-F4-2007-201924, EDICT Consortium. The work was also supported by the Agence Nationale de la Recherche (ANR Trans-MIT and MIT-2M projects). We thank P. Machillot for his technical help in the uptake experiments, N. Cherradi and O. Filhol-Cochet for access to the radioactivity room (CEA/iRTSV). The Grand Equipement National de Calcul Informatique and Centre Informatique National de l'Enseignement Supérieur are gratefully acknowledged for provision of computer time.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Andersen HC. 1983. A "velocity" version of the Shake algorithm for molecular dynamics calculations. J Comput Phys 52: 24–34.

Brustovetsky N, Bamberg E, Gropp T, Klingenberg M. 1997. Biochemical and physical parameters of the electrical currents measured with the ADP/ATP carrier by photolysis of caged ADP and ATP. Biochemistry 36:13865–13872.

Brustovetsky N, Becker A, Klingenberg M, Bamberg E. 1996. Electrical currents associated with nucleotide transport by the reconstituted mitochondrial ADP/ATP carrier. Proc Natl Acad Sci USA 93:664–668.

Dalbon P, Brandolin G, Boulay F, Hoppe J, Vignais PV. 1988. Mapping of the nucleotide-binding sites in the ADP/ATP carrier of beef heart mitochondria by photolabeling with 2-azido [alpha-32P]adenosine diphosphate. Biochemistry 27:5141–5149.

Darden T, York D, Pederson L. 1993. Article mesh Ewald: An Nlog(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092.

De Marcos Lousa C, Trézéguet V, Dianoux AC, Brandolin G, Lauquin GJ. 2002. The human mitochondrial ADP/ATP carriers: Kinetic properties and biogenesis of wild-type and mutant proteins in the yeast *S. cerevisiae*. Biochemistry 41:14412–14420.

Dehez F, Pebay-Peyroula E, Chipot C. 2008. Binding of ADP in the mitochondrial ADP/ATP carrier is driven by an electrostatic funnel. J Am Chem Soc 130:12725–12733.

Dianoux AC, Noel F, Fiore C, Trézéguet V, Kieffer S, Jaquinod M, et al. 2000. Two distinct regions of the yeast mitochondrial ADP/ATP carrier are photolabeled by a new ADP analogue: 2-azido-3′-O-naphthoyl-[beta-32P]ADP. Identification of the binding segments by mass spectrometry. Biochemistry 39:11477–11487.

Dolce V, Scarcia P, Iacopetta D, Palmieri F. 2005. A fourth ADP/ATP carrier isoform in man: Identification, bacterial expression, functional characterization and tissue distribution. FEBS Lett 579:633–637.

Driessen AJ, de Vrij W, Konings WN. 1985. Incorporation of beef heart cytochrome c oxidase as a proton-motive force-generating mechanism in bacterial membrane vesicles. Proc Natl Acad Sci USA 82:7555–7559.

Duée ED, Vignais PV. 1968. Atracyloside-sensitive translocation of phosphonic acid analogues of adenine nucleotides in mitochondria. Biochem Biophys Res Commun 30:546–553.

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

How Do Membrane Transporters Sense pH? The Case of the Mitochondrial ADP−ATP Carrier

Axel Bidon-Chanal,[†,‡,▽] Eva-Maria Krammer,[†,‡,▽] Delphine Blot,[§,⊥,#] Eva Pebay-Peyroula,[§,⊥,#] Christophe Chipot,[†,‡,∥] Stéphanie Ravaud,[*,§,⊥,#] and François Dehez[*,†,‡,∥]

[†]Université de Lorraine, BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France

[‡]CNRS, UMR 7565 SRSMC, 54506 Vandoeuvre-lès-Nancy, France

[§]Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), 6 rue Jules Horowitz -38000 Grenoble, France

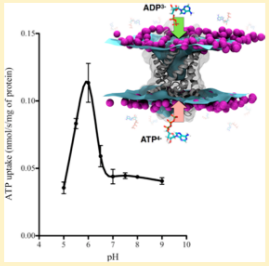
[⊥]CEA, DSV, IBS, 38000 Grenoble, France

[#]CNRS, IBS, 38000 Grenoble, France

[∥]Laboratoire International Associé CNRS/ University of Illinois at Urbana−Champaign, 54506 Vandoeuvre-les-Nancy, France

Supporting Information

ABSTRACT: The activity of many membrane proteins depends markedly on the pH. Pinpointing the amino acids forming the pH sensor domains of these proteins remains challenging for current experimental techniques. Combining molecular dynamics simulations and pK_a predictions with in vitro transport assays, we have revealed the molecular basis of the pH dependence of the mitochondrial carrier mediating the exchanges of ADP^{3-} and ATP^{4-} across the inner mitochondrial membrane. We have demonstrated that the transport activity of this mitochondrial carrier depends on the protonation state of both the substrate and a unique, highly conserved residue of the protein. The original strategy proposed here offers a convenient framework for identifying pH-sensitive residues in membrane proteins in such cases where one single amino acid is involved. Our findings are envisioned to help toward the rational design of active compounds ranging from drugs to biosensors.



SECTION: Biophysical Chemistry and Biomolecules

The question of the protonation state of titratable residues in membrane proteins and their connection to the function of the latter remains a daunting challenge for membrane proteins.[1,2] Recent studies have attempted to identify key residues in specific membrane proteins, the function of which depends upon the protonation state of the former and, hence, upon the pH.[3−5] Such studies can, however, only determine indirectly the regions of the protein that sense pH, but only rarely are individual residues in these regions determined. From a theoretical perspective, pinpointing pH-sensitive residues in membrane proteins represents an equal challenge, as attested to by the recent literature.[6−8] In the latter computational investigations, the identity of the titratable residues is known beforehand, and these residues are examined using a variety of methods ranging from a molecular mechanical to a quantum chemical treatment of a fragment of the membrane protein. The difficulties that these studies have brought to light emphasized the necessity for novel strategies. Here, we combined molecular dynamics simulations and pK_a predictions with in vitro transport assays to reveal the molecular basis of the pH dependence of the mitochondrial carrier mediating the exchanges of ADP^{3-} and ATP^{4-} across the inner mitochondrial membrane.

ATP^{4-}, the energy fuel of the cell, is synthesized from ADP^{3-} in the mitochondrial matrix. Optimal energetic balance in the cell requires continuous exchange of ATP^{4-} and ADP^{3-} across the inner mitochondrial membrane, alteration of this process leading to a variety of pathologies.[9−11] Transport of ADP^{3-} and ATP^{4-} is mediated by the highly substrate-specific ADP/ATP carrier (AAC), a member of the mitochondrial carrier family.[12−15] Due to its abundance in heart mitochondria, the AAC is the best-characterized mitochondrial carrier[16] and, to this date, the only one for which high-resolution structural data are available.[17,18] The structure, solved in the presence of the inhibitor carboxyatractyloside, reveals six transmembrane α-helical segments forming a compact bundle of pseudo-3-fold symmetry. The inhibitor occupies a deep internal cavity, occluded on the matrix side and wide open toward the intermembrane space. It was shown that conformational flexibility of the protein is a key element of the transport process.[10,19−21] Two patches of basic residues line this cavity. The upper basic patch is located at the mouth of AAC, and the lower one delineates the bottom of the cavity. These two

Received: August 30, 2013
Accepted: October 15, 2013
Published: October 15, 2013

patches are responsible for the unique electrostatic signature of the AAC, its funnel-shaped three-dimensional electrostatic potential driving rapidly the substrate down toward the bottom of the internal cavity.[22,23] Mutation or screening of the latter basic residues by high salt concentrations abolishes the transport activity of the AAC.[24,25] The AAC function has also been described to be pH-dependent. Using capacitive-current measurements combined with caged ATP^{4-} and ADP^{3-} photolysis in reconstituted liposomes, Broustovetsky et al. showed that ADP^{3-}/ATP^{4-} exchange is maximal at pHs lying between 6.5 and 7.5.[26] These authors attributed the origin of the pH dependence chiefly to the protonation state of the substrate while suggesting a possible implication of one or several basic amino acids of the AAC. In the present study, we combine synergistically experimental and theoretical approaches to demonstrate that the AAC activity depends not only on the protonation state of the substrate but also on that of the protein. Our investigation reveals the identity of a single, highly conserved amino acid, responsible for the pH dependence of ADP^{3-}/ATP^{4-} transport across the inner membrane of mitochondria.

Transport activity assays of [α-^{32}P]-labeled ATP^{4-} by the isoform 1 of human AAC (hAAC1) expressed in *E. coli* (see Supporting Information section 1) were carried out to follow the variation of the protein activity over a wide range of pH values (Figure 1). Our results agree with the pH dependence curves reported by Broustovetsky et al. for ADP^{3-} (see Figure 9 of ref 26). Protein activity is found to be optimal at a pH between 6 and 7.5, as a function of the experimental setup. The pH of the intermembrane space is about 6.9,[27] and the pK_a's of

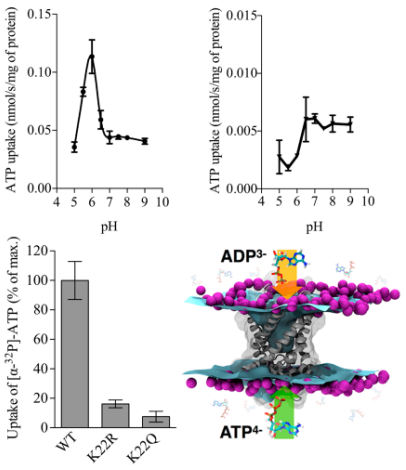


Figure 1. (Top left) pH dependence of [α-^{32}P]-ATP uptake into *E. coli* cells expressing hAAC1. (Top right) pH dependence of [α-^{32}P]-ATP uptake into *E. coli* cells expressing K22R mutant of hAAC1. (Bottom left) Transport activities of K22R and K22Q mutants compared to the wild-type. For clarity, the bovine heart AAC1 sequence numbering was systematically adopted; in hAAC, K22 corresponds to K23. (Bottom right) Schematic illustration of ADP^{3-}/ATP^{4-} exchange through the AAC. Violet spheres represent phospholipid head groups, and the ice-blue surface is an isocontour of the electrostatic potential.

ADP^{3-} and ATP^{4-} are equal to 6.3 and 6.7, respectively. At low pH, impairment of the substrate transport is rooted in the increase of the concentration of the protonated nucleotides, at the expense of their deprotonated counterparts, the latter being the only species transported by the AAC.[14] At high pH values, the substrates primarily exist in their deprotoned form, that is, ADP^{3-} and ATP^{4-}. Decrease of transport activity, therefore, cannot be rationalized by the sole titration state of the nucleotides but is likely to depend also on the protonation state of one or several amino acids of the protein.

To identify residues of the protein that can possibly sense the variation of the pH in a range nearing that of the maximum activity of the carrier, we systematically estimated the pK_a of all titratable residues of the protein with the PROPKA v3.1 software.[28,29] Although it is clear that pK_a prediction algorithms have not yet demonstrated unequivocally their robustness to be considered fully predictive, they, nonetheless, provide a guiding hand to help design experiments that will either confirm or disprove working hypotheses. The calculations were run over 500 representative conformations extracted from the 100 ns molecular dynamics trajectory generated with the apo wild-type bovine heart AAC embedded in a fully hydrated POPC bilayer (Figure 2) (see Supporting Information sections 2 and 3).

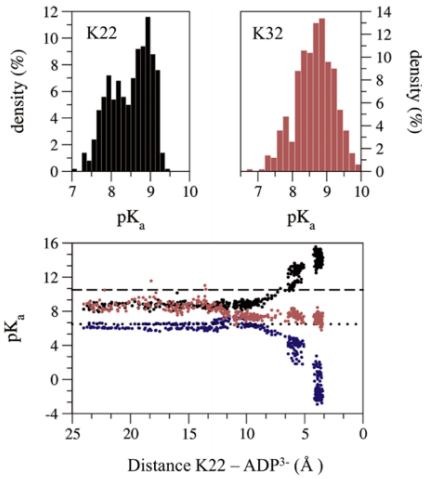


Figure 2. pK_a distributions computed with PROPKA over the last 50 ns of simulation of the apo wild-type AAC for the amino group in the side chains of K22 (top left) and K32 (top right). (Bottom) pK_a evolution of K22 (black dots) and K32 (brown dots) and ADP^{3-} (blue dots) as a function of the distance between the substrate and K22. The dashed and the dotted lines represent the experimental estimates of the pK_a of lysine and ADP^{3-} in aqueous solution.

From this exhaustive study, only two residues emerged, K22 and K32, exhibiting unusually low pK_a values, 8.5 ± 0.5 and 8.6 ± 0.6, respectively (Figure 2), suggestive of a sensitivity to the pH near the optimal value for AAC activity. Although both residues are located in the same region of the AAC, K32 is buried at the bottom of the internal cavity, its side chain forming long-lived interactions with the side chain of Q36 and D231, whereas K22 lies in the middle of the cavity, its side

chain not being engaged in any interaction with neighboring residues. On its way down to the binding site, ADP^{3-} establishes several interactions with the side chains lining the cavity of the protein, including K22 but not K32 (Figure 3).

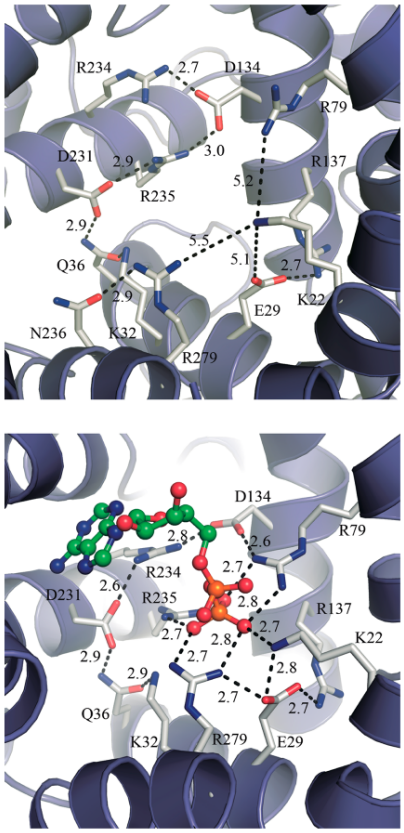


Figure 3. (Top) Close view of the residues surrounding K22 in the apo wild-type protein. (Bottom) Close view of the ADP^{3-} putative binding motif based on an all-atom simulation of AAC in the presence of the nucleotide. All reported distances are given in Å.

Lower pK_a values of 6.0 ± 0.6 and 6.5 ± 0.5 were inferred for K22 using, respectively, H++[30] and QMBP/GMCT,[31−33] methods aimed at predicting the protonation probabilities of titratable residues (see Supporting Information Figure 1). These approaches consistently indicate that the buried K32 cannot be deprotoned (see Supporting Information section 3). It should be stressed that our results do not depend on the protonation state assumed during the MD simulation (see Supporting Information Figure 2). Interestingly enough, the pK_a of ADP^{3-} and K22 (albeit not that of K32) varies as a function of the distance separating the nucleotide from the bottom of the protein cavity (see Figure 2). When the complex is steadily formed, ADP^{3-} and K22 are, respectively, deprotonated and protonated over a wide range of pH values. Put together, our data strongly suggest that K22 is involved in the dependence of AAC transport activity to the pH.

Numerical association assays (see Supporting Information section 1) were carried out to rationalize the effect of the pH on the binding of ADP^{3-} to the carrier. Three different assemblies were constructed, namely, protonated K22/protonated nucleotide ($ADPH^{2-}$), protonated K22/deprotonated nucleotide (ADP^{3-}), and deprotonated K22/deprotonated nucleotide (ADP^{3-}), reflecting acidic, physiological, and basic conditions, respectively. For each assembly, 100 ns molecular dynamics simulations starting from five different orientations of the substrate positioned at the mouth of the carrier were carried out.

Under physiological conditions, ADP^{3-} reached its binding site four times out of five within 25 ns. Going down the internal cavity, ADP^{3-} formed transient contacts with the upper basic patch, as described earlier.[22−24] In sharp contrast, binding was never observed for basic conditions and was monitored only two times for acidic ones, on the 100 ns time scale. In basic conditions, deprotonation of K22 lowered the total charge at the bottom of the cavity and, hence, modified the topology of the electrostatic field inside of the AAC. Such an alteration has been shown to impair systematically the transport activity of the carrier.[10,25] Furthermore, point mutation of K22 into alanine was demonstrated to abolish the activity of the protein.[13] In acidic conditions, the effect of protonation of ADP^{3-} on binding is not as clear-cut.

Using the adaptive biasing force algorithm[34,35] implemented in the NAMD package,[36] we generated two one-dimensional potentials of mean force (PMFs) measuring the reversible work required to translocate the nucleotide from the mouth of the carrier to the bottom of its internal cavity, in physiological and basic pH conditions (see Supporting Information section 4). The two PMFs depicted in Figure 4 ought to be seen as mere indicators of the affinity of the two nucleotides, ADP^{3-} and $ADPH^{2-}$, toward the internal cavity of the AAC. All things being equal, recognition of the deprotonated form of the substrate is to be far more favorable than that of its protonated counterpart. Maintaining the total charge of the endogenous substrates of the AAC, therefore, appears to be as important as preserving the native electrostatic signature of the protein to ensure optimal transport activity.

Uptake assays were performed for two mutants, namely, K22Q and K22R (in the hAAC1 sequence, K22 corresponds to K23). Mutating K22 into glutamine abolishes transport activity (Figure 1), emphasizing the importance of preserving the net charge of the internal cavity to guarantee the function of the protein. In contrast, the carrier activity, however markedly lowered, remained measurable upon mutation of K22 into arginine, the pK_a of the latter residue being unaffected by the cavity environment (see Supporting Information Figure 3). pH dependence experiments were carried out for K22R (Figure 1) and compared to the results obtained for the wild-type hAAC1. For both the wild-type and the mutant, the transport activity diminished when going from physiological to acidic conditions as a result of the increase of nontransportable protonated substrates, $ADPH^{2-}$. In the case of K22R, increasing the pH from physiological to basic conditions did not alter transport; the activity rather plateaus around a maximum value, in glaring contrast with the wild-type. The striking difference in the shape of the two pH dependence profiles cogently demonstrates the direct implication of K22 in the pH-sensing ability of hAAC1.

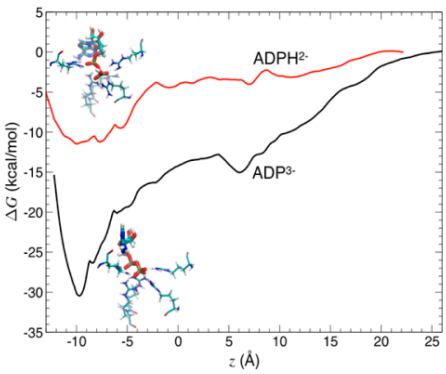


Figure 4. One-dimensional PMFs delineating the reversible association of ADP^{3-} (black curve) and $ADPH^{2-}$ (red curve) to the internal cavity of the AAC with protonated K22. The free-energy differences were determined using the adapting biasing force method. *z* corresponds to the Euclidian distance separating the last phosphate moiety of the nucleotides from the center of mass of the transmembrane helices, projected onto the longitudinal axis of the carrier. The total simulation times needed for mapping the free-energy landscapes of ADP^{3-} and $ADPH^{2-}$ binding the AAC amount to 624 and 580 ns, respectively.

To summarize, the present data reveal that the transport activity of AAC does not only depend on the protonation state of the substrate but also on that of the protein. Combining theoretical and experimental assays, we have shown that a single residue, K22, highly conserved in all AAC sequences, allows the protein to sense pH variations. By virtue of a pK_a value much lower than that expected in the bulk environment, this lysine residue is mainly deprotonated in basic conditions, a situation conducive to impair severely AAC activity. At variance, mutating K22 into arginine caused the knockout of the protein response to basic conditions. The original strategy put forth herein is envisioned to help in pinpointing amino acids composing pH sensor domains of membrane proteins.

■ ASSOCIATED CONTENT

Supporting Information

Experimental protocols and methodological details, including transport activity measurements, molecular dynamics simulations, pK_a calculations, and potential of mean force details, are given. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: Stephanie.Ravaud@ibs.fr (S.R.).
*E-mail: François.Dehez@univ-lorraine.fr (F.D.).

Author Contributions

[V]A.B.-C. and E.-M.K. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Agence National de La Recherche (ANR MIT-2M, N°ANR 2010 BLAN1518-01) and the Direction Régionale à la Recherche et à la Technologie de Lorraine. The Grand Equipement National de Calcul Informatique (GENCI) and the Centre Informatique National de l'Enseignement Supérieur (CINES) are gratefully acknowledged for the provision of computer time.

■ REFERENCES

(1) Frericks Schmidt, H. L.; Shah, G. J.; Sperling, L. J.; Rienstra, C. M. NMR Determination of Protein pK_a Values in the Solid State. *J. Phys. Chem. Lett.* **2010**, *1*, 1623−1628.
(2) Gleason, N. J.; Vostrikov, V. V.; Greathouse, D. V.; Koeppe, R. E. Buried Lysine, But Not Arginine, Titrates and Alters Transmembrane Helix Tilt. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 1692−1695.
(3) Thompson, A. N.; Posson, D. J.; Parsa, P. V.; Nimigean, C. M. Molecular Mechanism of pH Sensing in KcsA Potassium Channels. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6900−6905.
(4) Hu, F.; Schmidt-Rohr, K.; Hong, M. NMR Detection of pH-Dependent Histidine−Water Proton Exchange Reveals the Conduction Mechanism of a Transmembrane Proton Channel. *J. Am. Chem. Soc.* **2012**, *134*, 3703−3713.
(5) Hong, M.; DeGrado, W. F. Structural Basis for Proton Conduction and Inhibition by the Influenza M2 Protein. *Protein Sci.* **2012**, *21*, 1620−1633.
(6) Li, L.; Vorobyov, I.; Allen, T. W. Potential of Mean Force and pK_a Profile Calculation for a Lipid Membrane-Exposed Arginine Side Chain. *J. Phys. Chem. B* **2008**, *112*, 9574−9587.
(7) Goyal, P.; Ghosh, N.; Phatak, P.; Clemens, M.; Gaus, M.; Elstner, M.; Cui, Q. Proton Storage Site in Bacteriorhodopsin: New Insights from Quantum Mechanics/Molecular Mechanics Simulations of Microscopic pK_a and Infrared Spectra. *J. Am. Chem. Soc.* **2011**, *133*, 14981−14997.
(8) Zhu, S.; Brown, M. F.; Feller, S. E. Retinal Conformation Governs pK_a of Protonated Schiff Base in Rhodopsin Activation. *J. Am. Chem. Soc.* **2013**, *135*, 9391−9398.
(9) Palmieri, F. Diseases Caused by Defects of Mitochondrial Carriers: A Review. *Biochim. Biophys. Acta, Bioenerg.* **2008**, *1777*, 564−578.
(10) Ravaud, S.; Bidon-Chanal, A.; Blesneac, I.; Machillot, P.; Juillan-Binard, C.; Dehez, F.; Chipot, C.; Pebay-Peyroula, E. Impaired Transport of Nucleotides in a Mitochondrial Carrier Explains Severe Human Genetic Diseases. *ACS Chem. Biol.* **2012**, *7*, 1164−1169.
(11) Clémençon, B.; Babot, M.; Trézéguet, V. The Mitochondrial ADP/ATP Carrier (SLC25 Family): Pathological Implications of Its Dysfunction. *Mol. Aspects Med.* **2013**, *34*, 485−493.
(12) Kunji, E. R. The Role and Structure of Mitochondrial Carriers. *FEBS Lett.* **2004**, *564*, 239−244.
(13) Nury, H.; Dahout-Gonzalez, C.; Trézéguet, V.; Lauquin, G. J. M.; Brandolin, G.; Pebay-Peyroula, E. Relations Between Structure and Function of the Mitochondrial ADP/ATP Carrier. *Annu. Rev. Biochem.* **2006**, *75*, 713−741.
(14) Mifsud, J.; Ravaud, S.; Krammer, E.-M.; Chipot, C.; Kunji, E. R. S.; Pebay-Peyroula, E.; Dehez, F. The Substrate Specificity of the Human ADP/ATP Carrier AAC1. *Mol. Membr. Biol.* **2013**, *30*, 160−168.
(15) Monné, M.; Palmieri, F.; Kunji, E. R. S. The Substrate Specificity of Mitochondrial Carriers: Mutagenesis Revisited. *Mol. Membr. Biol.* **2013**, *30*, 149−159.
(16) Klingenberg, M. The ADP and ATP Transport in Mitochondria and Its Carrier. *Biochim. Biophys. Acta* **2008**, *1778*, 1978−2021.
(17) Pebay-Peyroula, E.; Dahout-Gonzalez, C.; Kahn, R.; Trézéguet, V.; Lauquin, G. J.-M.; Brandolin, G. Structure of Mitochondrial ADP/ATP Carrier in Complex with Carboxyatractyloside. *Nature* **2003**, *426*, 39−44.
(18) Nury, H.; Dahout-Gonzalez, C.; Trézéguet, V.; Lauquin, G.; Brandolin, G.; Pebay-Peyroula, E. Structural Basis for Lipid-Mediated Interactions Between Mitochondrial ADP/ATP Carrier Monomers. *FEBS Lett.* **2005**, *579*, 6031−6036.

(19) Johnston, J. M.; Khalid, S.; Sansom, M. S. P. Conformational Dynamics of the Mitochondrial ADP/ATP Carrier: A Simulation Study. *Mol. Membr. Biol.* **2008**, *25*, 506−517.
(20) Babot, M.; Blancard, C.; Pelosi, L.; Lauquin, G. J.-M.; Trézéguet, V. The Transmembrane Prolines of the Mitochondrial ADP/ATP Carrier Are Involved in Nucleotide Binding and Transport and Its Biogenesis. *J. Biol. Chem.* **2012**, *287*, 10368−10378.
(21) Babot, M.; Blancard, C.; Zeman, I.; Lauquin, G. J.-M.; Trézéguet, V. Mitochondrial ADP/ATP Carrier: Preventing Conformational Changes by Point Mutations Inactivates Nucleotide Transport Activity. *Biochemistry (Moscow)* **2012**, *51*, 7348−7356.
(22) Wang, Y.; Tajkhorshid, E. Electrostatic Funneling of Substrate in Mitochondrial Inner Membrane Carriers. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9598−9603.
(23) Dehez, F.; Pebay-Peyroula, E.; Chipot, C. Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel. *J. Am. Chem. Soc.* **2008**, *130*, 12725−12733.
(24) Krämer, R.; Kürzinger, G. The Reconstituted ADP/ATP Carrier from Mitochondria Is Both Inhibited and Activated by Anions. *Biochim. Biophys. Acta., Bioenerg.* **1984**, *765*, 353−362.
(25) Krammer, E.-M.; Ravaud, S.; Dehez, F.; Frelet-Barrand, A.; Pebay-Peyroula, E.; Chipot, C. High-Chloride Concentrations Abolish the Binding of Adenine Nucleotides in the Mitochondrial ADP/ATP Carrier Family. *Biophys. J.* **2009**, *97*, L25−L27.
(26) Broustovetsky, N.; Bamberg, E.; Gropp, T.; Klingenberg, M. Biochemical and Physical Parameters of the Electrical Currents Measured with the ADP/ATP Carrier by Photolysis of Caged ADP and ATP. *Biochemistry (Moscow)* **1997**, *36*, 13865−13872.
(27) Porcelli, A. M.; Ghelli, A.; Zanna, C.; Pinton, P.; Rizzuto, R.; Rugolo, M. pH Difference Across the Outer Mitochondrial Membrane Measured with a Green Fluorescent Protein Mutant. *Biochem. Biophys. Res. Commun.* **2005**, *326*, 799−804.
(28) Li, H.; Robertson, A. D.; Jensen, J. H. Very Fast Empirical Prediction and Rationalization of Protein pK_a Values. *Proteins: Struct., Funct., Bioinform.* **2005**, *61*, 704−721.
(29) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very Fast Prediction and Rationalization of pK_a Values for Protein−Ligand Complexes. *Proteins* **2008**, *73*, 765−783.
(30) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pK_as and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368−W371.
(31) Ullmann, R. T.; Ullmann, G. M. Coupling of Protonation, Reduction, and Conformational Change in Azurin from Pseudomonas Aeruginosa Investigated with Free Energy Measures of Cooperativity. *J. Phys. Chem. B* **2011**, *115*, 10346−10359.
(32) Ullmann, R. T.; Andrade, S. L. A.; Ullmann, G. M. Thermodynamics of Transport Through the Ammonium Transporter Amt-1 Investigated with Free Energy Calculations. *J. Phys. Chem. B* **2012**, *116*, 9690−9703.
(33) Ullmann, R. T.; Ullmann, G. M. GMCT: A Monte Carlo Simulation Package for Macromolecular Receptors. *J. Comput. Chem.* **2012**, *33*, 887−900.
(34) Darve, E.; Pohorille, A. Calculating Free Energies Using Average Force. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(35) Hénin, J.; Chipot, C. Overcoming Free Energy Barriers Using Unconstrained Molecular Dynamics Simulations. *J. Chem. Phys.* **2004**, *121*, 2904−2914.
(36) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

Feller SE, Mackerell AD Jr. 2000. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. J Phys Chem B104:7510–7515.
Heidkamper D, Müller V, Nelson DR, Klingenberg M. 1996. Probing the role of positive residues in the ADP/ATP carrier from yeast. The effect of six arginine mutations on transport and the four ATP versus ADP exchange modes. Biochemistry 35:16144–16152.
Hénin J, Shinoda W, Klein ML. 2008. United-atom acyl chains for CHARMM phospholipids. J Phys Chem B 112:7008–7015.
Humphrey W, Dalke A, Schulten K. 1996. VMD – visual molecular dynamics. J Mol Graphics 97:25–27.
Jasien P, Fitzgerald G. 1990. Molecular dipole moments and polarizabilities from local density functional calculations: Application to DNA base pairs. J Chem Phys 93:2554.
Klingenberg M. 2008. The ADP and ATP transport in mitochondria and its carrier. Biochim Biophys Acta 1778:1978–2021.
Krämer R, Klingenberg M. 1980. Enhancement of reconstituted ADP, ATP exchange activity by phosphatidylethanolamine and by anionic phospholipids. FEBS Lett 119:257–260.
Krammer EM, Ravaud S, Dehez F, Frelet-Barrand A, Pebay-Peyroula E, Chipot C. 2009. High-chloride concentrations abolish the binding of adenine nucleotides in the mitochondrial ADP/ATP carrier family. Biophys J 97:L25–L27.
Kunji ER. 2004. The role and structure of mitochondrial carriers. FEBS Lett 564:239–244.
Kunji ER, Slotboom DJ, Poolman B. 2003. *Lactococcus lactis* as host for overproduction of functional membrane proteins. Biochim Biophys Acta 1610:97–108.
Mackerell AD Jr, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B102:3586–3616.
Mackerell AD Jr, Feig M, Brooks CL 3rd. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25:1400–1415.
Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289–298.
Monne M, Chan KW, Slotboom DJ, Kunji ER. 2005. Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. Protein Sci 14:3048–3056.
Müller V, Heidkamper D, Nelson DR, Klingenberg M. 1997. Mutagenesis of some positive and negative residues occurring in repeat triad residues in the ADP/ATP carrier from yeast. Biochemistry 36:16008–16018.
Nury H, Dahout-Gonzalez C, Trézéguet V, Lauquin GJ, Brandolin G, Pebay-Peyroula E. 2006. Relations between structure and function of the mitochondrial ADP/ATP carrier. Annu Rev Biochem 75:713–741.
Palmieri F. 2004. The mitochondrial transporter family (SLC25): Physiological and pathological implications. Pflugers Arch 447:689–709.
Pebay-Peyroula E, Dahout-Gonzalez C, Kahn R, Trézéguet V, Lauquin GJ, Brandolin G. 2003. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. Nature 426:39–44.
Pfaff E, Klingenberg M. 1968. Adenine nucleotide translocation of mitochondria. 1. Specificity and control. Eur J Biochem 6: 66–79.
Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802.
Ravaud S, Bidon-Chanal A, Blesneac I, Machillot P, Juillan-Binard C, Dehez F, et al. 2012. Impaired transport of nucleotides in a mitochondrial carrier explains severe human genetic diseases. ACS Chem Biol 7:1164–1169.
Robinson AJ, Kunji ER. 2006. Mitochondrial carriers in the cytoplasmic state have a common substrate binding site. Proc Natl Acad Sci USA 103:2617–2622.
Robinson AJ, Overy C, Kunji ER. 2008. The mechanism of transport by mitochondrial carriers based on analysis of symmetry. Proc Natl Acad Sci USA 105:17766–17771.
Tuckerman M, Berne BJ, Martyna GJ. 1992. Reversible multiple time scale molecular dynamics. J Chem Phys 97:1990–1999.
Wang Y, Tajkhorshid E. 2008. Electrostatic funneling of substrate in mitochondrial inner membrane carriers. Proc Natl Acad Sci USA 105:9598–9603.

II. Effets de détergents sur les structures des protéines membranaires

Les protéines membranaires sont la porte d'entrée vers l'intérieur de la cellule. Malgré leur implication dans de nombreuses fonctions cellulaires majeures, la compréhension de leurs mécanismes de fonctionnement reste extrêmement limitée, en raison notamment de la difficulté d'accéder à des structures tridimensionnelles haute résolution. Outre la difficulté de produire les protéines membranaires dans des quantités suffisantes, celles-ci doivent généralement être extraites de leur environnement de membrane natif puis purifiées, et éventuellement reconstituées dans un environnement approprié. Les détergents jouent un rôle central dans toutes ces étapes. L'écrasante majorité des structures de protéines membranaires actuellement connues, déterminées par cristallographie aux rayons X, spectroscopie RMN en solution et Cryo-microscopie électronique a été obtenue à partir de protéines solubilisées en détergent. Les détergents à base de phoscholine zwitterionique (PC) sont parmi les plus utilisés dans les structures obtenues par RMN.[63] Une partie de mon activité de recherche consiste à étudier les effets des phoscholines sur les structures des protéines membranaires pour éprouver le sens physiologique de ces dernières.

II.1. Étude du transporteur mitochondrial UCP2 (UnCoupling Protein)

En 2011 est venu s'ajouter à la seule structure haute résolution connue d'un transporteur mitochondrial une structure basse résolution du transporteur UCP2.[64] Berardi et al. ont combiné des données RMN (Residual Dipolar Coupling et Paramagnetic Relaxation Enhancement) à une méthode de remplacement de fragment pour déterminer la position des atomes du squelette de la protéine. Comme nous l'avions fait pour AAC nous avons construit un modèle de UCP2 immergé dans une membrane POPC totalement hydratée (figure 16). Nous avons ensuite réalisé des simulations de dynamique moléculaire sur des échelles allant jusqu'à 400 à 500 ns pour équilibrer le système construit. Plusieurs protocoles ont été envisagés en incluant ou non une molécule de GDP, inhibiteur de UCP, présent dans les expériences de RMN. Les atomes du squelette ont dans un premier temps été restreints à leur position d'origine par un potentiel harmonique. Dans ces conditions nos simulations montrent que la structure d'UCP2 telle que résolue par Berardi et al. se comporte comme un large pore traversé par un flux massif de molécule d'eau (figure 16). Cette situation n'est bien sûr pas compatible avec le maintien d'un gradient proton de part et d'autre de la membrane interne de la mitochondrie, condition sine qua none à la survie de l'organelle.[65]

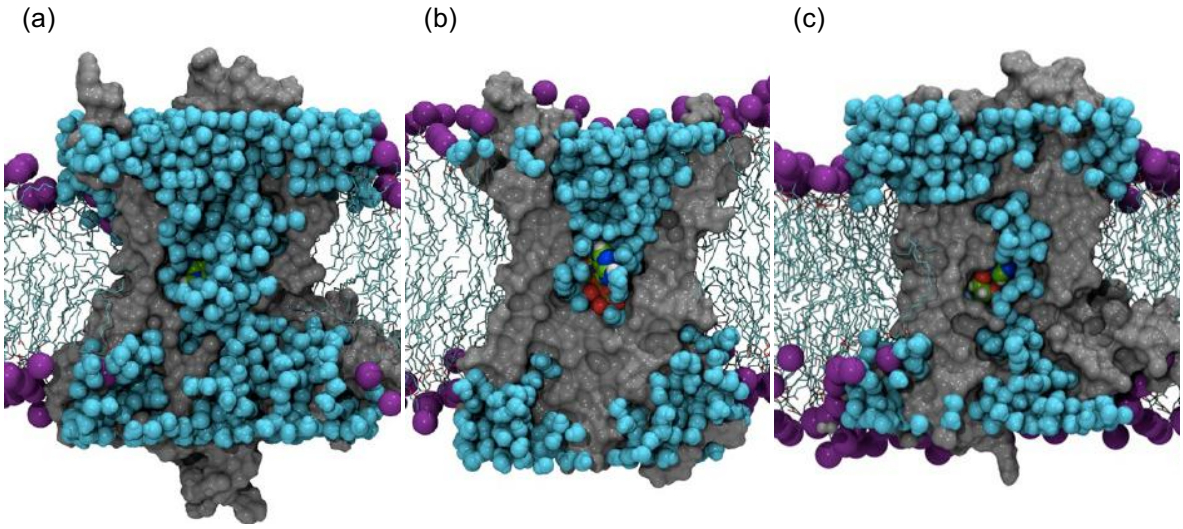


Figure 16. Plans de coupe de UCP2 et AAC inséré dans une bicouche de POPC (a) UCP2- GDP^{3-}. Les atomes du squelette sont contraints aux positions observées dans la structure RMN. Un flux d'eau massif connecte les deux côtés de la membrane (b) AAC -ADP3-. Le transporteur est complètement occlus coté matriciel (c) UCP2-GDP3- après relaxation de la structure de la protéine. La structure secondaire de la protéine s'effondre, conduisant à une structure beaucoup plus compacte que la structure RMN de référence

Les restreintes ont ensuite été diminuées progressivement jusqu'à leur annulation laissant le système s'équilibrer librement. Lorsque les restreintes sur les positions atomiques sont enlevées, la structure de UCP2 évolue de manière importante vers un état beaucoup plus compact, pratiquement imperméable et fermé côté espace intermembranaire (figure 16). Sur le côté matriciel, les hélices interfaciales se réorganisent également de manière importante conduisant à un RMSD stable mais important de l'ordre de 7 Å. Ce comportement est observé systématiquement quel que soit le protocole d'équilibration et quelle que soit la présence de GDP ou non. Ces résultats questionnent évidemment le sens de la structure résolue par RMN et démontrent que celle-ci n'est pas

compatible avec un environnement membranaire. Dans un premier temps nous avons pensé qu'à un jeu de contraintes RDC et de PRE pouvait correspondre plusieurs structures. Nous avons calculé ces contraintes à partir de notre structure relaxée. Les valeurs, sensiblement différentes de celle mesurées, ne sont pas particulièrement sensibles à l'évolution de la conformation au cours de la trajectoire. Nous avons donc fait l'hypothèse que, plus qu'une dégénérescence liée à la procédure de mise au point de la structure, le problème vient des conditions dénaturantes (DPC) dans lesquelles les expériences de Berardi et al. ont été faites.

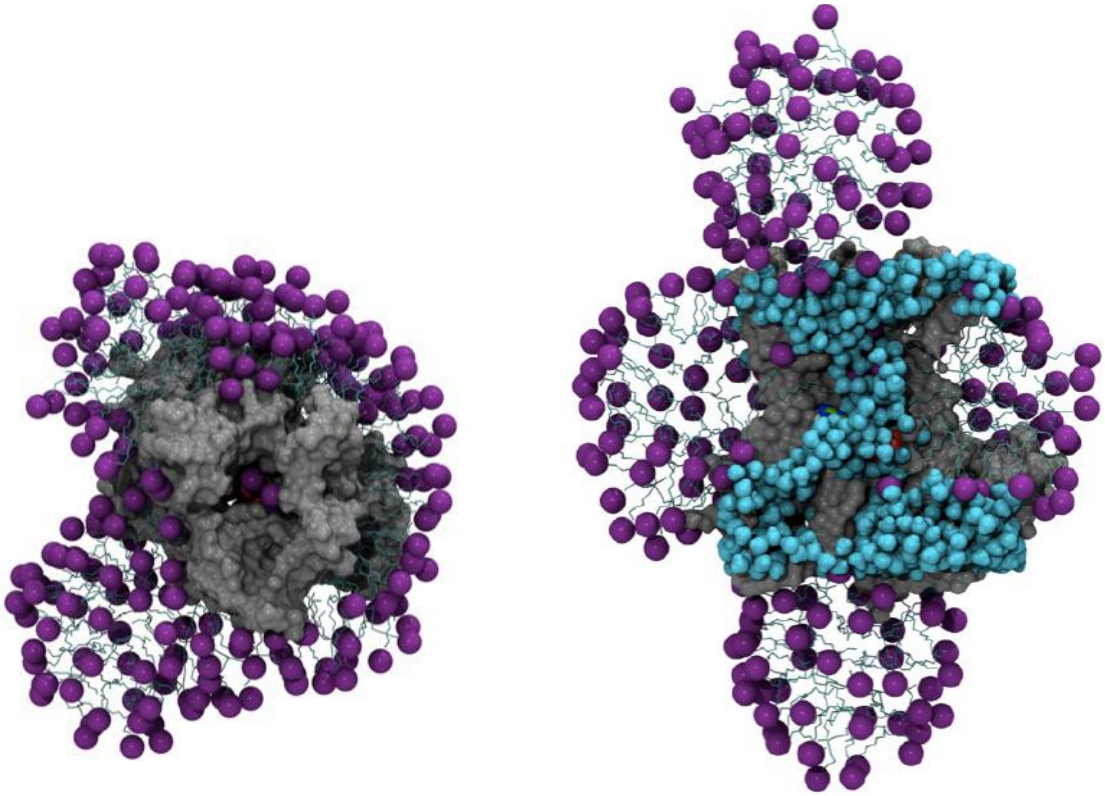


Figure 17. UCP2-GDP3- inséré dans des micelles de DPC micelles après relaxation complète de la protéine. Les molécules de détergent forment une bouée autour du cœur hydrophobe de la protéine. Deux micelles s'auto assemblent sur les interfaces côté matrice et espace intermembranaire, autour de zones amphiphiles. La cavité interne reste complètement ouverte telle que dans la structure obtenue par RMN et modélisation. (Gauche) Vue de dessus, pour plus de clarté la micelle côté espace intermembranaire n'est pas représentée. (Droite) plan de coupe du complexe UCP2/DPC.

Nous avons réalisé de multiples simulations de dynamique moléculaires gros-grains et tout atome pour étudier l'association de molécules de DPC à UCP2 à des concentrations similaires à celles employées dans les expériences de Berardi et al. Nos travaux démontrent que dans la structure RMN, les espaces entre hélices sont suffisants pour permettre aux molécules de DPC de s'insérer profondément dans la protéine. Par ailleurs les molécules de DPC s'auto-assemblent et tapissent les résidus hydrophobes situés en surface d'UCP2 (figure 17). Nos simulations de dynamique moléculaire sans restreintes, réalisées sur le complexe UCP2/DPC, montrent qu'en présence de détergent la structure prédite par RMN est stable (figure 17). Associé aux expériences de nos collègues de l'IBPC montrant l'inactivation de UCP2 par le DPC, nos simulations démontrent que la structure publiée par Berardi et al. est fortement altérée par les conditions d'extraction.[65]

Nos travaux, objet d'un highlight[66] dans la revue Journal of American Chemical Society, ont démontré que les simulations peuvent être employées non seulement pour étudier la dynamique des protéines membranaires au voisinage de leur structure d'équilibre mais également pour éprouver le sens physiologique de structure obtenues dans des environnements non natifs.

II.2. Effet du DPC sur la structure du transporteur de l'ADP/ATP

Grace à l'utilisation de variables collectives, j'ai intégré dans les simulations de dynamique moléculaire de l'AAC des contraintes de structure secondaires déduites de mesures RMN obtenues en DPC. Le programme TALOS+[67] a permis de prédire les angles de torsions associés au squelette des protéines à partir des déplacements chimiques mesures sur des échantillons de protéines en présence de DPC. Les mesures de déplacements chimiques faites chez nos collaborateurs à Grenoble (P. Schanda, IBS) sur des échantillons du transporteur ADP/ATP de levure nous ont permis de quantifier la déformation structurale asymétrique subit par la protéine en présence du détergent comme l'illustre la figure 18. Ces résultats nous ont permis de conclure que la dynamique µs-ms observée par RMN[68] n'était pas associée à la fonction mais était la conséquence de la déformation structurale, des expériences ont d'ailleurs montrées que cette dynamique non-spécifique perdurait en présence d'inhibiteur. [69,70]

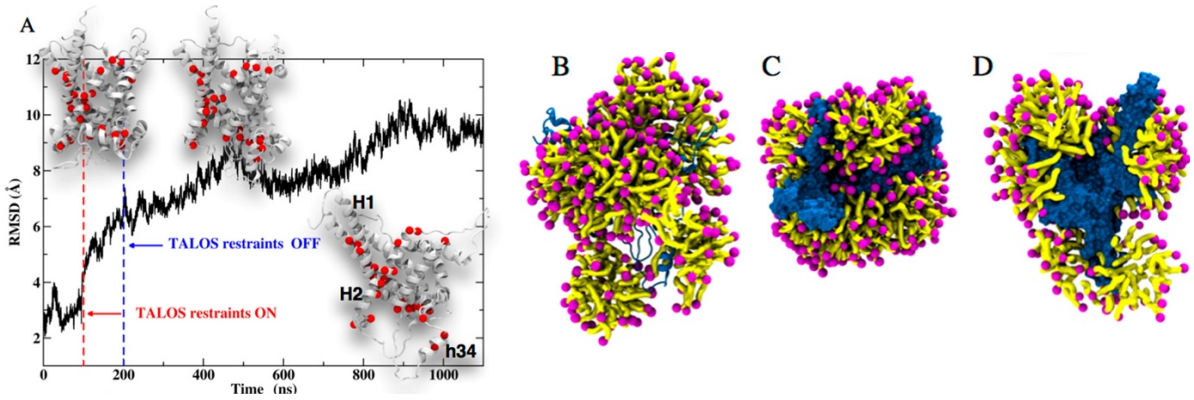


Figure 18. (A) Déplacement quadratique moyen des positions atomiques associée au transporteur AAC de levure (yAAC) au cours d'une trajectoire de dynamique moléculaire avant, pendant et après l'activation de restreintes déduites des déplacements chimiques mesures par RMN (prédiction TALOS+). (B,C,D) Agrégation de molécule de détergent (DPC) autour de yAAC après 1 µs de dynamique moléculaire, vues de côté (A), de dessus (B) et de coupe (D).

S'opposant à nos conclusions, le groupe de J. Chou a publié une étude suggérant que le DPC n'altérait pas la capacité du transporteur AAC à se lier de manière spécifique aux cardiolipines (des phospholipides essentiels au fonctionnement de AAC) et qu'en conséquence ce détergent était tout à fait adapté aux études structurales de cette famille de protéines.[71] Nous avons réalisé des dynamiques moléculaires en présence de 3 cardiolipines (CLs) placées à leurs sites de liaisons (tels qu'observées en cristallographie aux rayons X). Nos trajectoires montrent que les têtes polaires des CLs restent parfaitement ancrées à leurs sites de liaison là où les chaines grasses de ces mêmes lipides sont très flexibles et très mobiles (Figure 19 (A, B)). Dans les spectres NOE obtenus en DPC, les signaux montrent également l'ancrage des têtes polaires mais reflètent une association très spécifique des extrémités des chaines grasses au côté cytoplasmique de la protéine (à l'opposé

des sites des fixations des têtes polaires côté matriciel). La comparaison de nos trajectoires aux spectres NOE déterminés en DPC démontrent que pour être compatible, la structure de la protéine doit être déformée (Figure 19 (C)).

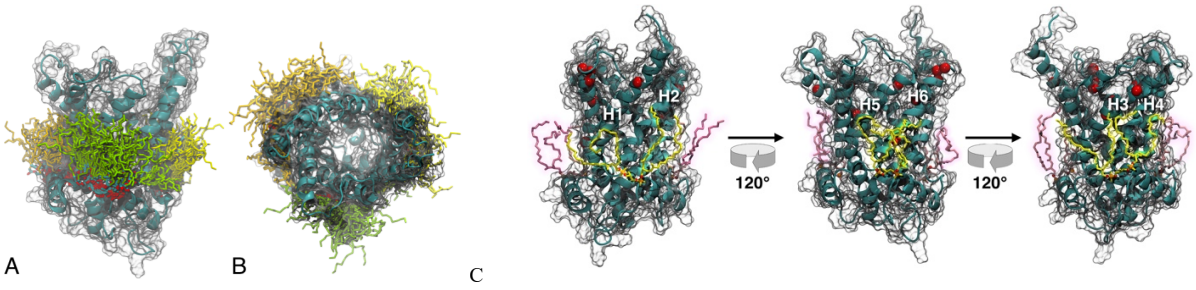


Figure 19. Snapshot de yAAC3 après ~150 ns d'une simulation de 200 ns. (A,B) Vue de côté et de dessus de la dynamique conformationnelle des cardiolipines (les structures des cardiolipines observées à des intervalles de temps réguliers sont superposées les unes aux autres). (C) Représentation des sites pour lesquels des crosspeaks NOE sont observés en DPC (I14, I100, G118, G127, F220, L282, S292, Q298, M299, et I300). Les distances entre les extrémités des chaines acyles et la protéine dans sa forme native en membrane sont clairement incompatible avec les signaux NOE mesurés en DPC.

II.3. Effet du DPC sur protéine p7 de l'appareil de réplication du virus HCV

La protéine p7 appartient à l'appareil de réplication du virus de l'hépatite C (HCV).[72] Celle-ci s'auto-assemble pour former des viroporines, canaux multimériques, assurant un transport ionique à travers la membrane du réticulum endoplasmique. A ce titre, p7 est une cible privilégiée pour le développement de médicaments capable d'endiguer la réplication du virus HCV. Les caractéristiques structurelles du monomère p7 ne sont toutefois pas complètement comprises et restent l'objet de nombreuses discussions. En particulier, l'étude structurale de p7 n'a été réalisée que dans des environnements détergents, rendant l'interprétation des résultats expérimentaux difficile. Dans la littérature deux modèles s'affrontent, un dans lequel le monomère de p7 est décrit comme une épingle à cheveux beta,[73] et l'autre dans lequel p7 a une conformation ouverte très inhabituelle.[74] Nous avons utilisé les simulations de dynamique pour étudier la structure et la dynamique des différentes conformations de p7 (formes monomériques et hexamèrique) avec l'objectif de caractériser leur sens physiologique respectif (figure 20). Pour cela nous avons dans un premier temps étudié les monomères en bicouche phospholipidique et en micelles de DPC (dodecylphosphocoline), le détergent utilisé par Ouyang et Chou[74] pour résoudre la structure ouverte de p7 par RMN/modélisation. Nos résultats montrent que la DPC favorise les structures ouvertes de p7, là où la conformation en épingle à cheveux est stabilisée par un environnement membranaire.[75] En collaboration avec le groupe de RMN de J. Schnell à Oxford nous avons finalement démontré que la structure hexamèrique du canal construite avec la forme ouverte de p7 était le résultat d'une deutération incomplète de l'échantillon utilisé pour les analyses RMN.[76,77]

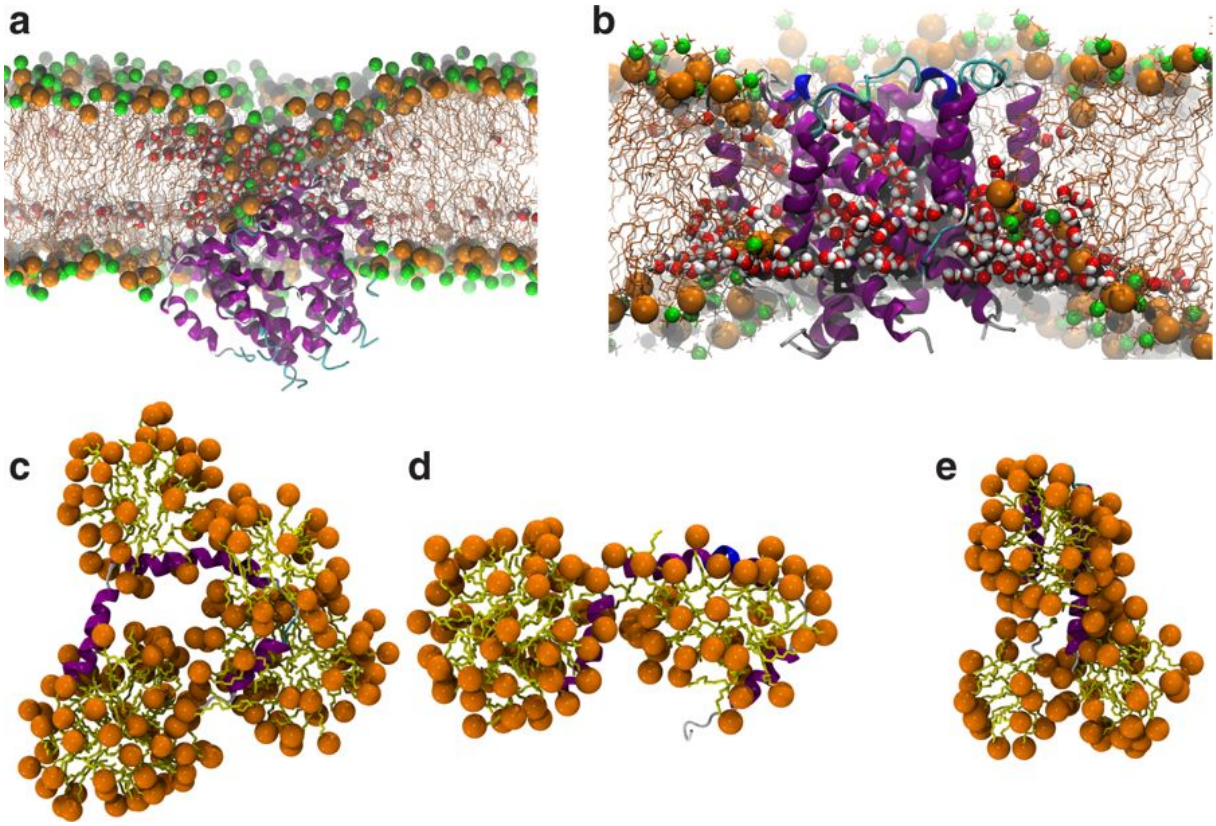


Figure 20. (a, b) Défauts d'insertion de la forme hexamèrique de p7 proposes par le groupe de J. Chou dans une bicouche lipidique. Prédiction d'insertion par le serveur MEMPROTMD dans une bicouche de DDPC (a). Structure de p7 insérée dans une bicouche de POPC obtenue après une trajectoire de 100 ns. (b). D'importantes déformations de la membrane peuvent être observées conduisant à la présence de nombreuses molécules d'eau dans la partie hydrophobe de la bicouche. (c-e) Simulations du monomère de p7 dans 300 mM de DPC (ratio protéine : détergent = 1). Deux simulations indépendantes de la forme entendue de p7 montrent que ~170 (c) and ~120 DPC (d) molécules de DPC complexent la protéine (c et e), contre ~100 molécules de DPC complexant la structure consensuelle en hairpine (e) (en accord avec les expériences de SEC-MALS (Size Exclusion Chomatography - Multi-Angle Light Scattering).

Ci-après sont reproduits 4 articles illustrant mes travaux dédiés aux effets de détergents sur les structures des protéines membranaires :

- Zoonens, M., Comer, J., Masscheleyn, S., Pebay-Peyroula, E., Chipot, C., Miroux, B. * & Dehez, F. * Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2. *J. Am. Chem. Soc.* **135**, 15174–15182 (2013)
- Holzmann, N., Chipot, C., Penin, F. & Dehez, F.* Assessing the physiological relevance of alternate architectures of the p7 protein of hepatitis C virus in different environments. *Bioorg. Med. Chem.* **24**, 4920–4927 (2016).
- Dehez, F., Schanda, P., King, M. S., Kunji, E. R. S. & Chipot, C. Mitochondrial ADP/ATP Carrier in Dodecylphosphocholine Binds Cardiolipins with Non-native Affinity. *Biophys. J.* **113**, 2311–2315 (2017).
- Kurauskas, V., Hessel, A., Ma, P., Lunetti, P., Weinhäupl, K., Imbert, L., Brutscher, B., King, M. S., Sounier, R., Dolce, V., Kunji, E. R. S., Capobianco, L., Chipot, C., Dehez, F., * Bersch, B. * & Schanda, P. * How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine. *J. Phys. Chem. Lett.* **9**, 933–938 (2018).

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

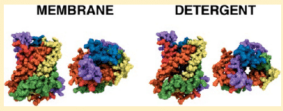
Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2

Manuela Zoonens,[†,‡] Jeffrey Comer,[§,⊥] Sandrine Masscheleyn,[†,‡] Eva Pebay-Peyroula,[∥,#,⊗] Christophe Chipot,[§,⊥,▽] Bruno Miroux,*[,†,‡] and François Dehez*[,§,⊥,▽]

[†]CNRS UMR 7099, Institut de Biologie Physico Chimique (IBPC), 75005 Paris, France
[‡]Université Paris-Diderot, IBPC, 75005 Paris, France
[§]CNRS UMR 7565, Structure et Réactivité des Systèmes Moléculaires Complexes (SRSMC), 54500 Vandoeuvre-les-Nancy, France
[⊥]Université de Lorraine, SRSMC, 54500 Vandoeuvre-les-Nancy, France
[∥]Université Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France
[#]CEA, DSV, IBS, F-38027 Grenoble, France
[⊗]CNRS, IBS, F-38027 Grenoble, France
[▽]Laboratoire International Associé CNRS and University of Illinois at Urbana−Champaign, 54506 Vandoeuvre-lès-Nancy, France

Supporting Information

ABSTRACT: The extraction of membrane proteins from their native environment by detergents is central to their biophysical characterization. Recent studies have emphasized that detergents may perturb the structure locally and modify the dynamics of membrane proteins. However, it remains challenging to determine whether these perturbations are negligible or could be responsible for misfolded conformations, altering the protein's function. In this work, we propose an original strategy combining functional studies and molecular simulations to address the physiological relevance of membrane protein structures obtained in the presence of detergents. We apply our strategy to a structure of isoform 2 of an uncoupling protein (UCP2) binding an inhibitor recently obtained in dodecylphosphocholine detergent micelles. Although this structure shares common traits with the ADP/ATP carrier, a member of the same protein family, its functional and biological significance remains to be addressed. In the present investigation, we demonstrate how dodecylphosphocholine severely alters the structure as well as the function of UCPs. The proposed original strategy opens new vistas for probing the physiological relevance of three-dimensional structures of membrane proteins obtained in non-native environments.



INTRODUCTION

Uncoupling proteins (UCPs) belong to the mitochondrial anion carrier family (MCF) (for reviews, see Rousset et al.[1] and Kunji[2]). Uncoupling protein 1 (UCP1) is exclusively found in mitochondria of brown adipocytes of mammals, where it increases the inner mitochondrial membrane permeability to protons. As a consequence, the energy from substrate oxidation stored in the electrochemical proton gradient does not promote ATP synthesis, but is dissipated. Both in vivo[3] and in vitro experiments[4−7] reflect the proton transport activity of UCP1 activated by free fatty acids and inhibited by purine nucleotides.[8] The mechanism of proton transport is, however, controversial. Several MCFs are able to perform free-fatty-acid cycling through the inner mitochondrial membrane by facilitating the pathway of fatty acids in their charged form from the internal to the external leaflet of the membrane (see Table 2 in Rial et al.[9]). Yet, the physiological relevance of this mechanism is still contentious, not only for UCP1, but also for UCP2 and UCP3, which are clearly not physiological uncouplers.[10,11] Patch-clamp measurements on isolated brown fat mitochondria have recently revealed that UCP1 is a fatty acid/proton symporter with an aborted transport for long-chain fatty acids.[12] This evidence is strongly in favor of a second model for UCP transport mechanism, in which fatty acids play a catalytic function.[13,14] In contrast to UCP1, the physiological function of UCP2 is linked with reactive oxygen species regulation in immune cells and neurons.[15−17] At the cellular level, UCP2 promotes fatty acid oxidation without detectable uncoupling of respiration, and its metabolite transport activity is still a matter of debate (for review, see Bouillaud[18]).

However challenging, structural investigation of the MCF is crucial for the understanding of transport mechanisms and regulation. Several strategies have been developed to overcome the low natural abundance of most of the carriers. Kunji and collaborators succeeded in producing the ADP/ATP carrier (AAC) from yeast mitochondria and in obtaining a two-dimensional projection map of the carrier.[20] A high-resolution structure of bovine AAC was obtained by three-dimensional crystallization of the protein isolated from natural sources.[21] The presence of a conserved motif repeated three times along the sequence between members of the MCF suggest a common

Received: July 19, 2013
Published: September 10, 2013

dx.doi.org/10.1021/ja407424v | J. Am. Chem. Soc. 2013, 135, 15174−15182

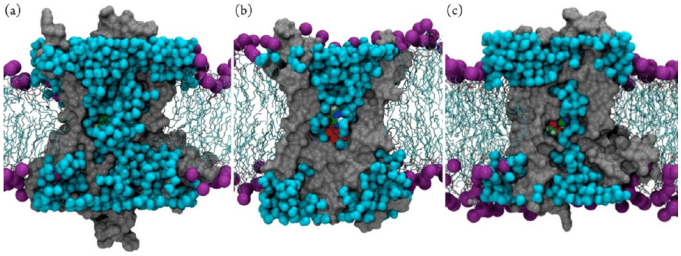


Figure 1. The UCP2 NMR structure acts as a large transmembrane channel. (a) Cross-sectional view of UCP2-GDP embedded in a POPC bilayer. Backbone atoms are constrained to their original positions in pdb:2LCK. A large water channel spanning the entire protein connects the two sides of the membrane. (b) Cross-sectional view of AAC-ADP^{3-} embedded in a POPC bilayer.[19] The protein is fully occluded on the matrix side and prevents water from crossing the bilayer. (c) Cross-sectional view of UCP2-GDP^{3-} embedded in a POPC bilayer after full relaxation of the entire protein. The large water channel observed in the backbone-constrained simulation collapses rapidly, leading to a markedly reduced internal cavity, occluded toward the intermembrane space. The protein surface is colored in gray. Violet and cyan spheres represent phospholipid headgroups and water molecules, respectively. Phosphorus, nitrogen, oxygen, and carbon atoms of GDP^{3-} and ADP^{3-} are colored in brown, blue, red, and green, respectively.

three-dimensional fold. This assumption was strengthened by the low-resolution structure of UCP2 obtained recently by a cutting-edge strategy based on solution-state NMR and molecular fragment replacement.[22] Even armed with such an advanced methodology, solving membrane protein structures remains challenging, and requires detergent to extract these biological macromolecules from their native environments. How different the structure of the membrane protein in detergents is from its membrane-bound state remains difficult to predict. Previous studies suggest that the replacement of natural lipids by detergents could modify to different extents the structure and/or the dynamics of membrane proteins.[23−30] The question of whether or not the detergents may favor non-native conformations, which in turn would alter the protein function, has remained hitherto elusive.

The comparison of the UCP2 structure with that of AAC indeed gives rise to some concerns related to structural differences possibly due to the surrounding detergent. The structures of AAC and UCP2 were solved in the presence of zwitterionic detergents, LAPAO and DPC respectively, and in the presence of inhibitors, namely carboxyatractyloside for AAC and GDP^{3-} for UCP2. Yet, if the structure of AAC is admittedly very compact, that of UCP2 appears looser and features unusual cavities between transmembrane spans with polar residues exposed at the surface of the hydrophobic region of the protein. Furthermore, the relative organization of glycine residues obliterates their participation in the helix packing, in stark contrast with AAC.[31] Given the unexpected NMR model of UCP2, we endeavored to investigate the conformational stability of UCP2 both in detergent and in a lipid bilayer by means of molecular dynamics (MD) simulations. The MD approach is particularly suited to unveil the dynamics of chemical and biological systems at the atomistic level. It is widely used to investigate the behavior of membrane proteins in a lipid environment and, hence, can complement structural data obtained in detergent solution. MD was recently employed to study the mitochondrial AAC, confirming that the structure solved in the presence of a specific inhibitor is also capable of binding spontaneously its endogenous nucleotide substrates.[19,32,33] Here, we show that UCP2 behaves as a pore in DPC, but rapidly collapses in a lipid environment, suggesting that the structure of UCP2 in DPC does not reflect the membrane-bound, functional conformation. To rationalize this finding, we tested the effect of DPC on UCP1 and UCP2 activities after purification from native and recombinant sources, respectively. Our results indicate that DPC is not a suitable detergent to maintain the activity of either UCP1 or UCP2, and suggest that it should be avoided in structural studies of mitochondrial uncoupling proteins.

■ RESULTS

Molecular Dynamics Simulation of UCP2 with Its Backbone Constrained to the NMR Structure in a Membrane-like Environment. The NMR structure of UCP2[22] (successively with and without GDP^{3-}) was embedded in a bilayer of pure 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidyl-choline (POPC). After proper equilibration of the lipids, water and GDP^{3-}, the side chains were progressively allowed to relax for 10 ns while the backbone atoms were kept fixed to their NMR positions (see Methods). The dynamics of water and the topology of the internal cavity of the protein were analyzed in the subsequent 10 ns (Figure 1a). The results were compared to the data accrued previously for AAC in the same environment[19] (Figure 1b). In its NMR structure, UCP2 mainly consists of a large water channel that spans the entire protein and connects both sides of the membrane, irrespective of the presence of GDP^{3-} (see Supporting Information, Figure S1). In the narrowest region of the protein, the effective radius of the internal cavity of the spatially constrained UCP2 is about 6.1 Å, at variance with AAC, which is fully occluded over approximately 10 Å on the matrix side[19,34] (see Figure 2). This large cavity radius resembles the size of constriction regions in pore-forming proteins, e.g., α-hemolysin,[35] and in bacterial outer-membrane proteins,[36] which convey in a nonselective fashion a variety of substrates across the membrane. To reinforce this analysis and exclude a putative intermediary transport state of UCP2, the osmotic permeability (P_f) for water of the constrained UCP2 was estimated following the method of Zhu et al.[37] For constrained UCP2 in the absence of GDP^{3-}, $P_f = (0.53 \pm 0.02) \times 10^{-12}$ cm^3/s, while it is $(0.32 \pm 0.02) \times 10^{-12}$ cm^3/s for constrained UCP2 binding GDP^{3-}. This value is only 4 times smaller than the corresponding quantity calculated for α-hemolysin, 1.9×10^{-12} cm^3/s.[35] The qualitative behavior of UCP2 NMR structure is, therefore,

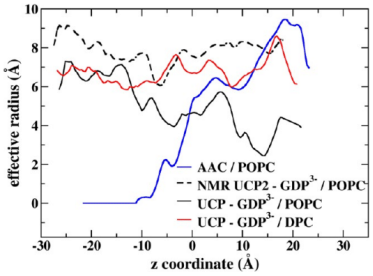


Figure 2. Effective radii of UCP2 and AAC internal cavities for (i) UCP2 binding GDP^{3-} embedded in POPC with the backbone frozen in its NMR conformation by means of holonomic constraints (dashed black line), (ii) UCP2 binding GDP^{3-} embedded in POPC with the backbone fully relaxed (black line), (iii) UCP2 binding GDP^{3-} embedded in DPC with the backbone fully relaxed (red line), and (iv) AAC embedded in POPC with the backbone fully relaxed (blue line), plotted as a function of the distance from the center of mass of the protein along the z axis.

closer to that of a porin than that of an occluded transmembrane protein like AAC, for which P_f is formally zero.

Molecular Dynamics Simulation of UCP2 Fully Relaxed in a Membrane-like Environment. For UCP2 associated to GDP^{3-}, the MD trajectory was extended further up to 175 ns without any positional holonomic constraint (a function that depends only on the atomic coordinates and the time, Figure 1c).

In sharp contrast with the majority of membrane proteins studied by MD, UCP2 quickly moves away from its initial reference structure as reflected in a large root-mean-square-deviation (RMSD) of the backbone atoms of about 6 Å (see Figure S2). Moreover, the decrease of both the radius of gyration (R_g) and the solvent-accessible surface area (SASA) measured in pdb:2LCK from 22.9 Å and 22 900 $Å^2$ to mean values of 20.9 Å and 16 300 $Å^2$, respectively, is representative of a collapse of UCP2 toward a more compact structure (see Figure 3b,c). This alteration is illustrated further by a pronounced decrease of the effective radius of the internal cavity (see Figure 2). The overall volume of the cavity is reduced by 60%, and, in its narrowest region, the effective radius is 2.4 Å. Although the fully relaxed structure appears much more occluded than the NMR reference, the water osmotic permeability still remains appreciable, $P_f = (0.11 \pm 0.01) \times 10^{-12}$ cm^3/s, suggesting that the collapsed structure is unlikely to correspond to the actual membrane-bound structure.

The UCP2 NMR structure is organized in a bundle of six integral helices, but unlike AAC,[21] packing of this bundle appears suboptimal, presenting numerous cavities between transmembrane spans (see Figure 3a). Their surface exposes not only hydrophobic residues to the core of the bilayer but also many polar amino acids (see Figure S3). In the constrained-backbone simulation, lipid tails fill interhelix cavities. When the constraints are eliminated, lipids rapidly move away from these interstices, allowing the helices to couple tightly, which in turn lowers the exposure of polar residues toward the membrane hydrophobic core and is responsible for the collapse of the initial NMR structure. It should be emphasized that brute-force MD simulations, due to current computational limitations, cannot capture slow processes like folding of large proteins. Our data, therefore, only highlight that the UCP2 NMR structure is not stable in a lipid bilayer environment (see Figure 1c and Figure S4). In the following section, we test the conclusions from the MD simulations by measuring the activities of UCP1 and UCP2 in liposomes after purification in DPC.

Loss of Activities of UCP1 and UCP2 after Being Purified in DPC. There is no abundant natural source of UCP2. Therefore native UCP1 was selected as a reference of native mitochondrial uncoupling protein. UCP1 is closely related to UCP2 with 57% of sequence identity. The effect of DPC was compared with that of Triton X-100, a relatively mild detergent known to preserve the activity of mitochondrial carriers.[38,39] Native UCP1 from brown adipose tissue (BAT) of mice was solubilized and purified in Triton X-100 and in DPC while recombinant UCP2 overexpressed in *E. coli* was solubilized from a pellet containing bacterial membranes and cell fragments and then purified in DPC (see Figure 4a). The proton-transport activity of UCP1 was measured after reconstitution in liposomes loaded with KCl. The ΔpH formed after addition of nigericin was converted into $\Delta\Psi$ when proton transport occurred through UCP1 (see Figure 4b). The electric potential was measured by recording the absorbance at 520 nm of the safranine O, which aggregates as a function of the membrane potential.[6] When purified in Triton X-100, proton-transport activity of UCP1 was strongly enhanced by the presence of 30 μM lauric acid and fully inhibited by addition of 50 μM GDP^{3-} (see Figure 5). In contrast, when UCP1 was purified in DPC, lauric acid had no significant activating effect on proton transport, and, as a consequence, no eventual inhibitory effect of GDP^{3-} could be measured. Next, the free-fatty-acid-dependent protonophoric activity of UCP2 was investigated. No proton-transport activity was observed when the protein was purified in DPC. Altogether, our results suggest that exposure to DPC severely compromises the proton-transport activity of both UCP1 and UCP2. In order to explain the irreversible inactivating effect of DPC, MD simulations in this detergent were performed on the structure of UCP2 obtained by NMR.

Molecular Dynamics Simulation of UCP2 Fully Relaxed in DPC. The interaction of DPC with the NMR structure of UCP2 was investigated by means of MD simulations. Bond et al.[25,40] have shown that DPC aggregation time around membrane proteins is short enough to be described by MD. Using all-atom and coarse-grained simulations, they were able to reproduce the self-assembly of detergents into a ring-shape bundle around the hydrophobic core of gramicidin A and OmpF, a common feature observed with membrane proteins.[41,42] All-atom and coarse-grained MD simulations were carried out starting from randomly disperse solutions of DPC surrounding a single UCP2 protein constrained to its NMR structure (see Supplementary Methods). Qualitatively, the ring-shape bundle of DPC systematically forms around UCP2 on a time scale consistent with previous studies.[25,40] Interestingly enough, protein amphiphilic segments exposed to the solvent on the cytoplasmic (residues 298−308) and the matrix (residues 29−59) sides catalyze the formation of two additional detergent micelles (see Figure S5).

Based on the information yielded by the MD simulations, an initial model was built, in which all monitored DPC-protein interactions were, whenever possible, integrated. It consists of a

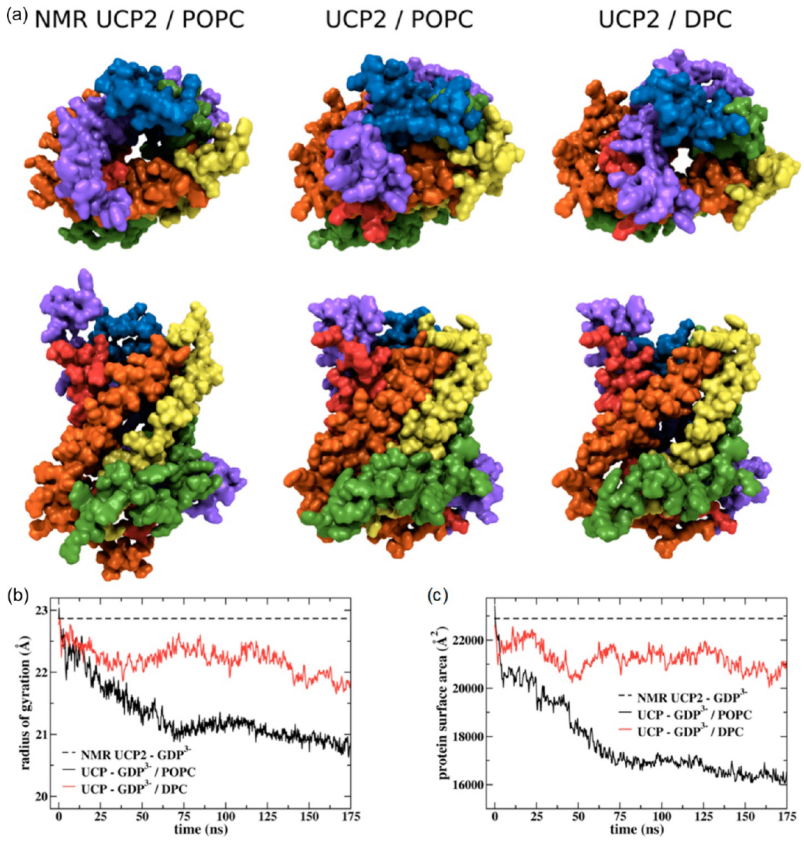


Figure 3. (a) Top and side views of three UCP2-GDP^{3-} complexes from simulations beginning with the pdb:2LCK structure. (Left) UCP2-GDP^{3-} in a POPC bilayer after 26 ns with its backbone holonomically constrained to the PDB structure. (Middle) UCP2-GDP^{3-} in a POPC bilayer after 175 ns, bereft of holonomic constraints or harmonic restraints. (Right) UCP2-GDP in the presence of 300 DPC detergent molecules after 175 ns, bereft of holonomic constraints or harmonic restraints. The top and bottom images show UCP2 from above the bilayer and in the plane of the bilayer, respectively. The protein DPC is aligned consistently with the views obtained for the constrained complex. For clarity, ions and water molecules, lipids, and detergents are not shown. (b) Radius of gyration of the UCP2 backbone atoms as a function of time for each one of the three systems. The constant value for the system where the backbone is constrained to that of the PDB structure is shown as a dashed black horizontal line. (c) Surface area of the protein as a function of time for each one of the three systems. The constant value for the system where the backbone is constrained to that of the PDB structure is shown as a dashed black horizontal line. In (a), each helix is colored differently.

200-μM solution of detergent, composed of 300 DPC molecules lying on the surface and in the internal cavity of UCP2. After proper equilibration of DPC around the constrained protein, the entire system was fully relaxed for 175 ns. The RMSD of backbone-atom positions converges toward a value of about 4.5 Å, corresponding to a lesser degree of reorganization than observed in the POPC simulation (see Figure S2). The topology of the overall NMR UCP2 structure is retained, as suggested by the effective radius of the internal cavity, in line with the NMR conformation (see Figure 2). Similarly, the water osmotic permeability, $(0.32 \pm 0.01) \times 10^{-12}$ cm^3/s, agrees nicely with the value calculated for UCP2 with backbone atoms constrained to their NMR positions. The evolution of R_g and SASA further confirms that in MD simulations in DPC, UCP2 retains a structure akin to the NMR one (see Figure 3). Interstitial spaces between helices are stabilized not only by the hydrophobic tails of DPC, but also by its polar head groups. The amphiphilic segments lying on both sides of the protein are steadily encapsulated into micelles of a size compatible with their solution counterpart[43] (see Figure 6). The small length of the DPC chain combined with its flexibility allows for precise targeting of every amphiphilic patches lying all over the exposed surface of the protein. About three DPC molecules are trapped at the edges of the internal cavity, in regions wide enough to not significantly affect the water osmotic permeability. Continuity of the ring-shape bundle of detergents around the protein is interrupted (see Figure 6). This disruption stems from a groove of polar

dx.doi.org/10.1021/ja407424v | J. Am. Chem. Soc. 2013, 135, 15174−15182

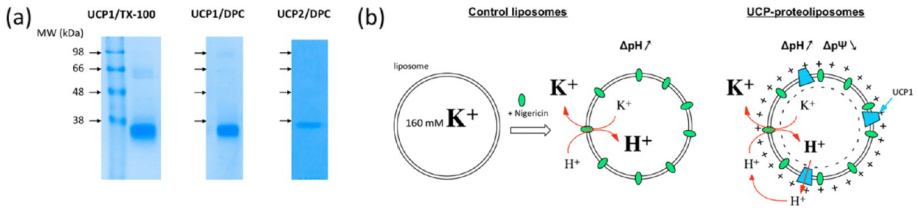


Figure 4. UCP1 and UCP2 samples after purification. (a) SDS-PAGE of UCP1 purified from mitochondria of brown adipocytes of mice in either Triton X-100 or DPC and mouse UCP2 overexpressed in *E. coli* and purified in DPC (see the Methods section). (b) Schematic representation of the liposomes assay setup. Liposomes are loaded with 160 mM potassium ions. The antibiotic nigericin is a potassium/proton ionophore. In the absence of UCP (control), a ΔpH is generated by exchanging potassium ions against protons. However, no signal is detected because the colorimetric probe (safranine O) present in solution is not pH-sensitive. In the presence of UCP1 (sample), the protons accumulated inside the vesicles are transported outside following the protons gradient created by nigericin. The ΔpH is converted into a ΔΨ, which is detected by the membrane potential-sensitive probe safranine O (for further details see Mozo et al.[6]).

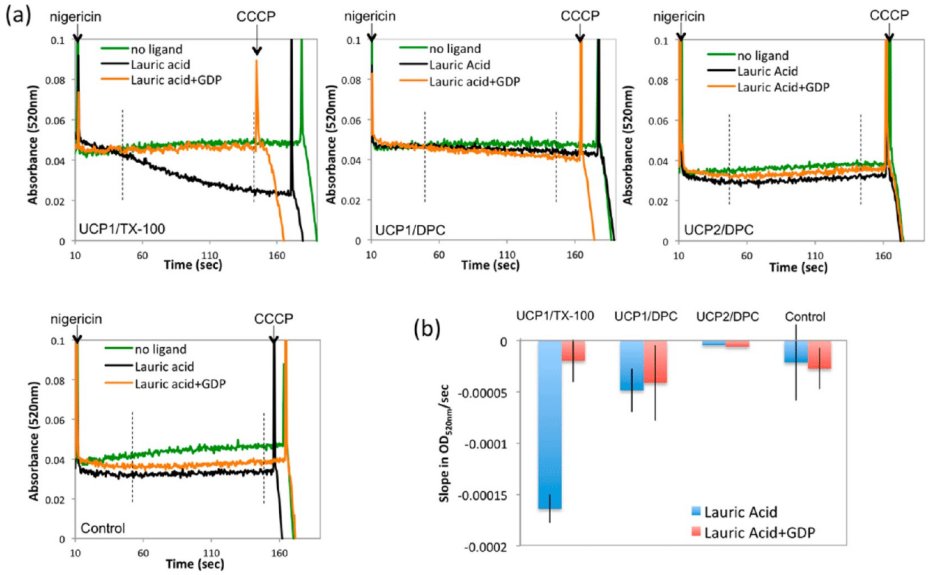


Figure 5. Functional assay of UCP1 and UCP2 after reconstitution in liposomes. (a) Kinetics curves of the safranine O polarization monitored at 520 nm. The arrows indicate the addition of nigericin and CCCP (the addition of ligands being done before nigericin). The delimited time scale for the slope determination is indicated on the graphs by dashed vertical lines. (b) Histogram reflecting the proton transport activity of UCP1 and UCP2. The bars show the differences in polarization slopes of the safranine O after subtraction of the background, i.e., no ligand. Activation of the proton transport was monitored in the presence of lauric acid whereas inhibition was monitored in the presence of both lauric acid and GDP (for details see the Methods section).

residues distributed along helix 2 (residues 76−110) and exposed toward the detergent bundle (see Figure S3).

■ DISCUSSION

Painting a detailed picture of mitochondrial communication relies on high-resolution structural information, the production of which remains a daunting challenge owing to the considerable difficulty to express and purify mitochondrial carriers, and ultimately determine the structure, either by crystallography or by NMR. Associating NMR data with molecular-fragment replacement, Berardi et al.[22] succeeded in obtaining a three-dimensional structure of an uncoupling protein, UCP2, binding an inhibitor, GDP^{3-}, solubilized by DPC (pdb:2LCK). The UCP2 structure surrounded by DPC molecules is stable in our MD simulations, suggesting that their model is relevant in this detergent.

Moreover, their study suggests that the structure of UCP2 closely resembles that of AAC. In its crystallographic structure, AAC is in a conformation open toward the intermembrane space, fully occluded on the matrix side and impervious to water or any other small substrate.[19,21,44] The MD simulations performed here enlighten, however, striking differences in the

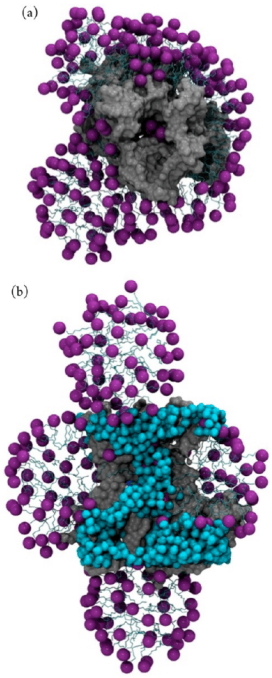


Figure 6. (a). Top view of UCP2-GDP^{3-} embedded in DPC micelles after full relaxation of the entire protein. For clarity, the DPC micelle lying on the cytoplasmic side is not shown. Violet and cyan spheres represent phospholipid headgroups and water molecules, respectively. Phosphorus, nitrogen, oxygen, and carbon of nucleotides are colored in brown, blue, red, and green, respectively. (b) Cross-sectional view of UCP2-GDP^{3-} embedded in DPC micelles after full relaxation of the entire protein. Detergent molecules are organized in a bundle around the hydrophobic core of the protein. Two extra micelles assemble on the matrix and cytoplasmic sides around amphiphilic patches of amino acids. The internal cavity of the protein is fully opened, a conformation akin to that observed in the backbone-constrained simulation

structure of the two proteins. Specifically, our results show that UCP2, wherein the backbone atoms are holonomically constrained, primarily consists of a large water channel connecting the two sides of a POPC membrane, irrespective of the presence of GDP^{3-}. This structural feature is shared by nonselective porins, but not by membrane transporters, and, hence, does not support an alternate access mechanism, in which the conformation of the transporting protein sequentially opens on each side of the membrane. The effective radius of the internal channel of UCP2, ∼6.1 Å, clearly compares to that of toxins or bacterial outer-membrane porins but is much larger than the one measured for aquaporins, a family of membrane channels that selectively conduct water while preventing proton permeation. The continuous water flow across UCP2 is thus likely to allow for proton leakage, a situation that could hardly preserve the proton gradient across the inner membrane of the mitochondria, an absolute requirement for the survival of the organelle. When performed without positional holonomic constraints, the molecular simulations revealed that the NMR structure of UCP2 is not stable in a lipid bilayer environment. The hydrophobic tails of POPC can barely stabilize the cavities found between transmembrane helices and lined with polar and hydrophobic residues. Consequently, the overall tertiary structure rapidly collapses toward a compact arrangement, which remains permeable to water. In contrast, MD simulations of the self-assembly of DPC around UCP2 demonstrate that small detergents can lodge themselves in these cavities and, hence, stabilize the overall NMR protein fold. The thermodynamic parameters of the interaction between two transmembrane helical spans can be affected by DPC as previously reported,[45] indicating that DPC does not behaves as an ideal membrane mimetic environment for some membrane proteins. It is noteworthy that, in spite of what appears to be a partial denaturation of UCP2 in detergent, the network of interactions formed by the protein and its inhibitor is preserved (see Figure S6). This observation is consistent with experimental data based on FRET experiments[22] and atomic force microscopy[46] suggesting that the binding site of UCP2 purified in DPC is preserved. However, retaining GDP^{3-} binding is not sufficient to prove that the protein is fully functional. While it is clearly demonstrated that the protein binds GDP^{3-}, it is difficult to draw definitive conclusions about its proton transport activity for several reasons. For example, if the concentration of activator ligand is above the usual concentration of fatty-acid required either to observe UCP1 activity[4] or to trigger free fatty-acid-mediated proton leakage on mitochondria,[9,47] an unspecific proton leakage may occur through the vesicle walls as exemplified in our proton transport assay (see Figure S7). Reproducing the purification procedure of the recombinant UCP2 in DPC, our functional assay showed no proton transport activity by UCP2. Several studies, performed on the better-characterized native UCP1, showed that the protein tolerates a limited variety of detergents, mainly Triton X-100 and some polyoxyethylene glycols, but not octylglucoside.[5] UCP1 purified in Triton X-100 was fully regulated as expected,[6] while purification of the protein in DPC abolished the activity of UCP1 and its regulation.

Our work highlights how critically important the proper choice of detergent is for membrane protein structure determination. Indeed, while detergents are useful for the solubilization of membrane proteins, they can compete with stabilizing intramolecular interactions within the protein, leading to inactivation.[48,49] The severity of the effect strongly depends on the nature of the membrane proteins, leading one to distinguish between "fragile" and "robust" proteins, which are generally α-helical and β-barrel proteins, and also, eukaryotic and prokaryotic proteins, respectively. As a general rule, the robustness of a membrane protein is correlated to the compactness of interaction networks, which generally increases with the living temperature of the organisms. This feature has created a bias in favor of structural determination of robust membrane proteins. Furthermore, designing a general rule for choosing the detergent suitable for fragile membrane proteins is difficult. Two molecular mechanisms of inactivation of membrane proteins by detergents can be considered. One is the presence of micelles acting as a hydrophobic sink into which stabilizing interacting partners such as lipids, cofactors or protein subunits can disperse, thereby leading to the loss of the protein integrity. Working close to the critical micellar concentration of the detergent can limit this effect. The second mechanism of destabilization, which is the most likely to occur

in the case of UCPs, is the intrusion of the hydrophobic tails of the detergents between the transmembrane helices. This mechanism can be restrained using a more rigid detergent. For example, novel DDM analogues with multiple, branched, or cyclic hydrophobic moieties presenting less flexibility than DDM have been shown to improve the stability of several membrane proteins.[50−52] Many other alternatives to replace detergents have been developed, among which are fluorinated detergents, nanodiscs, and amphipols (for review see Popot[53]). All these innovative surfactants improve the stability of membrane proteins, but require a previous solubilization step with conventional detergents. Finally, it should be stressed that genetic engineering of membrane proteins has been highly successful in improving the thermostability of G protein-coupled receptors.[54] This strategy might also benefit other classes of membrane proteins, including the mitochondrial carriers.

■ CONCLUSION

While acquiring precise structural information is crucial for a full understanding of the biological processes like transport mechanisms of mitochondrial carriers, assessing the physiological relevance of the newly solved three-dimensional structures is equally crucial. A recent study revealed indeed that detergents can cause structural perturbations and distortions leading to unsuitable interpretations as to the molecular mechanisms of some membrane proteins.[27] As exemplified in this study, a combination of theoretical and functional investigations represents an appealing and promising strategy, the frontiers of which go beyond the family of mitochondrial transporters to embrace most membrane proteins. This work further emphasizes that molecular simulations not only can be employed to rationalize and predict the properties of membrane proteins near their equilibrium conformation, but can also address the physiological relevance of structures obtained in non-native environments. Our data strongly suggest that the structure of UCP2 in DPC is not physiologically relevant and, therefore, that it may be misleading to use it as a basis for drawing inferences about the physiological behavior of members of the MCF. For membrane proteins like UCP2—which expose not only nonpolar contacts to the hydrophobic core of the membrane, but also a number of weakly polar contacts—it is preferable to avoid small, flexible detergents, which can easily target these two types of interactions simultaneously.

■ METHODS

Purification of UCP1 and UCP2. BAT was kindly provided by V. Lenoir (ICGM), Paris, France from mice cold-adapted for 2 days. The purification of UCP1 was carried out as described by Lin and Klingenberg.[38] Mitochondria were isolated from BAT and washed out with 3.2% Lubrol WX in buffer 20 mM MOPS, 20 mM Na_2SO_4, 0.16 mM EDTA, pH 6.9. The mitochondrial membranes were split in two. The total mitochondrial proteins were solubilized using either Triton X-100 with a w/w protein:detergent ratio of 1:1.6 or DPC with a ratio 1:1. The extracts were centrifuged and then applied to a hydroxyapatite column (BioRad). The flow-through fractions containing UCP1 were collected, and the purity was controlled by SDS-PAGE with 12% acrylamide gels (Invitrogen). The production of UCP2 was adapted from Berardi et al.[22] The cDNA of mouse UCP2 encoding for residues 14−309 was cloned into a pET-22b vector between the restriction sites *Nde*I and *Eco*RI. The protein was overexpressed in *E. coli* Rosetta λ(DE3). The concentration of total proteins present in the 50000*g* pellet was determined by a BCA assay. UCP2 was solubilized in DPC with a w/w protein:detergent ratio of 1:3 in buffer 20 mM Tris, 150 mM NaCl, pH 8. After 30 min of incubation at 4 °C, the sample was centrifuged at 100000*g* for 20 min. The supernatant was then loaded onto a Ni-NTA column (Thermo Scientific) and elution occurred with 300 mM imidazole. The fraction containing UCP2 was dialyzed and then applied to an ion exchange column (GE Healthcare) for further purification. The purity of the flow-through fractions was controlled by SDS-PAGE. The band corresponding to UCP2 on the gel was cut off for further analysis by MALDI-TOF mass spectrometry, confirming the expected UCP2 protein.

Functional Assay. The functional assay was carried out as described in Mozo et al.[6] Small unilamellar vesicles composed by phosphatidylcholine, phosphatidylethanolamine, and cardiolipin in a w/w/w ratio of 80:80:10 were prepared by sonication in buffer 20 mM KH_2PO_4, 70 mM K_2SO_4, pH 6.9, at a final concentration of 10 mg/mL. The proteoliposomes were prepared by mixing UCP samples, vesicles, and buffer in a v/v/v ratio of 1:8:4 and incubated for 1 h at room temperature under shaking. For UCP purified in DPC, the protein incorporation in liposomes was facilitated by supplying Triton X-100 (2%) to the sample. The detergent was removed by adsorption onto biobeads (BioRad), and the liposomes where then dialyzed three times for 1 h against a buffer containing 200 mM sucrose, 0.5 mM Hepes, and 0.5 mM EDTA, pH 6.8. For each measure, an aliquot of the liposomes suspension (100 μL) was mixed with 2 mL of 100 mM choline choride, 20 μM safranine O. The absorption of safranine O was recorded at 520 nm with a UV−visible spectrophotometer. All ligands are added first in order to avoid any perturbation of the absorbance baseline. The activation experiments were recorded by supplying lauric acid (30 μM) in the cuvette, while for inhibition experiments both lauric acid (30 μM) and GDP^{3-} (50 μM) were added. The proton transport was activated upon addition of nigericin (0.1 μM), which generates the pH gradient. At the end of the kinetics, CCCP (5 μM) was added to measure the complete polarization of liposomes, giving an indication of the internal volume of the liposomes. The slopes of the curves were determined between 50 and 150 s, and the background of the basal activity (in the absence of ligand) was subtracted in order to observe the ligand effects only. The activity measurements for UCP1 were repeated from three independent purifications and then averaged.

All-Atom Molecular Dynamics. Three different molecular assays were built based on the UCP2 NMR structure. A first setup was obtained by inserting the protein into a fully hydrated bilayer of 146 POPC lipids, using the membrane builder module of CHARMM-GUI (http://www.charmm-gui.org/).[55] Docking of a GDP^{3-} anion inside UCP2 internal cavity, at a position consistent with the NMR prediction, provided a second initial configuration. Finally, we performed several all-atom and coarse-grained MD simulations (see Supplementary Methods) to study self-aggregation of DPC at 200 μM around the protein at various detergent:protein ratios. We superimposed the final configuration of all trajectories and built an all-atom model, including GDP^{3-}, and integrating all observed DPC−protein interactions, for a total of 300 DPC molecules. Respectively 14 and 17 chloride anions were added to setups with and without GDP^{3-} to ensure electric neutrality. MD simulations were carried out with the NAMD[56] program. Setups and analyses were performed using VMD.[57] All trajectories were generated in the isobaric−isothermal ensemble, at 300 K under 1 atm using, respectively, Langevin dynamics[58] (damping coefficient, 1 ps^{-1}) and the Langevin piston[59] method. The particle mesh Ewald (PME) algorithm[60] was used to account for long-range electrostatic interactions. Covalent bonds involving hydrogen atoms were constrained to their equilibrium length by means of the Rattle algorithm.[53] The equations of motion were integrated by means of a multiple-time step algorithm[62] with a time step of 2 and 4 fs for short- and long-range interactions, respectively. The CHARMM force field[63−65] including CMAP[66] corrections was employed to model proteins, POPC lipids or DPC detergents, GDP^{3-}, and counterions; the TIP3P[67] model was used to describe water. Water, lipids or detergents, substrate, and counterions were first thermalized for 20 ns, while the entire protein was kept fixed. For the next 20 ns, the constraints on the side chains were removed, and only the backbone

was constrained to its NMR conformation. For UCP2 associated to GDP^{3-}, simulations in POPC and DPC were extended further up to 175 ns without any positional constraints.

Permeability Methods. To calculate the osmotic permeability of the UCP2 structures, we followed the algorithm described by Aksimentiev and Schulten.[35] The pore of the structure given by NMR had a more complex topology and structure than the α-hemolysin channel; we, therefore, calculated the permeability only over the central region of the pore where the topology was simpler and relatively stable among the simulations. The region was delineated by two approximately coplanar rings of Cα atoms, with one atom chosen from each of the six transmembrane helices. Let R_0 and R_1 be the centers of mass of the bottom and top rings, respectively, consisting of the Cα atoms of residues 34, 85, 137, 181, 239, and 274, and residues 20, 101, 120, 194, 227, and 288. The region considered had the form of a cylinder with a base at R_0 and axis lying along the vector $R_1 - R_0$. The length of the cylinder along the axis was L = 19.5 Å, equivalent to the distance $|R_1 - R_0|$ for the NMR structure. The radius of the cylinder was chosen to be 20 Å, which was large enough to encompass all water molecules between the two rings. The collective displacement of the water molecules within the protein at time $t + \Delta t$ of the simulation trajectory was calculated as

$$n(t + \Delta t) = n(t) + \sum_{i \in S(t,t+\Delta t)} \frac{\Delta r_i \cdot e}{L}$$

where $S(t,t + \Delta t)$ is the union of subsets of water molecules within the cylinder at time t and $t + \Delta t$, Δr_i is the displacement of water molecule i between time t and $t + \Delta t$, and e is the unit vector along $R_1 - R_0$. The displacements of water molecules entering or leaving the region between the frames were truncated at the boundary of the region. The calculation was performed with Δt = 10 ps. The collective diffusion coefficient, D_n, was obtained from $\langle n(t) \rangle^2 = 2D_n t$, where the average was performed over 100 subtrajectories each 10 ps in length. The permeability was calculated as $P_f = v_w D_n$, where v_w was the average volume of a water molecule. Further, following Aksimentiev and Schulten, the resulting permeabilities were scaled by 1/2.87 to account for the difference in the viscosity of the TIP3P water model and real water.[35]

■ ASSOCIATED CONTENT

Supporting Information

Methodological details and supplementary experimental and theoretical data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

bruno.miroux@ibpc.fr
francois.dehez@univ-lorraine.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank O. Ilioaia for helping in bacterial membrane preparation and V. Lenoir for providing BAT from cold-adapted mice. D. Warchawski, L. Catoire, K. Moncoq, and D. Picot are acknowledged for helpful discussions. We thank J.-L. Popot for insightful comments on the manuscript. This work was supported by the Agence National de La Recherche (ANR MIT-2M, No. ANR 2010 BLAN1518), the "Initiative d'excellence" program from the French State (DYNAMO, ANR-11-LABX-0011-01), and the Direction Régionale à la Recherche et à la Technologie de Lorraine. The Grand Equipement National de Calcul Informatique (GENCI) and the Centre Informatique National de l'Enseignement Supérieur (CINES) are gratefully acknowledged for the provision of computer time.

■ REFERENCES

(1) Rousset, S.; Alves-Guerra, M.-C.; Mozo, J.; Miroux, B.; Cassard-Doulcier, A.-M.; Bouillaud, F.; Ricquier, D. *Diabetes* **2004**, *53*, 130S−135.
(2) Kunji, E. R. S. *FEBS Lett.* **2004**, *564*, 239−244.
(3) Enerbäck, S.; Jacobsson, A.; Simpson, E. M.; Guerra, C.; Yamashita, H.; Harper, M.-E.; Kozak, L. P. *Nature* **1997**, *387*, 90−94.
(4) Winkler, E.; Klingenberg, M. *J. Biol. Chem.* **1994**, *269*, 2508−2515.
(5) Winkler, E.; Klingenberg, M. *Eur. J. Biochem.* **1992**, *207*, 135−145.
(6) Mozo, J.; Ferry, G.; Massche leyn, S.; Miroux, B.; Boutin, J. A.; Bouillaud, F. *Anal. Biochem.* **2006**, *351*, 201−206.
(7) Blesneac, I.; Ravaud, S.; Machillot, P.; Zoonens, M.; Masscheylen, S.; Miroux, B.; Vivaudou, M.; Pebay-Peyroula, E. *Eur. Biophys. J.* **2012**, *41*, 675−679.
(8) Rial, E.; Poustie, A.; Nicholls, D. G. *Eur. J. Biochem.* **1983**, *137*, 197−203.
(9) Rial, E.; Aguirregoitia, E.; Jiménez-Jiménez, J.; Ledesma, A. *Biochim. Biophys. Acta BBA—Bioenerg.* **2004**, *1608*, 122−130.
(10) Mozo, J.; Ferry, G.; Studeny, A.; Pecqueur, C.; Rodriguez, M.; Boutin, J. A.; Bouillaud, F. *Biochem. J.* **2006**, *393*, 431−439.
(11) Pecqueur, C.; Bui, T.; Gelly, C.; Hauchard, J.; Barbot, C.; Bouillaud, F.; Ricquier, D.; Miroux, B.; Thompson, C. B. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2008**, *22*, 9−18.
(12) Fedorenko, A.; Lishko, P. V.; Kirichok, Y. *Cell* **2012**, *151*, 400−413.
(13) Echtay, K. S.; Winkler, E.; Frischmuth, K.; Klingenberg, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1416−1421.
(14) González-Barroso, M. M.; Fleury, C.; Bouillaud, F.; Nicholls, D. G.; Rial, E. *J. Biol. Chem.* **1998**, *273*, 15528−15532.
(15) Emre, Y.; Hurtaud, C.; Karaca, M.; Nubel, T.; Zavala, F.; Ricquier, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19085−19090.
(16) Mattiasson, G.; Shamloo, M.; Gido, G.; Mathi, K.; Tomasevic, G.; Yi, S.; Warden, C. H.; Castilho, R. F.; Melcher, T.; Gonzalez-Zulueta, M.; Nikolich, K.; Wieloch, T. *Nat. Med.* **2003**, *9*, 1062−1068.
(17) Vogler, S.; Pahnke, J.; Rousset, S.; Ricquier, D.; Moch, H.; Miroux, B.; Ibrahim, S. M. *Am. J. Pathol.* **2006**, *168*, 1570−1575.
(18) Bouillaud, F. *BBA-Bioenerg.* **2009**, *1787*, 377−383.
(19) Dehez, F.; Pebay-Peyroula, E.; Chipot, C. *J. Am. Chem. Soc.* **2008**, *130*, 12725−12733.
(20) Kunji, E. R. S.; Harding, M. *J. Biol. Chem.* **2003**, *278*, 36985−36988.
(21) Pebay-Peyroula, E.; Dahout-Gonzalez, C.; Kahn, R.; Trézéguet, V.; Lauquin, G. J.-M.; Brandolin, G. *Nature* **2003**, *426*, 39−44.
(22) Berardi, M. J.; Shih, W. M.; Harrison, S. C.; Chou, J. J. *Nature* **2011**, *476*, 109−113.
(23) Cross, T. A.; Sharma, M.; Yi, M.; Zhou, H.-X. *Trends Biochem. Sci.* **2011**, *36*, 117−125.
(24) Smith, S. O.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S. *Biochemistry (Mosc.)* **2001**, *40*, 6553−6558.
(25) Bond, P. J.; Cuthbertson, J. M.; Deol, S. S.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2004**, *126*, 15948−15949.
(26) Bond, P. J.; Faraldo-Gómez, J. D.; Deol, S. S.; Sansom, M. S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9518−9523.
(27) Zhou, H.-X.; Cross, T. A. *Annu. Rev. Biophys.* **2013**, *42*, 361−392.
(28) Hong, H.; Bowie, J. U. *J. Am. Chem. Soc.* **2011**, *133*, 11389−11398.
(29) Shi, L.; Zheng, H.; Zheng, H.; Borkowski, B. A.; Shi, D.; Gonen, T.; Jiang, Q.-X. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 3369−3374.
(30) Lu, G. J.; Tian, Y.; Vora, N.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 9299−9302.
(31) Nury, H.; Dahout-Gonzalez, C.; Trézéguet, V.; Lauquin, G. J. M.; Brandolin, G.; Pebay-Peyroula, E. *Annu. Rev. Biochem.* **2006**, *75*, 713−741.

(32) Wang, Y.; Tajkhorshid, E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9598−9603.
(33) Mifsud, J.; Ravaud, S.; Krammer, E.-M.; Chipot, C.; Kunji, E. R. S.; Pebay-Peyroula, E.; Dehez, F. *Mol. Membr. Biol.* **2013**, *30*, 160−168.
(34) Pebay-Peyroula, E.; Brandolin, G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 420−425.
(35) Aksimentiev, A.; Schulten, K. *Biophys. J.* **2005**, *88*, 3745−3761.
(36) Robertson, K. M.; Tieleman, D. P. *FEBS Lett.* **2002**, *528*, 53−57.
(37) Zhu, F.; Tajkhorshid, E.; Schulten, K. *Phys. Rev. Lett.* **2004**, *93*, 224501−224504.
(38) Lin, C. S.; Klingenberg, M. *FEBS Lett.* **1980**, *113*, 299−303.
(39) Krämer, R.; Klingenberg, M. *FEBS Lett.* **1977**, *82*, 363−367.
(40) Bond, P. J.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2006**, *128*, 2697−2704.
(41) Timmins, P.; Pebay-Peyroula, E.; Welte, W. *Biophys. Chem.* **1994**, *53*, 27−36.
(42) Pebay-Peyroula, E.; Garavito, R.; Rosenbusch, J.; Zulauf, M.; Timmins, P. *Structure* **1995**, *3*, 1051−1059.
(43) Abel, S.; Dupradeau, F.-Y.; Marchi, M. *J. Chem. Theory Comput.* **2012**, *8*, 4610−4623.
(44) Krammer, E.-M.; Ravaud, S.; Dehez, F.; Frelet-Barrand, A.; Pebay-Peyroula, E.; Chipot, C. *Biophys. J.* **2009**, *97*, L25−L27.
(45) Mineev, K. S.; Khabibullina, N. F.; Lyukmanova, E. N.; Dolgikh, D. A.; Kirpichnikov, M. P.; Arseniev, A. S. *Biochim. Biophys. Acta BBA—Biomembr.* **2011**, *1808*, 2081−2088.
(46) Zhu, R.; Rupprecht, A.; Ebner, A.; Haselgrübler, T.; Gruber, H. J.; Hinterdorfer, P.; Pohl, E. E. *J. Am. Chem. Soc.* **2013**, *135*, 3640−3646.
(47) Jaburek, M. *J. Biol. Chem.* **2003**, *278*, 25825−25831.
(48) Bowie, J. U. *Curr. Opin. Struct. Biol.* **2001**, *11*, 397−402.
(49) Garavito, R. M.; Ferguson-Miller, S. *J. Biol. Chem.* **2001**, *276*, 32403−32406.
(50) Chae, P. S.; Rasmussen, S. G. F.; Rana, R. R.; Gotfryd, K.; Chandra, R.; Goren, M. A.; Kruse, A. C.; Nurva, S.; Loland, C. J.; Pierre, Y.; Drew, D.; Popot, J.-L.; Picot, D.; Fox, B. G.; Guan, L.; Gether, U.; Byrne, B.; Kobilka, B.; Gellman, S. H. *Nat. Methods* **2010**, *7*, 1003−1008.
(51) Hong, W.-X.; Baker, K. A.; Ma, X.; Stevens, R. C.; Yeager, M.; Zhang, Q. *Langmuir* **2010**, *26*, 8690−8696.
(52) Hovers, J.; Potschies, M.; Polidori, A.; Pucci, B.; Raynal, S.; Bonneté, F.; Serrano-Vega, M. J.; Tate, C. G.; Picot, D.; Pierre, Y.; Popot, J.-L.; Nehmé, R.; Bidet, M.; Mus-Veteau, I.; Bußkamp, H.; Jung, K.-H.; Marx, A.; Timmins, P. A.; Welte, W. *Mol. Membr. Biol.* **2011**, *28*, 171−181.
(53) Popot, J.-L. *Annu. Rev. Biochem.* **2010**, *79*, 737−775.
(54) Tate, C. G.; Schertler, G. F. *Curr. Opin. Struct. Biol.* **2009**, *19*, 386−395.
(55) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. *Biophys. J.* **2009**, *97*, 50−58.
(56) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(57) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−38.
(58) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
(59) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
(60) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(61) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24−34.
(62) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.
(63) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(64) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104.
(65) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. *J. Phys. Chem. B* **2010**, *114*, 7830−7843.
(66) MacKerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.
(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

Bioorganic & Medicinal Chemistry 24 (2016) 4920–4927

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Assessing the physiological relevance of alternate architectures of the p7 protein of hepatitis C virus in different environments

Nicole Holzmann [a,b], Christophe Chipot [a,b,c,*], François Penin [d,e], François Dehez [a,b,*]

[a] Unité Mixte de Recherche No. 7565, Université de Lorraine, B.P. 70239, 54506 Vandoeuvre-lès-Nancy cedex, France
[b] Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana–Champaign, CNRS, B.P. 70239, 54506 Vandoeuvre-lès-Nancy cedex, France
[c] Department of Physics, University of Illinois at Urbana–Champaign, 1110 West Green Street, Urbana, IL 61801, USA
[d] Bases Moléculaires et Structurales des Systèmes Infectieux, IBCP, UMR 5086, Université de Lyon, 69367 Lyon, France
[e] Bases Moléculaires et Structurales des Systèmes Infectieux, IBCP, UMR 5086, CNRS, 69367 Lyon, France

ARTICLE INFO

Article history:
Received 13 June 2016
Revised 27 July 2016
Accepted 28 July 2016
Available online 29 July 2016

Keywords:
Hepatitis C virus
Viroporin p7
Molecular dynamics
Channel conductance
Detergent

ABSTRACT

The viroporin p7 of the hepatitis C virus forms multimeric channels eligible for ion transport across the endoplasmic reticulum membrane. Currently the subject of many studies and discussion, the molecular assembly of the ion channel and the structural characteristics of the p7 monomer are not yet fully understood. Structural investigation of p7 has been carried out only in detergent environments, making the interpretation of the experimental results somewhat questionable. Here, we analyze by means of molecular dynamics simulations the structure of the p7 monomer as a function of its sequence, initial conformation and environment. We investigate the conductance properties of three models of a hexameric p7 ion channel by examining ion translocation in a pure lipid bilayer. It is noteworthy that although none of the models reflects the experimentally observed trend to conduct preferentially cations, we were able to identify the position and orientation of titratable acidic or basic residues playing a crucial role in ion selectivity and in the overall conductance of the channel. In addition, too compact a packing of the monomers leads to channel collapse rather than formation of a reasonable pore, amenable to ion translocation. The present findings are envisioned to help assess the physiological relevance of p7 ion channel models consisting of multimeric structures obtained in non-native environments.

1. Introduction

According to the World Health Organization (WHO), there are globally 130–150 million people currently suffering from a chronic hepatitis C infection.[1] Hepatitis C virus (HCV) is a major cause of liver cirrhosis and liver cancer and, thus, approximately 500,000 deaths each year are related to HCV infection.[2] Although novel direct-acting antivirals targeting HCV have dramatically improved antiviral treatment options,[3] no prophylactic vaccine exists to date.

The hepatitis C virus is a member of the Flaviviradae family of linear and single-stranded RNA genomes. It is a 9.6-kb positive-strand RNA genome encoding a single polyprotein precursor that is processed by cellular and viral proteases into ten mature proteins. The structural proteins, which form the viral particle, include the core protein and the envelope glycoproteins E1 and E2. The nonstructural proteins include the p7 viroporin, the NS2 protease, the NS3-4A complex harboring protease and NTPase/ RNA helicase activities, the NS4B and NS5A proteins, and the NS5B RNA-dependent RNA polymerase.[4,5]

The p7 protein belongs to the group of viroporins since it fulfills the major characteristics of this family, including its small size of 63 amino acids and its ability to form oligomeric, hydrophobic ion channels integral to the membrane.[6] In vitro analysis revealed that p7 is essential for HCV assembly and release, for viral replication.[7] The findings of Wozniak et al. suggest that p7 acts as an intracellular proton channel to prevent cellular acidification processes and to enable the right alkalized milieu for the virus to reproduce.[8] For these reasons, the p7 viroporin has been investigated as a potential druggable target in the prospect of fighting hepatitis C infection.[6,7]

During the last years, p7 was the center of attention of numerous experimental and theoretical studies.[3–7,9–24] NMR (nuclear magnetic resonance), CD (circular dichroism) and MD (molecular dynamics) investigations of p7 monomers of different strains and

* Corresponding authors. Tel.: +33 3 83 68 40 97 (C.C.), +33 3 83 68 40 98 (F.D.).
E-mail addresses: chipot@ks.uiuc.edu (C. Chipot), Francois.Dehez@univ-lorraine.fr (F. Dehez).

http://dx.doi.org/10.1016/j.bmc.2016.07.063
0968-0896/

different constructs have concordantly described the p7 structure as a hairpin in membrane environments,[9–13] where the N- and C-termini are facing the ER lumen and the central basic loop is oriented towards the cytoplasm.[14] Oligomerization of recombinant or chemically synthetized p7 has been reported to lead to heptameric[15] or hexameric,[16] cation selective[10,25] ion channels. However, the multimeric architecture of these ion channels is currently the subject of much debate (reviewed in Ref. 17). Zitzmann and co-workers dissolved synthetic p7 peptides in 1,2-diheptanoyl-sn-glycero-3-phosphocholine (DHPC) and obtained a low resolution three-dimensional structure of the ion channel, using single-particle electron microscopy. The resulting structure from the density map is flower-shaped, with six emerging petals oriented towards the endoplasmic reticulum (ER) lumen.[18] Similarly, Chou and co-workers renatured mutated p7 proteins in n-dodecyl-phosphocholine (DPC) and analyzed their multimeric assembly using NMR experiments. Although they also obtained a flower-shaped structure (Fig. 1), the aforementioned hairpin structure of the individual p7 monomers is not reflected and the individual monomers are interlocked, such that each subunit interacts with all the others.[19] However, the N- and C-termini of the structure of Chou and co-workers lie at the bottom of the flower, in stark contrast with the model of Zitzmann and co-workers, where the flower petals in the density map face the ER lumen. Based on their monomeric structure of recombinant p7 obtained by NMR in methanol and the ensuing structure calculation, Griffin and co-workers manually generated and minimized a heptameric architecture of the p7 channel.[9] In addition, molecular dynamics (MD) simulations of different heptameric and hexameric channels based on the p7 monomer model of Penin and co-workers[10] were carried out by Chandler et al.[20] in a pure palmitoyl-oleoyl-phosphatidyl-choline (POPC) bilayer. Similarly, Fischer and co-workers recently carried out MD simulations of hexameric assemblies based on previously equilibrated p7 monomers.[21]

In order to shed light on the architecture of p7 channels, we propose in this contribution a comparative study of different hexameric p7 assemblies in terms of their respective ion conduction ability. First, we investigated the effect of the environment (DPC or POPC) on the structure of p7 monomers. Next, we carried out MD simulations of the three p7 hexamers shown in Figure 1 and conductivity calculations at varying applied voltages. Hexamer I (strain HCV-J) corresponds to the most promising hexameric candidate in the study of Chandler et al.,[20] based on the p7 monomer described by Penin and co-workers.[10] Hexamer II (strain HCV-J) was modeled from the monomeric p7 structure of Griffin and co-workers[9] and the hexameric channel model of Chandler et al.,[20] but to facilitate comparison between the model hexamers and their relation with the available experimental data, the sequence of the p7 monomer of Penin and co-workers[10] was projected onto the structure of Hexamer I (strain HCV-J). Hexamer III (strain HCV-EUH1480) is the channel model put forth by Chou and co-workers.[19] As the protein sequence of the latter differs significantly from the sequence of the p7 monomers of Penin and co-workers[10] and of Griffin and co-workers,[9] we chose to keep it as is. For uniformity, we further chose to embed all three hexameric models in a fully hydrated POPC bilayer, the thickness of which is similar to that of the ER membrane.

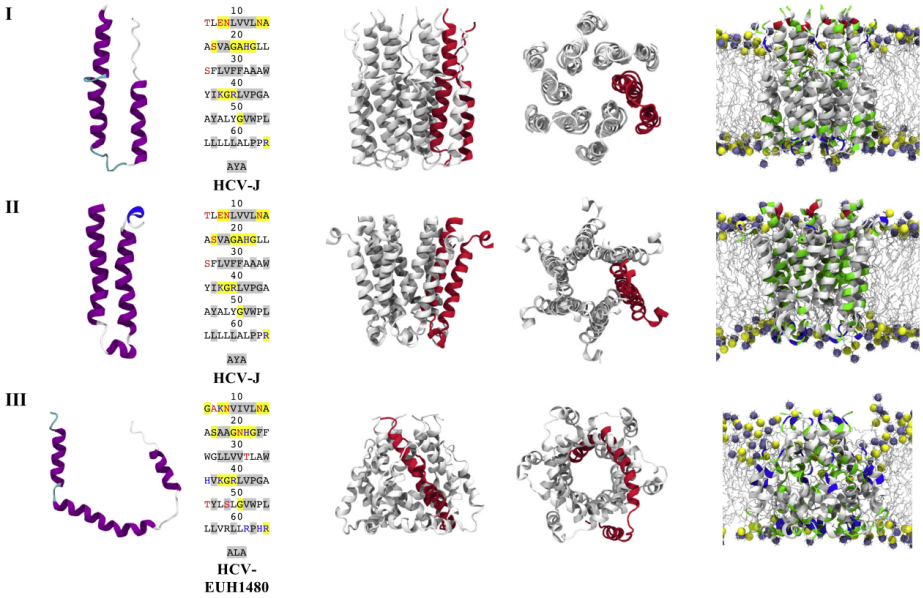


Figure 1. Monomeric and hexameric p7 models. (Left) Secondary structure and sequence information of p7 monomers as incorporated in the hexameric structures. In the secondary structure, ribbons in purple and blue denote α-helices and 3-turn–helices, cyan tubes correspond to turns and gray tubes to random coils. Conserved hydrophilic positions are highlighted in yellow and hydrophobic positions in gray. (Center) Side and bottom view of p7 hexamers, (Right) p7 hexamers embedded in membrane after membrane relaxation. Acidic residues are shown in red, basic residues in blue, polar residues in green and nonpolar residues in white. (I) Structure and sequence as postulated by Chandler and Penin, (II) Modeled structure on basis of the hexamer structure by Foster et al. with the sequence of Chandler and Penin, (III) Structure and sequence as postulated in hexamer by OuYang and Chou.

2. Materials and methods

2.1. Construction of the molecular assemblies

The hairpin conformation of the p7 monomer of strain HCV-J, genotype 1b, was considered as constructed by Penin and co-workers in light of NMR and MD experiments (Monomer I (strain HCV-J)),[10] while the 'open' Monomer III (strain HCV-EUH1480) corresponds to the conformation in a hexameric p7 channel proposed by Chou and co-workers (genotype 5a, strain HCV-EUH1480).[19]

To investigate the effect of the sequence on the p7 structure, Monomer I (strain HCV-EUH1480)) was built in a hairpin conformation, using the sequence of Chou and co-workers,[19] albeit with the backbone coordinates of the Chandler et al. structure.[20] Use was made of the psfgen structure building tool.[26] Similarly, Monomer III (strain HCV-J) was obtained by projecting the sequence of p7 strain HCV-J onto the 'open' structure mentioned previously.[19]

Correspondence between the monomeric and the hexameric structures investigated herein is as follows. Hexamer I represents the hexameric structure of Chandler et al.[20] Hexamer II was modeled based on the monomeric structure of p7 of Griffin and co-workers[9] using the sequence of Chandler.[20] Hexamer III corresponds to the solution NMR-structure reported by Chou and co-workers (pdb code: 2M6X) (strain HCV-EUH1480 bearing five mutations per monomer).[19]

2.2. Molecular dynamics simulations

All simulations were performed in the isobaric-isothermal ensemble using the NAMD simulation package[27] and the CHARMM36[28–30] force field. The pressure and the temperature were fixed at 1 atm and 300 K employing the Langevin piston algorithm and a Langevin thermostat.[31] Periodic boundary conditions were assumed. The equations of motion were integrated with a multiple-time stepping scheme and a time step of 2 fs. Short-range non-bonded interactions were truncated smoothly with a spherical cutoff radius of 12 Å and a switching distance of 10 Å. Long-range electrostatic interactions were computed using the particle-mesh Ewald method.[32]

The monomers and hexamers were immersed in a fully hydrated POPC lipid bilayer. In addition, the monomers were placed in DPC detergents. Histidines are all considered in their δ-tautomer. The total number of atoms in the simulations is as follows: 31234 for Monomer I (HCV-J) in POPC, 30991 for Monomer I (HCV-EUH1480) in POPC, 45326 for Monomer III (HCV-J) in POPC, 45371 for Monomer III (HCV-EUH1480) in POPC, 113399 for Monomer I (HCV-J) in DPC, 112902 for Monomer I (HCV-EUH1480) in DPC, 113399 for Monomer III (HCV-J) in DPC, 113411 for Monomer III (HCV-EUH1480) in DPC, 47480 for Hexamer I, 48239 for Hexamer II and 58339 for Hexamer III. Apart from POPC, which was treated utilizing a united-atom model,[33] all-atom representations were used. For equilibration of the membrane and detergent assays, harmonic restraints (5 kcal/mol/Å^2) were applied to the protein heavy atoms and slowly released after relaxation of the lipids and detergents. Additional symmetric restraints were employed during equilibration of Hexamer I and Hexamer II by averaging the positions of the C_α atoms in all six p7 fragments every 100th step of the simulation.

2.3. Ionic current and conductance calculations

In order to measure the conductance of the hexameric p7 ion channels, we applied uniform electric fields perpendicular to the membrane, i.e., along the z axis of the periodic box. Due to rearrangements of the ions and water in the vicinity of the membrane, the actual transmembrane voltage, V, is dependent on the applied electric field, E, and L_z, the system size in the direction perpendicular to the membrane:

$$V = -EL_z$$

The resulting ionic current, I, can then be obtained by counting the number of ion-crossing events, N, along the channel,

$$I = \frac{Nq}{\tau}$$

where q is the ion charge and τ the chosen time interval. The conductance is computed as I/V. The uncertainty on the ionic current and conductance is approximated as $(q\,N^{1/2})/V\,\tau$, assuming Poisson statistics of independent events. All biased simulations were carried out in the canonical ensemble.

To avoid collapse of the ion channel, collective variables[34] were defined to restrain the relevant degrees of freedom in the conductance simulations of Hexamer I. The root-mean-square displacement (RMSD) was calculated for the backbone atoms of the protein after every 500 steps of simulation. An upper boundary of 5 Å for the RMSD was applied.

2.4. Analyses and visualization

Visualization and analyses of the trajectories were carried out with VMD (visual molecular dynamics)[35] and Cpptraj.[36] Cubic regression was calculated using XMGrace.[37]

3. Results

3.1. Structural properties of the p7 monomers in POPC and DPC

In this section, the dynamics of the hairpin (Monomer I) and 'open' (Monomer III) p7 monomeric structures was investigated for strains HCV-J and HCV-EUH1480 in DPC and POPC. The distance RMSD of the backbone atoms determined throughout the 200 ns of simulation of Monomer I (strain HCV-J), Monomer III (strain HCV-EUH1480), Monomer I (strain HCV-EUH1480) and Monomer III (strain HCV-J) in a POPC bilayer and DPC micelles (see Supporting information, Fig. S1) indicates that the monomers undergo structural changes with respect to their respective starting conformation, but to an extent that the overall structural features are preserved and RMSD plateaus over time.

To analyze further the structural evolution of the monomers in the course of the simulation, we subdivided them into different structural subgroups as depicted in the Supplementary material (Fig. S2). Generally, the p7 monomer in its hairpin form is described as consisting of two antiparallel transmembrane units TM1 and TM2, where the N- and C-termini face the ER lumen, while the connecting loop is exposed to the cytoplasm. However, division of TM1 and TM2 into residues is not consistent throughout the literature (e.g., TM1 and TM2 stretch from residue 19–32 and 36–58, respectively, in Ref. 12, and from residues 13–32 and 39–57, respectively, in Ref. 34). Our arbitrary decomposition is loosely based on the classification of Cook and Opella,[12] who assign letters A and B to helical sections of TM1, and C and D to sections of TM2. Secondary structure analysis of the 200-ns trajectories with the STRIDE program[38] (see Supporting information, Fig. S2) reveals that the peptide forms stable helical stretches, which will serve as a basis to decompose p7 in four sections. We termed A and B the TM1 sections stretching from residues 4–14 and 20–31, respectively. These two α-helical segments are separated by a turn in the vicinity of Gly15, Ala16 and His17. For the TM2 segment we found discrepant secondary structure profiles as a function of the hairpin/'open' initial structure (Fig. S2). Starting from the hairpin structure of Penin and co-workers,[10] there

is a large stretch of α-helix conserved in all simulations roughly overlapping with the typical range of amino acids spanned by TM2, as described in the literature.[12] For the purpose of monitoring structural changes during the simulations, we divided TM2 into helix C (residues 39–47) and helix D (residues 47–54) for the hairpin structure of p7. For the 'open' structure of p7, the conserved TM2 helical range is much shorter and only starts around residue 47. We defined a section C* (residues 35–40), which, together with section B, organizes into an α-helix in the 'open' conformation. The secondary structure of the stretch separating sections B–C* and section D switches between α-helix, 3_{10}-helix and γ-turn in the simulations of Monomer I (strain HCV-EUH1480) and Monomer III (strain HCV-EUH1480) in POPC, and Monomer III (strain HCV-EUH1480) in DPC. Only Monomer I (strain HCV-EUH1480) in DPC shows a predominant population of γ-turn and random coil in this region (Fig. S2). It is worth mentioning that C* includes the basic residues Lys33 and Arg35. These residues are located in the loop region (exposed to the lumen) of the hairpin structure of p7.[14,22] Boundaries of sections A, B, C/C* and D are not intended to fully overlap with the p7 α-helices, but rather reflect secondary-structure stretches that are commonly preserved during the simulations of the four different monomers. The evolution of a selection of angles formed by α-helical sections together with snapshots of the protein are displayed in Figure S3.

As observed in previous molecular dynamics studies[21,39] p7 in a hairpin conformation, as described by Penin and co-workers,[10] (Monomer I (strain HCV-J)) does not undergo significant structural changes if simulated in a POPC bilayer environment. Only the angle between the N-terminal helix (A) and TM1 (B) diminishes slightly during the first half of the trajectory. Fluctuation of the (A,B) angle allows the N-terminus to adopt a variety of positions with respect to the NMR reference structure.[10]

Residues in the vicinity of the loop, which have been identified to stabilize the hairpin conformation, e.g., Phe26, Tyr45 and Trp30, with π–π stacking interactions, or Tyr31, Tyr42 and Tyr45, forming hydrogen bonds,[11] are only partially present in the strain employed by Chou and co-workers[19]—i.e., Monomer I (strain HCV-EUH1480) and Monomer III (strain HCV-EUH1480). However, the hairpin structure is largely retained in the simulation of Monomer I (strain HCV-EUH1480) in POPC. Starting from an 'open' structure, Monomer III (strain HCV-J) and Monomer III (strain HCV-EUH1480) embedded in POPC does not adopt a canonical hairpin conformation on the finite timescale of our simulations, but a slight compression of the protein towards the end of the trajectory is observed, wherein the α-helical regions formed by segments A, B and C* tilt in the POPC bilayer. The latter phenomenon was also observed by Fischer and co-workers[21] in a 100-ns simulation of Monomer III (strain HCV-EUH1480) in POPC.[21]

The large angle between TM1 (B) and C* of the open starting structure (143.5°) is not conserved in any of the simulations and decreases significantly. As can be seen in Figure S3, irrespective of the genotype, the hairpin structure is largely preserved throughout the simulation in POPC, while the same structures are significantly affected by the harsh DPC environment (albeit the monomer of strain HCV-EUH1480 appears to be slightly more so). However, a clear dependence to either the starting structure or the protein sequence cannot be observed.

The higher degree of flexibility in the simulations in DPC is also mirrored in the two-dimensional maps given in Figure S2, which shows distance RMSD of the protein C_α atoms determined for all pairs of conformations.

3.2. Structural properties of hexameric p7 assemblies

Hexamers I (strain HCV-J), II (strain HCV-J) and III (strain HCV-EUH1480) embedded in a POPC membrane were simulated for 150 ns after appropriate equilibration and protein relaxation, applying initially symmetry restraints. The distance RMSD after releasing all restraints (Fig. S1) varies minimally for Hexamer II and III. For Hexamer I, however, a sudden increase of the RMSD can be observed after half of the simulation. Analysis of the structure of Hexamer I in this portion of the trajectory shows a collapse and, thus, a closure of the ion channel. Such an event is not uncommon for long trajectories and were also previously observed in the 100 ns simulations of several p7 ion channels reported by Fischer and co-workers.[21]

The pore-radius profiles obtained with the program HOLE[40] are gathered in Figure 2. Although both initial structures of Hexamer I and Hexamer II possess similarly narrow pore radii as low as 2 Å, only Hexamer I is not water permeable during membrane equilibration when the protein coordinates are geometrically restrained. As previously described by Chandler et al.,[20] the pore is sealed in one position by a ring of hydrophobic residues, namely Phe25, and in a second position by Ile32. Water permeation only occurs when the harmonic restraints are released during the simulation. The Phe25 bottleneck is also apparent in Hexamer II, while Ile32 is now located in the loop region, facing outward. Yet, a single residue, i.e., Phe25, is obviously not sufficient to block the permeation of water during membrane equilibration. Water permeation is likely to be facilitated by the basic residue Arg35, which is located right at the bottom of the pore, inward-facing in the initial structure of Hexamer II. However, this arrangement is not conserved during the simulation, as Arg35 soon reorients towards the cytoplasm. In contrast with Hexamer I, the pore of which remains rather narrow throughout the simulation before eventually collapsing, Hexamer II forms a wide ion channel, with a pore radius hardly falling below 4 Å. As can be seen in the snapshot of the last frame in Figure 2, the initial funnel-like structure of Hexamer II is, however, lost and a straighter, cylindrical conformation is adopted. Consistently with the wider ion pore in Hexamer II, we observe a number density of water about 1.5 time larger than in both Hexamer I and Hexamer II (Table S4), thus suggesting a higher potential for permeability for this model.

Examination of the cross section of the electrostatic pore profile (Fig. S6) does not reveal any clear evidence for ion selectivity. However, this data indicates that ion approach from the cytoplasm side is expected to be disfavored electrostatically for Hexamer III.

Further analysis of the structures after 150 ns (after 75 ns for Hexamer I) reveals that the POPC bilayer adapts differently to the three hexamers. In Hexamer I, the N-terminus is significantly tilted with respect to the bilayer normal (Fig. 2). Hexamer II and the lipid bilayer match without any distortion of the protein or the membrane (Fig. 2).

Adaptation of Hexamer III to span the POPC bilayer is less pronounced. The channel remains in its flower-like conformation and does not stretch within the lipid environment. Furthermore, Hexamer III features two basic residues, namely Arg54 and Arg57, which are facing the membrane interior. In conjunction with these structural characteristics, the POPC bilayer shows a concave deformation towards the hexamer arising from the interaction of the hydrophilic head groups of single POPC lipids with residues Arg54 and Arg57 (Figs. 2 and S5). A direct comparison of the bilayer thickness in the vicinity and areas far from the protein (Table S3) reveals the best match for Hexamer I (between 36.5 and 38.5 Å). For Hexamer II, a slight stretch close to the protein up to 42.7 Å is observed, whereas for Hexamer III, the thickness can be as low as 22 Å, reflecting a severe mismatch between this model and a membrane environment.

Furthermore, we have analyzed the contact area between the six monomers of the channels based on the solvent accessible surface area (see Table S5). We found that the contact area between the monomers is comparable for Hexamer I and Hexamer II, but

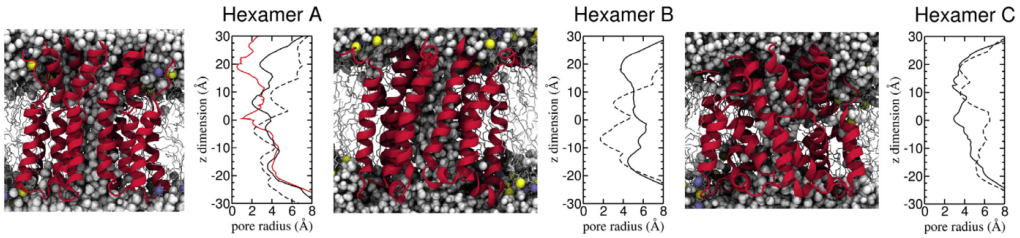


Figure 2. Molecular assemblies and pore radius profiles of Hexamer I, II and II in POPC. Left panels: side views of hexamers in ribbon representation. Water molecules are displayed as white van der Waals spheres. The model POPC layer is shown in white sticks and yellow and blue van der Waals spheres for phosphorus and nitrogen atoms, respectively. Right panels: radial profiles of the width of the pore along the longitudinal axis of the protein. Dashed lines correspond to the initial structures according to Figure 1. Solid lines represent the starting point of the conductivity calculations, thus the structure after 150 ns of simulation for Hexamer II and III and after 75 ns of simulation of Hexamer 1. The red line for Hexamer I corresponds to the collapsed ion channel after 128 ns of simulation. All radii were calculated using the HOLE program.

nearly twice as large for Hexamer III. This is not surprising and stems from the fact that the intramonomer 'hairpin' contact area is not featured in Hexamer III, the latter engaging in intermonomer contacts in its interlaced 'open' structure. Therefore, the higher intermonomer contact area in Hexamer III does not necessarily result in a more stable ion channel. At the same time, Hexamer III shows more exposure to the POPC bilayer as highlighted in Table S6. In particular, the aforementioned Arg57 residue interacts with the membrane, while both the loop region and the N- and C-termini interact with a large number of water molecules—largely exceeding comparable protein–water and protein–POPC interactions in Hexamer I and Hexamer II. A better accessibility of terminal and loop residues, just like in Hexamer II, might facilitate ion attraction and therewith ion permeation.

3.3. Conductance calculations in the hexameric ion channels

We applied electric fields corresponding to voltages of −1.5, −1.2, −0.9, −0.6, +0.6, 0.9, +1.2 and +1.5 V along the z axis to the ion channels and performed MD simulations over 60 ns, or until membrane poration by ions and water was observed. For Hexamer II and III, the final coordinates after 150 ns were used. In the case of Hexamer I, use was made of the coordinates after 75 ns and geometrical restraints were enforced (see Section 2) to avoid the collapse of the ion channel. Negative voltages correspond to an electric field along the z axis, from lumen to cytoplasm, and positive voltages to the opposite direction. Cations travel in the direction of the electric field and anions against it. Ionic currents and conductivities (Fig. 3) were calculated considering both chloride and sodium ions by counting ion translocation events across the membrane, which are listed in Table S1.

Strikingly, the experimentally reported cation selectivity[10] is not reflected in any of our simulations. Instead, we observe an overwhelming number of chloride ions permeating in all cases (see Table S1)—the largest being for Hexamer II (up to 7.4 Cl^- per ns), followed by Hexamer III and Hexamer I (up to 5.2 and 3.9 Cl^- per ns, respectively). In fact, for Hexamer I and Hexamer III, sodium ion translocation only occurs as single events. Hexamer II transports a significant number of up to 0.9 cations per nanosecond of simulation, but only when multiple chloride ions move through the channel in the opposite direction for positive voltages and about twelve times as many for negative voltages. Thus, despite its clear-cut anion selectivity, Hexamer II represents the only structure amid our models that does not preclude the possibility of cation transport. During protein equilibration, acidic residue Glu3 of Hexamer II orients towards the mouth of the ion channel and might, thus, facilitate sodium ion attraction and insertion. Glu3 is absent in Hexamer III and points towards the membrane domain in Hexamer I. Consistently with the pore radius size, Hexamer II is crossed by the largest amount of ions, followed by Hexamer III and Hexamer I, which mirrors the trend in pore diameter and in amount of water in the pores.

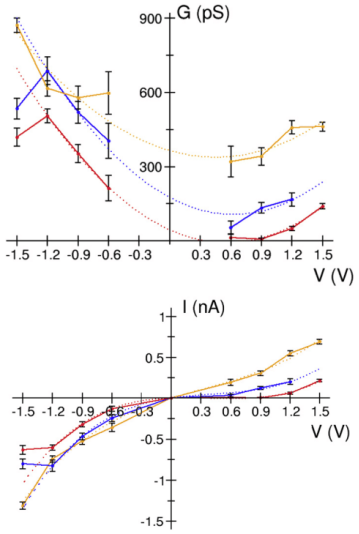


Figure 3. Calculated conductivity of p7 hexamers. Hexamer I is shown in red, Hexamer II in orange and Hexamer III in blue. (Top) I/V relationship based on channel transitions of ions for hexamers I, II and III. Solid lines: measured ionic currents at applied voltage biases from −1.5 to 1.5 V. Dotted lines: Cubic regression corresponding to functions in Table S2. (Bottom) Conductivity with respect to applied voltage biases.

If very large, negative voltages of −1.5 V are applied, all assays seem susceptible to membrane poration by water and ions. In the case of Hexamer I and Hexamer II, a pore formed distant from the ion channel after 15 and 50 ns, respectively. In all conductance assays, the membrane of Hexamer III was systematically porated, irrespective of the applied voltages of ±1 V or larger, and within the first 25 ns of the simulation. Poration occurred in direct vicinity of the ion channel (Fig. S5), where the membrane features a deformation towards the compact protein and its exposed basic residues. This observation can also be made from the electrostatic profile of the pore (Fig. S6), which clearly shows large deformations for different applied voltages. Aside from membrane poration, we

observed that for Hexamer I and Hexamer III, there is no significant change in ionic current at −1.5 V, compared with −1.2 V, which indicates that a conductance limit is reached at these voltages. The recorded measurements at −1.5 V for Hexamer I and Hexamer III are not considered for analysis in the following. Resulting ionic currents at negative voltages are of larger absolute values by comparison with the application of positive voltages, which was also observed in conductance calculations by Fischer and co-workers[23] and in the experimental measurements of Penin and co-workers,[10] by a factor of ∼2 to 3.5. For our conductivity calculations of Hexamer I we could not reproduce this trend and observed multiplicative factors in the range of 3.9–59.5 (Fig. 3). Hexamer II with a factor of 1.3–1.9 and Hexamer III with a factor of 3.9–7.6 seem to agree better with the experimental data, although the exact range is not reproduced. In addition, the internal structure pore, notably its diameter, does not appear to be affected by any of the applied voltages.

We performed a cubic fit of our data series (excluding the values diverging at the limit of −1.5 V for Hexamer I and Hexamer III). The results given in Table S2 show regression coefficients between 0.994 and 0.999, which verifies the integrity of our data. In order to compare with the experimental data of Penin and co-workers,[10] we extrapolated the equivalent conductivity values at +0.06 and +0.14 V based on our cubic regression curve. The experimental results of 22 ± 6 pS at +0.06 V and 41 ± 8 pS at +0.014 V are not properly reproduced by Hexamer II and Hexamer III, which yield one order of magnitude higher. The best agreement is obtained for Hexamer I with 34.4 and 22.2 pS, respectively. It should be noted that in contrast with the experimental data, our extrapolated values all show a negative resistance in this voltage range, hence, a lower conductivity at larger voltages. The values at +0.06 and +0.14 V obtained by Chandler et al. for Hexamer I of 9.4 pS and 22.0 pS follow the experimental trend, albeit they were extrapolated, assuming a linear behavior of ionic currents from data obtained at a voltage of +1.5 V. Monitoring permeation events over time (Fig. S7) obeys a linear behavior and thus supports the idea of a constant permeation flow of ions.

Integration of radial distribution functions (Fig. S8) does not reveal any significant change in ion solvation as it translocates from the bulk into the pore of the channel. Mechanistically both Hexamer I and Hexamer III show an ion-ion knock-on process when chloride ions are permeating in the +z direction. Similar to the mechanism described by Padhi et al.[41] (reported for different cations) the ions first enter the cavity at the bottom of the pore, interacting with residues Lys33, Leu36 and Arg35. After one ion passes the Phe25 barrier, it remains in the pore near residues His17 and Ser12 until the next ion leaves the cavity and knocks the first ion out towards the exit. The reversed path, in the –z direction, shows less knock-on, the chloride ions moving more freely. The anions also move in a less hampered fashion through the wider pore of Hexamer II, interacting less with the protein surface, compared to Hexamer I and Hexamer III. No knock-on mechanism is observed although the pore is often populated by more than one chloride at a time. Sodium flow is established as these ions travel in opposite direction along a chain of chloride ions. The process is observed for both the +z and −z directions.

4. Discussion

In this contribution we have analyzed the structural evolution and characteristics of several monomeric and hexameric p7 proteins of different sequences, assemblies and conditions, together with their ion-channel conductivity properties.

In simulations of p7 monomers of significantly different sequences and in either a membrane or a detergent environment we observed that in spite of a large flexibility of the proteins, the main secondary-structure features are retained over 200 ns. For instance, monomers with sequences (strain HCV-J) that were found in NMR studies to possess a hairpin conformation[10] were not able to adapt on the timescale of the simulation to the main secondary structure features of the hairpin, starting from an 'open' conformation, let alone interconvert into a hairpin structure. In light of this observation, it can be inferred that the flexibility of the monomers arises from the interhelical regions of the proteins. Monomer self-assembly into multimeric, functional ion channels, raises the question of how readily p7 monomers might form such channels, in which conformation of the monomer within the channel differs significantly from its native conformation.

Overall, we were able to show that p7 monomers embedded in the short-tail DPC environment experience severe structural changes, the flexibility of the proteins being more pronounced than in a POPC bilayer. In fact, except for the conserved secondary-structure features, the p7 monomers in the DPC simulations do not show any tendency towards preferred alignments in the micelles dependent of their sequence or initial conformation. In contrast, our simulations in POPC indicate that the long-tail lipids enable a more controlled structural evolution and a clear preference towards hairpin-like conformations. Hairpin monomers of both strains HCV-J and HCV-EUH1480 conserve their conformation throughout the majority of the simulation. Using the 'open' initial structure, monomers show a small degree of compression and tilt to align with the normal of the POPC bilayer.

The N-terminus of p7 was recently described to play a critical role in E2/p7 cleavage and, thus, in virus production.[24] Variation of the angle between the N-terminus and TM1 can be seen in all our monomer simulations, but is more pronounced when p7 is embedded in DPC. According to Cook and Opella,[12] tilts in the single helices of the TM1 segment (A and B) of p7 allow better match with the bilayer thickness, which is of importance for the longer TM1 strand in comparison to TM2 (C and D). These findings explain nicely the increased fluctuation of angle (A,B) and the decrease of its average value when p7 is immersed in a DPC environment. A nearly perpendicular kink between TM1 and the N-terminus segment was also previously found by Opella and co-workers as the lowest-energy structure calculated by Xplor-NIH refinement of NMR data.[13]

In Hexamer I, fluctuation of the angle between the N-terminus segment and TM1 facilitates the alignment of the six p7 monomers by slightly tilting the N-terminus segment. As demonstrated by Chandler et al.,[20] this enhances in turn the reorganization in short-tail detergents like DHPC of the initially cylindrical hexamer into a flower-like arrangement, consistent with the EM density maps determined by Zitzmann and co-workers.[18] Contrary to the structure of Chou and co-workers (Hexamer III), the N- and C-termini of Hexamer I point towards the flower petals and, thus, correspond to the description of Zitzmann and co-workers, who assigned this side to the ER lumen. Hexamer II aligns in the POPC bilayer by a tilt of the complete TM1 segment, including the N-terminal region. The compact arrangement of Hexamer III is based upon NMR studies of p7 in DPC.[19] The interlocked p7 monomers do not allow the hexamer to sacrifice its flower-like structure and stretch to span the POPC bilayer. Interlocked monomers and exposure of several basic residues towards the membrane lead to instability and induces membrane deformation. Thus, the proposed structure of Hexamer III might not be optimal in a POPC bilayer, as the protein is not able to adjust ideally to its environment. In addition, although Hexamer III corresponds to the flower-shape found in the EM density of Zitzmann and co-workers, its termini point towards the flower stem and the cytoplasm.

The experimentally observed cation selectivity could not be reproduced in any of our conductance calculations. Experimental

studies that show sodium ion permeation that are ten times higher than chloride permeation in conductance measurements[25] are in stark contrast with our results on the three hexameric models I-III. However, the theoretical I/V curves agree, at least qualitatively, with the experimental ones, notably the shape and the asymmetry with respect to positive and negative voltages.[10] We observed a constant flow of sodium ions only in Hexamer II and assigned this behavior to the attraction of cations by acidic residue Glu3 located at the pore entrance, at the top of the ion channel, and possibly the presence of basic residues, e.g., Arg32 and Arg35, at the bottom of the channel, pointing outwards.

It should be noted that a direct comparison with experimental values is not straightforward. In the experiments of Penin and co-workers,[10] 5 nmol of a 10 μM p7 stock solution was used, which presumably leads to the formation of more than one ion channel and, thus, to higher ionic currents. Indeed, the measured currents are about one order of magnitude higher than our calculated single-channel current. Since all of our assays reflect the shape of the experimental I/V curve, only the current ratio at negative and positive voltages should be considered to estimate the quality of hexameric models I, II and III. We can only observe reasonable agreement in Hexamers II and III, while Hexamer I features currents up to 59.5 times higher at applied negative voltages.

The anion selectivity observed herein arises from the presence of basic residues, notably in the cytosolic loop (Lys33 and Arg35) and the C-terminus (Arg60). Mutation of the aforementioned titrable cytosolic residues (to Gln33 and Gln35) leads to a channel that does not hamper cation transport, thereby providing a cogent demonstration of the implication of the basic residues in the observed anion selectivity (see Movie in the Supplementary information). Moreover, Hexamer II is devoid of acidic residues. Umbrella sampling studies by Padhi and Priyakumar based on this structure show free-energy barriers of 5.4 and 19.4 kcal/mol for the permeation of potassium and calcium ions, and, hence, suggest thermodynamically unfavorable cation translocation.[41,42] Last, to the question—which model best describes the hexameric p7 ion channel, we answer—Hexamer II. Although the latter does not reflect the cation selectivity observed experimentally,[10] it is the only model that transports a significant number of sodium ions. The shape of its I/V profile is consistent with experiment and the ratio of ionic current at negative and positive applied voltages is well reproduced in our calculations. Moreover, this model forms stable pores of reasonable size and aligns well in the POPC bilayer due to a tilt of the transmembrane helix strands. Hexamer III performs similarly well insofar as conductivity data are concerned (I/V curve shape and ratio of ionic currents). It does not feature, however, any acidic residue at the pore entrance, casting doubt on the organization of the monomers to allow cation translocation. Another weak point for this model lies in its inability to align along the POPC bilayer. The membrane deformation and additional titratable basic residues facing the hydrophobic interior of the bilayer lead to fast membrane poration at applied voltages exceeding ±1 V. Hexamer I equally fails as a good model p7 ion channel as it only forms very narrow channels that collapse easily.

It still remains that molecular simulations represent a convincing tool to discriminate between models of biological targets, notably membrane proteins, for which no or limited structural information is available. Moreover, as has been seen previously, availability of experimental three-dimensional structures of membrane proteins does not necessarily guarantee their physiological relevance, as their conformation stringently depends on the nature of the environment in which it has been determined—e.g., harsh detergent media are prone to modify the overall fold of the protein, thereby abolishing its function.[43–46]

Acknowledgments

The authors are indebted to the Grand Nancy and the Fonds Européen de Développement Économique et Régional (FEDER) for the support of this research. FP is supported by the French ANRS (France Recherche, Nord & Sud, Sida-HIV et Hépatites), an autonomous agency at INSERM, France. They also acknowledge fruitful discussions with Nicole Zitzmann (Oxford University). They gratefully thank Danielle Chandler (University of Illinois at Urbana-Champaign) for the simulation of p7 mutants.

Supplementary data

Distance root mean-square deviation (RMSD) plots for the monomers in POPC and in DPC, and the hexamers in POPC, two-dimensional RMSD maps for the monomers in POPC and in DPC, structural parameters and secondary structure of the monomers in POPC and in DPC, conductance data, ion coordination, permeation events, solvent-accessible surface area and inter-domainn contact analysis, water density, electrostatic potential cross-sections, membrane-thickness analysis and snapshots. Movie of the hexameric channel mutant K33Q/R35Q at −1.5 V.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2016.07.063.

References and notes

1. WHO|Hepatitis C http://www.who.int/mediacentre/factsheets/fs164/en/ (accessed 10 May 2016).
2. Lozano, R.; Naghavi, M.; Foreman, K.; Lim, S.; Shibuya, K.; Aboyans, V.; Abraham, J.; Adair, T.; Aggarwal, R.; Ahn, S. Y.; AlMazroa, M. A.; Alvarado, M.; Anderson, H. R.; Anderson, L. M.; Andrews, K. G.; Atkinson, C.; Baddour, L. M.; Barker-Collo, S.; Bartels, D. H.; Bell, M. L.; Benjamin, E. J.; Bennett, D.; Bhalla, K.; Bikbov, B.; Abdulhak, A. B.; Birbeck, G.; Blyth, F.; Bolliger, I.; Boufous, S.; Bucello, C.; Burch, M.; Burney, P.; Carapetis, J.; Chen, H.; Chou, D.; Chugh, S. S.; Coffeng, L. E.; Colan, S. D.; Colquhoun, S.; Colson, K. E.; Condon, J.; Connor, M. D.; Cooper, L. T.; Corriere, M.; Cortinovis, M.; de Vaccaro, K. C.; Couser, W.; Cowie, B. C.; Criqui, M. H.; Cross, M.; Dabhadkar, K. C.; Dahodwala, N.; De Leo, D.; Degenhardt, L.; Delossantos, A.; Denenberg, J.; Des Jarlais, D. C.; Dharmaratne, S. D.; Dorsey, E. R.; Driscoll, T.; Duber, H.; Ebel, B.; Erwin, P. J.; Espindola, P.; Ezzati, M.; Feigin, V.; Flaxman, A. D.; Forouzanfar, M. H.; Fowkes, F. G. R.; Franklin, R.; Fransen, M.; Freeman, M. K.; Gabriel, S. E.; Gakidou, E.; Gaspari, F.; Gillum, R. F.; Gonzalez-Medina, D.; Halasa, Y. A.; Haring, D.; Harrison, J. E.; Havmoeller, R.; Hay, R. J.; Hoen, B.; Hotez, P. J.; Hoy, D.; Jacobsen, K. H.; James, S. L.; Jasrasaria, R.; Jayaraman, S.; Johns, N.; Karthikeyan, G.; Kassebaum, N.; Keren, A.; Khoo, J.-P.; Knowlton, L. M.; Kobusingye, O.; Koranteng, A.; Krishnamurthi, R.; Lipnick, M.; Lipshultz, S. E.; Ohno, S. L.; Mabweijano, J.; MacIntyre, M. F.; Mallinger, L.; March, L.; Marks, G. B.; Marks, R.; Matsumori, A.; Matzopoulos, R.; Mayosi, B. M.; McAnulty, J. H.; McDermott, M. M.; McGrath, J.; Memish, Z. A.; Mensah, G. A.; Merriman, T. R.; Michaud, C.; Miller, M.; Miller, T. R.; Mock, C.; Mocumbi, A. O.; Mokdad, A. A.; Moran, A.; Mulholland, K.; Nair, M. N.; Naldi, L.; Narayan, K. M. V.; Nasseri, K.; Norman, P.; O'Donnell, M.; Omer, S. B.; Ortblad, K.; Osborne, R.; Ozgediz, D.; Pahari, B.; Pandian, J. D.; Rivero, A. P.; Padilla, R. P.; Perez-Ruiz, F.; Perico, N.; Phillips, D.; Pierce, K.; Pope, C. A., III; Porrini, E.; Pourmalek, F.; Raju, M.; Ranganathan, D.; Rehm, J. T.; Rein, D. B.; Remuzzi, G.; Rivara, F. P.; Roberts, T.; De León, F. R.; Rosenfeld, L. C.; Rushton, L.; Sacco, R. L.; Salomon, J. A.; Sampson, U.; Sanman, E.; Schwebel, D. C.; Segui-Gomez, M.; Shepard, D. S.; Singh, D.; Singleton, J.; Sliwa, K.; Smith, E.; Steer, A.; Taylor, J. A.; Thomas, B.; Tleyjeh, I. M.; Towbin, J. A.; Truelsen, T.; Undurraga, E. A.; Venketasubramanian, N.; Vijayakumar, L.; Vos, T.; Wagner, G. R.; Wang, M.; Wang, W.; Watt, K.; Weinstock, M. A.; Weintraub, R.; Wilkinson, J. D.; Woolf, A. D.; Wulf, S.; Yeh, P.-H.; Yip, P.; Zabetian, A.; Zheng, Z.-J.; Lopez, A. D.; Murray, C. J. *Lancet* **2012**, *380*, 2095.
3. Pawlotsky, J. *Gastroenterology* **2014**, *146*, 1176.
4. Penin, F.; Dubuisson, J.; Rey, F. A.; Moradpour, D.; Pawlotsky, J.-M. *Hepatology* **2004**, *39*, 5.
5. Moradpour, D.; Penin, F. In *Hepatitis C Virus: From Molecular Virology to Antiviral Therapy*; Bartenschlager, R., Ed.; Current Topics in Microbiology and Immunology; Springer: Berlin Heidelberg, 2013; pp 113–142.
6. Scott, C.; Griffin, S. *J. Gen. Virol.* **2015**, *96*, 2000.
7. Steinmann, E.; Penin, F.; Kallis, S.; Patel, A. H.; Bartenschlager, R.; Pietschmann, T. *PLoS Pathog.* **2007**, *3*, e103.
8. Wozniak, A. L.; Griffin, S.; Rowlands, D.; Harris, M.; Yi, M.; Lemon, S. M.; Weinman, S. A. *PLoS Pathog.* **2010**, *6*, e1001087.
9. Foster, T. L.; Thompson, G. S.; Kalverda, A. P.; Kankanala, J.; Bentham, M.; Wetherill, L. F.; Thompson, J.; Barker, A. M.; Clarke, D.; Noerenberg, M.;

Pearson, A. R.; Rowlands, D. J.; Homans, S. W.; Harris, M.; Foster, R.; Griffin, S. *Hepatology* **2014**, *59*, 408.
10. Montserret, R.; Saint, N.; Vanbelle, C.; Salvay, A. G.; Simorre, J.-P.; Ebel, C.; Sapay, N.; Renisio, J.-G.; Bockmann, A.; Steinmann, E.; Pietschmann, T.; Dubuisson, J.; Chipot, C.; Penin, F. *J. Biol. Chem.* **2010**, *285*, 31446.
11. Cook, G. A.; Opella, S. J. *Eur. Biophys. J.* **2010**, *39*, 1097.
12. Cook, G. A.; Opella, S. J. *Biochim. Biophys. Acta BBA - Biomembr.* **2011**, *1808*, 1448.
13. Cook, G. A.; Dawson, L. A.; Tian, Y.; Opella, S. J. *Biochemistry* **2013**, *52*, 5295.
14. Carrere-Kremer, S.; Montpellier-Pala, C.; Cocquerel, L.; Wychowski, C.; Penin, F.; Dubuisson, J. *J. Virol.* **2002**, *76*, 3720.
15. Clarke, D.; Griffin, S.; Beales, L.; Gelais, C. S.; Burgess, S.; Harris, M.; Rowlands, D. *J. Biol. Chem.* **2006**, *281*, 37057.
16. Griffin, S. D. C.; Beales, L. P.; Clarke, D. S.; Worsfold, O.; Evans, S. D.; Jaeger, J.; Harris, M. P. G.; Rowlands, D. J. *FEBS Lett.* **2003**, *535*, 34.
17. Madan, V.; Bartenschlager, R. *Viruses* **2015**, *7*, 4461.
18. Luik, P.; Chew, C.; Aittoniemi, J.; Chang, J.; Wentworth, P.; Dwek, R. A.; Biggin, P. C.; Venien-Bryan, C.; Zitzmann, N. *Proc. Natl. Acad. Sci.* **2009**, *106*, 12712.
19. OuYang, B.; Xie, S.; Berardi, M. J.; Zhao, X.; Dev, J.; Yu, W.; Sun, B.; Chou, J. J. *Nature* **2013**, *498*, 521.
20. Chandler, D. E.; Penin, F.; Schulten, K.; Chipot, C. *PLoS Comput. Biol.* **2012**, *8*, e1002702.
21. Kalita, M. M.; Griffin, S.; Chou, J. J.; Fischer, W. B. *Biochim. Biophys. Acta BBA - Biomembr.* **2015**, *1848*, 1383.
22. Patargias, G.; Zitzmann, N.; Dwek, R.; Fischer, W. B. *J. Med. Chem.* **2006**, *49*, 648.
23. Wang, Y.-T.; Schilling, R.; Fink, R. H. A.; Fischer, W. B. *Biophys. Chem.* **2014**, *192*, 33.
24. Scull, M. A.; Schneider, W. M.; Flatley, B. R.; Hayden, R.; Fung, C.; Jones, C. T.; van de Belt, M.; Penin, F.; Rice, C. M. *PLoS Pathog.* **2015**, *11*, e1005297.
25. Premkumar, A.; Wilson, L.; Ewart, G.; Gage, P. *FEBS Lett.* **2004**, *557*, 99.
26. VMD psfgen Plugin, Version 1.6.4 http://www.ks.uiuc.edu/Research/vmd/plugins/psfgen/ (accessed 7 Jun 2016).
27. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
28. MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
29. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 3257.
30. Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. *J. Phys. Chem. B* **2010**, *114*, 7830.
31. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
32. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
33. Lee, S.; Tran, A.; Allsopp, M.; Lim, J. B.; Hénin, J.; Klauda, J. B. *J. Phys. Chem. B* **2014**, *118*, 547.
34. Fiorin, G.; Klein, M. L.; Hénin, J. *Mol. Phys.* **2013**, *111*, 3345.
35. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33.
36. Roe, D. R.; Cheatham, T. E. *J. Chem. Theory Comput.* **2013**, *9*, 3084.
37. Grace Home http://plasma-gate.weizmann.ac.il/Grace/ (accessed 7 Jun 2016).
38. Frishman, D.; Argos, P. *Proteins Struct. Funct. Genet.* **1995**, *23*, 566.
39. Gentzsch, J.; Brohm, C.; Steinmann, E.; Friesland, M.; Menzel, N.; Vieyres, G.; Perin, P. M.; Frentzen, A.; Kaderali, L.; Pietschmann, T. *PLoS Pathog.* **2013**, *9*, e1003355.
40. Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P. *J. Mol. Graph.* **1996**, *14*, 354.
41. Padhi, S.; Priyakumar, U. D. *J. Phys. Chem. B* **2015**, *119*, 6204.
42. Padhi, S.; Priyakumar, U. D. *J. Phys. Chem. B* **2016**, *120*, 4351.
43. Cross, T. A.; Murray, D. T.; Watts, A. *Eur. Biophys. J.* **2013**, *42*, 731.
44. Zhou, H.-X.; Cross, T. A. *Annu. Rev. Biophys.* **2013**, *42*, 361.
45. Zoonens, M.; Comer, J.; Masscheleyn, S.; Pebay-Peyroula, E.; Chipot, C.; Miroux, B.; Dehez, F. *J. Am. Chem. Soc.* **2013**, *135*, 15174.
46. Murray, D. T.; Griffin, J.; Cross, T. A. *Biochemistry* **2014**, *53*, 2454.

Biophysical Journal
Biophysical Letter

Mitochondrial ADP/ATP Carrier in Dodecylphosphocholine Binds Cardiolipins with Non-native Affinity

François Dehez,[1] Paul Schanda,[2,*] Martin S. King,[3] Edmund R. S. Kunji,[3] and Christophe Chipot[1,4,*]
[1]Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana-Champaign, Unité Mixte de Recherche no. 7565, Université de Lorraine, Vandœuvre-lès-Nancy, France; [2]Université Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, Grenoble, France; [3]Medical Research Council, Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge, United Kingdom; and [4]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois

ABSTRACT Biophysical investigation of membrane proteins generally requires their extraction from native sources using detergents, a step that can lead, possibly irreversibly, to protein denaturation. The propensity of dodecylphosphocholine (DPC), a detergent widely utilized in NMR studies of membrane proteins, to distort their structure has been the subject of much controversy. It has been recently proposed that the binding specificity of the yeast mitochondrial ADP/ATP carrier (yAAC3) toward cardiolipins is preserved in DPC, thereby suggesting that DPC is a suitable environment in which to study membrane proteins. In this communication, we used all-atom molecular dynamics simulations to investigate the specific binding of cardiolipins to yAAC3. Our data demonstrate that the interaction interface observed in a native-like environment differs markedly from that inferred from an NMR investigation in DPC, implying that in this detergent, the protein structure is distorted. We further investigated yAAC3 solubilized in DPC and in the milder dodecylmaltoside with thermal-shift assays. The loss of thermal transition observed in DPC confirms that the protein is no longer properly folded in this environment.

In the context of membrane-protein structure determination, dodecylphosphocholine (DPC) has been the subject of much criticism for being a very harsh detergent, prone to induce protein denaturation (1–5). Yet ∼40% of NMR investigations (6) have utilized this detergent, leading to protein structures (7–9) possessing a three-dimensional fold at variance with that observed in a different, milder environment (10–14). Among these membrane proteins, mitochondrial carriers have been the object of several studies (2–4). In particular, the uncoupling protein (UCP2), for which a backbone fold has been determined (8), and the ADP/ATP carrier (AAC), have been thoroughly investigated in thermostability-shift assay (TSA) experiments (4,15). These experiments have shown that in a mild dodecylmaltoside (DDM) detergent environment, mitochondrial carriers extracted from native membranes show many of the expected features. Cooperative unfolding is observed, and addition of the inhibitor carboxyatractyloside (CATR) to AAC increases the stability and, hence, shifts the transition to a higher temperature. In contrast, when solubilized in DPC, the mitochondrial carrier UCP2 did not show any unfolding transition whatsoever (4), thereby strongly suggesting that in this harsh milieu, the protein is unfolded from the onset, or does not possess a stable tertiary structure. Transport assay experiments revealed that UCP2 purified in DPC (8) could not be resurrected into a functional state upon insertion into liposomes (3). Furthermore, molecular dynamics (MD) simulations indicated without any ambiguity that the putative structure of UCP2 determined by solution-state NMR in DPC (8), when inserted into membranes, collapses, suggesting that it is not a physiological state, and is possibly distorted beyond recovery by the harsh detergent environment. Recently, the functionality of the yeast mitochondrial AAC (yAAC3) has been investigated in DPC by looking at a very specific aspect of the carrier, namely, its ability to interact with cardiolipins (CLs) (16). CL is a highly abundant lipid in the inner mitochondrial membrane, and it has been shown to be important for the function of the transporter, albeit not strictly required (17). In the work of Zhao et al. (16), yAAC3 was prepared through a refolding procedure from inclusion bodies. The activity state was addressed indirectly by studying the well-characterized yAAC3-CL interactions observed in crystal structures extracted from the native membrane (13). Contacts of the protein with the surrounding CLs were identified by recording

Submitted July 7, 2017, and accepted for publication September 21, 2017.
*Correspondence: paul.schanda@ibs.fr or chipot@ks.uiuc.edu
Editor: Jose Faraldo-Gomez.
https://doi.org/10.1016/j.bpj.2017.09.019

nuclear Overhauser enhancement (NOE) spectra. The NOE data characteristic of the CL headgroups are consistent with the crystallographic binding sites (13). These data further indicate close spatial proximity between the CL acyl chains and residues that are located at the cytoplasmic side of the carrier, i.e., far from the CL-headgroup binding sites (see Fig. 1). The main message of the work of Zhao et al. (16) is that DPC preserves the ability of AAC to bind CLs in a specific fashion, which implies that it is a good detergent, compatible with functional mitochondrial carriers. In this contribution, we turn to MD simulations to assess whether or not the binding specificity of CLs toward yAAC3 observed in DPC holds in a native-like membrane environment. Furthermore, we investigate by means of TSAs the ability of two detergents, namely, DPC and dodecylmaltoside (DDM), to preserve the fold of the mitochondrial carrier. In light of our findings, we propose that DPC alters the fold of AAC and modifies its interaction with CLs.

Starting from the crystal structure (13), we carried out an all-atom, 200-ns simulation of yAAC3 binding three CLs in a palmitoyloleoylphosphatidylcholine (POPC) bilayer (see the Supporting Material for methodological details). As can be observed in Fig. 2, the root mean-square deviation (RMSD) for yAAC3 averages to ~3 Å for the entire protein, and to 2 Å for its secondary-structure elements, consistent with all other theoretical investigations of the mitochondrial AAC (18–21). In addition, the alkyl chains of the three CLs are markedly disordered (see Fig. 1), with a substantial number of *trans-gauche* defects, congruent with the behavior in a CL-containing lipid bilayer (22). This disorder and mobility of the CLs is mirrored in the distance RMSDs of Fig. 2, which can be as large as 10 Å for the entire lipids, reduced to 3 Å for their headgroup, firmly attached to the membrane carrier between its three amphipathic helices. Consistent with the recent work of Hedger et al. (23), our simulation indicates that CLs seldom extend above the middle of the lipid bilayer, let alone to the cytoplasmic side, at variance with the NMR observations in DPC (16).

As can be seen in Fig. 3, for two residues involved in NOEs (16), namely, I14 and L282, lying not too far from the CLs, the distance separating the terminal methyl groups from the nitrogen atom of the amino acid can be appreciably large—typically >10 Å on average. More importantly, regardless of the residue, the range over which this distance is distributed exceeds 10 Å, thereby supporting the view of highly flexible alkyl chains and the absence of specific binding to the carrier (see the Supporting Material). From Figs. S2 and S3, it is clear that in a native-like membrane environment, the CL acyl chains cannot interact with the residues on the cytoplasmic side of the protein and hence are unlikely to give rise to NOEs, in stark contrast to observations in DPC. Our simulation demonstrates unambiguously that the CL acyl chains evolve far from many residues giving rise to NOE signals, hence suggesting that to satisfy the NOE constraints, the mitochondrial carrier must adopt a conformation in DPC distinct from the crystallographic one (10,13). To ascertain that the dynamics of the CL acyl chains does not depend on the environment, we performed a separate MD simulation, wherein yAAC3 binding the three CLs is embedded in a DPC micelle (see Fig. S1). Just like in a POPC bilayer, the 100-ns trajectory reveals that the CL acyl chains extend too far from the residues involved in NOEs to rationalize the measured signals (16) (see the Supporting Material).

To probe protein unfolding induced by DPC, we have applied two different thermal-shift assays (24) in DPC and DDM using 1) a maleimide coumarin fluorophore, 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM), and 2) dye-free differential scanning fluorimetry (nanoDSF) (see Supporting Material). As a control, we have added the specific inhibitor CATR, which brings the carrier in an

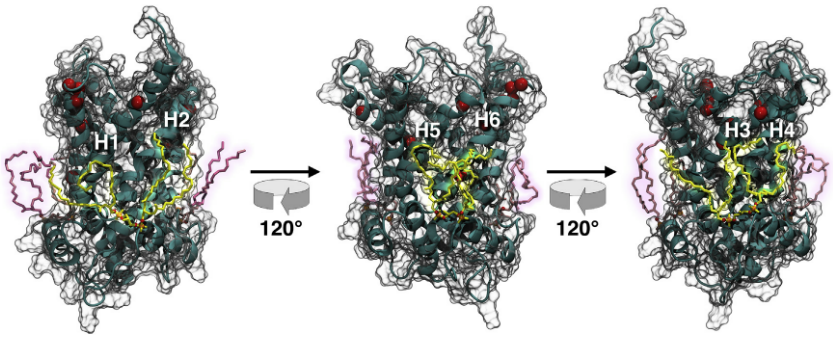


FIGURE 1 Snapshot of yAAC3 after ~150 ns of a 200 ns MD simulation. The membrane carrier is represented as a transparent solvent-accessible surface with its secondary structure inside. The amide sites for which NOE crosspeaks to CL atoms were observed (16), namely, I14, I100, G118, G127, F220, L282, S292, Q298, M299, and I300, are highlighted as red van der Waals spheres. The three CLs binding yAAC3 are shown in yellow and pink to distinguish the region of the protein being examined. To see this figure in color, go online.

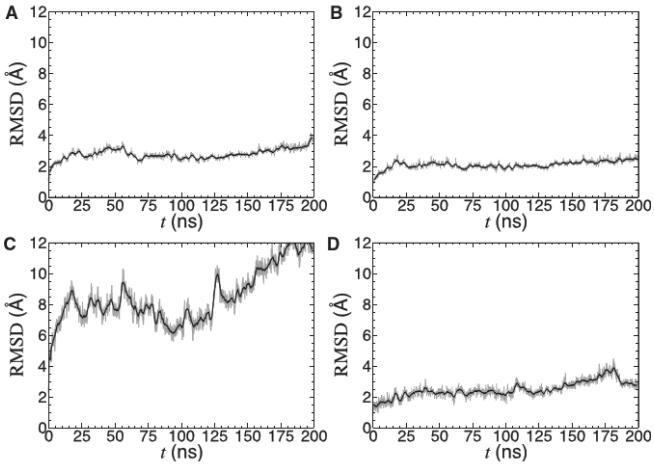


FIGURE 2 Distance root mean-square deviation (RMSD) from the starting configuration computed over the heavy atoms of yAAC3 (*A*), the α-helical content of yAAC3 (*B*), the acyl-chain heavy atoms of the three CLs binding the mitochondrial carrier (*C*), and the heavy atoms of the CL headgroups (*D*). The instantaneous values of the RMSD are shown in gray and the running average in black.

aborted cytoplasmic state (13), leading to enhanced stability. Purified yAAC3 in DDM displayed a typical protein melting curve consistent with thermal denaturation of a folded protein, whether CPM or nanoDSF was used, giving apparent melting temperatures of 47.8 or 50.2°C, respectively (see Fig. 4). In the presence of CATR, a large shift in the melting temperature is observed, leading to apparent melting temperatures of 80.4 and 81.0°C, respectively. However, when the same yAAC3 preparation was diluted into DPC, high fluorescence signals at the start of the experiment were observed in both assays, showing no thermal transition, indicative of an unliganded AAC3 in an unfolded state. Addition of CATR did not alter the profiles, underscoring that yAAC3 in DPC has lost the ability to bind the inhibitor.

The apparent discrepancy between the reported experimental results (16) and the distances between the same atoms measured in our simulation, which are unlikely to produce detectable NOEs, calls into question the fold adopted by yAAC3 in DPC and the CL binding, which necessarily must be different from those in the crystal structures obtained from a native protein. There is a further hint for a different interaction in a native protein and in the refolded protein in DPC. In a native protein extracted from the

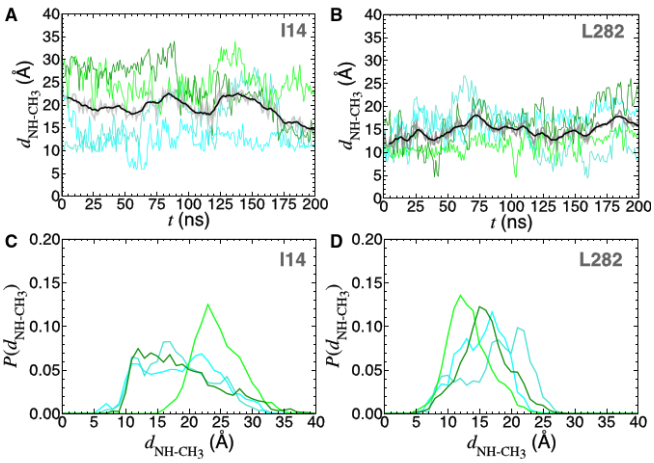


FIGURE 3 Distance separating the –NH hydrogen atom of residues I14 (*A*) and L282 (*B*) from the four terminal methyl groups of the closest CL. The four distances are shown as cyan, turquoise, light green, and dark green curves, alongside their mean, represented by a thick gray curve, and the running average thereof, represented as a black curve. Probability distributions of the four aforementioned hydrogen-methyl distances for residues I14 (*C*) and L282 (*D*). To see this figure in color, go online.

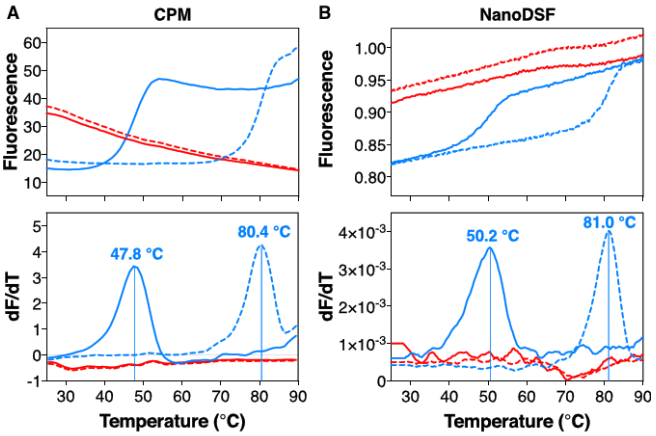


FIGURE 4 Thermostability of yAAC3 in 0.1% DDM (*blue line*) or 0.1% DPC (*red line*), in the presence (*dashed line*) and absence (*straight line*) of CATR. Thermostability was monitored using (*A*) the maleimide coumarin fluorophore CPM (24) and (*B*) nanoDSF (see Supporting Material). The top panels show the changes in fluorescence with temperature, whereas the bottom panels show the derivatives and the apparent melting temperatures, where they could be determined. To see this figure in color, go online.

membrane, CL molecules are very tightly bound, as they cannot be removed by extensive washes with detergent solutions (17,25). The samples used for the crystal structures were obtained with yAAC3 purified from yeast mitochondria, and in this purification and crystallization protocol, no CL was present in the buffers, yet a clear density was discernible for the CL headgroups. This feature strongly suggests that the affinity of CL to native yAAC3 is very high. This property is shared with CATR, which binds with nanomolar affinity and also remains bound through the purification stages, though it is absent from the buffer. In such cases of high-affinity binding, addition of the binding partner, i.e., CLs, is expected to lead to disappearance of the NMR resonances corresponding to the free state and appearance of peaks characteristic of the bound state, which corresponds to a so-called slow-exchange regime. In sharp contrast, the NMR results in DPC reveal a clear signature of fast exchange upon addition of CLs, with gradually shifting crosspeaks and titration curves characteristic of low-affinity binding, often associated with nonspecific interactions, which might arise due to complementary charges on the protein and CL.

Binding of the inhibitor CATR to yAAC3 has also been shown to be distinct in DPC, compared to other environments, with a dissociation constant, K_d, of ~150 μM (26). However, this inhibitor is extremely toxic, and all studies published hitherto have shown that K_d values fall within the low-nanomolar range (27–30), in line with its toxicity, inhibiting the transporter in an aborted state. Interestingly enough, the most recent biochemical investigation reported a value of K_d equal to 192 μM (31). As it turns out, this surprising value was a typographical error, later corrected to 192 nM (32).

Taken together, our observations indicate that the binding features of CLs to yAAC3 witnessed in DPC are markedly distinct from those characteristic of a native-like membrane environment. Specifically, NOEs in DPC are not in agreement with the distance distributions observed from state-of-the-art simulations of the native structure embedded in a lipid bilayer. In fact, our simulation strongly suggests that to satisfy the NOE data, the mitochondrial carrier in DPC must be distorted. In other words, the interactions at play reflect nonspecific binding, likely to be driven by electrostatic interactions of the CL headgroups with an unfolded mitochondrial carrier. This claim is supported by TSAs, which, in the case of yAAC3 in DPC, reveals no temperature transition, in stark contrast with the same protein in the milder DDM environment.

SUPPORTING MATERIAL

Supporting Materials and Methods and four figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)31031-7.

AUTHOR CONTRIBUTIONS

F.D., P.S., E.R.S.K., and C.C. designed the research. F.D., P.S., M.S.K., E.R.S.K., and C.C. performed the research. F.D., P.S., E.R.S.K., and C.C. wrote the article; and all authors participated in the discussions and commented on the article.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Laboratoire International Associé Centre National de la Recherche Scientifique and University of Illinois at Urbana-Champaign. They are indebted to the Grand Équipement National de Calcul Intensif (GENCI) and the Centre Informatique National de l'Enseignement Supérieur (CINES) for generous provision of computer time. This work was supported by the European Research Council (ERC-StG-2012-311318-ProtDyn2Function) to P.S., and by program U105663139 of the UK Medical Research Council to M.S.K. and E.R.S.K.

REFERENCES

1. Maffeo, C., and A. Aksimentiev. 2009. Structure, dynamics, and ion conductance of the phospholamban pentamer. *Biophys. J.* 96:4853–4865.
2. Zhou, H. X., and T. A. Cross. 2013. Influences of membrane mimetic environments on membrane protein structures. *Annu. Rev. Biophys.* 42:361–392.
3. Zoonens, M., S. Masscheleyn, …, F. Dehez. 2013. Dangerous liaisons between detergents and membrane proteins. The case of mitochondrial uncoupling protein 2. *J. Am. Chem. Soc.* 135:15174–15182.
4. Crichton, P. G., Y. Lee, …, E. R. S. Kunji. 2015. Trends in thermostability provide information on the nature of substrate, inhibitor, and lipid interactions with mitochondrial carriers. *J. Biol. Chem.* 290:8206–8217.
5. Holzmann, N., C. Chipot, …, F. Dehez. 2016. Assessing the physiological relevance of alternate architectures of the p7 protein of hepatitis C virus in different environments. *Bioorg. Med. Chem.* 24:4920–4927.
6. Warschawski, D. E. 2015. Membrane Proteins of Known Structure Determined by NMR. http://www.drorlist.com/nmr/MPNMR.html.
7. Oxenoid, K., and J. J. Chou. 2005. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl. Acad. Sci. USA.* 102:10870–10875.
8. Berardi, M. J., W. M. Shih, …, J. J. Chou. 2011. Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature.* 476:109–113.
9. OuYang, B., S. Xie, …, J. J. Chou. 2013. Unusual architecture of the p7 channel from hepatitis C virus. *Nature.* 498:521–525.
10. Pebay-Peyroula, E., C. Dahout-Gonzalez, …, G. Brandolin. 2003. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature.* 426:39–44.
11. Verardi, R., L. Shi, …, G. Veglia. 2011. Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proc. Natl. Acad. Sci. USA.* 108:9101–9106.
12. Cook, G. A., and S. J. Opella. 2011. Secondary structure, dynamics, and architecture of the p7 membrane protein from hepatitis C virus by NMR spectroscopy. *Biochim. Biophys. Acta.* 1808:1448–1453.
13. Ruprecht, J. J., A. M. Hellawell, …, E. R. S. Kunji. 2014. Structures of yeast mitochondrial ADP/ATP carriers support a domain-based alternating-access transport mechanism. *Proc. Natl. Acad. Sci. USA.* 111:E426–E434.
14. Guo, Y., R. C. Kalathur, …, W. A. Hendrickson. 2015. Protein structure. Structure and activity of tryptophan-rich TSPO proteins. *Science.* 347:551–555.
15. King, M. S., M. Kerr, …, E. R. S. Kunji. 2016. Formation of a cytoplasmic salt bridge network in the matrix state is a fundamental step in the transport mechanism of the mitochondrial ADP/ATP carrier. *Biochim. Biophys. Acta.* 1857:14–22.
16. Zhao, L., S. Wang, …, J. J. Chou. 2016. Specific lipid binding of membrane proteins in detergent micelles characterized by NMR and molecular dynamics. *Biochemistry.* 55:5317–5320.
17. Klingenberg, M. 2009. Cardiolipin and mitochondrial carriers. *Biochim. Biophys. Acta.* 1788:2048–2058.
18. Falconi, M., G. Chillemi, …, A. Desideri. 2006. Structural dynamics of the mitochondrial ADP/ATP carrier revealed by molecular dynamics simulation studies. *Proteins.* 65:681–691.
19. Dehez, F., E. Pebay-Peyroula, and C. Chipot. 2008. Binding of ADP in the mitochondrial ADP/ATP carrier is driven by an electrostatic funnel. *J. Am. Chem. Soc.* 130:12725–12733.
20. Wang, Y., and E. Tajkhorshid. 2008. Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc. Natl. Acad. Sci. USA.* 105:9598–9603.
21. Johnston, J. M., S. Khalid, and M. S. P. Sansom. 2008. Conformational dynamics of the mitochondrial ADP/ATP carrier: a simulation study. *Mol. Membr. Biol.* 25:506–517.
22. Aguayo, D., F. D. González-Nilo, and C. Chipot. 2012. Insight into the properties of cardiolipin containing bilayers from molecular dynamics simulations, using a hybrid all-atom/united-atom force field. *J. Chem. Theory Comput.* 8:1765–1773.
23. Hedger, G., S. L. Rouse, …, M. S. P. Sansom. 2016. Lipid-loving ANTs: molecular simulations of cardiolipin interactions and the organization of the adenine nucleotide translocase in model mitochondrial membranes. *Biochemistry.* 55:6238–6249.
24. Alexandrov, A. I., M. Mileni, …, R. C. Stevens. 2008. Microscale fluorescent thermal stability assay for membrane proteins. *Structure.* 16:351–359.
25. Bamber, L., M. Harding, …, E. R. S. Kunji. 2006. Yeast mitochondrial ADP/ATP carriers are monomeric in detergents. *Proc. Natl. Acad. Sci. USA.* 103:16224–16229.
26. Brüschweiler, S., Q. Yang, …, J. J. Chou. 2015. Substrate-modulated ADP/ATP-transporter dynamics revealed by NMR relaxation dispersion. *Nat. Struct. Mol. Biol.* 22:636–641.
27. Vignais, P. V., P. M. Vignais, and G. Defaye. 1973. Adenosine diphosphate translocation in mitochondria. Nature of the receptor site for carboxyatractyloside (gummiferin). *Biochemistry.* 12:1508–1519.
28. Klingenberg, M., K. Grebe, and B. Scherer. 1975. The binding of atractylate and carboxy-atractylate to mitochondria. *Eur. J. Biochem.* 52:351–363.
29. Krämer, R., and M. Klingenberg. 1977. Reconstitution of inhibitor binding properties of the isolated adenosine 5′-diphosphate,adenosine 5′-triphosphate carrier-linked binding protein. *Biochemistry.* 16:4954–4961.
30. Krämer, R. 1983. Interaction of membrane surface charges with the reconstituted ADP/ATP-carrier from mitochondria. *Biochim. Biophys. Acta.* 735:145–159.
31. Babot, M., C. Blancard, …, V. Trézéguet. 2012. Mitochondrial ADP/ATP carrier: preventing conformational changes by point mutations inactivates nucleotide transport activity. *Biochemistry.* 51:7348–7356.
32. Babot, M., C. Blancard, …, V. Trézéguet. 2016. Correction to Mitochondrial ADP/ATP carrier: preventing conformational changes by point mutations inactivates nucleotide transport activity. *Biochemistry.* 55:2422.

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

Cite This: J. Phys. Chem. Lett. 2018, 9, 933−938

pubs.acs.org/JPCL

How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine

Vilius Kurauskas,[†] Audrey Hessel,[†] Peixiang Ma,[†] Paola Lunetti,[‡] Katharina Weinhäupl,[†] Lionel Imbert,[†] Bernhard Brutscher,[†] Martin S. King,[§] Rémy Sounier,[∥] Vincenza Dolce,[⊥] Edmund R. S. Kunji,[§] Loredana Capobianco,*[,‡] Christophe Chipot,[#] François Dehez,*[,#] Beate Bersch,[†] and Paul Schanda*[,†]

[†]Université Grenoble Alpes, CNRS, CEA, IBS, 38000 Grenoble, France

[‡]Department of Biological and Environmental Sciences and Technologies, University of Salento, 73100 Lecce, Italy

[§]MRC-MBU, University of Cambridge, Cambridge CB2 0XY, United Kingdom

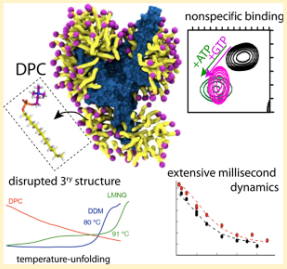
[∥]CNRS, INSERM, Université de Montpellier, 34094 Montpellier, France

[⊥]Dept of Pharmacy, University of Calabria, 87036 Arcavacata di Rende, Italy

[#]LPCT, UMR 7019 Université de Lorraine, CNRS and Laboratoire International Associé & University of Illinois at Urbana−Champaign, F-54500 Vandoeuvre-lès-Nancy, France

Supporting Information

ABSTRACT: Characterizing the structure of membrane proteins (MPs) generally requires extraction from their native environment, most commonly with detergents. Yet, the physicochemical properties of detergent micelles and lipid bilayers differ markedly and could alter the structural organization of MPs, albeit without general rules. Dodecylphosphocholine (DPC) is the most widely used detergent for MP structure determination by NMR, but the physiological relevance of several prominent structures has been questioned, though indirectly, by other biophysical techniques, e.g., functional/thermostability assay (TSA) and molecular dynamics (MD) simulations. Here, we resolve unambiguously this controversy by probing the functional relevance of three different mitochondrial carriers (MCs) in DPC at the atomic level, using an exhaustive set of solution-NMR experiments, complemented by functional/TSA and MD data. Our results provide atomic-level insight into the structure, substrate interaction and dynamics of the detergent−membrane protein complexes and demonstrates cogently that, while high-resolution NMR signals can be obtained for MCs in DPC, they systematically correspond to nonfunctional states.



Mitochondrial carriers (MCs) constitute a large family of transport proteins of ca. 30−35 kDa molecular weight that play important roles in the intracellular translocation of metabolites, nucleotides, and coenzymes.[1] The transport mechanism involves the alternating exposure of a central cavity to the cytosolic side (c-state) or the mitochondrial matrix (m-state).[2] Crystallographic structures of the most prominent member, the adenosine diphosphate (ADP)/adenosine triphosphate (ATP) carrier (AAC), extracted from the native mitochondrial membrane[3,4] could only be obtained in the presence of the strong inhibitor carboxyatractyloside (CATR), which locks the carrier in an aborted c-state. These structures revealed a central cavity, in which CATR is bound, surrounded by six transmembrane helices. No other atomic-resolution structures of MCs are available, underscoring the difficulty of obtaining crystals of these dynamic proteins. Solution-NMR may overcome the difficulties that crystallography faces for dynamic MPs. Substrate, lipid, and inhibitor binding[5−10] as well as dynamics,[8,11] detected by NMR have been reported and related to functional mechanisms of this class of membrane proteins. Other biophysical techniques, e.g., functional/thermostability assays and molecular-dynamics simulations, have challenged some of these interpretations.[12−14] Solving this apparent controversy has been hampered by the lack of direct experimental atomic-resolution data and stringent control experiments. Here, we resort to an extensive set of NMR experiments to address the structure, substrate interactions and dynamics of MCs in dodecylphosphocholine (DPC) alongside with functional/thermostability assays and molecular-dynamics simulations and conclude on their physiological significance.

To probe the structural integrity of the yeast ADP/ATP carrier 3 (AAC3) in different detergents, we first investigated its thermal stability, consistently with our previous investigations.[13,14] We produced AAC3 in the yeast mitochondrial membrane, extracted it with the mild detergent dodecylmaltoside (DDM), diluted it into solutions of either DDM, lauryl maltoside neopentyl glycol (LMNG) or DPC, and performed

Received: January 26, 2018
Accepted: February 2, 2018
Published: February 3, 2018

thermostability shift assay (TSA) experiments.[14,15] In DDM and LMNG, a typical protein melting curve reveals a cooperative thermal denaturation event with an apparent melting temperature of ca. 49 °C (Figure 1, blue/green

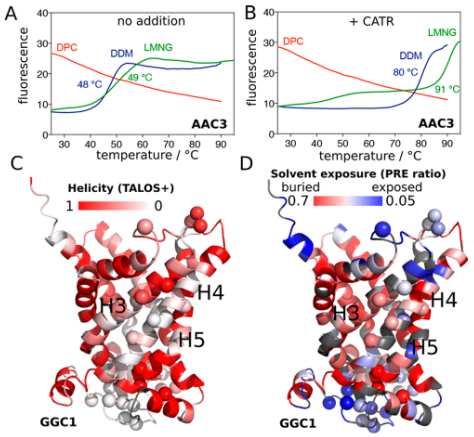


Figure 1. Stability of secondary and tertiary structures in MCs. (A, B) Thermostability of AAC3, extracted from yeast mitochondria and purified in DDM, diluted 20-fold into 0.1% DDM (blue line), 0.1% LMNG (green line) or 0.1% DPC (red line) in the absence (A) or presence (B) of CATR, measured with TSA experiments (see Supplementary Methods). (C) Residue-wise helix propensity in GGC1, determined from NMR chemical shifts using the program TALOS+,[16] plotted onto a structural model of GGC1. Residues in white color are not in a helical conformation. (D) Residue-wise solvent accessibility in GGC1, as probed with the paramagnetic agent gadodiamide. Residues shown in white-to-blue colors are accessible to solvent. In panels C and D, amide sites undergoing μs−ms dynamics (discussed further below) are indicated by spheres.

lines). In the CATR-bound state, the melting temperature is dramatically increased to 80−90 °C. In contrast, the same protein preparation diluted into DPC already shows a high fluorescence signal at low temperature, and no transition in the course of the temperature scan, irrespective of the presence of CATR. These observations indicate that AAC3 in DPC does not form a stable tertiary structure, nor binds CATR, in line with previous findings.[13] To understand this behavior at the atomic level, we turned to solution-NMR of three members of the MC family, namely, AAC3, guanine diphosphate (GDP)/guanosine triphosphate (GTP) carrier (GGC1), and ornithine carrier (ORC1). Given their low sequence identity (<32%, Figure S1), common traits observed in DPC are likely shared by the entire MC family. Samples of AAC3, GGC1, and ORC1, prepared through refolding from inclusion bodies, just like in previous NMR studies of MCs,[5,7,8,11] are all homogeneous (as shown by size-exclusion chromatograms and analytical ultracentrifugation; Figure S2). Consistently with previous NMR studies,[5−9,11] all these samples yield high-quality NMR spectra (Figure S3), a criterion that is commonly employed to assess the functional relevance of MP samples and motivate structural studies. To reconcile these data with TSA experiments and to characterize the structural organization of MCs that is responsible for high-quality NMR spectra, we performed an

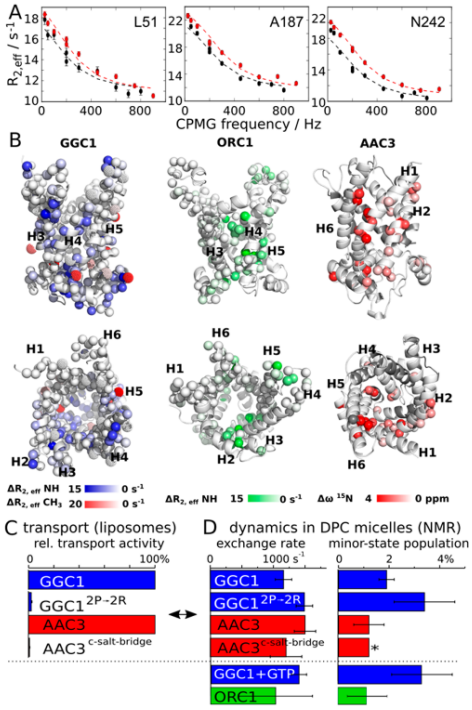


Figure 2. Millisecond dynamics in MCs in DPC detergent micelles. (A) Representative 3D-HNCO-detected CPMG RD profiles of ^{15}N amide sites in GGC1 (WT, no substrate). All nonflat RD profiles of this study are shown in Figures S7 to S17. (B) Location of residues with nonflat CPMG RD curves in GGC1, ORC1, and AAC3 plotted onto structural models based on the crystal structure of AAC3. For GGC1 and ORC1, the color code represents the extent of relaxation-dispersion, as determined from the difference in $R_{2,eff}$ of the first and last points of the RD profiles; for AAC3, the color code represents $\Delta\omega$ values reported in ref 8. A qualitatively similar picture is obtained for two mutants studied here, GGC1$^{2P\rightarrow 2R}$ and AAC3$^{-c\text{-}saltbridge}$; see main text. (C) Transport activity and (D) NMR-derived dynamics parameters in MCs. The p_B value denoted with an asterisk was fixed to the value found for AAC3, which resulted in satisfactory fits. All transport activity and dynamics data reported in this study are provided in Tables S1−S3.

NMR-observed temperature series up to ca. 70 °C (Figure S3). One expects a discontinuity in the NMR parameters if collective unfolding arises. By contrast, we observe linear chemical-shift changes up to high temperatures, mirroring the TSA data, and suggesting that these MCs undergo a gradual change of the structural ensemble without defined denaturation transition.

We probed the structural and dynamical properties of MCs in DPC further with three different NMR observables. First, the residue-specific chemical-shift assignments were used with the TALOS+ software[16] to derive local backbone geometry at the level of each residue (Figures 1C and S4). While there is an overall good agreement between the residues that, according to

DOI: 10.1021/acs.jpclett.8b00269
J. Phys. Chem. Lett. 2018, 9, 933−938

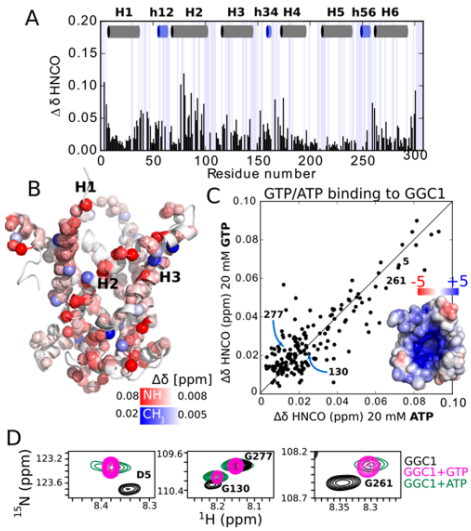


Figure 3. GGC1 in DPC lacks the expected binding specificity. (A) Chemical-shift perturbations in GGC1 upon addition of GTP. (B) Plot of amide (H/N/CO) and methyl (H/C) CSPs onto a structural model of GGC1. (C,D) GTP and ATP produces very similar CSPs in GGC1, as exemplified with extracts from HNCO spectra (D) and shown as a correlation plot in (C). The inset shows the electrostatic surface of a AAC-derived structural model of GGC1. Equivalent data for AAC3 are shown in Figure S21.

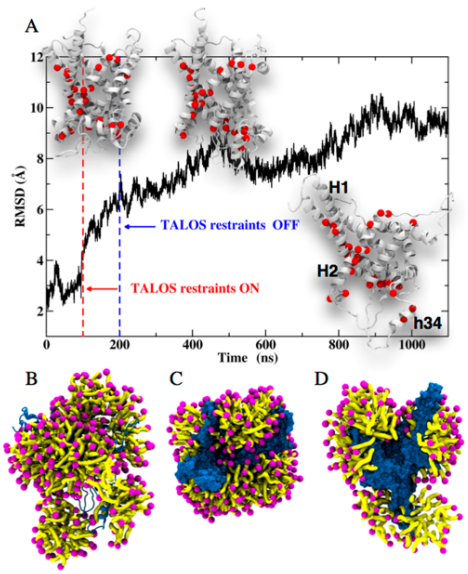


Figure 4. All-atom MD simulations of AAC3 in DPC micelles. (A) Time evolution of the root-mean-square displacement (RMSD) of yAAC3 backbone atomic positions with respect to those in the crystal structure. From 0 to 100 ns, the protein embedded in DPC is free to move. From 100 to 200 ns, a subset of dihedral angles is restrained to TALOS+ values inferred from NMR chemical shifts (see Materials and Methods and Table S4). At 200 ns, restraints are turned off. The secondary structures of the protein before and after applying the TALOS+ restraints and at the end of the trajectory are depicted chronologically in the insets. Amide sites for which μs−ms motion is found in CPMG experiments are highlighted with red spheres. (B−D) Snapshots at 1 μs simulation time showing the DPC (purple/yellow) organization around yAAC3 from side view (B), top view (C), and cut-open side view (D). Supplementary Movie 1 and Movie 2 show the time evolution of AAC3 and its surrounding micelle.

their chemical shifts, have α-helical conformation and the location of helices in the crystal structures of AAC3, several stretches of residues in the TM helices appear disordered in DPC, particularly in helices H4 (GGC1), and H2/H3 (AAC3). Independent support for the flexible nature of these parts comes from ^{15}N transverse relaxation rate constants (R_2), which reveal enhanced flexibility in GGC1's H4 (Figure S5). Second, we performed solvent-paramagnetic relaxation-enhancement (sPRE) experiments with the soluble hydrophilic paramagnetic complex Gd-DTPA-BMA (gadodiamide), which probe the solvent-accessibility of each amide site (Figure 1D). As expected, the most exposed parts are located in loops of GGC1 and AAC3. Unexpectedly, however, sizable sPRE effects are also found in TM helices, in particular those for which TALOS+ finds a loop conformation. Third, we measured amide hydrogen/deuterium exchange (HDX) by dilution of GGC1 from H_2O- into D_2O-buffer. The observed very fast HDX kinetics (<1 min) indicates that backbone hydrogen bonds are marginally stable (Figure S6).

To gain further insight into the properties and possibly functional relevance of MCs in DPC, we investigated the details of their dynamics. Functional turnover of MCs in the membrane occurs at a rate of ca. 500 s^{-1}.[17] Functional MCs are, thus, expected to undergo motions on this time scale, and these motions are expected to be sensitive to substrates or mutations. Carr−Purcell−Meiboom−Gill (CPMG) relaxation-dispersion (RD) NMR experiments[18] are ideally suited to probe μs-ms motions at the level of individual residues. In this study we performed extensive CPMG RD measurements on a total of 10 different samples, including wild-type (WT) and mutant proteins, as well as samples with substrates and lipids, and on up to three magnetic field strengths per sample (see Table S1 and Figures S7−S17); in addition, previous work reported CPMG data for WT AAC3.[8] In AAC3, GGC1 and ORC1, we find extensive μs−ms motions for about 20% of the ^{15}N backbone amide sites, and in all cases the exchanging residues are clustered on one side of the molecule (Figure 2). We have performed additional ^{13}C CPMG experiments on methyl groups of Ala, Val, and Leu residues in GGC1 (Table S1 and Figure S8), independently confirming the presence of μs−ms dynamics in the region revealed by ^{15}N CPMG experiments (red in Figure 2B, left). In order to obtain more insight into this process, we fitted a two-state exchange model to the CPMG RD data. Statistical analyses, shown in Figure S18, indicate that this simple model describes the data satisfyingly. The fitted parameters are similar for all three proteins, with an exchange rate constant, $k_{ex} = k_{AB} + k_{BA}$, on the order of 1000−2000 s^{-1}, and populations of the minor state, p_B, of ca. 1−3% (Figure 2D and Table S1).

DOI: 10.1021/acs.jpclett.8b00269
J. Phys. Chem. Lett. 2018, 9, 933−938

A major challenge in the interpretation of relaxation dispersion data is to understand the underlying motional process structurally. An intriguing observation is that the time scale of motion observed in all three MCs (ca. 1 ms) is similar to the kinetics of transmembrane transport,[17] suggesting a possible role in functional turnover. To test this hypothesis, we investigated whether the introduction of function-abolishing mutations or the presence of substrates and lipids alter the dynamics. We designed two mutants that address different structural/functional aspects of the mitochondrial carriers, reasoning that if they inactivate the proteins, they would do so for different reasons. In a first mutant, termed GGC1$^{2P\rightarrow 2R}$, two highly conserved Pro residues of the MC signature motifs (PX[ED]XX[KR]) in H3 and H5 were mutated to Arg. Given the importance of these Pro kinks in the structure and mechanism,[2] and the unfavorable energetics of placing Arg residues in the core of TM helices,[19] we expected significant changes in the activity and dynamics. A second mutant targets Asp, Lys, and Gln residues on the cytosolic side of the TM helices of AAC3, which have been proposed to form a salt bridge network in the elusive m-state.[4] In the AAC3$^{-c\text{-}saltbridge}$ mutant, we inverted charges such that the putative salt-bridge network cannot form any more. Liposome transport assays show that the residual transport activity in GGC1$^{2P\rightarrow 2R}$ and AAC3$^{-c\text{-}saltbridge}$ is below 3%, i.e., that the mutations render the proteins essentially nonfunctional (Figure 2C and Table S3).

Strikingly, however, the characteristics of the mutants in DPC micelles are hardly different from those of the WT protein samples. NMR spectra of these mutants are highly similar to the respective WT proteins. Chemical-shift differences are small, and restricted to residues in the immediate vicinity of the mutation sites (Figure S19). If the helices were in close contact to each other, as in the crystal structure, one would expect pronounced chemical-shift changes in the neighboring helices in these mutants. Another surprising observation is that the μs−ms dynamics in these nonfunctional mutants is very similar to the one of WT proteins, even though the mutants are nonfunctional, likely suggesting that the μs−ms motions are not directly related to function (Figures 2D, S12−S14 and Table S1). Along the same lines, we find that the presence of substrate (GTP) does not lead to a detectable difference in the dynamics compared to apo-GGC1 (Figure 2D), indicating that substrate cannot influence this dynamic process. Put together, the observed dynamics does not appear to be related to function. We reasoned that the functionally important[2] lipid cardiolipin (CL) might resurrect the native properties of the protein. However, CPMG experiments with CL reveal an identical behavior for WT GGC1 and GGC1$^{2P\rightarrow 2R}$ (see Supplementary Discussion and Figures S9, S11, and S20), showing that these MCs in DPC are nonfunctional, irrespective of added lipids.

An important criterion to assess functionality of a MP is its ability to interact with substrates. MCs are highly substrate-specific transporters. In liposome transport assays, we find that recombinantly expressed and reconstituted AAC3 and GGC1 bind and transport specifically their substrate (ATP/GTP, respectively), while they do not bind the other nucleotide (GTP/ATP; see Table S3). Furthermore, CATR inhibits AAC3, but not GGC1 (see Supplementary Discussion). Testing the binding specificity is, thus, a convenient way to investigate the structural integrity of MCs in DPC. Addition of the substrate GTP to GGC1 leads to significant chemical shift perturbations (CSPs) for many amides and Ala, Leu, and Val methyls (Figure 3A). The largest CSPs are observed for residues lining H2, and on the cytoplasmic side of H1 and H6. Weaker yet significant effects are observed in the three amphipathic matrix helices (Figure 3B); these CSPs are in good agreement with previously reported data.[11] However, addition of ATP, which does not interact with GGC1 in membranes, unexpectedly produces essentially the same CSPs as GTP (Figure 3C,D). Likewise, addition of GTP or ATP to AAC3 produces CSPs that are nearly identical to each other (Figure S21). Thus, while AAC3 and GGC1 in their lipid-bilayer environment can discriminate between GTP and ATP, this ability is lost in DPC micelles. We repeated GTP-titration experiments with the nonfunctional GGC1$^{2P\rightarrow 2R}$ mutant. Even though this protein is nonfunctional and likely distorted through the Arg in the TM part, the GTP-induced CSPs are very similar to those observed in WT protein (Figure S22). This observation suggests that the observed GTP interaction is not dependent on structural integrity, and is thus nonspecific.

We then titrated WT GGC1 with CATR, a doubly negatively charged inhibitor of AAC, but not of GGC1.[20] Despite the absence of GGC1/CATR interaction in membranes, the addition of CATR to GGC1 in DPC produces significant CSPs. The affected set of residues partly coincides with the sites that are modulated by the presence of GTP and ATP, and the effects are similar to those observed in AAC3 upon addition of CATR (Figure S23). These observations lead us to propose that the reported interactions of CATR with GGC1 or AAC3 in DPC are nonspecific and not reminiscent of binding in biological membranes. Indeed, the affinity of CATR to AAC3 in DPC (K_d = 20−150 μM[8]) is ca. 3 orders of magnitude weaker than the AAC/CATR interaction in lipid bilayers (K_d = 10−20 nM[21]) or of AAC extracted from mitochondria with LAPAO detergent (K_d = 310 nM[21]). A plausible driving force for all these nonspecific low-affinity interactions is the electrostatic attraction between the negatively charged solutes (nucleotides, CATR) and the positively charged cavity of GGC1 and AAC3 (Figures 3C and S21D). In agreement with this view, ORC1, which has a significantly lower electrostatic charge, does not show significant CSPs upon addition of its substrate, L-ornithine (Figure S24).

To obtain a mechanistic understanding of the structural organization of the protein and detergent, we turned to all-atom molecular dynamics (MD) simulations in explicit DPC. MD simulations have been applied previously to MCs in a membrane environment[22−24] and in DPC.[12] We specifically sought a description of the structural organization of AAC3 in DPC, which would reflect our experimental data. In initial unrestrained all-atom simulations over 0.7 μs, AAC3 remains close to the conformation adopted in crystals,[4] which is in disagreement with the TALOS-derived secondary structure in DPC. Presumably, the time scale accessible to simulation is too short to allow the transition of the compact conformation in crystals to the one in DPC micelles (see Supplementary Discussion and Figures S25, S26). To better reflect the experimental data, we, thus, used chemical-shift derived backbone torsional angles (limited to the ϕ,ψ angles for which the TALOS+ result was unambiguous; see Materials and Methods), and introduced a soft harmonic potential to these residues over a simulation period of 0.1 μs. This strategy, which rests on the use of NMR restraints to guide conformational analyses of proteins, is similar in spirit to that proposed by others.[25,26] When the torsional angles coincide with those inferred from the NMR data, the overall structure of AAC3 is

DOI: 10.1021/acs.jpclett.8b00269
J. Phys. Chem. Lett. 2018, 9, 933−938

markedly distorted, corresponding to a shift from ca. 3 to 6 Å of the root-mean-square deviation (RMSD) of atomic positions (Figure 4A). The evolution of the tertiary structure upon introduction of the TALOS+ restraints is asymmetrical and chiefly involves the interfacial helix h34 (connecting H3 and H4) and the cytoplasmic side of H1, H2, and H3 (Figure 4 and S27). In a last step, the restraints were removed, and the simulation was extended to 1.1 μs. Over this time scale, AAC3 remains distorted. The cytoplasmic side of the TM helices, H1 and H2, and the interfacial helix h34 undergo relatively slow dynamics on the simulation time scale (see Movie 1). Interestingly, this region coincides with the segments that were found experimentally to undergo μs−ms motions (red spheres in Figure 4). We speculate that the motions detected experimentally might correspond to the kind of slow rearrangements seen by MD. Most importantly, the simulations provide a view of how the detergent may organize around the protein. Figures 4B−D show that DPC molecules do not only form a corona around the hydrophobic core of the protein, as one might expect, but also form micelles interacting with both sides of the protein (cf. Movie 2). DPC molecules protrude between the TM helices and penetrate deep within the protein cavity (Figure 4C,D and Movie 2).

DPC has been employed for solving about 40% of all known NMR structures, but less than 1% of all crystal structures (Figure S28). Our data provide possible clues about the origin of this intriguing track record. In our hands, DPC is the only detergent able to maintain MCs in solution for extended time periods (a prerequisite for solution-NMR), and the embedded proteins exhibit well-resolved spectra and solute interactions that may be interpreted as signs of a functionally relevant sample. Our study reveals that these observations do not necessarily point to a functionally relevant protein sample, and stresses the importance of stringent control experiments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.8b00269.

Movie 1: 1.1 s molecular-dynamics simulation of yeast AAC3 in DPC (MPG)

Movie 2: Same as Movie 1 including the explicit representation of DPC organization around yAAC3 (AVI)

Detailed experimental materials and methods, tables summarizing the results of transport activity measurements, CPMG fits and assigned chemical shifts, and figures with supporting experimental data (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: paul.schanda@ibs.fr.
*E-mail: francois.dehez@univ-lorraine.fr.
*E-mail: loredana.capobianco@unisalento.it.

ORCID

Bernhard Brutscher: 0000-0001-7652-7384
Christophe Chipot: 0000-0002-9122-1698
Paul Schanda: 0000-0002-9350-7606

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Klaus Zangger for providing gadodiamide, and Astrid C. Sivertsen, Guillaume Bouvignies, Eva Pebay-Peyroula, Stéphanie Ravaud, Ewen Lescop, Cécile Breyton, Sven Brüschweiler, Christine Ebel and Aline LeRoy for insightful discussions, Jason R. Schnell for help in preparing Figure S28 and Michael Kohlhaas for inspiring thoughts. This work was supported by the European Research Council (ERC-Stg-311318) to P. S. and the UK Medical Research Council (U105663139) to E.R.S.K. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

■ REFERENCES

(1) Palmieri, F.; Monné, M. Discoveries, Metabolic Roles and Diseases of Mitochondrial Carriers: A Review. *Biochim. Biophys. Acta, Mol. Cell Res.* **2016**, *1863*, 2362−2378.

(2) Kunji, E. R. S.; Aleksandrova, A.; King, M. S.; Majd, H.; Ashton, V. L.; Cerson, E.; Springett, R.; Kibalchenko, M.; Tavoulari, S.; Crichton, P. G.; et al. The Transport Mechanism of the Mitochondrial ADP/ATP Carrier. *Biochim. Biophys. Acta, Mol. Cell Res.* **2016**, *1863*, 2379−2393.

(3) Pebay-Peyroula, E.; Dahout-Gonzalez, C.; Kahn, R.; Trezeguet, V.; Lauquin, G.; Brandolin, G. Structure of Mitochondrial ADP/ATP Carrier in Complex with Carboxyatractyloside. *Nature* **2003**, *426*, 39−44.

(4) Ruprecht, J. J.; Hellawell, A. M.; Harding, M.; Crichton, P. G.; McCoy, A. J.; Kunji, E. R. S. Structures of Yeast Mitochondrial ADP/ATP Carriers Support a Domain-Based Alternating-Access Transport Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E426−34.

(5) Run, C.; Yang, Q.; Liu, Z.; OuYang, B.; Chou, J. J. Molecular Basis of MgATP Selectivity of the Mitochondrial SCaMC Carrier. *Structure* **2015**, *23*, 1394−1403.

(6) Zhao, L.; Wang, S.; Zhu, Q.; Wu, B.; Liu, Z.; OuYang, B.; Chou, J. J. Specific Interaction of the Human Mitochondrial Uncoupling Protein 1 with Free Long-Chain Fatty Acid. *Structure* **2017**, *25*, 1371−1379.e3.

(7) Zhao, L.; Wang, S.; Run, C.; OuYang, B.; Chou, J. J. Specific Lipid Binding of Membrane Proteins in Detergent Micelles Characterized by NMR and Molecular Dynamics. *Biochemistry* **2016**, *55*, 5317−5320.

(8) Brüschweiler, S.; Yang, Q.; Run, C.; Chou, J. J. Substrate-Modulated ADP/ATP-Transporter Dynamics Revealed by NMR Relaxation Dispersion. *Nat. Struct. Mol. Biol.* **2015**, *22*, 636−641.

(9) Berardi, M. J.; Shih, W. M.; Harrison, S. C.; Chou, J. J. Mitochondrial Uncoupling Protein 2 Structure Determined by NMR Molecular Fragment Searching. *Nature* **2011**, *476*, 109−113.

(10) Berardi, M. J.; Chou, J. J. Fatty Acid Flippase Activity of UCP2 Is Essential for Its Proton Transport in Mitochondria. *Cell Metab.* **2014**, *20*, 541−552.

(11) Sounier, R.; Bellot, G.; Chou, J. J. Mapping Conformational Heterogeneity of Mitochondrial Nucleotide Transporter in Uninhibited States. *Angew. Chem., Int. Ed.* **2015**, *54*, 2436−2441.

(12) Zoonens, M.; Comer, J.; Masscheleyn, S.; Pebay-Peyroula, E.; Chipot, C.; Miroux, B.; Dehez, F. Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2. *J. Am. Chem. Soc.* **2013**, *135*, 15174−15182.

(13) Dehez, F.; Schanda, P.; King, M. S.; Kunji, E. R. S.; Chipot, C. Mitochondrial ADP/ATP Carrier in Dodecylphosphocholine Binds Cardiolipins with Non-Native Affinity. *Biophys. J.* **2017**, *113*, 2311−2315.

(14) Crichton, P. G.; Lee, Y.; Ruprecht, J. J.; Cerson, E.; Thangaratnarajah, C.; King, M. S.; Kunji, E. R. S. Trends in Thermostability Provide Information on the Nature of Substrate,

Inhibitor, and Lipid Interactions with Mitochondrial Carriers. *J. Biol. Chem.* **2015**, *290*, 8206−8217.
(15) Alexandrov, A. I.; Mileni, M.; Chien, E. Y. T.; Hanson, M. A.; Stevens, R. C. Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. *Structure* **2008**, *16*, 351−359.
(16) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. *J. Biomol. NMR* **2009**, *44*, 213−223.
(17) Gropp, T.; Brustovetsky, N.; Klingenberg, M.; Muller, V.; Fendler, K.; Bamberg, E. Kinetics of Electrogenic Transport by the ADP/ATP Carrier. *Biophys. J.* **1999**, *77*, 714−726.
(18) Mittermaier, A. K.; Kay, L. E. Observing Biological Dynamics at Atomic Resolution Using NMR. *Trends Biochem. Sci.* **2009**, *34*, 601−611.
(19) Zhou, H.-X.; Cross, T. A. Influences of Membrane Mimetic Environments on Membrane Protein Structures. *Annu. Rev. Biophys.* **2013**, *42*, 361−392.
(20) Vozza, A.; Blanco, E.; Palmieri, L.; Palmieri, F. Identification of the Mitochondrial GTP/GDP Transporter in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **2004**, *279*, 20850−20857.
(21) Krämer, R.; Klingenberg, M. Reconstitution of Inhibitor Binding Properties of the Isolated Adenosine 5′-diphosphate, Adenosine 5′-triphosphate Carrier-Linked Binding Protein. *Biochemistry* **1977**, *16*, 4954−49561.
(22) Wang, Y.; Tajkhorshid, E. Electrostatic Funneling of Substrate in Mitochondrial Inner Membrane Carriers. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9598−9603.
(23) Dehez, F.; Pebay-Peyroula, E.; Chipot, C. Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel. *J. Am. Chem. Soc.* **2008**, *130*, 12725−12733.
(24) Hedger, G.; Rouse, S. L.; Domanski, J.; Chavent, M.; Koldso, H.; Sansom, M. S. P. Lipid-Loving ANTs: Molecular Simulations of Cardiolipin Interactions and the Organization of the Adenine Nucleotide Translocase in Model Mitochondrial Membranes. *Biochemistry* **2016**, *55*, 6238−6249.
(25) Vallurupalli, P.; Hansen, D. F.; Kay, L. E. Structures of Invisible, Excited Protein States by Relaxation Dispersion NMR Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11766−11771.
(26) Robustelli, P.; Kohlhoff, K.; Cavalli, A.; Vendruscolo, M. Using NMR Chemical Shifts as Structural Restraints in Molecular Dynamics Simulations of Proteins. *Structure* **2010**, *18*, 923−933.

DOI: 10.1021/acs.jpclett.8b00269
J. Phys. Chem. Lett. 2018, 9, 933−938

L'ensemble des travaux de simulations présentés dans cette partie a permis d'établir de manière claire l'effet délétère du DPC sur les protéines membranaires composées d'hélices alpha et a permis de discriminer les structures non physiologiques de certaines protéines membranaires présentes dans la PDB. Un article de revue discutant de l'ensemble de cette problématique a été publié dans Chemical Review.

III. Les récepteurs canaux pentamériques activés par des ligands

Le récepteur de la sérotonine 5-HT3 est un canal ionique pentamérique activé par un ligand (pLGIC).[78] Il appartient à la grande famille de récepteurs qui fonctionnent comme des transducteurs de signal allostériques au travers de la membrane plasmique suite à la liaison de molécules neurotransmettrices aux sites extracellulaires.[79] Ces récepteurs subissent des transitions conformationnelles complexes qui se traduisent par une ouverture transitoire d'un pore perméable aux ions.[80,81] Les récepteurs 5-HT3 sont des cibles thérapeutiques pour les vomissements et les nausées, le syndrome du côlon irritable et la dépression.[82] Les récepteurs 5-HT3 sont structurellement et fonctionnellement distincts des six autres classes de récepteurs 5-HT, qui appartiennent tous à la famille des récepteurs activés par les protéines G.[83] En dépit de plusieurs structures pLGIC publiées dans la littérature, une vision d'ensemble claire des transitions conformationnelles impliquées dans la synchronisation du canal manque cruellement. Le groupe de Hugues Nury à l'IBS à grâce à la Cryo-microscopie électronique a résolu un ensemble de 4 structures du récepteur complet de souris 5-HT3 en complexe avec le tropisétron (un antiémétique) avec la sérotonine, et avec la sérotonine et un modulateur allostérique positif, à des résolutions allant de 3.2 à 4.5 Å (Figure 21). La structure liée au tropisétron ressemble à celles obtenues en présence d'un inhibiteur (nanobody) ou sans ligand. Les autres structures comprennent un état « ouvert » et deux états « intermédiaires » liés à un ligand. Grâce aux simulations numériques, nous avons pu accéder à des informations sur la dynamique de ces structures. Nous avons notamment établi les profils d'hydratation des pores et les profils d'énergie libre associés au passage d'ions, ce qui a conduit à la caractérisation des mouvements au niveau de la porte du canal et l'accessibilité des cations dans les pores. Ces résultats ont permis d'établir à un niveau de détail inédit les mécanismes d'activation des pLGIC ainsi que la topologie du site de liaison du ligand (Figure 22).[84]

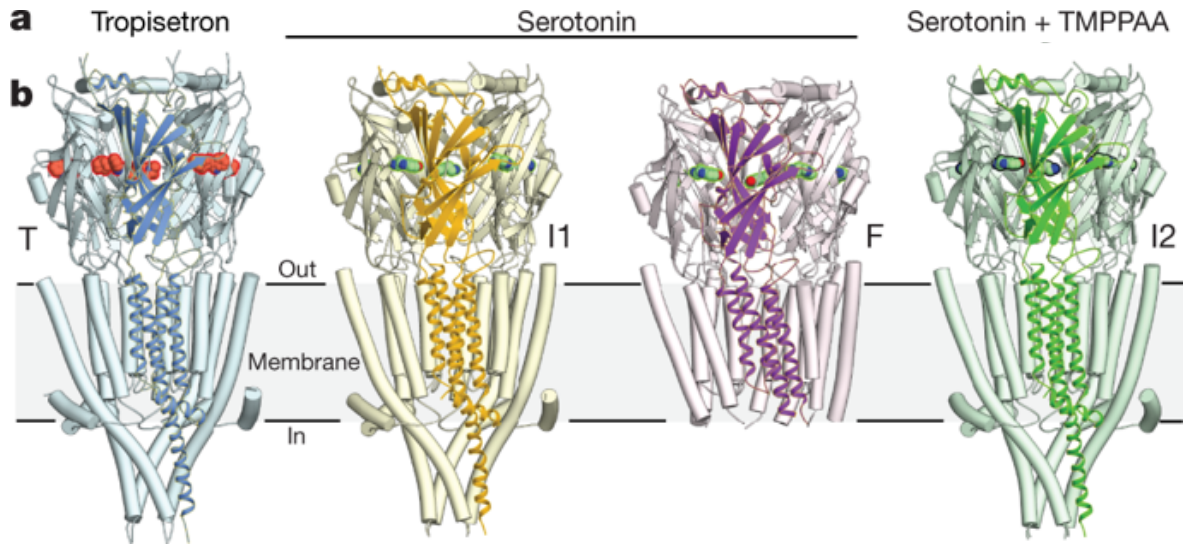


Figure 21. Structures représentant différents états conformationnels du récepteur 5-HT3 le long de son chemin d'activation, obtenues par Cryo-microscopie électronique. Une structure inhibée par le tropisétron (T), une structure « ouverte » par la sérotonine (ligand naturel de 5-HT3) (F) et deux structures intermédiaires (I1 et I2) observées en présence de sérotonine et d'un mélange sérotonine/TMPPAA (modulateur allostérique).

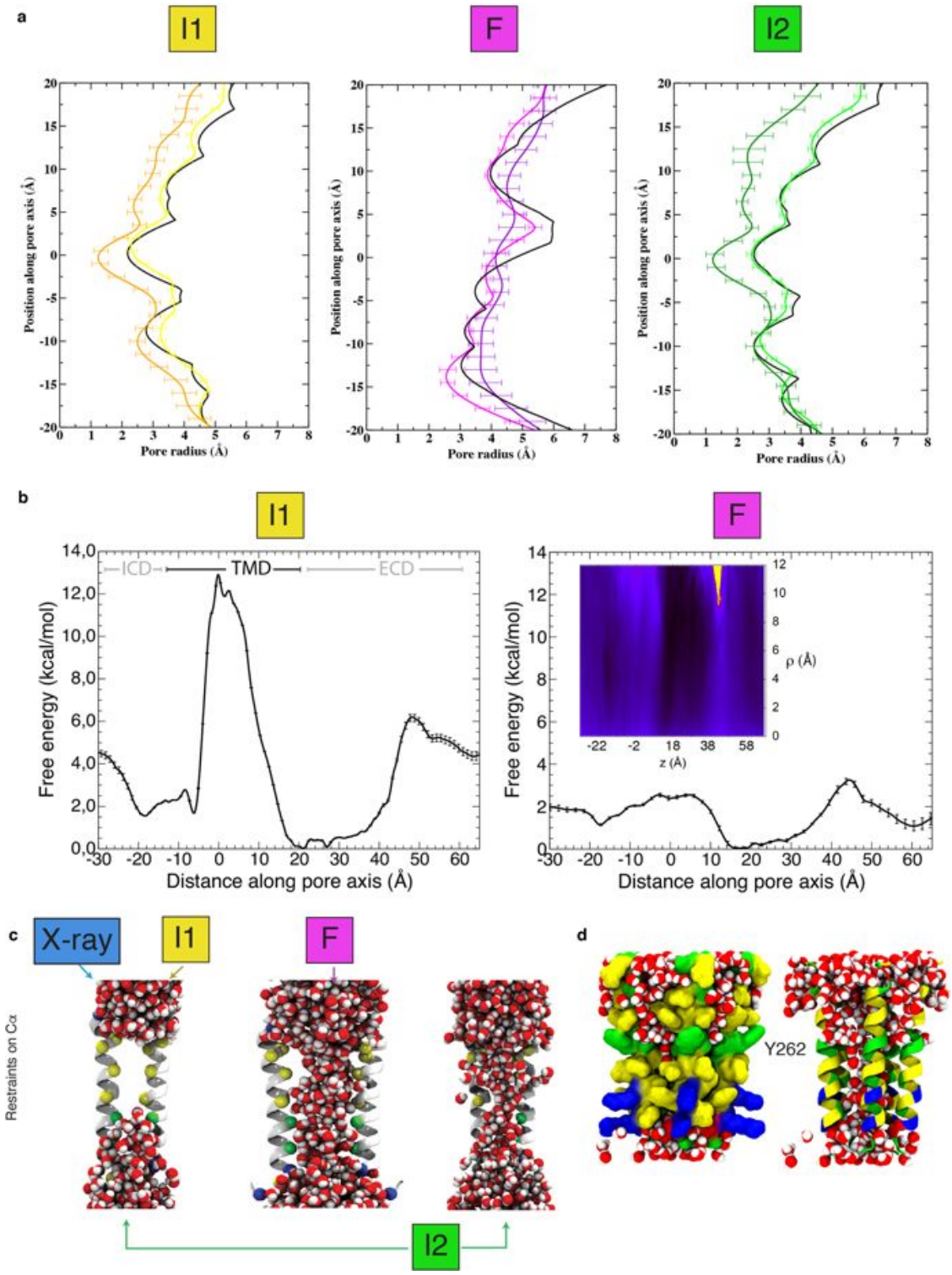


Figure 22. Caractérisation par dynamique moléculaire des différentes structures de 5-HT3 observées par Cryo-microscopie électronique. (a) Profils de pores transmembranaires (HOLE) pour les trajectoires à partir de I1, F et I2. Dans chaque graphique, le profil de la structure Cryo-EM est représenté en noir, le profil moyen pendant la partie restreinte dans la couleur plus claire et le profil moyen au cours de la partie non restreinte (60-1 000 ns) de couleur plus foncée. (b) Potentiels de force moyenne associes à la translocation des ions K^+ le long de l'axe des pores. L'encart représente le paysage 2D d'énergie libre en fonction de la position le long de l'axe des pores et de la direction radiale, orthogonale à celui-ci. Le profil d'énergie libre représenté a été moyenné sur 2,16 µs pour la structure I1 et sur 3,04 µs pour la structure F. (c) Instantanés représentatifs de l'hydratation des pores lors de la première partie restreinte (30 à 60 ns) des simulations. Un pore déshydraté est observé pour les trajectoires T et I1, un pore complètement hydraté

pour la trajectoire F, alors que pour la trajectoire I2 est observé une alternance d'hydratation/déshydratation. (d) Instantané prélevé au cours de la trajectoire I2 montrant l'hydratation à l'extérieur du pore (représentation en surface (gauche) et en hélice (droite)) corrélée à l'hydratation à l'intérieur du canal.

Ci-après sont reproduits 2 articles illustrant mes travaux dédiés aux canaux pentamériques :

- Polovinkin, L., Hassaine, G., Perot, J., Neumann, E., Jensen, A. A., Lefebvre, S. N., Corringer, P.-J., Neyton, J., Chipot, C., Dehez, F., Schoehn, G. & Nury, H. Conformational transitions of the serotonin 5-HT 3 receptor. *Nature* **563**, 275–279 (2018).
- Zarkadas, E., Zhang, H., Cai, W., Effantin, G., Perot, J., Neyton, J., Chipot, C., Schoehn, G., Dehez, F. * & Nury, H. * The binding of palonosetron and other antiemetic drugs to the serotonin 5-HT3 receptor. *Structure*, **28**, 1131-1140 (2020)

LETTER

https://doi.org/10.1038/s41586-018-0672-3

Conformational transitions of the serotonin 5-HT_3 receptor

Lucie Polovinkin[1], Ghérici Hassaine[2], Jonathan Perot[1], Emmanuelle Neumann[1], Anders A. Jensen[3], Solène N. Lefebvre[4], Pierre-Jean Corringer[4], Jacques Neyton[1]*, Christophe Chipot[5,6,7], Francois Dehez[5,6], Guy Schoehn[1] & Hugues Nury[1]*

The serotonin 5-HT_3 receptor is a pentameric ligand-gated ion channel (pLGIC). It belongs to a large family of receptors that function as allosteric signal transducers across the plasma membrane[1,2]; upon binding of neurotransmitter molecules to extracellular sites, the receptors undergo complex conformational transitions that result in transient opening of a pore permeable to ions. 5-HT_3 receptors are therapeutic targets for emesis and nausea, irritable bowel syndrome and depression[3]. In spite of several reported pLGIC structures[4–8], no clear unifying view has emerged on the conformational transitions involved in channel gating. Here we report four cryo-electron microscopy structures of the full-length mouse 5-HT_3 receptor in complex with the anti-emetic drug tropisetron, with serotonin, and with serotonin and a positive allosteric modulator, at resolutions ranging from 3.2 Å to 4.5 Å. The tropisetron-bound structure resembles those obtained with an inhibitory nanobody[5] or without ligand[9]. The other structures include an 'open' state and two ligand-bound states. We present computational insights into the dynamics of the structures, their pore hydration and free-energy profiles, and characterize movements at the gate level and cation accessibility in the pore. Together, these data deepen our understanding of the gating mechanism of pLGICs and capture ligand binding in unprecedented detail.

A decade after the structure of the *Torpedo marmorata* nicotinic acetylcholine receptor[10] (nAChR), the set of known pLGIC structures is rapidly expanding and reflects the diversity of this protein family. The structures share a conserved architecture, in which subunits are arranged around a central five-fold pseudo-symmetry axis. Together they have clarified details of ligand binding, selectivity and allosteric modulation. They have also revealed a complex landscape of conformations, raising questions of how to relate structures to the wealth of data that established the existence of multiple agonist-bound preactive intermediate states[11–13], of distinct open states[14] and of multiple desensitized states[15].

Mouse homomeric 5-HT_{3A} receptors, with their entire intracellular domain (ICD), were solubilized with the detergent C12E9 and purified. We first performed cryo-electron microscopy (cryo-EM) in the presence of the potent antagonist tropisetron, and obtained a 4.5 Å structure (Fig. 1b, Extended Data Figs. 1, 2, Extended Data Table 1), hereafter referred to as T. T is globally similar to the structure previously solved by X-ray crystallography[5] (root mean square deviation (r.m.s.d.) of 0.6 Å), the pore of which was shown by molecular dynamics to be occluded[16]. Tropisetron fits in a peanut-shaped density present in the neurotransmitter pocket (Extended Data Fig. 3d–f). The ICD contains a region of about 60 residues, which is averaged out (also in the other reconstructions, see below) because of its intrinsic flexibility[1,2]. T resembles the 4.5 Å cryo-EM structure of the apo 5-HT_3 receptor[9] (r.m.s.d. of 1.15 Å), with small differences in the lipid-exposed helices M3, MX and M4.

We then sought to identify agonist-elicited conformations of the 5-HT_3 receptor, and performed cryo-EM imaging in the presence of serotonin. A first reconstruction presented heterogeneity in the membrane domain. Further focused 3D classification allowed two subsets of particles to be separated, which yielded reconstructions at 4.2 Å and 4.1 Å resolution, corresponding to two different conformations (Fig. 1b, Extended Data Fig. 4). The maps offer a variable level of information: most side chains in the extracellular domain are resolved, whereas some parts of the transmembrane domain (TMD) do not have side-chain information and some have limited information in the main chain position (Extended Data Figs. 4c, 5), reflecting the receptor dynamics. In the two refined structures, the extracellular domains (ECDs) have undergone an equivalent transition from the T state and serotonin could be modelled in the neurotransmitter site, whereas the TMDs differed markedly (Fig. 2, Extended Data Fig. 6). We call the first structure I1 for intermediate 1 and the second structure F for full, on the basis of the extent of movements compared to the inhibited state. I1 exhibits only limited displacements in the upper part of M1 and M2, and a rearrangement of the M2–M3 loop (Supplementary Video 1). F is characterized by a pronounced reorganization of the transmembrane helices, which can be described by a rigid-body movement of the four-helix bundle coupled to a rearrangement of M4 (and of M3 to a lesser extent) sliding along M1 and M2 (Supplementary Videos 2, 3). F also features a very dynamic ICD, beyond the intrinsically disordered region, in which model building was not possible even though the data showed incomplete densities for MX and M4 (Extended Data Fig. 4a–c).

Finally, we collected a dataset in the presence of serotonin and trans-3-(4-methoxyphenyl)-N-(pentan-3-yl)acrylamide (TMPPAA, a compound exhibiting agonist and positive allosteric modulator activity on the human receptor[17]), a combination that yields weakly desensitizing currents (Extended Data Fig. 1f). From this dataset, we reconstructed a 3.2 Å resolution structure (Extended Data Fig. 2d–g), which provides non-ambiguous side-chain information for nearly the entire receptor. The refined structure has an ECD conformation essentially equivalent to that of I1 and F. The membrane domain is similar to that of I1, albeit with a slightly more expanded top section and pore (Fig. 2). We call this structure I2 for intermediate 2.

Serotonin can be unambiguously positioned in the neurotransmitter site of I2. It fits tightly within its binding pocket (Extended Data Fig. 3a–c) in an orientation consistent with functional and binding studies[1]. Surrounded by obligatory aromatic residues (F199 and Y207 on the principal side, Y126 and W63 on the complementary side), it is positioned to form a cation-π interaction with W156 and hydrogen bonds with the main chain of S155 and Y64. The C loop is positioned moderately inward relative to the inhibited conformations, its position locked by a salt bridge between D202 and R65. A hallmark of allosteric activation is the subunit–subunit rearrangement (Extended Data Fig. 3d), which affects the site volume and geometry.

[1]CNRS, Université Grenoble Alpes, CEA, IBS, Grenoble, France. [2]Theranyx, Marseille, France. [3]Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [4]Channel Receptors Unit, CNRS UMR 3571, Institut Pasteur, Paris, France. [5]Université de Lorraine, CNRS, LPCT, Nancy, France. [6]Laboratoire International Associé CNRS and University of Illinois at Urbana-Champaign, Vandoeuvre-les-Nancy, France. [7]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA. *e-mail: jacques.neyton@ibs.fr; hugues.nury@ibs.fr

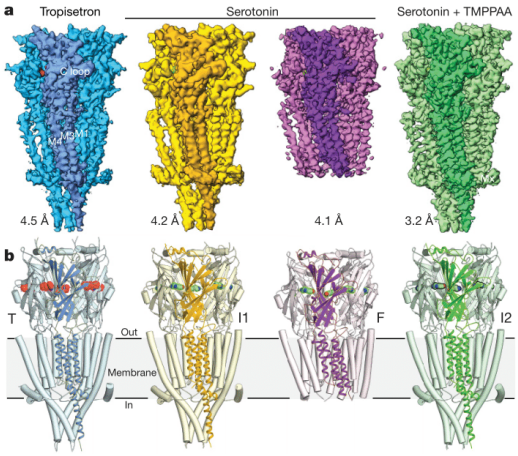


Fig. 1 | Three-dimensional reconstructions and structures of homomeric 5-HT$_{3A}$ receptor. a, **b**, Reconstructions (**a**) and structures (**b**) for: the tropisetron dataset (protein in blue and ligand in red), the serotonin and Ca^{2+} dataset (I1 in yellow, F in purple and ligand in green), and the serotonin and TMPPAA dataset (I2 and ligand in green). Resolutions are shown according to the Fourier shell coefficient 0.143 criterion.

TMPPAA has previously been proposed to bind to an allosteric site in the TMD[17], but there is no clearly interpretable density in our data to model the compound. We tested TMPPAA agonist activity on a set of around 45 single-point mutants of the human receptor, which collectively reveal that the drug binds between M4, M1 and M3 into an intra-subunit cavity skirted by lipids of the outer leaflet in the upper part of the TMD (Extended Data Fig. 7, Supplementary Table 1), where endogenous steroids bind to nAChRs[18]. More generally, several allosteric druggable sites have been identified in pLGICs, which bind diverse compounds including general anaesthetics such as propofol and flavourings such as citral or eucalyptol[3]. Allosteric sites in the 5-HT$_3$ receptor change in both shape and volume during transitions (Extended Data Figs. 3i, 7a).

At the ECD–TMD interface, a set of conserved residues that are essential for gating[19] form a structural motif that is common to all pLGIC structures, the location of which may correlate with the state of the channel[5]. This motif consists of charged residues (E53, D145, E186 and R218) sandwiched between conserved aromatic residues: W187 at the top and the 142-FPF-144 motif of the Cys loop, plus Y223 at the bottom (Extended Data Fig. 8). The FPF motif itself penetrates the transmembrane domain similar to a wedge; its position differs in each conformation. When superimposing structures on a TMD subunit, the wedge lies close to M2 in the tropisetron-bound structure and moves towards M4–M1 in the I1, I2 and F structures (which is possible because the conserved P230 allows the upper M1 to kink or straighten). A marked downward concerted movement of the wedge and of the β8–β9 loop (containing E186 and W187) occurs in the F structure, pushing on the M2–M3 loop, and may be responsible for the marked reorganization of the TMD observed in that state.

pLGIC pores are lined with side chains of residues from the five M2 helices. In direct agreement with results from substituted cysteine accessibility mutagenesis[1] (SCAM), our structures show that in all conformations, positions −1′, 2′, 6′, 9′, 13′, 16′, 17′ and 20′ of the M2 α-helix are exposed to the pore lumen (Fig. 3a). Positions 12′ and 15′, which are also accessible in SCAM, are partly exposed to solvent on the rear of the M2 helices. Superimposition of a single M2 helix underscores its flexibility at both ends; superimposition of the five M2 helices highlights the crucial role of movements of the hydrophobic side chain at the 9′ position (Fig. 3b, c). Minimum pore diameter is often a key element in the assignment of experimental structures to physiological states. Pore profiles are compatible with closed hydrophobic gates (rings at positions 9′, 13′, 16′ and 17′) in the T and I1 structures, with an open channel in the F structure, whereas the I2 structure presents an intermediate profile. However, pore profiles are influenced by the resolution, symmetry and rotameric state of the side chains pointing into the pore lumen. Some positions, such as the key negative charge in −1′ are often poorly resolved in density maps, and are known to adopt alternate conformations[8,20]. Moreover, pore profiles are not informative about hydrophobicity, wetting or dynamics, which have key roles in permeability[21,22]. We performed molecular dynamics simulations to better characterize permeation. In the microsecond-long trajectories starting from the inhibited X-ray structure or from I1, no water or ions cross the pore, and the hydrophobic 9′, 13′ and 16′ rings establish a de-wetted hydrophobic gate that is tighter than in the starting structures (Fig. 3d, Extended Data Fig. 9). By sharp contrast, a simulation starting from F features an open pore that is accessible to ions and water throughout the trajectory (Supplementary Video 4). During the initial part of the simulations starting from I2, when the C_α atoms are positionally restrained, wetting and de-wetting events of the pore occur as if its conformation were on the open–closed verge (Extended Data Fig. 9). Wetting is linked to the presence of transiently hydrated grooves at the back of M2 helices, down to the polar residues Y11′–S12′, affecting the electrostatic landscape inside the pore. Wetting also correlates with rotation of L9′ out of the pore lumen. Once the geometric restraints are removed, the I2 pore relaxes to a closed conformation similar to that observed in the I1 trajectory. The absence of TMPPAA in the simulation may rationalize the closure. Potentials of mean force for the translocation of a K^+ ion reveal an insurmountable 12 kcal mol^{-1} barrier in the case of I1 (representative of closed hydrophobic gate conformations), and an essentially flat free-energy landscape in the case of F (Fig. 3e).

We asked whether the structures could be assigned to physiological states. T and F are straightforward to assign, whereas I1 and I2 are less so. T typifies an inhibited state, with resting-like ECD stabilized by tropisetron, and a closed pore, resembling the apo state. F represents an open state, with an activated ECD with bound serotonin, an activated TMD and a wide open pore. Two assignments are possible for the closely related I1 and I2 conformations (Fig. 4a). In a first scheme, I1 exemplifies a serotonin-bound, pre-active closed state, in which the ECD and the ECD–TMD interface—but not the TMD—have undergone a transition. This is consistent with single-channel analysis of the 5-HT$_3$ receptor, which yields kinetic models in which opening can occur from a ligand-bound pre-active state[14]. In a second scheme, I1 represents a closed desensitized state that occurs downstream from the open state. In both schemes, I2 is best described as in a state close to I1 (that is, close to either pre-active or desensitized) wherein the slightly wider pore promotes wetting, which could enable ion passage or merely favour the switch to a fully open state, consistent with the TMPPAA-induced modulation.

A distinctive feature of the second scheme is that the activation gate, consisting of rings of hydrophobic residues in the upper pore, would open in the active state and close in the desensitized state. This implies a marked movement of the upper pore during desensitization, and no ion access from the extracellular compartment to the lower pore in the desensitized state. Functional experiments on anionic receptors indicate that they have distinct activation and desensitization gates, the latter being located at the cytoplasmic end of the pore[23]. Moreover, structures of the GABA$_A$ receptor β3 homopentamer and of the α4β2 nAChR, with pores constricted at positions −2′ and −1′, have been assigned to desensitized states[6,8]. These receptors, however, were engineered close to the constriction, a feature that can alter desensitization in 5-HT$_3$ receptors[24].

To challenge the two schemes, we measured the movement of the gate region and the accessibility of the pore to organic cations. First, we

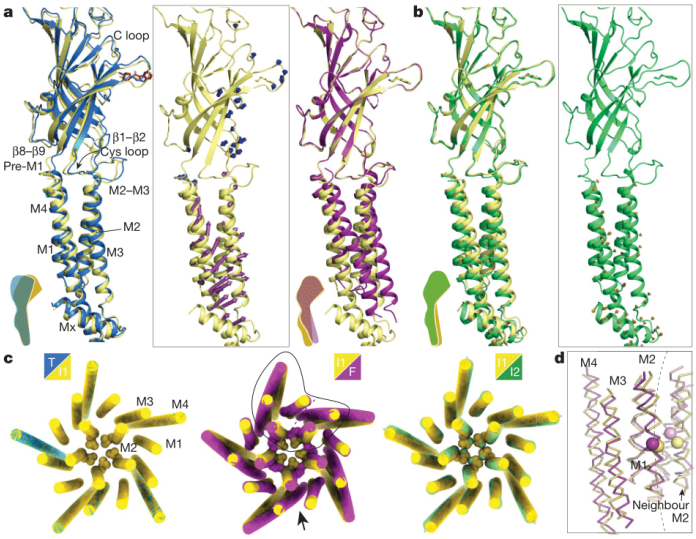


Fig. 2 | Transitions between the tropisetron-bound, the serotonin-bound, and the serotonin and TMPPA-bound states. a, View parallel to the membrane of one subunit of the serotonin-bound conformation I1 (yellow) overlaid with the tropisetron-bound conformation (blue, TMD superposition, left) or with the serotonin-bound conformation F (purple, ECD superposition, right). Inset, vectors indicate the local amplitude of movements, sampled on C_α atoms (T to I1 in blue, I1 to F in purple). **b**, Overlay of the I1 (yellow) and the serotonin and TMPPAA-bound I2 (green) conformations (left); vector representations of the I1 to I2 transition (ECD superposition). **c**, Pairwise overlays of the TMD illustrating transitions at the quaternary level. Structures were superimposed on the ECD pentamer. L9′ (L260) residues are shown as spheres. The line and arrow on the middle overlay indicate the region depicted in **d** and the orientation of the view, respectively. **d**, Tertiary reorganization within a subunit TMD. Overlay in ribbon representation of I2 and F with L260 C_α as spheres (TMD superposition within a single subunit). Note the vertical shift of M3–M4 relative to M1–M2, distortions on the extracellular halves of M1–M2, and interface re-arrangement with the neighbouring subunit helices (on the right of the dotted line).

used voltage-clamp fluorometry (VCF) to probe the local conformational changes in the upper pore at the 19′ position[25,26] (Extended Data Fig. 10). We labelled S19′C mutants, expressed at the surface of *Xenopus laevis* oocytes, with 5-carboxytetramethylrhodamine methanethiosulfonate (MTS-TAMRA). Transient stimulation with serotonin elicited simultaneous changes in current and fluorescence with parallel rise

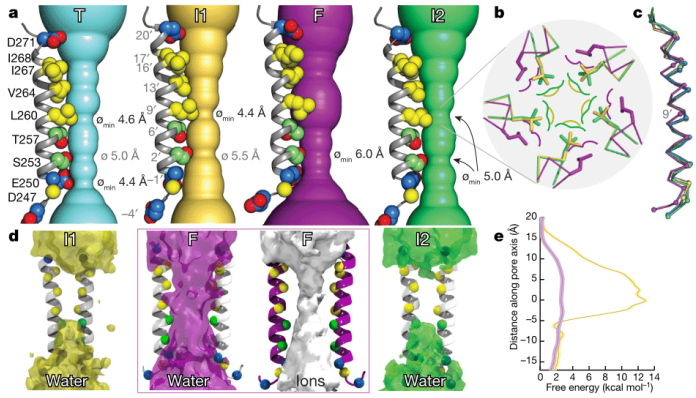


Fig. 3 | Geometry, wetting and energy landscape of the transmembrane pore. a, Static pore geometry; the accessible pathway through the pore is represented as a solid surface for each structure. Diameters (∅) of the constriction zones are noted. One M2 helix is depicted as cartoon with pore-exposed side chains as spheres (polar in green, hydrophobic in yellow, charged in blue). The view is parallel to the plane of the membrane. **b**, Zoom on the hydrophobic gate constriction, formed at the level of L260 (sticks), which has a small movement backward in the I2 state and rotates outwards in the F state. Coloured lines indicate molecular surfaces. The view is perpendicular to the membrane plane. **c**, Superposition of one M2 helix in the four conformations. **d**, Water densities of the pore region during the unrestrained part of simulations. Densities are depicted as transparent surfaces at the same contouring level. The density for K^+ ions is also included for the F trajectory featuring an open pore. **e**, Potentials of mean force of K^+ ions as a function of the position along the pore axis.

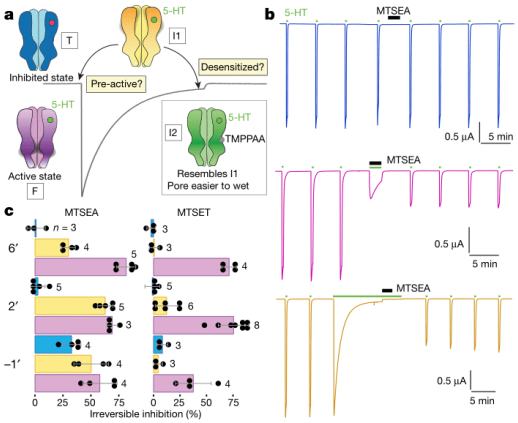


Fig. 4 | Putative molecular mechanisms of operation. a, T (blue) represents an inhibited state, stabilized by tropisetron (red circle). I1 (yellow) represents either a closed pre-active serotonin-bound (green circle) state or a desensitized state. F (purple) represents an open active state. The grey line illustrates the electrical response to serotonin recorded in an oocyte expressing 5-HT_{3A} receptors. **b**, Recordings of MTSEA (1mM) modification on serotonin-evoked current in the S2′C mutant. MTSEA is applied in the resting (blue), the desensitized (orange), or active (purple) state. The protocol includes 10-s test serotonin applications before and after 2-min MTSEA modification. **c**, Changes in current after MTSEA (1 mM) or MTSET (1 mM) modification, for −1′C, S2′C and T6′C mutants, in the resting (blue), active (purple) or desensitized (yellow) state.

and decay, whereas prolonged (7-min) exposure to serotonin resulted in similar signals at the onset, followed by a slow decay for the current signal without a change in fluorescence. Both signals returned to baseline when serotonin was removed. The VCF results argue against the second scheme, because the probe environment changes upon activation but not during desensitization. Second, we performed SCAM in the resting (absence of ligand), open (during transient serotonin application) and desensitized (after prolonged serotonin application) states (Fig. 4b). MTSEA (2-aminoethyl methanethiosulfonate) and MTSET (2-2-(trimethylammonium)ethyl methanethiosulfonate) are organic cations that react with free cysteines through their methanethiosulfonate moiety. The aminoethyl head group of MTSEA is small, which enables it to access narrow spaces. MTSEA is, however, also known to cross membranes in its uncharged form. Application of 1 mM extracellular MTSEA in the resting state yielded no modification of currents in T6′C and S2′C mutants but inhibited currents (by 32%) in E1′C mutants (Fig. 4c). In the resting state, the compound can probably access position −1′ from the internal compartment, but it is unable to access the 2′ or 6′ positions. Similar applications in the open state produced irreversible inhibitions at 6′, 2′ and −1′ positions (82%, 69% and 58%, respectively). MTSEA can therefore reach 6′ and 2′ from the extracellular compartment in the open state, consistent with previous studies[1] and with the opening of the activation gate in F. Applications of extracellular MTSEA in the desensitized state also produced irreversible inhibitions at 6′, 2′ and −1′ positions (30%, 63% and 50%, respectively). From these results, we infer that positions 6′ and 2′, located below the hydrophobic gate seen in T and I1, are accessible from the extracellular side in the desensitized state. No conclusions can be drawn for position −1′, as its labelling in the resting state precludes further interpretation. The results suggest that the activation gate is open in the desensitized conformation(s). MTSET possesses a bulkier trimethylammonium headgroup and cannot cross membranes. We observed that MTSET labels cysteine mutants at 6′, 2′ and −1′ positions when applied in the open state (71%, 75% and 46%, respectively). Application of MTSET in the desensitized state yielded a small variable inhibition at 2′ (12 ± 10%) and none at 6′, indicating that these positions are more accessible for MTSEA than for the bulkier MTSET. On the basis of these results, we favour the scheme in which I1 is assigned to a pre-active state. Nevertheless, desensitization is a complex process involving several distinct states[15], and we cannot rule out the possibility that I1 represents a desensitized state.

The challenge of matching structures to states without ambiguity transcends the present study and pertains to the whole field of pLGIC structures[2,27,28]. This challenge arises from diverse factors: limited resolution, putative influence of detergent, crystal packing and receptor engineering, and the possibility that ensembles of multiple related conformations are necessary to properly depict a physiological state. Bearing in mind the ambiguities on state assignment, we compared the 5-HT_3 transitions to those observed for the *Gloeobacter violaceus* receptor (GLIC)[4], the worm glutamate-gated receptor (GluCl)[29] and the glycine α1 receptor (GlyR)[7]. Common agonist-induced features emerge, such as a global twist, quaternary reorganization of the ECD, rearrangement of the interface between domains involving the conserved sandwich motif, and local movements of the upper TMD. Differences also appear: the extent of TMD reorganization seen in F, with M4 sliding on other helices, is not observed in the other receptors; the ECD reorganization is well-described as 'un-blooming'[4] for GLIC and GluCl, but not for the 5-HT_3 and glycine receptors. The open pore of F is wider than that of GLIC and narrower than that of GlyR. Bacterial, animal anionic and animal cationic channels may have evolved distinct sets of conformations for a given physiological state, as they belong to separate branches of the pLGIC family. Our 5-HT_3 receptor structures highlight several transitions in the cationic branch. They also contribute to knowledge on other important aspects of pLGIC research that are only alluded to in this report, such as the role of M4 in gating, the pharmacology of allosteric sites and ICD dynamics. Further work with better resolution, structures of mutant receptors and structures of receptors in complex with other ligands will complement and increase mechanistic insights, but knowledge of this area may nevertheless remain incomplete until kinetic structural experiments come of age.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0672-3.

Received: 7 February 2018; Accepted: 7 September 2018;
Published online 31 October 2018.

1. Thompson, A. J., Lester, H. A. & Lummis, S. C. R. The structural basis of function in Cys-loop receptors. *Q. Rev. Biophys.* **43**, 449–499 (2010).
2. Nemecz, Á., Prevost, M. S., Menny, A. & Corringer, P.-J. Emerging molecular mechanisms of signal transduction in pentameric ligand-gated ion channels. *Neuron* **90**, 452–470 (2016).
3. Sparling, B. A. & DiMauro, E. F. Progress in the discovery of small molecule modulators of the Cys-loop superfamily receptors. *Bioorg. Med. Chem. Lett.* **27**, 3207–3218 (2017).
4. Sauguet, L. et al. Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proc. Natl Acad. Sci. USA.* **111**, 966–971 (2014).
5. Hassaïne, G. et al. X-ray structure of the mouse serotonin 5-HT_3 receptor. *Nature* **512**, 276–281 (2014).
6. Miller, P. S. & Aricescu, A. R. Crystal structure of a human $GABA_A$ receptor. *Nature* **512**, 270–275 (2014).
7. Du, J., Lü, W., Wu, S., Cheng, Y. & Gouaux, E. Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526**, 224–229 (2015).
8. Morales-Perez, C. L., Noviello, C. M. & Hibbs, R. E. X-ray structure of the human α4β2 nicotinic receptor. *Nature* **538**, 411–415 (2016).
9. Basak, S. et al. Cryo-EM structure of 5-HT_{3A} receptor in its resting conformation. *Nat. Commun.* **9**, 514 (2018).
10. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J. Mol. Biol.* **346**, 967–989 (2005).
11. Mukhtasimova, N., Lee, W., Wang, H. & Sine, S. Detection and trapping of intermediate states priming nicotinic receptor channel opening. *Nature* **459**, 451–454 (2009).
12. Lape, R., Colquhoun, D. & Sivilotti, L. G. On the nature of partial agonism in the nicotinic receptor superfamily. *Nature* **454**, 722–727 (2008).
13. Menny, A. et al. Identification of a pre-active conformation of a pentameric channel receptor. *eLife* **6**, e23955 (2017).
14. Corradi, J., Gumilar, F. & Bouzat, C. Single-channel kinetic analysis for activation and desensitization of homomeric 5-HT_{3A} receptors. *Biophys. J.* **97**, 1335–1345 (2009).
15. Sakmann, B., Patlak, J. & Neher, E. Single acetylcholine-activated channels show burst-kinetics in presence of desensitizing concentrations of agonist. *Nature* **286**, 71–73 (1980).
16. Trick, J. L. et al. Functional annotation of ion channel structures by molecular simulation. *Structure* **6**, 2207–2216 (2016).
17. Gasiorek, A. et al. Delineation of the functional properties and the mechanism of action of TMPPAA, an allosteric agonist and positive allosteric modulator of 5-HT_3 receptors. *Biochem. Pharmacol.* **110–111**, 92–108 (2016).
18. Jin, X. & Steinbach, J. H. A portable site: a binding element for 17β-estradiol can be placed on any subunit of a nicotinic α4β2 receptor. *J. Neurosci.* **31**, 5045–5054 (2011).
19. Mukhtasimova, N. & Sine, S. M. Nicotinic receptor transduction zone: invariant arginine couples to multiple electron-rich residues. *Biophys. J.* **104**, 355–367 (2013).
20. Cymes, G. D. & Grosman, C. The unanticipated complexity of the selectivity-filter glutamates of nicotinic receptors. *Nat. Chem. Biol.* **8**, 975–981 (2012).
21. Zhu, F. & Hummer, G. Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction. *Biophys. J.* **103**, 219–227 (2012).
22. Aryal, P., Sansom, M. S. P. & Tucker, S. J. Hydrophobic gating in ion channels. *J. Mol. Biol.* **427**, 121–130 (2015).
23. Gielen, M., Thomas, P. & Smart, T. G. The desensitization gate of inhibitory Cys-loop receptors. *Nat. Commun.* **6**, 6829 (2015).
24. McKinnon, N. K., Bali, M. & Akabas, M. H. Length and amino acid sequence of peptides substituted for the 5-HT_{3A} receptor M3M4 loop may affect channel expression and desensitization. *PLoS ONE* **7**, e35563 (2012).
25. Pless, S. A., Dibas, M. I., Lester, H. A. & Lynch, J. W. Conformational variability of the glycine receptor M2 domain in response to activation by different agonists. *J. Biol. Chem.* **282**, 36057–36067 (2007).
26. Dahan, D. S. et al. A fluorophore attached to nicotinic acetylcholine receptor βM2 detects productive binding of agonist to the αδ site. *Proc. Natl Acad. Sci. USA* **101**, 10195–10200 (2004).
27. daCosta, C. J. B. & Baenziger, J. E. Gating of pentameric ligand-gated ion channels: structural insights and ambiguities. *Structure* **21**, 1271–1283 (2013).
28. Gonzalez-Gutierrez, G., Wang, Y., Cymes, G. D., Tajkhorshid, E. & Grosman, C. Chasing the open-state structure of pentameric ligand-gated ion channels. *J. Gen. Physiol.* **149**, 1119–1138 (2017).
29. Hibbs, R. E. & Gouaux, E. Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474**, 54–60 (2011).

Acknowledgements We acknowledge access to the C-CINA and ESRF Krios microscopes, and thank M. Chami, L. Kovacik, H. Stahlberg, G. Effantin, E. Kandiah and M. Hons for support. We thank L. Estrozi, M. Bacia, G. Effantin and A. Desfosses for advice on cryo-EM; the Vivaudou and Moreau laboratories for providing high-quality oocytes and help with electrophysiology; members of the Nury laboratory, M. Gielen and E. Pebay-Peyroula for discussions. The work was funded by the Marie Curie CIG NeuroPenta and ERC Starting grant 637733 (to H.N.). It used the platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10 INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble PSB. The electron microscopy facility is supported by the Rhône-Alpes Region, the FRM, the FEDER and the GIS-IBISA.

Reviewer information *Nature* thanks S. Sine, A. Sobolevsky and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions L.P., J.P. and G.H. performed sample preparation. L.P., H.N, E.N. and G.S. optimized and collected microscopy data. L.P. and H.N. built the models. F.D. and C.C. conducted molecular dynamics simulations. J.N. performed electrophysiology experiments. S.N.L. and P.-J.C. performed VCF experiments. A.A.J. performed FLIPR experiments. All authors extensively discussed the data, H.N. and J.N. wrote the manuscript with inputs from all authors.

Competing Interests G.H. is employed by Theranyx. The remaining authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0672-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0672-3.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to J.N. or H.N.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression. The wild-type mouse 5-HT_{3A} receptor was expressed in a stable, inducible cell line derived from HEK T-REx 293 cells (Thermo Fisher), as previously described[5,30,31]. The cells were cultured in suspension in flasks in an orbital incubator (typical culture size, 5 l). The protein expression was induced when cells reached 2×10^6 cells/ml. Valproic acid was added one day later and cells were cultured for one more day. Cells were then pelleted by low-speed centrifugation, frozen and stored at −80°C.

Protein purification. In a typical purification batch, 20 g of cells were resuspended in buffer A (10 mM HEPES pH 7.4, 1 mM EDTA, antiprotease cocktail; 10 ml buffer per gram of cells) mechanically lysed (Ultraturrax T20, 6 × 30 s) and membranes were collected by ultracentrifugation (100,000*g* for 1 h). All steps were carried out at 4 °C. Membranes were resuspended in buffer B (50 mM Tris pH 8, 500 mM NaCl, antiprotease cocktail, 25 ml buffer per gram of membrane) and the solution was supplemented with 0.15% of C12E9 for solubilisation using gentle stirring (1.5 h). The insoluble material was removed by ultracentrifugation (100,000*g* for 45 min). Solubilized proteins were purified by affinity chromatography using gravity flow Strep-Tactin resin (IBA, typically 25 ml resin), eluted in buffer C (50 mM Tris pH 7.5, 125 mM NaCl, 0.01% C12E9) and concentrated to ~0.5 mg/ml using Millipore 100-kDa cut-off filters. The purification tag was cleaved, and carbohydrates were digested by addition of 0.04 mg TEV protease and 0.1 mg PNGase F per 1 mg protein with gentle stirring overnight. The protein was further concentrated and then applied to a Superose 6 column (GE healthcare) equilibrated in buffer C.

Electron microscopy. The most homogeneous fractions of 5-HT_3 receptor following size-exclusion chromatography were concentrated to ~1.5 mg/ml (in the best cases, no concentration was required). The sample was mixed with lipids (0.01% phosphatidic acid, 0.01% cholesterol hemisuccinate, 0.01% brain phosphatidylcholine; Avanti Polar Lipids) and ligands: 2 mM tropisetron (Tocris), or 50 μM serotonin and 2 mM calcium (conditions known to promote fast desensitization[32,33]); or 30 μM serotonin and 100 mM TMPPAA (Sigma-Aldrich). Samples were incubated for 10–30 min on ice. 3.5 μl were deposited on a glow-discharged (30 mA, 50 s) Quantifoil copper-rhodium 1.2/1.3 grid, blotted for 10 s at force 0 using a Mark IV Vitrobot and plunge-frozen in liquid ethane. Between four and ten grids were screened during each data collection, as ice thickness varied between grids. Optimization was performed on an in-house Polara electron microscope. Datasets were recorded on Titan Krios electron microscopes with K2 cameras at C-CINA (Basel) or at the ESRF (Grenoble). Details of data collections are shown in Extended Data Table 1.

Image processing. At the Basel Krios microscope, the data collection was monitored online and good images were selected using Focus[34]; images were acquired in super-resolution mode and binned by Fourier-space cropping during the drift correction. At the ESRF Krios microscope, the counted mode was used. Drift was corrected with MotionCor2[35] and dose-weighted sums were used for subsequent processing, except for CTF correction, which was performed using GCTF[36] on non-dose-weighted sums. Picking was performed with Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) using average from 2D classes as templates. Subsequent steps were performed using Relion[37] on a GPU workstation. Typically, two rounds of 2D classifications with 20–30 classes were performed, followed by a round of 3D classification without imposing symmetry (3–6 classes) with a low-pass-filtered initial model of the receptor. Particles presenting five-fold symmetry were selected, submitted to 3D classifications (classifications and data processing are further described in Extended Data Figs. 2, 4) and the best sets of particles were subjected to 3D auto-refinement. In the post-processing step, a soft mask was calculated and applied to the two half maps before the Fourier shell coefficient (FSC) was calculated. Map sharpening (*B*-factor fixed at −100 $Å^2$ for I1 and F, automatic estimation for T and I2) was also performed in the post-processing step. We tried the Phenix auto-sharpen program[38], which improved only the F map. The quality of the final reconstructions is shown in Extended Data Fig. 5.

Model refinement and structure analysis. Refinement was performed with the Phenix suite[39]. Cycles of real-space refinement were performed using global minimization, rigid body fit and local rotamer fitting (and *B*-factor refinement in late stages), alternating with manual rebuilding in Coot[40]. NCS and secondary structure restraints were enabled. The 4 models comprise residues 10–307 and 426–460. T, I1 and I2 also have MX and MA residues 308–330 and 399–426, which—owing to flexibility—could not be built in F. Tropisetron (numbered 902) was placed in the orthosteric side of T, and serotonin (numbered 901) was placed in the site of I1, F and I2. The densities for serotonin in I1 and F enable the ligand to be placed in several equivalent orientations, and we used the unambiguous density in I2 to choose the same ligand pose in these 3 structures. The stereochemical properties of the final models, analysed with the Molprobity server (http://molprobity.biochem.duke.edu/), are reported in Extended Data Table 1. Pore profiles were plotted using HOLE[41], r.m.s.d. values were calculated with 'superpose' in the CCP4 suite[42]. Figures were prepared with the PyMOL Molecular Graphics System (Schrodinger), Chimera[43] or CueMol.

Molecular dynamics. Molecular assays were built for the four conformational states of 5-HT_3 described in the manuscript. For F, I1 and I2, we used the reported cryo-EM structures, whereas for the inhibited conformation we used the crystal structure (RCSB Protein Data Bank code (PDB): 4PIR) because of its higher resolution. The co-crystallized nanobodies were removed and the protein was modelled without serotonin. The five serotonins present in F, I1 and I2 structures were kept in the models. For I2, because no obvious densities were observed for TMPPAA, it was not represented in the model. The intracellular domain of the F conformation (not resolved in the cryo-EM density) was not included in the model.

Using the CHARMM-GUI web interface[44,45], each structure was embedded in a fully hydrated palmitoyl-oleyl-phosphatidylcholine (POPC) bilayer consisting of around 240 lipid units and about 30,000 for F and 42,000 water molecules for T, I1 and I2. K^+Cl^- (150 mM) was explicitly added to each system, while ensuring their electric neutrality. The all-atom CHARMM36 force field[46] and revision thereof for lipids[47] were used to describe the system and CMAP corrections were introduced for the protein[48]. For water, we used the TIP3P model[49]. A subset of mass of the heavy atoms of lipids, protein and serotonin were transferred on the hydrogens atoms to which they are bound, to reach a hydrogen mass of 3.024 Da. Using such a mass repartitioning scheme, the equation of motions can be integrated with a time step of up to 4 fs without modifying the dynamics and thermodynamics of the system[50].

All simulations were carried out with the NAMD package v.2.12[51]. Simulations were performed in the isothermal–isobaric ensemble at $T = 300$K and $P = 1$ atm with anisotropic scaling of the simulation cell[52], long-range electrostatic interactions were treated with the particle-mesh Ewald method[53], and short-range electrostatics and Lennard-Jones interactions were smoothly truncated. The equations of motion were integrated with a time step of 4 and 8 fs for short- and long-range forces, respectively, using the Verlet r-RESPA multiple time-step propagator[54]. Covalent bonds involving hydrogen atoms were constrained to their equilibrium length by means of the 'rattle/shake algorithm'[55,56] and the 'settle' algorithm was used for water[57]. For each system, a smooth equilibration along which the positions of the heavy atoms of the protein were restrained harmonically, was carried out for 60 ns. After releasing the restraints, the simulations were extended up to 1 μs. All analyses and molecular rendering were achieved with VMD[58]. Pore radii were inferred using the program HOLE[41].

The potentials of mean force underlying the translocation of a potassium ion in the I1 and F conformations of the 5-HT_3 receptor were determined using a multiple-walker version[59] of the adaptive biasing force algorithm[60]. For the I1 conformation, the reaction coordinate model was chosen as the Euclidian distance between the ion and the centre of mass of the protein, projected onto its longitudinal axis; that is, the *z*-direction of Cartesian space. In the case of the F conformation, a two-dimensional free-energy landscape was generated, exploring ion diffusion in the pore not only longitudinally by means of the aforementioned projected Euclidian distance, but also radially. The potential of mean force along the pore was recovered by integration of the marginal law in the radial direction. The reaction pathway was broken down into 14 and 12 windows for the I1 and F conformations, respectively. The free-energy landscapes were explored by four walkers, syncing gradients every 500 molecular-dynamics steps. The total simulation time for the I1 and F conformations amounted to 2.16 and 3.04 μs, respectively, wherein the last 0.8 μs was used to estimate the error bars associated with the potentials of mean force, based on an independent mapping of the free-energy landscape by the walkers.

Electrophysiology. Electrical recordings were obtained by two-electrode voltage-clamp (TEVC) on *Xenopus* oocytes expressing either wild-type or mutated homomeric mouse 5-HT_{3A} receptors. Mutants were obtained using the QuickChange Lightning (Stratagene) site-directed mutagenesis kit and oligonucleotides (Supplementary Table 2) to introduce point mutations into the pcDNA5/TO-m5-HT_{3A} plasmid. All the mutations were verified by sequencing.

Oocytes were prepared as previously reported[61] using procedures that conformed to European regulations for animal handling and experiments, and were approved by governmental services (authorization no. D 38 185 10 001 for the animal facility delivered by the Prefect of Isère) and the Institutional Ethics Committee (ethics approval N° 12-040 granted to C. Moreau by the Ethics Committee of Commissariat à l'Energie Atomique et aux Energies Alternatives). Difolliculated oocytes were injected with 30 nl plasmid DNA (1–10 ng/μl) coding for the desired 5-HT_{3A} subunit (subcloned into pcDNA5 vector). Microinjected oocytes were incubated for 1 to 5 days at 19 °C in Barth's solution (in mM: 1 KCl, 0.82 $MgSO_4$, 88 NaCl, 2.4 $NaHCO_3$, 0.41 $CaCl_2$, 16 HEPES, pH 7.4) supplemented with 100 U ml^{-1} penicillin and streptomycin.

Whole-cell TEVC recordings were obtained using an OC-725C Oocyte Clamp amplifier (Warner Instruments) at a constant holding potential of −50 mV.

Macroscopic currents were filtered at 1 kHz, digitized at 2 kHz with a Digidata 1440 analogue-to-digital interface and analysed with Clampfit (Molecular Devices). During the recordings, oocytes were constantly perfused with a 0 Ca-ND96 solution (containing in mM: 91 NaCl, 2 KCl, 1 $MgCl_2$, 5 HEPES, 1 EGTA, pH 7.4) at a rate of 1.3 ml/min, which allowed the application of serotonin and other compounds in a time range of seconds (20-fold concentration change of applied compounds in 5 s). Throughout this study, electrophysiological responses were induced by 10 μM serotonin (a saturating dose for all tested constructions, not shown) and peaked in less than 3 s. In all constructions tested, 10-min applications of serotonin induced almost complete desensitization (to less than 3% of the peak current with $t_{1/2}$ in the min range). Full recovery of desensitization was observed after washing serotonin for 10 min. Ca^{2+} ions are permeant through 5-HT_3 receptors[62]. Calcium was therefore carefully removed from the recording perfusion solution to avoid contamination of the serotonin-induced responses by endogenous oocyte current secondarily activated by Ca^{2+} ions entering the oocyte through the serotonin receptors.

MTSEA (2-aminoethyl methanethiosulfonate bromide) and MTSET (2-trimethylammonium-ethyl methanethiosulfonate bromide) were purchased from Interchim and prepared immediately before perfusion from stock solutions in water stored at −20 °C. The effect of MTS compounds on pore cysteine-mutants was studied using the following protocol: (i) checking stability of the response with a train of three to five applications of 10 μM serotonin for 10 s, every 5 min (I_{pre}); (ii) applying MTS compounds for 2 min at 1 mM, either in the absence of 5-HT (to probe cysteine accessibility in the resting state), simultaneously with 10 μM 5-HT (to probe the open state) or simultaneously with 10 μM 5-HT after a 10 min pre-application of 5-HT alone allowing for complete desensitization; (iii) following the effect of MTS compounds during a second set of five applications of 5-HT for 10 s every 5 min (I_{post}). Irreversible effects of MTS compounds were quantified by measuring I_{post} 25 min after removal of the MTS compound. The percentage of inhibition or potentiation was calculated as $(I_{pre}-I_{post})/I_{pre} \times 100$. MTS compounds (applied at 1 mM simultaneously with 10 μM serotonin) have no detectable effect on wild-type receptors (not shown).

Voltage-clamp fluorometry. VCF recordings were performed on *Xenopus* oocytes provided by Ecocyte Bioscience that were injected with 50 ng/μl plasmid DNA encoding for the mouse 5-HT_{3A} receptor S19′C mutant, after 2 to 6 days of expression in Barth solution at 17 °C. They were labelled with MTS-TAMRA (Toronto Chemicals) for 5 min at 17 °C in ND96 buffer without $CaCl_2$ containing 10 μM 5-hydroxytryptamine (5-HT), then rinsed and stored at 17 °C for up to 4 h before recording in ND96. Recording were made in a TEVC setup as described[13], adapted for fluorescence recording. In brief, recordings were made with a VCF dedicated chamber with only a fraction of the oocyte perfused and illuminated by a LED (coolLED PE-4000). Light was collected using a fluorescence microscope (Zeiss Axiovert135) equipped with a 40× objective (Plan Neofluar), a TRITC filter set and a photomultiplier tube (Hamamatsu Photonics). Recordings were made at −60 mV clamp, 500-Hz sampling rate with a 550-nm excitation wavelength. Data were filtered, corrected for baseline and photobleaching where necessary, and analysed using pClamp and Axograph. Dose–response curves were fitted using Prism to the Hill equation: $I/I_{max} = 1/(1 + (EC_{50}/[5\text{-HT}])^{nH})$ in which I is the response at a given [5-HT] (serotonin concentration), I_{max} is the maximal response, EC_{50} is half maximal effective concentration and nH is the Hill coefficient. Serotonin dose–response relations were shifted to the left with half-responses for fluorescence and current at 40 and 240 nM, respectively (the EC_{50} of the wild-type receptor is 800 nM).

Fluorescence imaging plate reader membrane potential blue assay. The agonist properties of 5-HT and TMPPAA (Sigma-Aldrich) were characterized at human wild-type or mutant 5-HT_{3A} receptors transiently expressed in tsA201 cells in the fluorescence imaging plate reader (FLIPR) membrane potential blue (FMP; fluoresence-based membrane potential) assay. The generation of some of the human 5-HT_{3A} mutants have previously been described[63]. Other mutants were constructed by introduction of point mutations into the h5-HT3A-pCIneo plasmid using Quikchange II XL site-directed mutagenesis (Stratagene) and oligonucleotides (TAG Copenhagen). The absence of unwanted mutations in all cDNAs created by PCR was verified by sequencing (Eurofins MWG Operon). The cells were cultured in Dulbecco's Modified Eagle Medium supplemented with penicillin (100 U/ml), streptomycin (100 mg/ml) and 10% fetal bovine serum in a humidified atmosphere of 5% CO_2 and 95% air at 37 °C. Cells (1.2×10^6) were split into a 6-cm tissue culture plate, transfected the following day with 4 μg cDNA (wild-type or mutant h5-HT3A-pCIneo) using Polyfect (Qiagen) as transfection reagent, and split into poly-D-lysine-coated black 96-well plates (8×10^4 cells per well) with clear bottom (BD Biosciences) the following day. After 20–24 h following transfection, the medium was aspirated, and the cells were washed with 100 μl Krebs buffer (140 mM NaCl, 4.7 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgCl_2$, 11 mM HEPES, 10 mM D-glucose, pH 7.4). Then 100 μl Krebs buffer supplemented with 0.5 mg/ml FMP assay dye (Molecular Devices) was added to each well, and the 96-well plate was assayed at 37 °C in a FLEXStation3 Benchtop Multi-Mode Microplate Reader (Molecular Devices) measuring emission (in fluorescence units (FU)) at 565 nm caused by excitation at 530 nm before and up to 90 s after addition of 33.3 μl assay buffer supplemented with agonist (5-HT or TMPPAA). The experiments were performed in duplicate at least three times for each agonist at all receptors. Concentration–response curves for the agonists were constructed based on the difference in the fluorescence units (rFU) between the maximal fluorescence recording made before and after addition of the agonists at different concentrations. The curves were generated by non-weighted least-squares fits using the program KaleidaGraph (Synergy Software).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Atomic coordinates of the four conformations have been deposited in the Protein Data Bank with accession numbers 6HIN, 6HIO, 6HIQ and 6HIS for conformations F, I1, I2 and T, respectively. The cryo-EM density maps have been deposited in the Electron Microscopy Data Bank with accession numbers EMD-0225, EMD-0226, EMD-0227 and EMD-0228 for conformations F, I1, I2 and T, respectively.

30. Hassaïne, G. et al. Large scale expression and purification of the mouse 5-HT_3 receptor. *BBA Biomembranes* **1828**, 2544–2552 (2013).
31. Hassaïne, G. et al. Expression, biochemistry, and stabilization with camel antibodies of membrane proteins: case study of the mouse 5-HT_3 receptor. *Methods Mol. Biol.* **1635**, 139–168 (2017).
32. Yakel, J. L., Lagrutta, A., Adelman, J. P. & North, R. A. Single amino acid substitution affects desensitization of the 5-hydroxytryptamine type 3 receptor expressed in *Xenopus* oocytes. *Proc. Natl Acad. Sci. USA* **90**, 5030–5033 (1993).
33. Hu, X.-Q. & Lovinger, D. M. Role of aspartate 298 in mouse 5-HT_{3A} receptor gating and modulation by extracellular Ca^{2+}. *J. Physiol. (Lond.)* **568**, 381–396 (2005).
34. Biyani, N. et al. Focus: The interface between data collection and data processing in cryo-EM. *J. Struct. Biol.* **198**, 124–133 (2017).
35. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
36. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
37. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
38. Terwilliger, T. C., Sobolev, O. V., Afonine, P. V. & Adams, P. D. Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr. D* **74**, 545–559 (2018).
39. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
41. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360 (1996).
42. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
43. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
44. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
45. Wu, E. L. et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
46. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
47. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
48. Best, R. B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
50. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).
51. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
52. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
53. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an *N*·log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
54. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**, 1990–2001 (1992).
55. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
56. Andersen, H. C. Rattle: A 'velocity' version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **52**, 24–34 (1983).

57. Miyamoto, S. & Kollman, P. A. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).
58. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
59. Comer, J., Phillips, J. C., Schulten, K. & Chipot, C. Multiple-replica strategies for free-energy calculations in NAMD: multiple-walker adaptive biasing force and walker selection rules. *J. Chem. Theory Comput.* **10**, 5276–5285 (2014).
60. Comer, J. et al. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* **119**, 1129–1151 (2015).
61. Moreau, C. J., Niescierowicz, K., Caro, L. N., Revilloud, J. & Vivaudou, M. Ion channel reporter for monitoring the activity of engineered GPCRs. *Methods Enzymol.* **556**, 425–454 (2015).
62. Yang, J. Ion permeation through 5-hydroxytryptamine-gated channels in neuroblastoma N18 cells. *J. Gen. Physiol.* **96**, 1177–1198 (1990).
63. Trattnig, S. M. et al. Discovery of a novel allosteric modulator of 5-HT_3 receptors: inhibition and potentiation of Cys-loop receptor signaling through a conserved transmembrane intersubunit site. *J. Biol. Chem.* **287**, 25241–25254 (2012).
64. Suryanarayanan, A. et al. The loop C region of the murine 5-HT_{3A} receptor contributes to the differential actions of 5-hydroxytryptamine and *m*-chlorophenylbiguanide. *Biochemistry* **44**, 9140–9149 (2005).
65. Yan, D., Schulte, M. K., Bloom, K. E. & White, M. M. Structural features of the ligand-binding domain of the serotonin $5HT_3$ receptor. *J. Biol. Chem.* **274**, 5537–5541 (1999).
66. Ruepp, M.-D., Wei, H., Leuenberger, M., Lochner, M. & Thompson, A. J. The binding orientations of structurally-related ligands can differ; a cautionary note. *Neuropharmacology* **119**, 48–61 (2017).
67. Kesters, D. et al. Structural basis of ligand recognition in 5-HT_3 receptors. *EMBO Rep.* **14**, 49–56 (2013).
68. Hibbs, R. E. et al. Structural determinants for interaction of partial agonists with acetylcholine binding protein and neuronal α7 nicotinic acetylcholine receptor. *EMBO J.* **28**, 3040–3051 (2009).
69. Spurny, R. et al. Pentameric ligand-gated ion channel ELIC is activated by GABA and modulated by benzodiazepines. *Proc. Natl Acad. Sci. USA* **109**, E3028–E3034 (2012).
70. Delbart, F. et al. An allosteric binding site of the α7 nicotinic acetylcholine receptor revealed in a humanized acetylcholine-binding protein. *J. Biol. Chem.* **293**, 2534–2545 (2018).
71. Steward, L. J. et al. Importance of phenylalanine 107 in Agonist recognition by the 5-hydroxytryptamine$_{3A}$ receptor. *Mol. Pharmacol.* **57**, 1249–1255 (2000).
72. Purohit, P. & Auerbach, A. Acetylcholine receptor gating at extracellular transmembrane domain interface: the 'pre-M1' linker. *J. Gen. Physiol.* **130**, 559–568 (2007).
73. Lee, W. Y. & Sine, S. M. Principal pathway coupling agonist binding to channel gating in nicotinic receptors. *Nature* **438**, 243–247 (2005).
74. Hu, X.-Q., Zhang, L., Stewart, R. R. & Weight, F. F. Arginine 222 in the pre-transmembrane domain 1 of 5-HT_{3A} receptors links agonist binding to channel gating. *J. Biol. Chem.* **278**, 46583–46589 (2003).

Structure

CellPress

Article

The Binding of Palonosetron and Other Antiemetic Drugs to the Serotonin 5-HT3 Receptor

Eleftherios Zarkadas,[1,6] Hong Zhang,[2,6] Wensheng Cai,[2] Gregory Effantin,[1] Jonathan Perot,[1] Jacques Neyton,[1] Christophe Chipot,[3,4,5] Guy Schoehn,[1] Francois Dehez,[3,4,*] and Hugues Nury[1,7,*]

[1]CNRS, Univ. Grenoble Alpes, CEA, IBS, 38000 Grenoble, France

[2]Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin Key Laboratory of Biosensing and Molecular Recognition, Tianjin 300071, People's Republic of China

[3]Université de Lorraine, CNRS, LPCT, 54000 Nancy, France

[4]Laboratoire International Associé CNRS, Vandoeuvre-les-Nancy, France

[5]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[6]These authors contributed equally

[7]Lead Contact

*Correspondence: francois.dehez@univ-lorraine.fr (F.D.), hugues.nury@ibs.fr (H.N.)

https://doi.org/10.1016/j.str.2020.07.004

SUMMARY

Inaccurately perceived as niche drugs, antiemetics are key elements of cancer treatment alleviating the most dreaded side effect of chemotherapy. Serotonin 5-HT3 receptor antagonists are the most commonly prescribed class of drugs to control chemotherapy-induced nausea and vomiting. These antagonists have been clinically successful drugs since the 1980s, yet our understanding of how they operate at the molecular level has been hampered by the difficulty of obtaining structures of drug-receptor complexes. Here, we report the cryoelectron microscopy structure of the palonosetron-bound 5-HT3 receptor. We investigate the binding of palonosetron, granisetron, dolasetron, ondansetron, and cilansetron using molecular dynamics, covering the whole set of antagonists used in clinical practice. The structural and computational results yield detailed atomic insight into the binding modes of the drugs. In light of our data, we establish a comprehensive framework underlying the inhibition mechanism by the -setron drug family.

INTRODUCTION

In the realm of serotonin receptors, the 5-HT3 receptor stands apart as the only ion channel (Lummis, 2012). It belongs to the family of pentameric ligand-gated ion channels (pLGICs) (Nemecz et al., 2016). The pLGIC family comprises, in mammals, members activated by serotonin, acetylcholine, glycine, and ɣ-aminobutyric acid. They share a common architecture: neurotransmitters bind to extracellular interfaces between subunits; the transmembrane domain (TMD) possesses 5 × 4 helices and features a central gated pore; the dispensable intracellular domain is for a good part intrinsically disordered.

The interest in 5-HT3 receptor pharmacology surged when chemotherapies became first available (e.g., cisplatin approved in 1978 for ovarian and testicular cancers). At the time of their discovery in the 1980s, the first competitive antagonists were characterized as the "*most potent drugs of any pharmacological class so far described*" (Richardson et al., 1985), and less than a decade later the antiemetic research on 5-HT3 receptors was considered done by the pharma industry with numerous clinically and commercially successful drugs (coined -setrons) available (Sanger and Andrews, 2018). Beyond emesis, fundamental and pre-clinical research indicate that the 5-HT3 receptor plays a role in depression (Fakhfouri et al., 2019) and in body weight control (Oh et al., 2015; Weber et al., 2009); alosetron is used clinically to treat irritable bowel syndrome (Chey et al., 2015), vortioxetine is an antidepressant non-selectively targeting the serotonin transporter and receptors (Bang-Andersen et al., 2011).

The molecular mode of action of antiemetics has been explored extensively through binding, mutagenesis, and electrophysiology studies, while structural studies on soluble model proteins gave approximate poses for the bound drugs (Kesters et al., 2013; Price et al., 2016). Recently cryoelectron microscopy (cryo-EM) structures of the 5-HT3 receptor have revealed the conformations of the apo state, of the tropisetron-bound inhibited, and of two serotonin-bound states (Basak et al., 2018a, 2018b; Polovinkin et al., 2018). However, the limited resolution of the tropisetron-bound structure precluded a thorough analysis of its binding.

Here, we report a cryo-EM structure of the 5-HT3 receptor in its inhibited state, binding palonosetron, and solved at a 2.8-Å resolution, germane for structural pharmacology. Palonosetron possesses the highest affinity in the -setron family, a slow dissociation rate, and nanomolar half maximal inhibitory concentration values (~0.2 nM when pre-incubated, ~80 nM when co-applied) (Lummis and Thompson, 2013). Leaning on this structure, and that recently reported for the granisetron-bound 5-HT3 receptor (Basak et al., 2019), we set out to provide a comprehensive structural framework that established the mode of action of

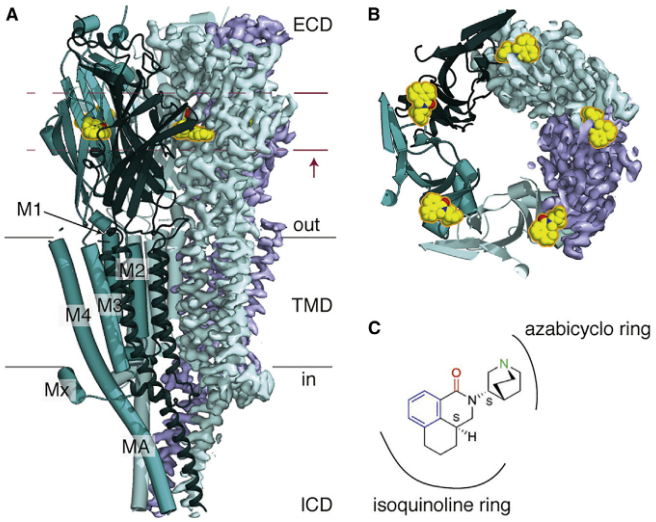


Figure 1. Cryo-EM Structure of the Mouse 5-HT3 Receptor in Complex with Palonosetron

(A) View parallel to the membrane plane. Three subunits are represented as cartoons while the cryo-EM map is depicted for the two last subunits. Palonosetron molecules are depicted as yellow spheres.

(B) View of a slab (fuchsia lines) perpendicular to the membrane (as indicated by the arrow).

(C) Chemical formula of palonosetron, with its aromatic moiety in blue, its ring-embedded protonated nitrogen in green, and its H bond acceptor oxygen group in red. The "s" letters next to chiral carbon atoms indicate stereochemistry.

antiemetics. Our palonosetron-bound structure served as a template for molecular dynamics (MD) simulations not only with palonosetron, but also with cilansetron, dolasetron, granisetron, and ondansetron to elucidate the general rules that underlie binding of the setron family to the 5-HT3 receptor. We further establish a relationship between the body of biochemical and pharmacological data available in light of our findings.

RESULTS AND DISCUSSION

Global Conformation

Cryo-EM imaging of the mouse homomeric 5-HT3A receptor, in the presence of palonosetron, yielded a 2.8-Å reconstruction with well-defined densities throughout the three domains of the protein (Figures 1 and S1)—except at the level of the intrinsically disordered region of the intracellular domain (residues 329–399, see Figure S1 for topology and numbering), for which the density is averaged out. The global conformation corresponds to that of previously determined inhibited states (all-atom root-mean-square deviation [RMSD] = 1.29 Å with the nanobody-inhibited 4PIR structure [Hassaine et al., 2014], of 1.41 Å with the tropisetron-bound 6HIS one [Polovinkin et al., 2018], and of 1.31 Å with the granisetron-bound 6NP0 one [Basak et al., 2019]). Compared with the previous structures, this one and the granisetron-bound one stand out in quality and may be used as templates in future modeling work, where the inhibited state is relevant. Non-proteinic densities in the orthosteric sites could be attributed non-ambiguously to palonosetron molecules (Figure 2B).

General Features of the -Setron Binding Site

Palonosetron binds tightly in the classical neurotransmitter site at the subunit/subunit interface (Figure 1). The occluded surface area between two subunits, for the extracellular domain (ECD) only, amounts to ~1,400 Å^2. Palonosetron acts as a molecular glue, each molecule locking the conformation of the interface and shielding an extra ~170 Å^2 of the principal subunit and ~210 Å^2 of the complementary subunit from the solvent. MD simulations essentially provide the same picture (Figure S2). The conformational dynamics of the drugs within their binding sites induce variability in the shielding of the principal subunit, concomitant with the dynamics of the H-bond network (see below).

Like every -setron drug, palonosetron possesses two moieties (Table S1). The azabicyclo ring penetrates deep into the binding pocket, while the isoquinoline ring sits in a cavity at the surface of the complementary subunit, capped by loop C of the principal subunit (Figure 2B). The drug remains partly solvent-accessible between loop C and loop F (and one could imagine derivatives with probes grafted there).

Four aromatic residues (W156, Y126, Y207, and W63) forming the so-called aromatic cage tightly wrap around the azabicyclo ring (Figures 2B and 3A). Mutants harboring non-aromatic side chains at any of these four positions show either much decreased inhibition by palonosetron, or simply no response to serotonin (Del Cadia et al., 2013). Even conservative mutations, such as W156Y or Y126F result in a 180- and 20-fold loss of inhibition, respectively (Del Cadia et al., 2013). Throughout the report, mutants are simply noted residue-number-residue to facilitate the reading, without specifying if they are mutants of the human or of the mouse 5-HT3 receptor (as long as both receptors have the same amino acid at a given position of course).

The palonosetron isoquinoline moiety is constrained by hydrophobic interactions with Y126, W63, I44, and I201 (Figures 2B and 3A). It also interacts with R65, the conformation of which is particularly interesting: the arginine extends over palonosetron, possibly engaging in a cation-π interaction, as reported for granisetron (Kesters et al., 2013) (the conformation of loop D residues is further discussed in Figure S3).

A dense network of interactions stabilizes the side chains of Y126, R65, Y67, D42, R169, and D177 on the complementary side (Figures 3 and S4). During simulations of all drugs, the D42-R65 and D42-R169 salt bridges are always established,

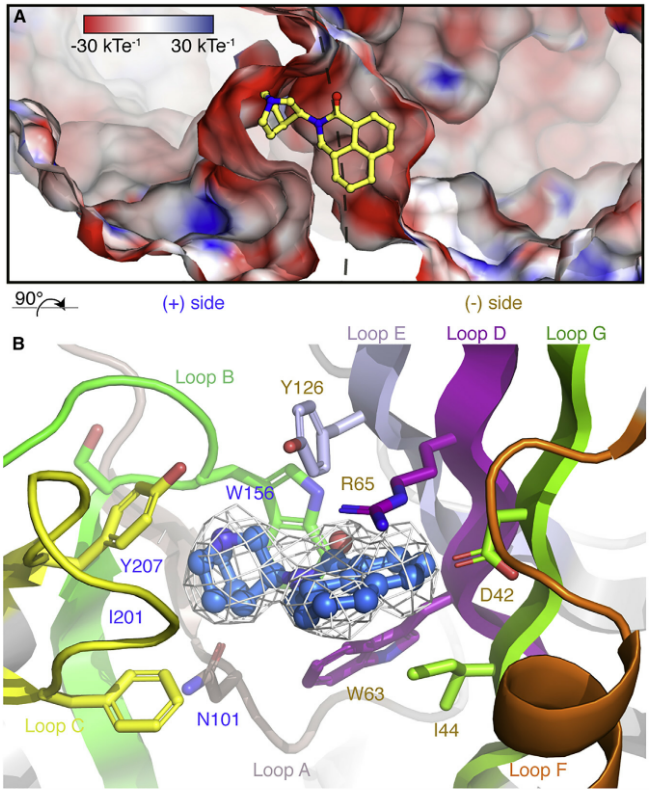


Figure 2. Close Up of the Palonosetron Binding Site

(A) Electrostatic potential mapped on the molecular surface of the 5-HT3 receptor, with stick representation of palonosetron in the neurotransmitter binding site. The view is from outside, approximately perpendicular to the membrane plane.

(B) Cartoon and stick representation, approximately parallel to the membrane. Side chains within 4 Å of the drug are depicted as sticks. The protein is colored using the same code throughout the report (on the principal side: loop A, maroon; loop B, green; loop C, yellow; on the complementary side: loop D, purple; loop E, orchid; loop F, orange; loop G, forest green). The cryo-EM density for palonosetron is depicted as a gray mesh.

whereas the D177-R169 one is occasionally disrupted. While the side-chain conformations in the binding loops of the complementary subunit strikingly resemble those of the serotonin-bound receptor, it is worth noting that both D42-R65 and D177-R169 (to a lesser extent) can break in simulations of the apo or serotonin-bound receptor (Figures S4E and S4F).

Conformational Dynamics of the Drugs, H-Bonding Patterns

The positively charged nitrogen of the azabicyclo ring forms an H bond with the main chain carbonyl of W156 (Figure 6B). This H bond is consistently observed throughout the MD trajectory for four palonosetron molecules, whereas the fifth one engages in a long-lived H bond with N101 (Figures 4B and 4C). The equivalent nitrogen of granisetron steadily interacts with W156 in four subunits and is left unpaired in the last subunit (Figure S5). In trajectories generated for ondansetron and cilansetron, the nitrogen also forms H bonds with W156 and N101, with a preference for the latter residue (Figure S6). This trend is more pronounced for ondansetron, whereas cilansetron also occasionally forms an H bond with S155 (Figure S6). Dolasetron engages in a long-lived H bond with N101 through its ring nitrogen. It is also transiently H bonded to W156 but through its carboxyl oxygen, at variance with the other -setrons. Interestingly, dolasetron also forms an H bond between the nitrogen of its aromatic moiety and W63 (Figure S5).

The H-bond dynamics observed at the binding sites along the MD trajectories seem correlated with the conformational dynamics of the inhibitors (Figures 4A, S5, and S6). It essentially involves a rotation around one (two for cilansetron) dihedral angle connecting the two rigid moieties of the -setron drugs, resulting in distinct conformations, with populations possibly modulated by the environment, which are observed for all drugs in solution or bound to the 5-HT3 receptor. The free-energy profile along the dihedral connecting the palonosetron rigid moieties, computed both in water and in the protein, revealed four minima (labeled A–D in Figure 4A) separated by barriers amounting up to 4 kcal/mol (Figure S7). It echoes the 1993 report on the discovery of palonosetron, where the authors noted: "*An affinity of this magnitude could only be associated with a compound that binds at or near its global minimum*" (Clark et al., 1993).

What Interacts with the Ligand H Bond Acceptor?

The presence of an H bond acceptor in between the aromatic moiety and the cycle-embedded nitrogen moiety seems a defining feature of 5-HT3 receptor antagonists (Thompson, 2013). Yet, no obvious interaction for this chemical group is immediately visible when observing the palonosetron-bound cryo-EM structure (or the granisetron-bound one). In all simulations, we observed the formation of an H bond between a water molecule and the oxygen of the ligand central carbonyl. This water molecule also engages in two H bonds with the protein, one with the backbone carbonyl of Y64 and one with the -NH group of the W156 indole. In experimental densities, a faint

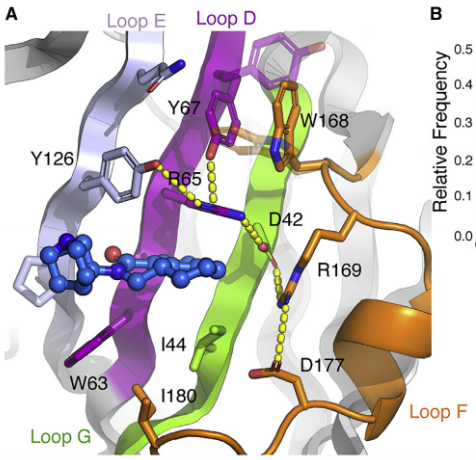


Figure 3. Network of Side-Chain Interactions on the Complementary Subunit

(A) Cartoon and stick representation, approximately parallel to the membrane. Side chains within 4 Å of the drug are depicted as sticks. Yellow dashes indicate H bonds or salt bridges. The dual conformation of residues Y67 and W168 (see the STAR Methods) is depicted as transparent sticks. (B) Distance distributions between atoms forming the three salt bridges D42-R65, D42-R169, and D177-R169 (black, red, and transparent cyan, respectively) sampled during the MD simulation in the presence of palonosetron.

peak is recorded and may correspond to a water molecule, this signal being stronger in the granisetron structure (visual inspection of the 6NP0 density map). We placed a water molecule and conducted new MD simulations. Indeed, this molecule remained trapped most of the time in the case of palonosetron (Figure 5), and at various levels for the other drugs (Figure S7).

A Structural Framework for Functional Data

Over nearly three decades, numerous studies have probed how mutations in, or around, the orthosteric binding site influenced antiemetics inhibition of the 5-HT3 receptor. We briefly discuss the literature, in light of our structure and simulations, for each stretch of the receptor participating in the binding site.

Loop A

Among loop A residues, only N101 lies close enough to palonosetron (and to other ligands in other structures) for a direct interaction (Figure 6A). In our simulations, the distance distributions of N101 with -setron drugs feature a main peak at 5 Å with two additional peaks at 2.5 and ~8 Å, indicating transient H bonding of the ligands with this residue (Figures S5 and S6). Palonosetron is 10-fold more potent with the N101A mutant, while N101Q remains similar to the wild type (Price et al., 2016), in line with the idea that the absence of H-bonding capacity would stabilize the azabicyclo moiety in the conformation where the N interacts with W156 only. Interestingly, mutations of N101 into a variety of amino acids (A, D, E, L, Q, R, and V) does not affect [^{3}H]granisetron binding (Sullivan et al., 2006). Accordingly, simulations show an absence of H bond with the drug at this residue.

[^{3}H]Granisetron-binding experiments illuminate a drastic effect of mutations of E102, with a complete loss of binding at the subnanomolar range for A, D, G, K, and N (data obtained on oocytes) (Price et al., 2008); or 60- to 80-fold decrease in K_d for E102D/N (data obtained on HEK cells) (Boess et al., 1997). E102 is not oriented toward the binding site and interacts with S150 and T152 (β7), while its main chain carboxyl forms an H bond with A134 (β6) (Figure 6A). Thus, one possible role for E102 is to pack the β sandwich and to position properly loops A and B. In addition, E102 creates an electronegative patch next to the binding site entrance, which might participate in ligand attraction, in line with the observation that the deleterious effect of introducing a net positive charge in loop A (e.g., N101K, E102Q) can be assuaged by removal of the positive charge at K211 (regardless of H-bonding ability at position 211) (Mosesso and Dougherty, 2018). At position 103, granisetron binding appears only marginally affected by F103Y/N substitutions (Steward et al., 2000), while it was decreased ~5-fold by F103A/W mutations (Sullivan et al., 2006).

Loop B

In most pLGICs, loop B harbors a pivotal tryptophan residue (here W156), which forms the back wall of the binding site (Figure 6B) and establishes cation-π interactions with ligands (Beene et al., 2002). Unsurprisingly, alanine substitution abolishes granisetron and palonosetron binding (Del Cadia et al., 2013; Thompson et al., 2008), while a more conservative mutation to tyrosine increases the respective K_d of these drugs by 10- and 180-fold. Nearby alanine substitutions are relatively innocuous. L157A has a 10-fold effect on granisetron binding, maybe because the smaller side chain does not hinder the possible motions of Y126. L157I is wild type-like.

Loop C

Loop C appears to be the most variable loop (with loop F, see below) in multiple sequence alignments of pentameric ligand-gated channels. It undergoes deformation upon ligand binding, which remains modest in the 5-HT3 receptor structures. As a heritage from studies of stiff acetylcholine binding proteins, where loop C was the most evident moving element (Celie et al., 2004; Gao et al., 2005), its position is often referred to as a hallmark of the receptor state. Legitimate emphasis on loop C should not blunt other relevant markers of states (see below). Yet, it remains worth underscoring the simultaneous motion of the flexible loop C and of the ECD/TMD interface loops. The coupling between ligand binding and pore opening is a complex process (Lee and Sine, 2005; Purohit and Auerbach, 2013). A structurally intuitive hypothesis is that motion of loop C in response to ligand binding might be transmitted to the ECD/TMD interface by means of a central β sheet rigid hinge.

I201 and Y207 are the loop C residues that point toward palonosetron (Figure 6C). Only aromatic mutations are tolerated at

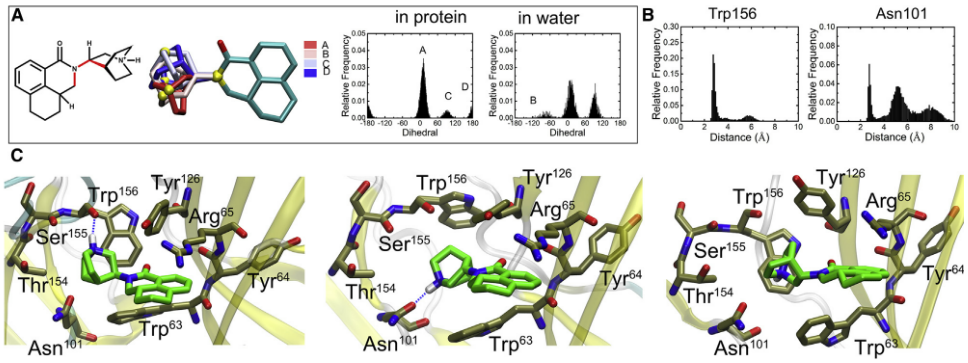


Figure 4. H-Bonding Pattern and Dynamics of Palonosetron during MD Simulations

(A) The inset shows the distributions of the dihedral angle between the two moieties of palonosetron (drawn in red), either bound to the receptor or free in water. The position of the aromatic part is relatively conserved in the protein indicating that the conformations sampled result from the conformational dynamics of the azabicyclo ring in the binding site.

(B) The two distance distributions depict the H bonds between the azabicyclo ring -NH group and the backbone carbonyl of W156 (the major interactant) and the side-chain carbonyl of N101 (note the sharp peak at 2.5 Å distance) monitored over the MD trajectory for the five subunits.

(C) Three snapshots illustrate the main conformations sampled by palonosetron.

Y207, with Y207F yielding a 3-fold increase of K_d (Del Cadia et al., 2013) (11-fold increase of the granisetron K_d [Suryanarayanan et al., 2005]). Interestingly an extra density is present at the tip of Y207 (at low contour level), which might correspond to a water molecule linking this residue to the complementary side via Y126 and Q124. I201 is not conserved in the human receptor. It corresponds to M228, the mutation to alanine of which produces a 6-fold decrease of K_d (Del Cadia et al., 2013), in line with the van de Waals interaction observed in the structure (I201A also deteriorates 5-fold granisetron binding in the mouse receptor [Suryanarayanan et al., 2005]). From a structural standpoint, loop C is stabilized by interactions of E198 and S206, and the presence of a negative charge at its tip (D202) might participate in attracting the azabicyclo moiety during binding. In line with a much broader spatial rearrangement of loop C following serotonin binding, mutations of loop C residues tend to have stronger effects on serotonin binding and activation than on binding of granisetron or palonosetron (for instance, F199Y does not change the binding of both antagonists, but decreases the serotonin K_i by 180-fold, while D202A entails a 140 K_i decrease and a 12-fold increase in half maximal effective concentration).

Loop D (and G)

Two residues of loop D, W63 and R65, are in direct contact with palonosetron, sandwiching its isoquinoline ring (Figure 3). As mentioned previously, W63A is nonfunctional, while W63Y exhibits wild-type properties both for serotonin activation and palonosetron inhibition (Price et al., 2016). The conformation of R65 seems in line with a cation-π interaction both in our palonosetron-bound structure and in the granisetron-bound one (Basak et al., 2019), and this interaction has been demonstrated experimentally with fluorinated derivatives of granisetron (Kesters et al., 2013). It is noteworthy that both R65A and R65C mutations yield different effects for different antiemetics. R65A reduces the binding affinity for granisetron, but produces no significant change for palonosetron (Price et al., 2016; Yan et al., 1999). R65C abolishes granisetron binding, yet marginally affects binding of tropisetron (Ruepp et al., 2017). The differential effect was linked to putatively different orientations of the aromatic moieties of the drug, which is not the case for palonosetron and granisetron. In our previous, lower-resolution structure in complex with tropisetron (Polovinkin et al., 2018), we modeled its aromatic ring perpendicular to the ones of palonosetron or granisetron. However, ligand building in 4.5-Å maps is tricky, and better data would be required to ascertain this difference in orientation.

In loop G, D42C abolishes tropisetron binding, albeit that it barely modifies that of granisetron.

Loop E

Mutations at Y126 affect both granisetron (Y126A, 8-fold) (Venkataraman et al., 2002) and palonosetron binding (Y126A/F, 20- to 30-fold) (Del Cadia et al., 2013), in line with the proximity of the side chain to palonosetron and the implication of the terminal hydroxyl in intra- and inter-subunit interactions (Figure 3). The necessity of H-bonding capacity at Y126 has also been established by nonnatural amino acid mutagenesis (Beene et al., 2004).

In addition, mutagenesis suggests roles of neighboring residues in shaping the binding site properly, as alanine mutation of any residues in the 124–127 range produces a 6- to 8-fold decrease in the granisetron K_d (Venkataraman et al., 2002).

Loop F

The extensive network of interactions on the complementary side, together with mutagenesis, points to an essential role of the precise positioning of loop F. Mutants at positions W168(Y), P171(H), L167(A), W168(A), R169(A), and K178(A) yield >5-fold increase of granisetron K_d, while D177(A) shows no affinity (Thompson et al., 2006) (Figure 6D). The interaction of D42, R169, and D177 readily explains the mutant phenotypes at these

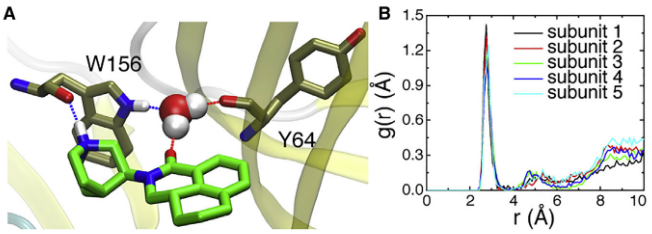


Figure 5. A Trapped Water Molecule Links Palonosetron H Bond Acceptor and the Receptor

(A) Representative snapshot illustrating the water-mediated H-bond network linking the palonosetron central carbonyl to protein residues W156 and Y64.

(B) Pair radial distribution functions between the oxygen of the palonosetron central carbonyl and the water oxygen, characterizing the probability density to find the two species at a given distance. The peak located at ~3 Å indicates that for every subunit there is always one water molecule H bonded to palonosetron.

positions: if disrupted, free charged side chains would impede the drug binding. Interestingly enough, D177N/E mutations, preserving the H-bonding capacity, yield unchanged granisetron binding, while D177A abolishes it. Serotonin activation is also lost for D177A. On top of its interaction with R169, D177 also positions the bottom of loop F through an interaction with the main chain of S179. This idea that the precise position of loop F is crucial for function is further reinforced by the effect at K178, which anchors the loop to the β9 strand at the main chain of L190. Conversely, mutation of R175 does not produce any change, in line with its orientation toward the solvent.

Would Palonosetron Bind at Heteromeric Interfaces?

Five homologous subunits of the 5-HT3 receptor are encoded by the human genome, while rodents possess only two of them. In heteromeric receptors, the activation interface is usually thought to be only the A(+)/A(−) one, i.e., the one where palonosetron fits tightly in the present structure. Earlier studies have concluded that some ligands bind differently to different subpopulations of receptors (Tang and Tang, 2012), and one wonders whether palonosetron may also bind to other interfaces. Mutations at position W156, R65, and Y126, yield no or only poor binding of palonosetron. Thus, one can infer that interfaces involving a B(+), C(+), D(+), or E(+) subunit would not bind the drug (they all lack the tryptophan in loop B, at the position equivalent to W156). Similarly B(−), C(−), or E(−) interfaces lack an arginine at positions equivalent to R65, and D(−) interfaces lack a tyrosine at position equivalent to Y126. Put together, we hypothesize that only A(+)/A(−) interfaces are responsible for palonosetron binding and inhibition.

The Singularity of Palonosetron?

Discovered in 1993 (Clark et al., 1993) and marketed in 2003, palonosetron was coined a "second-generation antagonist," obviously because it came after granisetron, tropisetron, and ondansetron, but also because, unlike these molecules, it was reported to prevent not only acute, but also delayed, chemotherapy-induced nausea and vomiting (CINV) in clinical studies (Grunberg et al., 2004). Distinctive features of palonosetron (higher potency [Wong et al., 1995] and longer plasma half-life [Stoltz et al., 2004]) cannot in principle explain this unique effect on delayed CINV: other drugs administered more often and at higher doses would also prevent it. In attempts to identify other singularities that might rationalize the advertised clinical outcome, it was also proposed that palonosetron (1) shows cooperativity and binds to allosteric sites (Rojas et al., 2008), (2) triggers 5-HT3 receptor internalization (Rojas et al., 2010a), and (3) influences the cross-talk between 5-HT3 receptors and the neurokinin 1 receptor in a way that elicits a synergistic effect of antagonists at each receptor (neurokinin 1 inhibitors, such as aprepitant are widely used in combination with -setron drugs for CINV) (Rojas et al., 2010b).

Our data unequivocally show that palonosetron binds to the orthosteric site; no secondary site was apparent from the inspection of the residual densities, in line with mutagenesis experiments disproving the existence of an allosteric site underneath the neurotransmitter one (Del Cadia et al., 2013). Granisetron, tropisetron, and palonosetron bind to the same site and apparently stabilize the same conformation of the homomeric 5-HT3A receptor: in this respect, the clinical singularity of palonosetron is not apparent in current structural studies.

We have, however, underscored some of the unique features of palonosetron in terms of binding and dynamics: the preeminence of its conformation of minimal energy (Figures 4A and S7), and the continuous presence of a water molecule next to its H bond acceptor (Figure 5). The notion of "*restricted*" conformational freedom of the aromatic moiety guided the synthesis of palonosetron (Clark et al., 1993). This restriction is clearly echoed here: the distribution of the angle between the aromatic moiety and the main axis of the protein is narrower for palonosetron compared with other drugs (Figure S7).

Serotonin-Bound ECD Conformations Clearly Differ from Antagonist-Bound Ones

The orthosteric site of pentameric ligand-gated channels lies at the interface between subunits, and quaternary reorganization of ECDs is key to their activation. To get a sense of the spread of quaternary arrangements of available 5-HT3 receptor structures, we superimposed their complementary subunits and looked at their principal subunits (Figure 7). Two groups are clearly distinguished. The average pairwise RMSD within the group of inhibited structures is 0.49 ± 0.06 Å, while the average pairwise RMSD between inhibited and activated structures is 0.91 ± 0.22 Å. At the binding site level, the activated structures possess tighter cavities where -setron binding would not be possible (Figure 7C).

The apo structure rather lies within the cluster of inhibited ones, following the allosteric theory hypothesis stating that ligands stabilize preexisting conformations (Monod et al., 1965). However, as already noted by Basak et al. (2019), the apo structure presents moderate deviations compared with the inhibited

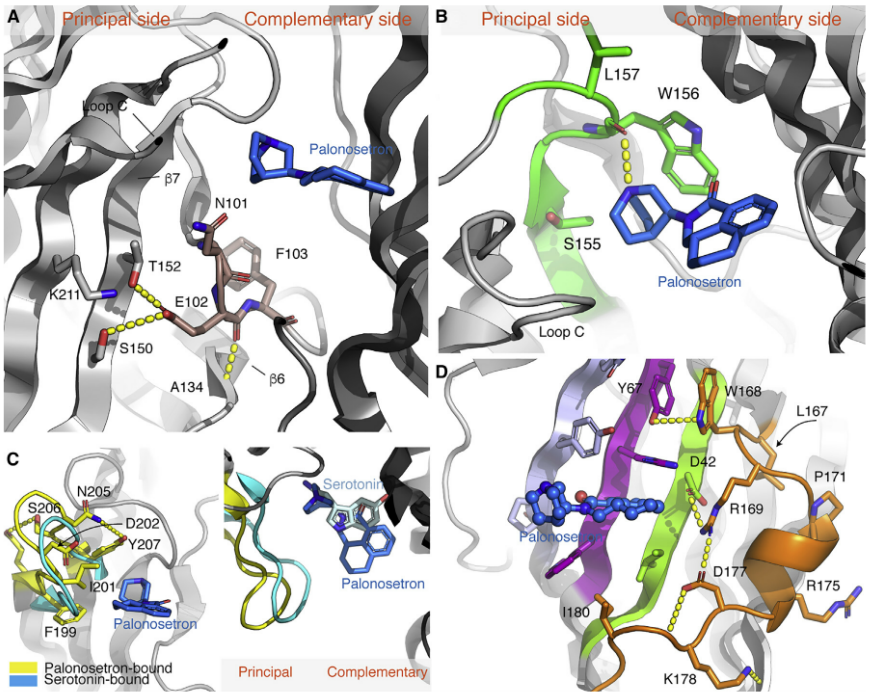


Figure 6. Interactions of Palonosetron with Binding Loops A, B, C and F

(A) The three most important residues of loop A are shown as sticks, together with neighbors of E102. The interactions of E102 described in the text are represented as dashes.

(B) Residues of loop B facing palonosetron are shown as sticks. The interaction of W156 main chain with the positively charged nitrogen of the ligand is depicted as a dashed line.

(C) The serotonin-bound activated structure is superimposed to the palonosetron inhibited one (on the complementary subunit) to highlight the yellow→cyan capping motion of loop C. Left panel, view parallel to the membrane plane; right panel, view perpendicular to the membrane plane.

(D) Loop F residues (orange), the mutations of which are discussed in the text, are represented as sticks.

ones, which are most apparent at the level of strands β8 and β9, and of loop F. In line with the closed phenotype in the absence of agonist, no deviation is observed at the ECD/TMD interface, a key region for the transduction of the signal from the orthosteric site to the pore gate.

5-HT3 Receptor Inhibitors beyond Emesis?

5-HT3 receptor antagonists are widely used for the control of emesis. Their potential in other therapeutic areas, by themselves or as adjunct drugs, has been explored in pre-clinical and clinical studies with equivocal results. A significant fraction of these studies were carried out more than two decades ago and it is worth combining older (Costall and Naylor, 2004) and more recent reviews (Fakhfouri et al., 2019; Gupta et al., 2016; Sanger and Andrews, 2018) to get a sense of the yet unmet promises of 5-HT3 receptor inhibition in areas, such as depression, anxiety, schizophrenia, and drug withdrawal.

In parallel, roles of 5-HT3 receptors in a variety of tissues continue to be uncovered, for example, in the lacrimal gland (Imada et al., 2017), in the development of the urinary system (Ritter et al., 2017), as well as in adipocytes (Oh et al., 2015). These and future similar studies may one day help reinterpret older observations from (pre)clinical studies. In our era of drug repurposing, -setron molecules, which enjoy a good safety profile and a relative absence of adverse effects, might spur renewed interest.

Conclusion

Detailed knowledge of the mode of action of antiemetics with the 5-HT3 receptor has remained hitherto fragmentary. We have solved, using cryo-EM, the palonosetron-bound structure of a mouse 5-HT3 receptor, which, amid the very few related structures available, provides valuable information toward understanding inhibitory mechanisms at the atomic level. Combining our structural and computational data, we put forth a convenient

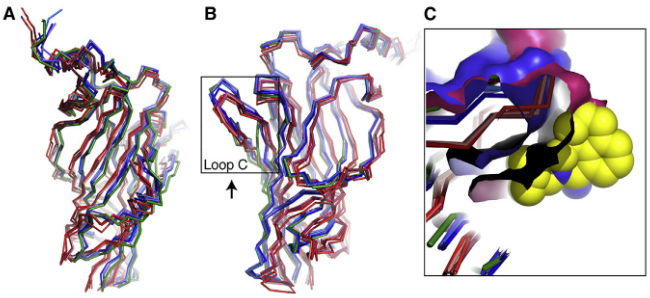


Figure 7. The Inhibited Conformations Form a Group that Differs from the Active Ones

Because the quaternary organization is key to the gating of the 5-HT3 receptor, we here represent the principal subunit ECD after alignment of the complementary subunit ECD. Active conformations are depicted in red, inhibited in blue, while the apo one is in green. All the inhibited conformations are clearly grouped, all the active ones as well. The inhibited group is closer to the apo conformation, and similar at the ECD/TMD interface while differences are seen in β8 and β9.

(A and B) Two different orientations where the subunit is seen from the point of view of its two neighbors.

(C) Palonosetron (yellow spheres) fits tightly in its binding site (surface of loop C in blue) but would clash with the binding site of the active conformation (red surface of loop C).

framework for rationalizing the host of biochemical and pharmacological information accrued since the first -setron drug synthesis. This framework is well-suited for correlating specific mutations to observed phenotypes, and for establishing a general structure-activity relationship for antiemetic drugs. As more ligand-bound structures of serotonin 5-HT3 receptors will likely be obtained at increasing resolutions, our grasp of pharmacology and conformational equilibrium will deepen. In the coming years we also anticipate that experimental data on asymmetry (either in homomeric or heteromeric receptors) and on conformational variability will continue to reshape the current view on the modus operandi of these fascinating allosteric neuroreceptors.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cultured Cells
- METHOD DETAILS
  - Protein Production and Purification
  - Electron Microscopy and Image Analysis
  - Molecular Dynamics Simulations Details
  - Model Refinement and Structure Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2020.07.004.

ACKNOWLEDGMENTS

We acknowledge access to ESRF CM01 Krios microscope. We thank L. Estrozi and A. Desfosses for advice in image analysis; members of the Nury group for discussions. The work was funded by ERC Starting grant 637733 Pentabrain. It used the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The electron microscopy facility is supported by the Rhône-Alpes Region, the FRM, the FEDER, and the GIS IBiSA. F.D. acknowledge the State-Region Plan "Technological Innovations, Modeling and Personalized Medical Support" (IT2MP) and the European Regional Development Funds (ERDF) for generous support. H.Z. is grateful for the financial support from China Scholarship Council (no. 201806200074).

AUTHOR CONTRIBUTIONS

Conceptualization, E.Z., H.Z., F.D., and H.N.; Methodology and Investigation, all authors; Writing – Original Draft, F.D. and H.N.; Writing – Review & Editing, all authors; Funding Acquisition, C.C., F.D., and H.N.; Methodology and Resources, G.E. and G.S.; Supervision, F.D. and H.N.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 24, 2020
Revised: June 18, 2020
Accepted: July 8, 2020
Published: July 28, 2020

SUPPORTING CITATIONS

The following reference appears in the Supplemental Information: Kelley et al., 2003

REFERENCES

Andersen, H.C. (1983). Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. J. Comput. Phys. *52*, 24–34.

Bang-Andersen, B., Ruhland, T., Jørgensen, M., Smith, G., Frederiksen, K., Jensen, K.G., Zhong, H., Nielsen, S.M., Hogg, S., Mørk, A., et al. (2011). Discovery of 1-[2-(2,4-dimethylphenylsulfanyl)phenyl]piperazine (Lu AA21004): a novel multimodal compound for the treatment of major depressive disorder. J. Med. Chem. *54*, 3206–3221.

Basak, S., Gicheru, Y., Samanta, A., Molugu, S.K., Huang, W., Fuente, M., Hughes, T., Taylor, D.J., Nieman, M.T., Moiseenkova-Bell, V., et al. (2018a). Cryo-EM structure of 5-HT3A receptor in its resting conformation. Nat. Commun. *9*, 514.

Basak, S., Gicheru, Y., Rao, S., Sansom, M.S.P., and Chakrapani, S. (2018b). Cryo-EM reveals two distinct serotonin-bound conformations of full-length 5-HT3A receptor. Nature *563*, 270–274.

Basak, S., Gicheru, Y., Kapoor, A., Mayer, M.L., Filizola, M., and Chakrapani, S. (2019). Molecular mechanism of setron-mediated inhibition of full-length 5-HT3A receptor. Nat. Commun. *10*, 3225.

Beene, D.L., Brandt, G.S., Zhong, W., Zacharias, N.M., Lester, H.A., and Dougherty, D.A. (2002). Cation-pi interactions in ligand recognition by serotonergic (5-HT3A) and nicotinic acetylcholine receptors: the anomalous binding properties of nicotine. Biochemistry *41*, 10262–10269.

Beene, D.L., Price, K.L., Lester, H.A., Dougherty, D.A., and Lummis, S.C.R. (2004). Tyrosine residues that control binding and gating in the 5-hydroxytryptamine3 receptor revealed by unnatural amino acid mutagenesis. J. Neurosci. *24*, 9097–9104.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., and MacKerell, A.D. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi1$ and $\chi2$ dihedral angles. J. Chem. Theory Comput. *8*, 3257–3273.

Boess, F.G., Steward, L.J., Steele, J.A., Liu, D., Reid, J., Glencorse, T.A., and Martin, I.L. (1997). Analysis of the ligand binding site of the 5-HT3 receptor using site directed mutagenesis: importance of glutamate 106. Neuropharmacology *36*, 637–647.

Celie, P.H.N., van Rossum-Fikkert, S.E., van Dijk, W.J., Brejc, K., Smit, A.B., and Sixma, T.K. (2004). Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. Neuron *41*, 907–914.

Chey, W.D., Kurlander, J., and Eswaran, S. (2015). Irritable bowel syndrome: a clinical review. JAMA *313*, 949–958.

Clark, R.D., Miller, A.B., Berger, J., Repke, D.B., Weinhardt, K.K., Kowalczyk, B.A., Eglen, R.M., Bonhaus, D.W., and Lee, C.H. (1993). 2-(Quinuclidin-3-yl) pyrido[4,3-b]indol-1-ones and isoquinolin-1-ones. Potent conformationally restricted 5-HT3 receptor antagonists. J. Med. Chem. *36*, 2645–2657.

Costall, B., and Naylor, R.J. (2004). 5-HT3 receptors. Curr. Drug Targets CNS Neurol. Disord. *3*, 27–37.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Del Cadia, M., De Rienzo, F., Weston, D.A., Thompson, A.J., Menziani, M.C., and Lummis, S.C.R. (2013). Exploring a potential palonosetron allosteric binding site in the 5-HT(3) receptor. Bioorg. Med. Chem. *21*, 7523–7528.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Estrozi, L.F., and Navaza, J. (2008). Fast projection matching for cryo-electron microscopy image reconstruction. J. Struct. Biol. *162*, 324–334.

Fakhfouri, G., Rahimian, R., Dyhrfjeld-Johnsen, J., Zirak, M.R., and Beaulieu, J.-M. (2019). 5-HT3 receptor antagonists in neurologic and neuropsychiatric disorders: the iceberg still lies beneath the surface. Pharmacol. Rev. *71*, 383–412.

Feller, S.E., Zhang, Y., Pastor, R.W., and Brooks, B.R. (1995). Constant pressure molecular dynamics simulation: the Langevin piston method. J. Chem. Phys. *103*, 4613–4621.

Gao, F., Bren, N., Burghardt, T.P., Hansen, S., Henchman, R.H., Taylor, P., McCammon, J.A., and Sine, S.M. (2005). Agonist-mediated conformational changes in acetylcholine-binding protein revealed by simulation and intrinsic tryptophan fluorescence. J. Biol. Chem. *280*, 8443–8451.

Grunberg, S.M., Vanden Burgt, J.A., Berry, S., Rubenstein, E.B., and Berry, D. (2004). Prevention of delayed nausea and vomiting (D-CINV): carryover effect analysis of pooled data from 2 phase III studies of palonosetron (PALO). J. Clin. Oncol. *22*, 8051.

Gupta, D., Prabhakar, V., and Radhakrishnan, M. (2016). 5HT3 receptors: target for new antidepressant drugs. Neurosci. Biobehav. Rev. *64*, 311–325.

Hassaine, G., Deluz, C., Tol, M.B., Li, X.-D., Graff, A., Vogel, H., and Nury, H. (2013). Large scale expression and purification of the mouse 5-HT3 receptor. Biochim. Biophys. Acta *1828*, 2544–2552.

Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M.B., Hovius, R., Graff, A., Stahlberg, H., Tomizaki, T., Desmyter, A., et al. (2014). X-ray structure of the mouse serotonin 5-HT3 receptor. Nature *512*, 276–281.

Hassaïne, G., Deluz, C., Grasso, L., Wyss, R., Hovius, R., Stahlberg, H., Tomizaki, T., Desmyter, A., Moreau, C., Peclinovska, L., et al. (2017). Expression, biochemistry, and stabilization with camel antibodies of membrane proteins: case study of the mouse 5-HT3 receptor. Methods Mol. Biol. *1635*, 139–168.

Hopkins, C.W., Le Grand, S., Walker, R.C., and Roitberg, A.E. (2015). Long-time-step molecular dynamics through hydrogen mass repartitioning. J. Chem. Theory Comput. *11*, 1864–1874.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38, 27–8.

Imada, T., Nakamura, S., Hisamura, R., Izuta, Y., Jin, K., Ito, M., Kitamura, N., Tanaka, K.F., Mimura, M., Shibuya, I., et al. (2017). Serotonin hormonally regulates lacrimal gland secretory function via the serotonin type 3a receptor. Sci. Rep. *7*, 6965.

Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. J. Comput. Chem. *29*, 1859–1865.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Kandiah, E., Giraud, T., de Maria Antolinos, A., Dobias, F., Effantin, G., Flot, D., Hons, M., Schoehn, G., Susini, J., Svensson, O., et al. (2019). CM01: a facility for cryo-electron microscopy at the European Synchrotron. Acta Crystallogr. D Struct. Biol. *75*, 528–535.

Kelley, S.P., Dunlop, J.I., Kirkness, E.F., Lambert, J.J., and Peters, J.A. (2003). A cytoplasmic region determines single-channel conductance in 5-HT3 receptors. Nature *424*, 321–324.

Kesters, D., Thompson, A.J., Brams, M., van Elk, R., Spurny, R., Geitmann, M., Villalgordo, J.M., Guskov, A., Helena Danielson, U., Lummis, S.C.R., et al. (2013). Structural basis of ligand recognition in 5-HT(3) receptors. EMBO Rep. *14*, 49–56.

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A.D., Jr., and Pastor, R.W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J. Phys. Chem. B *114*, 7830–7843.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. *60*, 2256–2268.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods *11*, 63–65.

Lee, W.Y., and Sine, S.M. (2005). Principal pathway coupling agonist binding to channel gating in nicotinic receptors. Nature *438*, 243–247.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877.

Lummis, S.C.R. (2012). 5-HT3 receptors. J. Biol. Chem. *287*, 40239–40245.

Lummis, S.C.R., and Thompson, A.J. (2013). Agonists and antagonists induce different palonosetron dissociation rates in 5-HT_3A and 5-HT_3AB receptors. Neuropharmacology *73*, 241–246.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

Miyamoto, S., and Kollman, P.A. (1992). Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J. Comput. Chem. *13*, 952–962.

Monod, J., Wyman, J., and Changeux, J.P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Mosesso, R., and Dougherty, D.A. (2018). A triad of residues is functionally transferrable between 5-HT3 serotonin receptors and nicotinic acetylcholine receptors. J. Biol. Chem. *293*, 2903–2914.

Nemecz, À., Prevost, M.S., Menny, A., and Corringer, P.-J. (2016). Emerging molecular mechanisms of signal transduction in pentameric ligand-gated ion channels. Neuron *90*, 452–470.

Oh, C.-M., Namkung, J., Go, Y., Shong, K.E., Kim, K., Kim, H., Park, B.-Y., Lee, H.W., Jeon, Y.H., Song, J., et al. (2015). Regulation of systemic energy homeostasis by serotonin in adipose tissues. Nat. Commun. *6*, 6794.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Polovinkin, L., Hassaine, G., Perot, J., Neumann, E., Jensen, A.A., Lefebvre, S.N., Corringer, P.-J., Neyton, J., Chipot, C., Dehez, F., et al. (2018). Conformational transitions of the serotonin 5-HT3 receptor. Nature *563*, 275–279.

Price, K.L., Bower, K.S., Thompson, A.J., Lester, H.A., Dougherty, D.A., and Lummis, S.C.R. (2008). A hydrogen bond in loop A is critical for the binding and function of the 5-HT3 receptor. Biochemistry *47*, 6370–6377.

Price, K.L., Lillestol, R.K., Ulens, C., and Lummis, S.C.R. (2016). Palonosetron-5-HT3 receptor interactions as shown by a binding protein cocrystal structure. ACS Chem. Neurosci. *7*, 1641–1646.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Purohit, P., and Auerbach, A. (2013). Loop C and the mechanism of acetylcholine receptor-channel gating. J. Gen. Physiol. *141*, 467–478.

Richardson, B.P., Engel, G., Donatsch, P., and Stadler, P.A. (1985). Identification of serotonin M-receptor subtypes and their specific blockade by a new class of drugs. Nature *316*, 126–131.

Ritter, K.E., Wang, Z., Vezina, C.M., Bjorling, D.E., and Southard-Smith, E.M. (2017). Serotonin receptor 5-HT3A affects development of bladder innervation and urinary bladder function. Front. Neurosci. *11*, 690.

Rojas, C., Stathis, M., Thomas, A.G., Massuda, E.B., Alt, J., Zhang, J., Rubenstein, E., Sebastiani, S., Cantoreggi, S., Snyder, S.H., et al. (2008). Palonosetron exhibits unique molecular interactions with the 5-HT3 receptor. Anesth. Analg. *107*, 469–478.

Rojas, C., Thomas, A.G., Alt, J., Stathis, M., Zhang, J., Rubenstein, E.B., Sebastiani, S., Cantoreggi, S., and Slusher, B.S. (2010a). Palonosetron triggers 5-HT(3) receptor internalization and causes prolonged inhibition of receptor function. Eur. J. Pharmacol. *626*, 193–199.

Rojas, C., Li, Y., Zhang, J., Stathis, M., Alt, J., Thomas, A.G., Cantoreggi, S., Sebastiani, S., Pietra, C., and Slusher, B.S. (2010b). The antiemetic 5-HT3 receptor antagonist Palonosetron inhibits substance P-mediated responses in vitro and in vivo. J. Pharmacol. Exp. Ther. *335*, 362–368.

Ruepp, M.-D., Wei, H., Leuenberger, M., Lochner, M., and Thompson, A.J. (2017). The binding orientations of structurally-related ligands can differ; a cautionary note. Neuropharmacology *119*, 48–61.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. *23*, 327–341.

Sanger, G.J., and Andrews, P.L.R. (2018). A history of drug discovery for treatment of nausea and vomiting and the implications for future research. Front. Pharmacol. *9*, 913.

Steward, L.J., Boess, F.G., Steele, J.A., Liu, D., Wong, N., and Martin, I.L. (2000). Importance of phenylalanine 107 in agonist recognition by the 5-hydroxytryptamine(3A) receptor. Mol. Pharmacol. *57*, 1249–1255.

Stoltz, R., Cyong, J.C., and Shah, A. (2004). Pharmacokinetic and safety evaluation of palonosetron, a 5-hydroxytryptamine-3 receptor antagonist, in US and Japanese healthy subjects. J. Clin. Pharmacol. *44*, 520–531.

Sullivan, N.L., Thompson, A.J., Price, K.L., and Lummis, S.C.R. (2006). Defining the roles of Asn-128, Glu-129 and Phe-130 in loop A of the 5-HT3 receptor. Mol. Membr. Biol. *23*, 442–451.

Suryanarayanan, A., Joshi, P.R., Bikádi, Z., Mani, M., Kulkarni, T.R., Gaines, C., and Schulte, M.K. (2005). The loop C region of the murine 5-HT3A receptor contributes to the differential actions of 5-hydroxytryptamine and m-chlorophenylbiguanide. Biochemistry *44*, 9140–9149.

Tang, L.C., and Tang, S.J. (2012). Neurochemistry in Clinical Application (Springer Science & Business Media).

Thompson, A.J. (2013). Recent developments in 5-HT3 receptor pharmacology. Structure *34*, 100–109.

Thompson, A.J., Padgett, C.L., and Lummis, S.C.R. (2006). Mutagenesis and molecular modeling reveal the importance of the 5-HT3 receptor F-loop. J. Biol. Chem. *281*, 16576–16582.

Thompson, A.J., Lochner, M., and Lummis, S.C.R. (2008). Loop B is a major structural component of the 5-HT3 receptor. Biophys. J. *95*, 5728–5736.

Tuckerman, M., Berne, B.J., and Martyna, G.J. (1992). Reversible multiple time scale molecular dynamics. J. Chem. Phys. *97*, 1990–2001.

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., et al. (2010). CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. *31*, 671–690.

Venkataraman, P., Venkatachalan, S.P., Joshi, P.R., Muthalagi, M., and Schulte, M.K. (2002). Identification of critical residues in loop E in the 5-HT3ASR binding site. BMC Biochem. *3*, 15.

Weber, S., Volynets, V., Kanuri, G., Bergheim, I., and Bischoff, S.C. (2009). Treatment with the 5-HT3 antagonist tropisetron modulates glucose-induced obesity in mice. Int. J. Obes. *33*, 1339–1347.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. *27*, 293–315.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Wong, E.H., Clark, R., Leung, E., Loury, D., Bonhaus, D.W., Jakeman, L., Parnes, H., Whiting, R.L., and Eglen, R.M. (1995). The interaction of RS 25259-197, a potent and selective antagonist, with 5-HT3 receptors, in vitro. Br. J. Pharmacol. *114*, 851–859.

Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V., Venable, R.M., et al. (2014). CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. J. Comput. Chem. *35*, 1997–2004.

Yan, D., Schulte, M.K., Bloom, K.E., and White, M.M. (1999). Structural features of the ligand-binding domain of the serotonin 5HT3 receptor. J. Biol. Chem. *274*, 5537–5541.

Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

IV. Vectorisation Membranaire

Un autre aspect essentiel du transport membranaire est la vectorisation de principes actifs au sein des cellules dans le but d'agir sur telle ou telle cible protéique.[85] Les mécanismes d'association, de permutation de relargage à l'intérieur de la cellule restent à élucider.[86] La diffusion de composés soit sous forme libre, soit attachés à des nanovecteurs à travers la membrane dépend de leurs propriétés physicochimiques, mais également de leur état de protonation. Ce dernier évolue en fonction de l'environnement immédiat du composé, et est largement méconnu dans le cas d'acides aminés approchant la membrane.

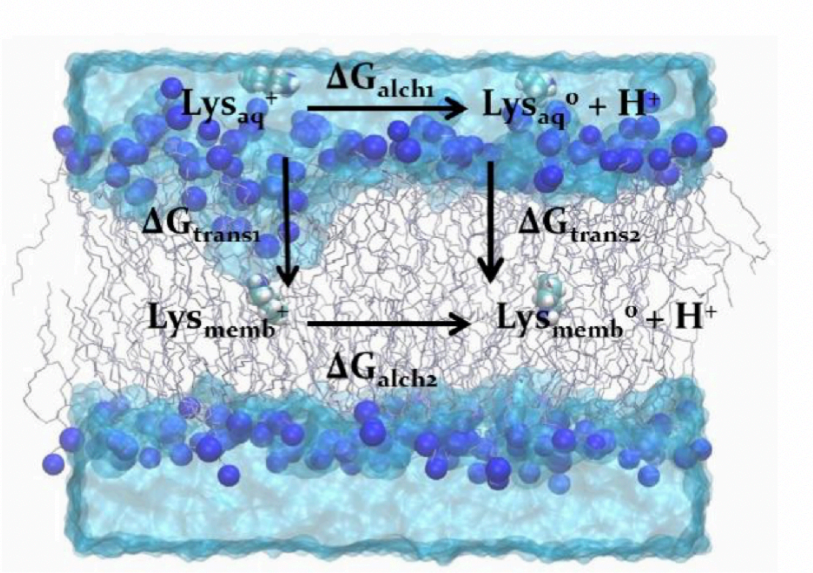


Figure 23. Cycle thermodynamique utilisé pour déterminer le pKa d'un analogue de lysine à différentes profondeurs d'insertion dans la membrane. Les chemins horizontaux correspondent aux transformations alchimiques d'un analogue de lysine chargée en sa forme neutre. Les chemins verticaux correspondent à la diffusion de la forme chargée ou neutre le long de la normale à la bicouche.

Nous avons grâce à des calculs d'énergie libre (potentiels de force moyenne et transformations alchimiques) prédit la probabilité de transfert d'analogues d'acides aminés en fonction de leur état de protonation (figure 23). Nous avons également caractérisé l'influence la nature de la membrane (effet des têtes polaires et des chaines grasses) sur le pKA des substrats et déterminé à quelle distance du centre de la membrane peut se produire la déprotonation d'un substrat donné. Nous avons ainsi pu proposer schéma complet qui permet de comprendre sous quelles conditions certaines fonctions chargées peuvent traverser des milieux hydrophobes tels que les membranes biologiques. L'étude thermodynamique a été complétée par une étude cinétique qui a permis de montrer que le transfert d'espèces comme les analogues de lysine bien que rare était un évènement rapide.[87,88]

Ci-après est reproduit 1 article illustrant mes travaux dédiés à la vectorisation:

- Bonhenry, D., Tarek, M.* & Dehez, F.* Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane. *J. Chem. Theory Comput.* **9**, 5675–5684 (2013).

V. Champs électriques et membranes

V.1. Modélisation d'une membrane biologique dans des « conditions physiologiques »

L'objectif est de modéliser une membrane biologique sous l'effet d'un champ électrique généré par des fluctuations locales de concentrations ioniques, reproduisant ainsi un potentiel transmembranaire tel que ressenti par les membranes des cellules.[89–91] Les quelques études publiées sur ce sujet utilisent en particulier un système de double bicouche permettant de casser la périodicité du système par la présence de bains séparés par les deux bicouches, et ainsi créer un défaut de charge dans un bain par rapport à l'autre.[92] Une des faiblesses de cette approche est notamment de générer deux champs électriques de sens opposés agissant respectivement sur chacune des deux doubles couches. De plus, celle-ci nécessite l'utilisation de cellules de modélisation de tailles importantes couteuses en temps de calcul. Nous avons choisi de développer une approche plus séduisante pour modéliser ce type d'effet permettant notamment de s'affranchir de la présence de deux bicouches.[93]

Nous avons dans un premier temps cassé la périodicité du système par l'introduction de « murs » de potentiel parallèles à la bicouche, perméables à l'eau mais pas aux ions, et situés à une distance égale de part et d'autre de la bicouche. De cette manière, la création de deux zones distinctes permet de générer deux bains où la répartition des charges peut être différente. Le potentiel transmembranaire (V) créé doit alors dépendre de la différence de charge nette (Q) entre les deux bains exprimés par unité de surface et de la capacitance (C) de la membrane étudiée : $V = Q/C$. Cette stratégie, bien qu'au départ prometteuse, s'est finalement avérée inefficace, le bain d'eau contenu en dehors de la zone des ions jouant le rôle de court-circuit et interdisant de fait à la membrane de se comporter comme un condensateur. La solution a été d'isoler les deux bains d'ions de part et d'autre de la membrane en introduisant une lame d'air suffisamment grande pour séparer les deux milieux. Un tel système permet de repartir les charges différemment de part et d'autre de la membrane sans qu'il y ait d'échange possible et sans que le « condensateur » - membrane ne soit court-circuité (figure 24).

Nous avons ainsi construit plusieurs systèmes de ce type en faisant varier la concentration en ions (sodium et chlore) de part et d'autre de la membrane ainsi que la différence de charge nette entre les deux bains, la charge totale du système restant quant à elle nulle. Plusieurs dynamiques moléculaires de quelques nanosecondes chacune ont été réalisées. L'analyse des résultats montre que le potentiel transmembranaire calculé à partir des trajectoires est en accord avec celui prédit par la formule $V = Q/C$. Une des caractéristiques intéressantes des systèmes est la formation de couches parallèles à la membrane pour lesquelles les répartitions en ions sont très spécifiques. Cette méthodologie également appliquée à un système composé d'un nanotube de peptide cyclique inséré dans une membrane a permis d'observer un flux d'ions au travers du pore sans aucune application d'une force externe.[93]

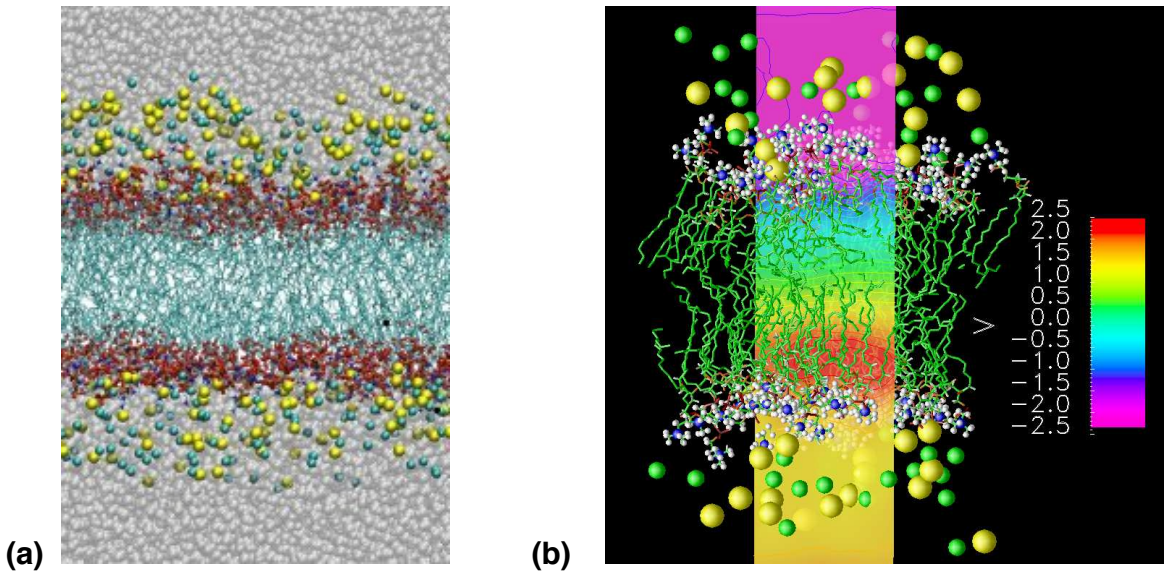


Figure 24. Les ions au voisinage de la bicouche sont contenus par des « murs » de potentiels imperméables aux ions. (b) Carte de potentiel électrostatique générée pour un système dans lequel, les deux bains ioniques, portant une charge nette différente, sont séparés par une lame d'air.

V.2. Électroporation

L'électroporation, l'application de champs électriques pour modifier la perméabilité des membranes biologiques, est récemment devenu un outil clinique pour le traitement de divers cancers.[94] La théorie actuelle d'électroporation suppose que la membrane est perméabilisée par la formation de pores hydrophiles conducteurs, stabilisées par un réarrangement de groupes de tête lipidiques.[95] Nous avons effectué des simulations de dynamique moléculaire de bicouches lipidiques chargées négativement soumis à des tensions enlevées transmembranaires. Nos données révèlent un processus d'électroporation jusque-là inconnu dans lequel de grandes colonnes d'eau conductrices d'ions sont formées au travers de membranes sans qu'elles ne soient stabilisées par des groupes de tête lipidiques à l'intérieur du noyau hydrophobe de la bicouche (figure 25).[96] Bien que les étapes initiales de l'électroporation soient les mêmes pour des pores hydrophiles et hydrophobes, les seconds présentent un seuil d'électroporation plus élevé et une durée de vie réduite. L'existence de ces pores hydrophobes conteste la description théorique standard décrivant la création de pores dans les membranes lipidiques. Nos résultats montrent que le processus d'électroporation est modulé par la nature de la membrane ; celle-ci devant être prise en compte pour calibrer les protocoles médicaux reposant sur l'électroporation.

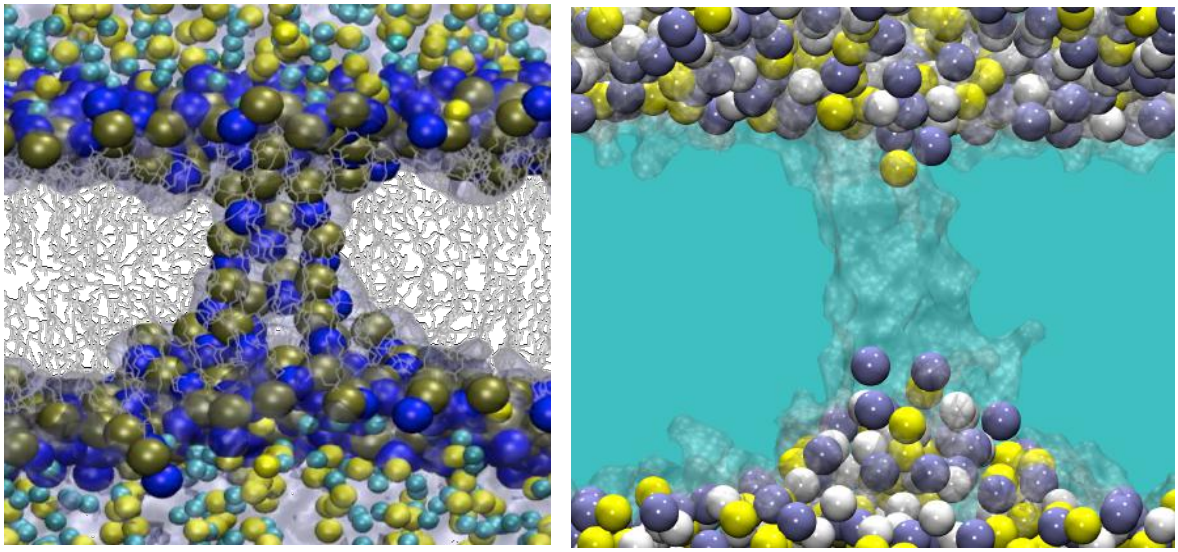


Figure 25. Vues instantanées de la formation de pores hydrophiles et hydrophobes sous l'effet d'un potentiel transmembranaire dans respectivement (gauche) une membrane de POPC (lipides zwitterioniques) et (droite) une membrane de POPS (lipides chargés négativement).

Ci-après est reproduit 1 article illustrant mes travaux dédiés aux effets des champs électrique sur les membranes :

- Dehez, F. *, Delemotte, L., Kramar, P., Miklavčič, D. & Tarek, M. * Evidence of Conducting Hydrophobic Nanopores Across Membranes in Response to an Electric Field. *J. Phys. Chem. C* **118**, 6752–6757 (2014).

THE JOURNAL OF PHYSICAL CHEMISTRY C

Article

pubs.acs.org/JPCC

Evidence of Conducting Hydrophobic Nanopores Across Membranes in Response to an Electric Field

François Dehez,*[,‡,§] Lucie Delemotte,[‡,§,†] Peter Kramar,[⊥] Damijan Miklavčič,[⊥] and Mounir Tarek*[,‡,§]

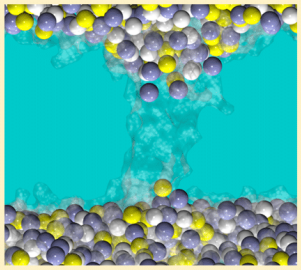
[‡]CNRS, UMR 7565, SRSMC, F-54506 Vandœuvre-lès-Nancy, France

[§]Université de Lorraine, UMR 7565, SRSMC, F-54506 Vandœuvre-lès-Nancy, France

[⊥]Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Supporting Information

ABSTRACT: Electroporation, the application of electric fields to alter the permeability of biological membranes, has recently become a clinical tool for the electrochemotherapy treatment of various cancers. Current electroporation theory assumes that the membrane is permeabilized through the formation of conducting hydrophilic pores, stabilized by rearrangement of lipid head groups. Here we have performed molecular dynamics simulations of negatively charged lipid bilayers subject to high transmembrane voltages together with electroporation experiments on planar bilayers. Our data reveal a hitherto unknown electroporation process in which large ion-conducting water columns not stabilized by lipid head groups are formed within the bilayer's hydrophobic core. The existence of such hydrophobic pores challenges the standard theoretical description of pore creation in lipid membranes. Our findings open a new vista toward fine-tuning of electroporation-based treatments and biotechnical applications, and, in general, for enhancing the import of various substrates in liposomes or cells.



INTRODUCTION

Electroporation (EP) is a process in which lipid membranes, the cell envelopes, are permeabilized when subjected to high enough electric fields.[1] The technique also known as electropermeabilization[2,3] is nowadays widely used in various applications ranging from in vitro DNA and siRNA cell delivery[4,5] to clinical electrochemotherapy where delivery of drugs to cancer cells is enhanced.[6−8] Experimental evidence suggests that the electric fields act primarily in the lipid domain of the membrane, and the key features of EP are based on theories involving the formation of stochastic pores.[9] In erythrocyte membranes, large pores could be observed by using electron microscopy,[10] but in general, direct observation of nanosized pores is not possible with conventional techniques. It is, however, largely accepted that high transmembrane (TM) voltages produce aqueous-filled pores in the lipid bilayer[10−14] as evidenced by early molecular-dynamics simulation[15,16] and recent experiments.[17,18] The properties of pores, e.g., size, density, and lifetime (after the electric pulse shutdown), are key factors that determine how leaky a membrane can become.[19]

Traditionally, electroporation is triggered by applying long (microseconds) electric pulses of low magnitude (kV/m). The resulting current causes an accumulation of electrical charges at both sides of the cell membrane over a time that depends on the electrical parameters of the system, often in the order of 100s of ns.[20−24] As the membrane behaves as a capacitor, this charge accumulation gives rise to a TM voltage. Simulation protocols have so far been developed to mimic *in silico* the effect of low magnitude microsecond pulses (msEP) on planar bilayers[25] by imposing a net charge imbalance across the membrane. Such a protocol has been used to monitor the effect of high voltages on zwitterionic membranes.[26−31] The simulations show that it triggers rearrangements of the membrane components, i.e., water and lipids. When the TM voltage is above a threshold value characteristic of the lipid composition, the membrane loses its integrity. This enables the ionic and molecular transport through the otherwise impermeable membranes to increase substantially.[18] Electroporation starts with the formation of water fingers that protrude inside the hydrophobic core of the membrane. Within nanoseconds, water wires bridging the two sides of the lipid bilayer appear. If the simulations are further extended, lipid head groups migrate along one wire and form a hydrophilic connected pathway and if voltage is maintained, ions start to flow through this pathway. Moreover, when the TM voltage is lowered, or is reduced due to ionic flow across the membrane, the pores collapse to reach a nonsolvated and nonconductive state.[25,27−29] It was shown that full recovery of the membrane integrity (migration of the lipid head groups back to the lipid water interface) requires few tens of nanoseconds.[29,32]

Although a considerable effort has been devoted to characterize basic processes of classical electroporation at a molecular level better, the details of the phenomenon in

Received: November 22, 2013
Revised: February 21, 2014
Published: February 24, 2014

membranes of increasing complexity still remain to be elucidated. Indeed, to date, most simulations have been performed on membranes formed by zwitteroinic lipids (phosphatidyl choline (PC) head groups) with the exception of a few that considered for instance the presence of negatively charged lipids to study their externalization,[31,33−35] a cholesterol fraction,[36] or *Escherichia coli* and *Staphylococcus aureus* model membranes.[37] Membranes which are formed by a variety of lipids can form several phases, with domains of various compositions. Today we still lack information about the response of such membranes to external electrical stress such as that leading to electroporation. This is precisely what this study aims at. We characterize here in particular electroporation of lipid bilayers formed by negatively charged lipids.

■ COMPUTATIONAL DETAILS

We considered a fully hydrated palmitoyl-oleyl-phosphatidylserine (POPS) lipid bilayer embedded in a ∼250 mM NaCl solution. The system consisted of 392 lipids, 39 200 water molecules, 576 sodium and 184 chloride ions (a total of 144 624 atoms). The bilayer was first replicated in all three dimensions and equilibrated at constant pressure (1 atm) and constant temperature (300 K) (103 × 100 × 160 Å^3). At the temperature set for the study, i.e., 300 K, the bilayer is in the biologically relevant liquid crystal Lα phase. The area per lipid reached at equilibration was ∼52.5 Å^2, close to the 55 Å^2 estimated for a POPS bilayer using a different force field.[38]

The MD simulations presented here were carried out with use of the program NAMD targeted for massively parallel architectures.[39] The systems were examined in the NPT (1 atm and 300 K) or NVT (300 K) ensembles employing the Langevin dynamics and Langevin piston method. The equations of motion were integrated by using a multiple time step algorithm. A time step of 2.0 fs was employed. Short- and long-range forces were calculated every 1 and 2 time steps, respectively. Chemical bonds between hydrogen and heavy atoms were constrained to their equilibrium value. Long-range, electrostatics forces were taken into account by using a fast implementation of the particle mesh Ewald (PME) approach,[40,41] with a direct space sum tolerance of 10^{-6} and a spherical truncation of 11 Å. The water molecules were described by using the TIP3P model[42] and a united-atom representation was adopted for the acyl chains of the POPS lipid molecules[43] while keeping intact the CHARMM27 description of the head groups.

The system was first equilibrated at constant temperature (300 K) and constant pressure (1 atm), as in the simulations of multilamellar stacks of lipids. The simulation box was then extended in the z direction perpendicular to the bilayer (to 360 Å) in order to create air (vacuum)−water interfaces, enabling a charge imbalance (ΔQ) between the solutions on either side of the membrane to be imposed.[25,44] Including a vacuum slab between replicas (use of 3D boundary conditions) along the membrane normal in order to prevent communication between baths is an alternative workaround to the more costly double bilayer setup (see the Supporting Information (SI), Figure S1).[26] As far as water slabs that are thicker than 25−30 Å, the presence of air water interfaces has no incidence on the lipid bilayer properties and the membrane "feels" as if it is embedded in infinite baths whose characteristics are those of the modeled finite solutions. In such a setup, the simulation is carried out at constant volume, and therefore a surface tension builds in as the bilayer undergoes electroporation.

The charge imbalance setup enables generating a transmembrane voltage by imposing a charge imbalance between the solutions on either side of the POPS bilayer. Such protocols aim at mimicking the effects of low-field pulses of microsecond or millisecond duration applied to this cell that result in an accumulation of electrical charges at both sides of the cell membrane. Note that here the "charging" of the membrane is not modeled. Rather, the simulations are initiated by assuming that the charging has already taken place, i.e., with a system already set at a specific (selected) charge imbalance. A charge imbalance ΔQ between the two sides of the bilayer is generated by simply displacing at time t = 0 an adequate number of ions from one side to the other.

Eventually, a set of MD simulations was carried out at varying transmembrane voltages (see the SI, Table S1). The lower voltage, giving rise to electroporation on a time scale accessible to MD simulations, amounts to 4 V. Statistics were thus gathered specifically at the threshold voltage of 4 V (10 simulations of 20 to 40 ns). Three simulations were also conducted at higher TM voltages (see the SI, Table S1).

■ EXPERIMENTAL METHODS

The salt solution consists of 0.1 M KCl and 0.01 M Hepes in the same proportion. NaOH (1 M) was added to obtain a neutral pH (7.4). POPS (1-pamitoyl 2-oleoyl phosphatidylserine) lipids in powder form (Avanti Polar-Lipids Inc. USA) were melted in a 9:1 hexane/ethanol solution. A 3:7 mixture of hexadecane and pentane was used for torus forming.

The chamber where planar lipid bilayers are formed consists of two 5.3 cm^3 Teflon reservoirs separated by a thin Teflon sheet with a 117 μm diameter round aperture[45] on which bilayers are built by using the Montal−Muller method.[46] The measuring system[47] was comprised of four Ag-AgCl electrodes (2 current and 2 reference electrodes) and consisted of two modules allowing the measurement of the lipid bilayer capacitance and the voltage across the bilayer, and a potentiostat-galvanostat for the current control. The specific membrane (bilayer) capacitance (c_{BLM}) normalized to the surface area of the bilayer was measured with capacitance to period conversion.[17] Current clamp measurements were performed by using a measuring system, as described by Kalinowski et al.[48,49] and by using four Ag-AgCl electrodes (two current electrodes and two reference electrodes). The measuring system consists of two modules. The first is a capacity to period converter, used for measuring the bilayer's capacitance. The second is a potentiostat-galvanostat for current clamp planar lipid bilayer studies. Both modules were controlled with a personal computer.

Measuring protocols consist of two parts: capacitance measurement and lipid bilayer breakdown voltage measurement. The membrane capacitance was measured with capacitance to period converting measuring principle described in detail by Kalinowski et al. and then normalized to the surface area of the bilayer to calculate the specific capacitance (c_{BLM}). We then determined breakdown voltage (U_{br}) of each lipid bilayer by applying linear rising current signals (of slope k) (see the SI, Figure S2). Six different slopes were selected: 0.5, 1, 4, 8, 10, 20 μA/s. U_{br} was defined as the voltage at t_{br}, when a sudden drop in the voltage was detected (see the SI, Figure S2). Voltage breakdown U_{br} was measured for 34 planar lipid bilayer constructs (see the SI, Table S2). The average membrane capacitance was 0.41 ± 0.13 μF/cm^2. Voltage drops such as

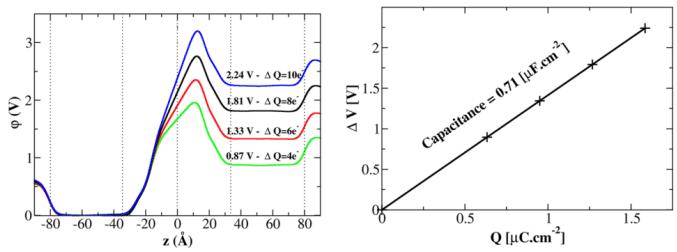


Figure 1. (Left) Electrostatic potential across a POPS lipid bilayer for different net charge imbalances ΔQ between the upper and lower electrolytes from MD simulations considering the setup of Figure S1 in the SI. $\varphi(z)$ is estimated as an in-plane average of the EP distributions (see the SI for calculation details). As a reference it was set to zero in the lower electrolyte. (Right) TM potential ΔV as a function of the charge imbalance Q per unit area. The capacitance C of the bilayer can be derived from the slope of the curve.

observed in previous studies on POPC were not present in any of experiment on POPS planar lipid bilayers.

■ RESULTS AND DISCUSSION

MD simulations performed at different TM voltages (ΔQ) showed that, as demonstrated previously for POPC (palmitoyl-oleyl-phosphatidyl-choline) membranes, the POPS bilayer behaves as a capacitor (see Figure 1). Its capacitance of ~0.71 μF/cm^2 (see Figure 1) is 16% lower than the value found for the POPC membrane (0.85 μF/cm^2) under similar conditions (salt concentration, FF parameters). This is consistent with the 20% decrease found experimentally (0.51 μF/cm^2 for POPC[50] and 0.41 μF/cm^2 for POPS), which tends to indicate that the setup and the MD force field we use are reasonable for comparison between the two bilayers.

As in the case of zwitterionic membranes, the process of electroporation starts with the protrusion of water wires from either one or both sides of the membrane (see Figure 2). As far as only water is concerned, the latter process is reminiscent of the electric-field-induced transition from a water droplet to a water wire described by Cramer et al.[51] Starting from a water droplet absorbed on a polar surface droplet, the authors demonstrated that the rise of the water pillar is induced by an electrostatic pressure that overcomes surface tension at a critical field strength. Water defects appear within a few nanoseconds (1−16 ns with a majority ~5 ns; see the SI, Table S1). This then results in the water wires joining in the membrane core. Such a defect, hereafter called a "hydrophobic pore" to indicate that the water columns are in contact with the hydrophobic lipid tails, then expands. In a few occurrences (simulations nos. 4, 5, and 12 of Table S1, SI), as in the case of the POPC, some lipid head groups dive partially into the pore to stabilize its structure ("hydrophilic" pore, see Table S1, SI). However, surprisingly, in a majority of the trajectories, the pores remained "hydrophobic" (see Figures 1 and 2, well as Figure S3 in the SI). Regardless of the pore nature, ions were then driven along the electrical gradient through the pore resulting in a rapid drop of the TM voltage. Once the TM voltage reaches a few hundreds of millivolts (0.15 to 0.98 V), conduction stops and the pores collapse. While full recovery of the membrane is not observed for tens of nanoseconds in the case of hydrophilic pores,[29,32] resealing of the hydrophobic pores occurred on the other hand over a much shorter time scale, leading to a few nanoseconds pore lifetime. Electroporation was observed in each of the 13 simulations conducted.

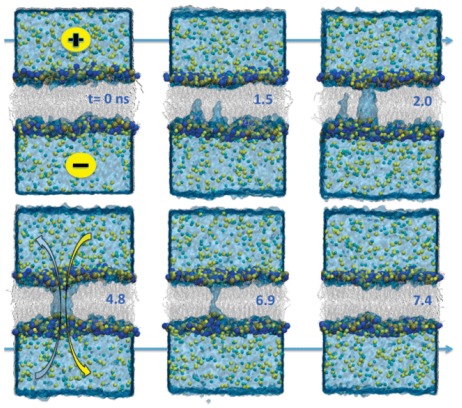


Figure 2. Snapshots along a representative MD trajectory of the formation of a hydrophobic pore within a POPS bilayer subject to a TM voltage of ~4 V. Yellow and cyan spheres depict Na^+ and Cl^- ions, respectively. Note that the phospholipid head groups (phosphatidyl: dark blue spheres; serine: tan spheres) remain localized at the lipid−water interface, or at the pore entrance, while water molecules (blue) flood the pore allowing for ion conduction.

To further investigate and corroborate these observations that suggest a drastic difference between the pores formed in charged (POPS) and those in zwitterionic (POPC) bilayers, we conducted experiments under current controlled conditions (linear rising current) similar to those recently reported for POPC bilayers.[17]

We studied bilayers prepared from POPS at 100 mM KCl concentration. Thirty-four planar bilayers were examined. The average membrane capacitance was 0.41 ± 0.13 μF/cm^2, ~20% lower than that of POPC (0.51 μF/cm^2).[50] Six linear rising current signal slopes k were used: 0.5, 1.0, 4.0, 8.0, 10.0, and 20.0 μA/s. As in previous investigations, for each system, the voltage breakdown U_{br} was defined as the voltage at the time t_{br}, when an abrupt voltage drop due to lipid bilayer rupture was detected (see Figure 3). The values of U_{br} were not statistically different for the various k values. The average breakdown voltage U_{br} was, however, much higher for POPS bilayers (398 ± 108 mV) then for POPC (307 ± 78 mV). Typical traces previously recorded for POPC are reported in Figure 4.

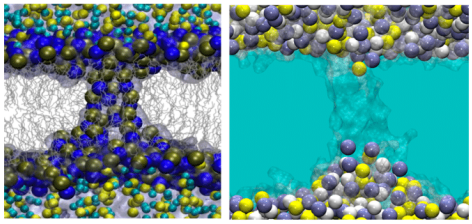


Figure 3. Snapshots of the formation of pores along representative MD trajectories within a POPS (right) and a POPC (left) bilayer subject to TM voltages leading to electroporation. Note the hydrophilic (POPC) vs hydrophobic (POPS) nature of the conducting pores. Same color code as Figure 1.

The voltage traces present transient small voltage drops often followed by a voltage rerise within a fraction of a second. Modeling the observed phenomenon by equivalent electric circuits showed that these events relate to opening and closing of conducting pores through the bilayer.[17] The most peculiar observation from the extended set of experiments on POPS bilayers performed here is that no transient voltage drop similar to those present in the POPC traces was observed in any of the 34 experiments on POPS membrane.

The high number of ion translocations monitored for each of the 13 simulations carried out in this work (see Table S1, SI) rule out the hypothesis of a low conductance pore to explain the lack of voltage drops in our experiments. In light of what has been found in the MD simulations, no detection of "drop events" in POPS bilayers may in fact be due to a recovery or resealing of pores formed as in the case of the POPC bilayers, which is too fast to be detected with actual measurement time resolution (0.01s) Accordingly, the traces of the experiments are consistent with the results from simulations.

■ CONCLUSION

This study shows for the first time a difference in the electroporation process between a model zwitterionic membrane and a model anionic membrane. While the initial steps of the electroporation are similar in both cases, several other parameters are different, among which an increased electroporation threshold, a hydrophobic nature of most pores formed, and a reduced lifetime of the latter. Since the 1970s the physical mechanism of electroporation (or pore creation) has been described by the Smoluchowski equation[11] where it is assumed that short-lived hydrophobic pores (i.e., pore with edges formed by hydrocarbon chains of the lipids) are first created at a rate $S(t)$ that depends exponentially on the square of the transmembrane potential $\sim V$. Pores with radii r above a critical radius ($r^* \approx 1$ nm) convert then spontaneously to long-lived stable hydrophilic pores. The overall process is governed by $S(t)$ and by the energy of the hydrophilic pore $W(r)$.[13,52] The possible existence of stable hydrophobic conductive pores challenges this interpretation.

Overall the present study indicates that the electroporation process is modulated by the nature of the membrane and may be different than so far assumed in "real" membranes made of a complex mixture of lipids, proteins, and sugars, possibly presenting domains with different lipid phases. Accordingly, the impact of our findings ought to be considered in biomedical applications of electroporation. On the biolotechnological front, protocols using electroporation to trigger drug-release from smart-liposome-based nanocarriers are being devised.[53,54] Our study suggests that such approaches can be fine-tuned by selecting specific lipid composition, including negatively charged species.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details, electrostatic calculation method, and movies extracted from the MD simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: francois.dehez@univ-lorraine.fr.
*E-mail: Mounir.Tarek@univ-lorraine.fr.

Present Address

†Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania, United States.

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency and by a bilateral cooperation program between France and Slovenia (PROTEUS). This manuscript is a result of the

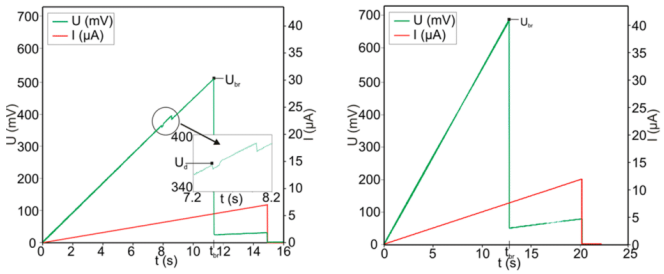


Figure 4. Typical traces of the measured TM voltage (U) under linearly rising current (I) conditions. (Left) POPC bilayers.[17] U_{br} is the voltage at which the bilayer breakdown occurs, and the arrow points at a transient voltage drop. After this event, the bilayer appears to recover completely within a fraction of a second.[17] (Right) For POPS planar lipid bilayer the voltage drops are not observed in any of the measurement.

networking efforts of the COST Action TD1104 (http://www.electroporation.net). The research was conducted in the scope of the EBAM European Associated Laboratory (LEA). Calculations were performed at the CINES, Montpellier, France, under grant no. 76434 (2010-2011). The authors thank the ANR Intcell program (ANR-10-BLAN-096).

■ REFERENCES

(1) Neumann, E.; Rosenheck, K. Permeability Changes Induced by Electric Impulses in Vesicular Membranes. *J. Membr. Biol.* **1972**, *10*, 279−290.
(2) Mir, L. M.; Banoun, H.; Paoletti, C. Introduction of Definite Amounts of Nonpermeant Molecules into Living Cells after Electropermeabilization: Direct Access to the Cytosol. *Exp. Cell Res.* **1988**, *175*, 15−25.
(3) Teissié, J.; Eynard, N.; Gabriel, B.; Rols, M. Electropermeabilization of Cell Membranes. *Adv. Drug Delivery Rev.* **1999**, *35*, 3−19.
(4) Golzio, M.; Teissié, J.; Rols, M.-P. Direct Visualization at the Single-Cell Level of Electrically Mediated Gene Delivery. *Proc. Natl. Acad. Sci.* **2002**, *99*, 1292−1297.
(5) Villemejane, J.; Mir, L. M. Physical Methods of Nucleic Acid Transfer: General Concepts and Applications. *Br. J. Pharmacol.* **2009**, *157*, 207−219.
(6) Mir, L. M.; Orlowski, S.; Belehradek, J., Jr.; Teissié, J.; Rols, M. P.; Serša, G.; Miklavčič, D.; Gilbert, R.; Heller, R. Biomedical Applications of Electric Pulses with Special Emphasis on Antitumor Electrochemotherapy. *Bioelectrochem. Bioenerg.* **1995**, *38*, 203−207.
(7) Serša, G.; Čemažar, M.; Miklavčič, D. Antitumor Effectiveness of Electrochemotherapy with Cis-Diamminedichloroplatinum (II) in Mice. *Cancer Res.* **1995**, *55*, 3450−3455.
(8) Heller, R.; Gilbert, R.; Jaroszeski, M. J. Clinical Applications of Electrochemotherapy. *Adv. Drug Delivery Rev.* **1999**, *35*, 119−129.
(9) Chang, D. C.; Chassy, B. M.; Saunders, J. A.; Sowers, A. E. *Guide to Electroporation and Electrofusion*; Academic Press: New York, NY, 1992.
(10) Weaver, J. C. Electroporation of Biological Membranes from Multicellular to Nano Scales. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 754−768.
(11) Abidor, I. G.; Arakelyan, V. B.; Chernomordik, L. V.; Chizmadzhev, Y. A.; Pastushenko, V. F.; Tarasevich, M. P. Electric Breakdown of Bilayer Lipid Membranes: I. The Main Experimental Facts and Their Qualitative Discussion. *J. Electroanal. Chem. Interfacial Electrochem.* **1979**, *104*, 37−52.
(12) Benz, R.; Beckers, F.; Zimmermann, U. Reversible Electrical Breakdown of Lipid Bilayer Membranes: A Charge-Pulse Relaxation Study. *J. Membr. Biol.* **1979**, *48*, 181−204.
(13) Weaver, J. C.; Chizmadzhev, Y. A. Theory of Electroporation: A Review. *Bioelectrochem. Bioenerg.* **1996**, *41*, 135−160.
(14) Chen, C.; Smye, S. W.; Robinson, M. P.; Evans, J. A. Membrane Electroporation Theories: A Review. *Med. Biol. Eng. Comput.* **2006**, *44*, 5−14.
(15) Tieleman, D. P.; Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Simulation of Pore Formation in Lipid Bilayers by Mechanical Stress and Electric Fields. *J. Am. Chem. Soc.* **2003**, *125*, 6382−6383.
(16) Tarek, M. Membrane Electroporation: A Molecular Dynamics Simulation. *Biophys. J.* **2005**, *88*, 4045−4053.
(17) Kramar, P.; Delemotte, L.; Maček Lebar, A.; Kotulska, M.; Tarek, M.; Miklavčič, D. Molecular-Level Characterization of Lipid Membrane Electroporation Using Linearly Rising Current. *J. Membr. Biol.* **2012**, *245*, 651−659.
(18) Breton, M.; Delemotte, L.; Silve, A.; Mir, L. M.; Tarek, M. Transport of siRNA through Lipid Membranes Driven by Nanosecond Electric Pulses: An Experimental and Computational Study. *J. Am. Chem. Soc.* **2012**, *134*, 13938−13941.
(19) Ho, M.-C.; Casciola, M.; Levine, Z. A.; Vernier, P. T. Molecular Dynamics Simulations of Ion Conductance in Field-Stabilized Nanoscale Lipid Electropores. *J. Phys. Chem. B* **2013**, *117*, 11633−11640.
(20) Hu, Q.; Viswanadham, S.; Joshi, R. P.; Schoenbach, K. H.; Beebe, S. J.; Blackmore, P. F. Simulations of Transient Membrane Behavior in Cells Subjected to a High-Intensity Ultrashort Electric Pulse. *Phys. Rev. E* **2005**, *71*, 031914.
(21) Beebe, S. J.; Schoenbach, K. H. Nanosecond Pulsed Electric Fields: A New Stimulus to Activate Intracellular Signaling. *J. Biomed. Biotechnol.* **2005**, *2005*, 297−300.
(22) Vasilkoski, Z.; Esser, A. T.; Gowrishankar, T. R.; Weaver, J. C. Membrane Electroporation: The Absolute Rate Equation and Nanosecond Time Scale Pore Creation. *Phys. Rev. E* **2006**, *74*, 021904.
(23) Sundararajan, R. Nanosecond Electroporation: Another Look. *Mol. Biotechnol.* **2009**, *41*, 69−82.
(24) Deng, J.; Schoenbach, K. H.; Stephen Buescher, E.; Hair, P. S.; Fox, P. M.; Beebe, S. J. The Effects of Intense Submicrosecond Electrical Pulses on Cells. *Biophys. J.* **2003**, *84*, 2709−2714.
(25) Delemotte, L.; Tarek, M. Molecular Dynamics Simulations of Lipid Membrane Electroporation. *J. Membr. Biol.* **2012**, *245*, 531−543.
(26) Sachs, J. N.; Crozier, P. S.; Woolf, T. B. Atomistic Simulations of Biologically Realistic Transmembrane Potential Gradients. *J. Chem. Phys.* **2004**, *121*, 10847−10851.
(27) Gurtovenko, A. A.; Vattulainen, I. Pore Formation Coupled to Ion Transport through Lipid Membranes as Induced by Transmembrane Ionic Charge Imbalance: Atomistic Molecular Dynamics Study. *J. Am. Chem. Soc.* **2005**, *127*, 17570−17571.
(28) Kandasamy, S. K.; Larson, R. G. Cation and Anion Transport through Hydrophilic Pores in Lipid Bilayers. *J. Chem. Phys.* **2006**, *125*, 074901−074901−9.
(29) Gurtovenko, A. A.; Anwar, J.; Vattulainen, I. Defect-Mediated Trafficking across Cell Membranes: Insights from in Silico Modeling. *Chem. Rev.* **2010**, *110*, 6077−6103.
(30) Venturini, A.; Zerbetto, F. Dynamics of a Lipid Bilayer Induced by Electric Fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9216−9222.
(31) Polak, A.; Bonhenry, D.; Dehez, F.; Kramar, P.; Miklavčič, D.; Tarek, M. On the Electroporation Thresholds of Lipid Bilayers: Molecular Dynamics Simulation Investigations. *J. Membr. Biol.* **2013**, *246*, 843−850.
(32) Vernier, P. T.; Sun, Y.; Gundersen, M. A. Nanoelectropulse-Driven Membrane Perturbation and Small Molecule Permeabilization. *BMC Cell Biol.* **2006**, *7*, 37.
(33) Vernier, P. T.; Ziegler, M. J.; Sun, Y.; Chang, W. V.; Gundersen, M. A.; Tieleman, D. P. Nanopore Formation and Phosphatidylserine Externalization in a Phospholipid Bilayer at High Transmembrane Potential. *J. Am. Chem. Soc.* **2006**, *128*, 6288−6289.
(34) Vernier, P. T.; Ziegler, M. J.; Sun, Y.; Gundersen, M. A.; Tieleman, D. P. Nanopore-Facilitated, Voltage-Driven Phosphatidylserine Translocation in Lipid Bilayers—in Cells and in Silico. *Phys. Biol.* **2006**, *3*, 233.
(35) Vernier, P. T.; Ziegler, M. J. Nanosecond Field Alignment of Head Group and Water Dipoles in Electroporating Phospholipid Bilayers. *J. Phys. Chem. B* **2007**, *111*, 12993−12996.
(36) Fernández, M. L.; Marshall, G.; Sagués, F.; Reigada, R. Structural and Kinetic Molecular Dynamics Study of Electroporation in Cholesterol-Containing Bilayers. *J. Phys. Chem. B* **2010**, *114*, 6855−6865.
(37) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Electroporation of the E. Coli and S. Aureus Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes. *J. Phys. Chem. B* **2011**, *115*, 13381−13388.
(38) Mukhopadhyay, P.; Monticelli, L.; Tieleman, D. P. Molecular Dynamics Simulation of a Palmitoyl-Oleoyl Phosphatidylserine Bilayer with Na^+ Counterions and NaCl. *Biophys. J.* **2004**, *86*, 1601−1609.
(39) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089.

(41) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(43) Hénin, J.; Shinoda, W.; Klein, M. L. United-Atom Acyl Chains for CHARMM Phospholipids. *J. Phys. Chem. B* **2008**, *112*, 7008−7015.
(44) Delemotte, L.; Dehez, F.; Treptow, W.; Tarek, M. Modeling Membranes under a Transmembrane Potential. *J. Phys. Chem. B* **2008**, *112*, 5547−5550.
(45) Kramar, P.; Miklavcic, D.; Lebar, A. M. A System for the Determination of Planar Lipid Bilayer Breakdown Voltage and Its Applications. *IEEE Trans. NanoBioscience* **2009**, *8*, 132−138.
(46) Montal, M.; Mueller, P. Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties. *Proc. Natl. Acad. Sci.* **1972**, *69*, 3561−3566.
(47) Kalinowski, S.; Ibron, G.; Bryl, K.; Figaszewski, Z. Chronopotentiometric Studies of Electroporation of Bilayer Lipid Membranes. *Biochim. Biophys. Acta BBA, Biomembr.* **1998**, *1369*, 204−212.
(48) Kalinowski, S.; Figaszewski, Z. A Four-Electrode System for Measurement of Bilayer Lipid Membrane Capacitance. *Meas. Sci. Technol.* **1995**, *6*, 1043.
(49) Kalinowski, S.; Figaszewski, Z. A Four-Electrode Potentiostat-Galvanostat for Studies of Bilayer Lipid Membranes. *Meas. Sci. Technol.* **1995**, *6*, 1050.
(50) Sabotin, I.; Lebar, A. M.; Miklavcic, D.; Kramar, P. Measurement Protocol for Planar Lipid Bilayer Viscoelastic Properties. *IEEE Trans. Dielectr. Electr. Insul.* **2009**, *16*, 1236−1242.
(51) Cramer, T.; Zerbetto, F.; García, R. Molecular Mechanism of Water Bridge Buildup: Field-Induced Formation of Nanoscale Menisci. *Langmuir* **2008**, *24*, 6116−6120.
(52) Neu, J. C.; Krassowska, W. Asymptotic Model of Electroporation. *Phys. Rev. E* **1999**, *59*, 3471−3482.
(53) ElBayoumi, T. A.; Torchilin, V. P. Current Trends in Liposome Research. In *Liposomes*; Weissig, V., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2010; pp 1−27.
(54) Napotnik, T. B.; Reberšek, M.; Kotnik, T.; Lebrasseur, E.; Cabodevila, G.; Miklavčič, D. Electropermeabilization of Endocytotic Vesicles in B16 F1Mouse Melanoma Cells. *Med. Biol. Eng. Comput.* **2010**, *48*, 407−413.

Chapitre 4 - Lésions et réparation de l'ADN

Maintenir l'intégrité du génome des organismes vivants est fondamental pour assurer le bon fonctionnement des cellules. Cependant, l'ADN est constamment exposé à des menaces externes, telles que la lumière UV, les rayonnements ionisants, le stress oxydatif ou l'exposition à des agents chimiques externes qui peuvent générer une variété de lésions complexes de l'ADN, y compris la dimérisation de bases, l'oxydation ou les ruptures de brin. L'accumulation de lésions d'ADN dans les cellules peut in fine être responsable de la mort cellulaire et du développement de cancers ou encore de maladies associées au vieillissement. Les cellules ont développé des systèmes de réparation spécifiques qui reconnaissent la présence d'une lésion dans un grand bassin de nucléobases intactes et ensuite catalysent le processus de réparation. Il existe divers mécanismes de réparation sophistiques, comme la voie de réparation de l'excision nucléotidique (NER), adaptée à l'élimination des lésions complexes et volumineuses. Comparé aux mécanismes de réparation de l'excision de base (BER), le NER qui implique des complexes multiprotéiques est beaucoup moins caractérisé.

I. Recombinaison homologue

La réparation de l'ADN est essentielle pour la survie des cellules et le maintien de l'intégrité du génome.[97] La recombinaison homologue est responsable de la réparation dans les cellules des cassures des double-brins d'ADN.[98] Dans les bactéries, cette voie de réparation peut être effectuée par l'un ou l'autre des systèmes protéiques RecBCD et RecFOR.[99] Un des complexes régulateurs important de la voie RecFOR est le complexe RecOR qui aide au recrutement de l'ADN par la protéine RecA, une protéine qui catalyse l'hybridation d'ADN simple brin avec de l'ADN double brin.[100] En 2007 Timmins et al. ont résolu en l'absence d'ADN, la structure hexamèrique de RecOR de Deinococcus radiodurans.[101] Toutefois, l'architecture compacte du complexe cristallisé ne permettait pas de faire d'hypothèse quant au mode d'association de l'ADN simple brin à la structure. Récemment les mêmes auteurs ont isolé dans des conditions de cristallisation identique une conformation « ouverte » de RecOR (figure 26). Grace à des simulations de dynamique, nous avons pu montrer qu'un simple brin d'ADN pouvait s'ancrer de manière robuste à l'intérieur de cette structure ouverte. Sur la base des simulations nous avons pu décrire le site de liaison de l'ADN et proposer une série des mutants dont la réalisation expérimentale a démontré la robustesse de notre prédiction. En l'absence d'ADN dans la structure nos résultats montrent que le complexe « respire » en passant spontanément de la forme « ouverte » à une forme de la structure « fermée ».[101]

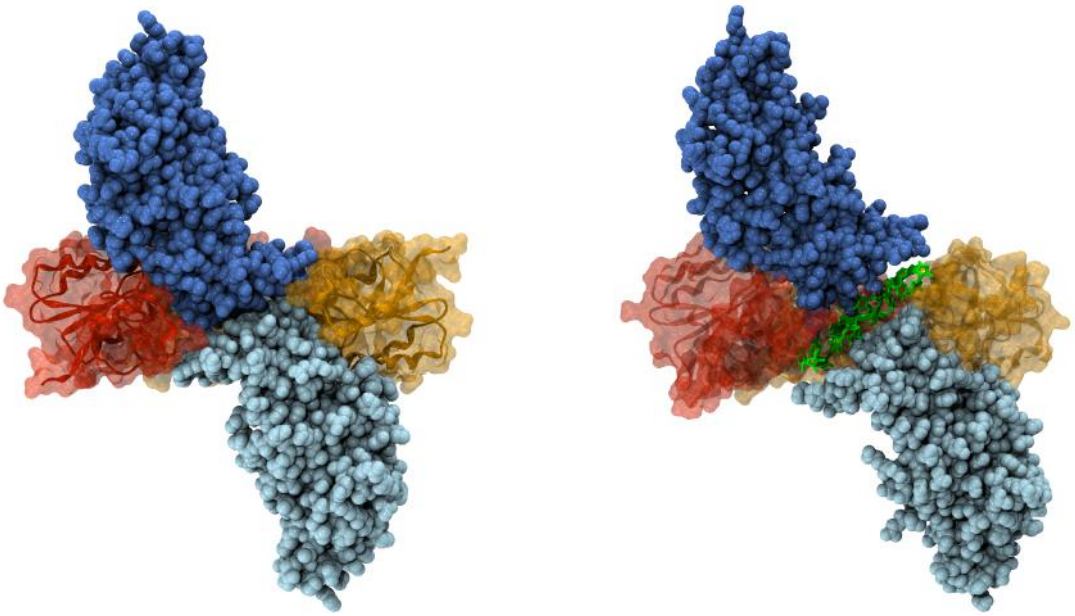


Figure 26. Conformation (gauche) « fermée » et « ouverte » (droite) du complexe RecOR. Contrairement à la forme fermée, la forme ouverte peut aisément accueillir un simple brin d'ADN. Les deux protéines RecO sont représentées en sphères de van der Waals. Seulement deux des protéines constituant l'homotetramère de RecR sont représentées (surface transparente).

II. Lésions complexes de l'ADN

Les lésions cluster de type apurinique/apyrimidinique (AP ; abasiques) de l'ADN produites par des radiations ionisantes sont de loin beaucoup plus cytotoxiques que ne le sont les lésions AP isolées.[102] Des études expérimentales menées in vitro montrent pour des petits oligonucléotides que l'efficacité de réparation par les endonucléases du système BER (Base Excision Repair) varie fortement selon les positions relatives des sites abasiques sur les doubles brins d'ADN. Par ailleurs les protéines bactérienne (Nfo) (qui n'interagit que par l'intermédiaire d'un simple contact avec le site abasique de l'ADN) et humaine (APE1) (qui reconnait le site abasique de façon spécifique par l'intermédiaire d'une pince, et présente une interface d'interaction entendue avec l'ADN composés de résidus basiques) ont également un comportement très différent vis-à-vis des lésions AP (Figure 27).[103,104]

Pour rationaliser les efficacités de réparation des différentes lésions clusters, nous avons modélisé de façon systématique toutes les lésions étudiées expérimentalement. Nous avons réalisé des simulations de dynamique moléculaire non-biaisées, pour l'ensemble des brins doublement lésés, isoles dans un premier temps et en présence de la protéine APE1 dans un second temps. Pour les brins isoles nos résultats montrent que la déformation de l'ADN en réponse aux lésions cluster varie de faible à importante selon la position relative des deux sites abasiques. Pour seq-3, seq-1 et seq0, les simulations mettent en évidence des déformations importantes du double brin qui corrèlent parfaitement avec la faible réparation de ces lésions clusters observées expérimentalement (figure 27).[105]

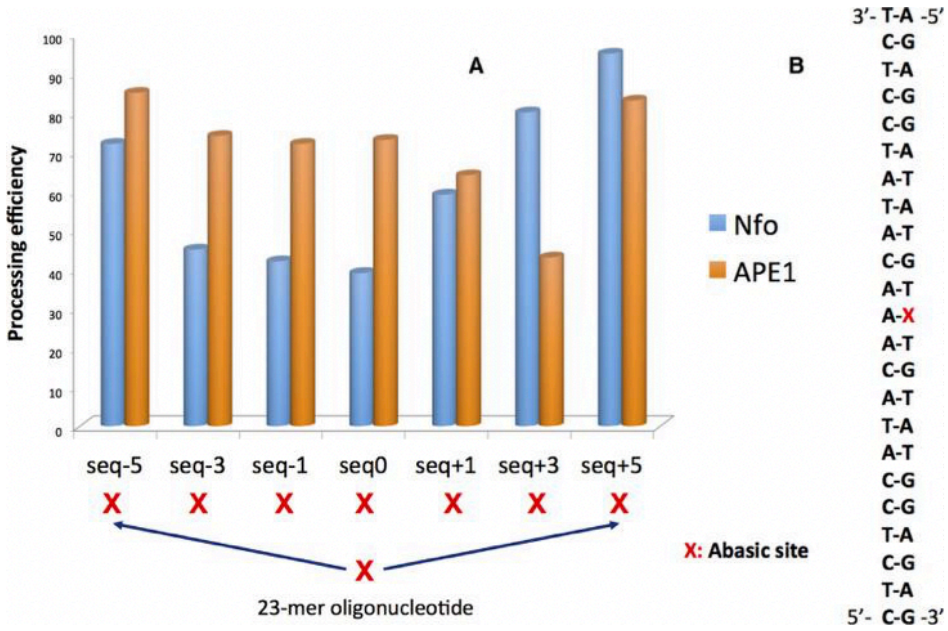


Figure 27. Taux de réparation de lésions cluster composées de deux sites abasiques par les protéines de réparation bactérienne (Nfo) et eukaryote (APE1). Seq-5, seq-3, ... repèrent les différentes positions relatives des sites abasiques.

Pour la protéine humaine, le comportement est sensiblement différent et seule la séquence +3 est mal réparée. Les simulations de dynamique moléculaire des complexes impliquant la protéine APE1 et les différentes lésions clusters nous ont permis de comprendre la différence de comportement avec Nfo. Pour toutes les lésions clusters, sauf seq+3, la protéine APE1 en raison son interface spécifique avec l'ADN est capable de prévenir la déformation du double brin et de le maintenir dans une conformation adéquate pour la réparation. Dans le cas de seq+3, la position de la deuxième lésion induit une déstabilisation de la protéine qui se traduit par une ouverture de la pince reconnaissant le site abasique (figure 28) et une déstabilisation globale du complexe.

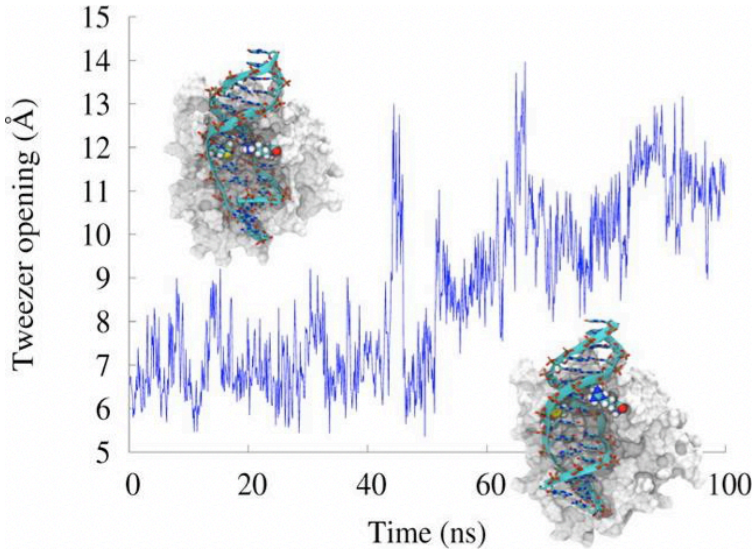


Figure 28. Évolution temporelle de l'ouverture de la pince de la protéine APE1 en réponse à la lésion cluster seq+3, illustrant la déformation partielle de la protéine et la perte d'interactions stabilisantes.

III. Déformations structurales de l'ADN en réponse aux lésions

Au niveau structural, la présence de lésions peut entrainer des modifications de la structure en double brins de l'ADN, modifications dont l'importance dépend fortement du type de lésion.[106] Dans ce contexte nous nous sommes intéresses à deux photolésions (une classe particulièrement délétère de défauts moléculaires) responsables de l'apparition d'une grande majorité de tumeurs malignes de la peau : le dimère de cyclobutane pyrimidine (CPD) résultant de la dimérisation de deux thymines ou cytosines adjacentes et la pyrimidine (6-4) pyrimidone (64-PP).[107] CPD est une lésion bien plus abondante que 64-PP mais elle est dans le même temps beaucoup moins bien réparée par le NER, connu pour prendre en charge les lésions très déformantes. 64-PP dont la réparation est efficace est en revanche source d'erreurs de réplications. Les deux lésions entant toutes les deux très génotoxiques.

Nous avons réalisé de longues simulations de dynamique moléculaire (~2 µs) en utilisant un protocole d'échange de répliques (*parallel tempering*)[108] pour échantillonner la structure d'un brin de 16 paires de base d'ADN en présence des deux lésions CPD et 64-PP. En particulier, nous avons mis en évidence une déformation structurelle très limitée induite par CPD, alors que 64- PP est caractérisée par un polymorphisme structurel complexe (Figure 29).[109] Nos simulations permettent d'unifier les résultats structuraux expérimentaux ambigus obtenus par résonance magnétique nucléaire ou méthode de transfert d'énergie par résonance de Förster[110] pour les deux types de lésions et de rationaliser la résistance à la réparation de CPD d'un côté et les erreurs de réplication liée à 64-PP.

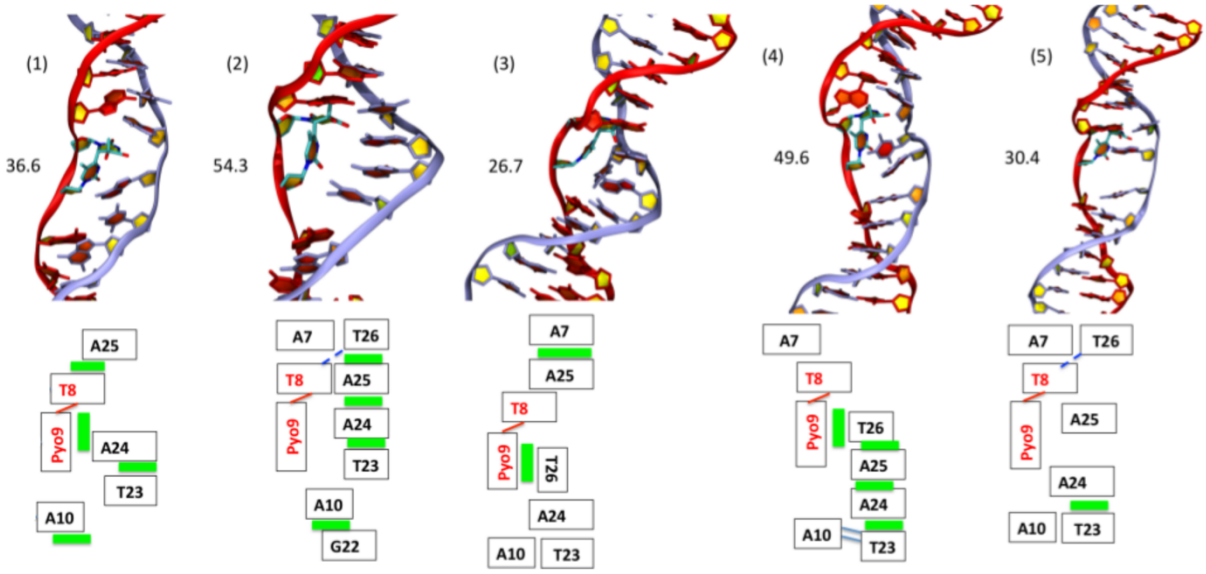


Figure 29. Représentation de cinq conformations représentatives observées le long de la trajectoire de 2 µs d'un brin d'ADN en présence de la lésion 64-PP. La représentation schématique des d'interactions est donnée dans les figures du bas pour chaque conformation significative du polymorphisme conformationnel engendré par cette lésion (les cases vertes représentent les interactions π-stacking et les lignes bleues les liaisons hydrogènes). La valeur de l'angle de *bending* pour chaque conformère est indiquée entre parenthèses.

Ci-après sont reproduits 2 articles illustrant mes travaux dédiés aux effets des champs électrique sur les membranes :

- Gattuso, H., Durand, E., Bignon, E., Morell, C., Georgakilas, A. G., Dumont, E., Chipot, C., Dehez*, F. & Monari*, A. Repair Rate of Clustered Abasic DNA Lesions by Human Endonuclease: Molecular Bases of Sequence Specificity. *J. Phys. Chem. Lett.* **7**, 3760–3765 (2016).
- Dehez, F.*, Gattuso, H., Bignon, E., Morell, C., Dumont, E.* & Monari, A.* Conformational polymorphism or structural invariance in DNA photoinduced lesions: implications for repair rates. *Nucleic Acids Res.* **45**, 3654–3662 (2017).

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

Repair Rate of Clustered Abasic DNA Lesions by Human Endonuclease: Molecular Bases of Sequence Specificity

Hugo Gattuso,[†,‡] Elodie Durand,[†,‡] Emmanuelle Bignon,[§,∥] Christophe Morell,[∥] Alexandros G. Georgakilas,[⊥] Elise Dumont,[*,§] Christophe Chipot,[†,‡,▽,⊗] François Dehez,[*,†,‡,⊗] and Antonio Monari[*,†,‡]

[†]Université de Lorraine—Nancy, Theory-Modeling-Simulation SRSMC, 54000 Vandoeuvre-lès-Nancy, France

[‡]CNRS, Theory-Modeling-Simulation SRSMC, 54000 Vandoeuvre-lès-Nancy, France

[§]Univ Lyon, Ens de Lyon, CNRS UMR 5182, Université Claude Bernard Lyon 1, Laboratoire de Chimie, F-69342 Lyon, France

[∥]Université de Lyon, Institut des Sciences Analytiques UMR 5280, CNRS, Université de Lyon 1, ENS Lyon 5 rue de la Doua, F-69100 Villeurbanne, France

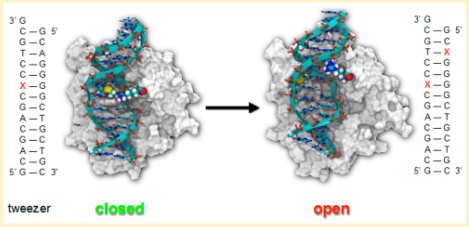
[⊥]DNA damage laboratory, Physics Department, School of Applied Mathematical and Physical Sciences, National Technical University of Athens (NTUA), Zografou, Athens 15780, Greece

[▽]Department of Physics, University of Illinois at Urbana—Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States

[⊗]Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana−Champaign

Supporting Information

ABSTRACT: In the present contribution, the interaction between damaged DNA and repair enzymes is examined by means of molecular dynamics simulations. More specifically, we consider clustered abasic DNA lesions processed by the primary human apurinic/apyrimidinic (AP) endonuclease, APE1. Our results show that, in stark contrast with the corresponding bacterial endonucleases, human APE1 imposes strong geometrical constraints on the DNA duplex. As a consequence, the level of recognition and, hence, the repair rate is higher. Important features that guide the DNA/protein interactions are the presence of an extended positively charged region and of a molecular tweezers that strongly constrains DNA. Our results are on very good agreement with the experimentally determined repair rate of clustered abasic lesions. The lack of repair for one particular arrangement of the two abasic sites is also explained considering the peculiar destabilizing interaction between the recognition region and the second lesion, resulting in a partial opening of the molecular tweezers and, thus, a less stable complex. This contribution cogently establishes the molecular bases for the recognition and repair of clustered DNA lesions by means of human endonucleases.



Some of the most common, and nonetheless toxic, DNA lesions are the abasic or apurinic/apyrimidic sites (AP). AP sites are produced upon cleavage of the DNA glycosidic bond, thus leaving only the uncapped deoxyribose moiety in the DNA strand. Noncoding abasic (AP) sites can form in the DNA either spontaneously or as a result to the exposure to reactive oxygen/nitrogen species (ROS/RNS), ionizing radiations or chemotherapeutic drugs like bleomycin.[1] Although the residual AP moiety is in equilibrium between a cyclic sugar and an open form (see Figure S1 in Supporting Information (SI)), the cyclic conformer is absolutely predominant, accounting for about 99% of the occurrence[2,3] and hence it is the only one considered in the following simulations. Molecular modeling[4] and NMR structural determination[5] agree in underlying a strong local deformation of the DNA helix as a response to the presence of the AP site. In particular, one can pinpoint a marked preference of the AP site to adopt an extrahelical position that correlates with the presence of a bulge in the DNA backbone. Interestingly, the latter structural deformation can be seen as extremely relevant for lesion repair.[6−8]

AP intermediates can be the most common toxic or mutagenic DNA lesions in vivo. The human APE1[9,10] and bacterial Nfo[11] repair enzymes epitomize the two well-conserved 5′-P endonuclease families but are structurally nonhomologous. Moreover, they do not present similar interaction patterns with DNA and do not require the same metal ion.[12] The processing of AP lesions[13] involves recognition by the enzyme followed by a cleavage of the

Received: July 30, 2016
Accepted: September 9, 2016
Published: September 9, 2016

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

phosphodiester bond mediated by a Mg^{2+} ion in the case of APE1, Zn^{2+} for Nfo and a water molecule. In the case of human APE1[14] the active site is embedded in the enzyme and exhibits an extended contact region between the positively charged residues and the negatively charged DNA backbone. On the other hand, in the case of bacterial *Escherichia coli* endonuclease IV (Nfo), the contact interface with DNA is much smaller, and resumes essentially to the few residues composing the AP recognition site (Figure 1). As a consequence, the interaction with APE1 induces much larger structural deformations to the damaged DNA fragment, whereas its native structure is globally preserved by Nfo.

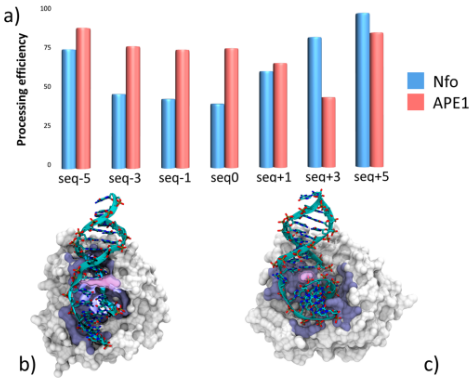


Figure 1. (a) Repair efficiency (cleavage) of different clustered AP lesions in different DNA sequences by human and bacterial endonucleases based on previously published data.[7,8] (b) Representation of the crystal structure of bacterial Nfo interacting with a DNA sequence (PDB code 4K1G).[11] (c) Representation of the crystal structure of human APE1 interacting with a DNA sequence (PDB code 1DEW).[10] Recognition area is highlighted in mauve and the contact area close to the AP site are further highlighted, notice the presence of the tweezers in APE1 holding the lesioned DNA, whereas for Nfo, only a small finger interacting with the lesioned DNA is visible. See Figure S2 in SI for a detailed definition of the oligomers' sequence nomenclature.

The chemical mechanism of recognition and repair of AP sites is interesting per se, also due to the high occurrence of this lesion (~5000−10000 lesions per cell daily).[15] Moreover, it is nowadays established that exposure to ionizing radiations (γ- and X-rays), like in the case of radiotherapy, may lead to the accumulation of two or more AP sites, as well as oxidized bases, double-strand breaks (DSBs) and non-DSB lesions within 2 or 3 DNA helical turns.[16,17] Two closely spaced DNA lesions constitute indeed the minimal size cluster damage.[18] Clustering of DNA lesions may become extremely threatening for the cells[19,7,20,21] because of their well-established repair resistance and the high possibility of non-DSB clustered lesions to be readily converted into DSBs during base excision repair (BER).[22]

Furthermore, a strong interplay between DSBs and cluster sites may also enhance their biological harmfulness; indeed, the processing of DSB can be strongly limited, or totally inhibited, by the presence of neighboring AP sites.[23]

Numerical simulation techniques have steadily become an invaluable tool to characterize with an atomistic and molecular resolution the outcome of damages on the structural and dynamical properties of different DNA sequences as well as their interaction with repair enzymes.[24−29] Molecular modeling has been applied to damaged oligonucleotides, such as abasic sites,[4,30,31] bulky adducts,[32,33] cross-links,[34] or photolesions,[20,35] and to duplexes in interaction with repair enzymes.[36] MD explorations of clustered 8-oxoguanine and abasic sites have been recently published.[37,38] We recently reported[39] the molecular modeling and simulation of clustered AP sites that may be induced in DNA duplexes by low-dose radiation.[40] The latter class of lesions have been produced in plasmid or phage DNA,[7,21,41] and isolated DNA fibers.[42,43] By exploiting the DNA repair rate measurements[44,45] clustered AP sites have been confirmed up to nucleosomal DNA[46] and human cells.[8]

The repair rate of clustered AP sites has been studied[7,8] in the case of both human APE1 and bacterial (*Escherichia coli*) Nfo (Figure 1, see also SI for the definition of the sequences and the nomenclature). Nfo features a global lower repair rate and strong dependence on the AP position as similarly suggested by independent works.[47] We have correlated the effect of the relative position of the two AP sites with the structural deformation of the DNA structure induced by the lesions, remarkably our finding also provide a rationale for the observed processing rate of bacterial Nfo. The low processing rate for seq0 and especially seq-3 (see SI and Figure 1) is characterized by a strong local deformation with shrinking of the DNA duplex for seq0 and a complicated stabilization pattern involving the ejection of nonlesioned purines for seq-3.[39] This inference is also justified by the observation that the bacterial enzyme contact with DNA is quite local (Figure 1) and, as a consequence, is not able to induce strong structural deformation and mechanical constraints on the DNA duplex.

The behavior of APE1 enzyme is singularly different. First, the processing rate is globally higher compared to that of the bacterial enzyme. Second, the ease of repair in humans, and, in general in all eukaryotic organisms, is consistent with a different evolutionary pathway and with the necessity to preserve genome integrity in complex organisms. The sequence dependence is also far less pronounced, all the lesions being repaired at a rate going from 70 to 80%; the only exception is the seq+3 lesion for which a mediocre repair rate of about 45% is observed.[8] Interestingly, the latter was also one of the best-repaired lesions by the bacterial Nfo. One, therefore, can immediately surmise very different recognition patterns between the two organisms and, as a consequence, the correlation between DNA structural deformation and repair ought to be revised. Toward this end, we explore in present work the behavior of clustered AP sites containing a DNA double strand in solution and in interaction with the human APE1 enzyme. The structural evolution of the DNA-enzyme complex has been simulated using all-atom molecular dynamics (MD) with the parm99bsc1 force field.[48,29,49] The simulations have been carried out in cubic water boxes with the NAMD package.[50] Production runs of 100 ns have been carried out in the NPT ensemble (see SI for a full description of the methodology). As will be illustrated in the following 100 ns appear sufficient to explore the conformational space of the DNA/APE1 complex in particular due to the strong interactions and high rigidity of the aggregates. Cluster AP lesions have been produced by specific mutations of the crystallographic DNA strands.[10]

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

A molecular tweezers, capped at its end respectively by one arginine and one methionine residues (see also SI), stabilizes further the DNA−protein complex. In the case of the isolated enzyme (Figure 2a), this molecular tweezers is open and the

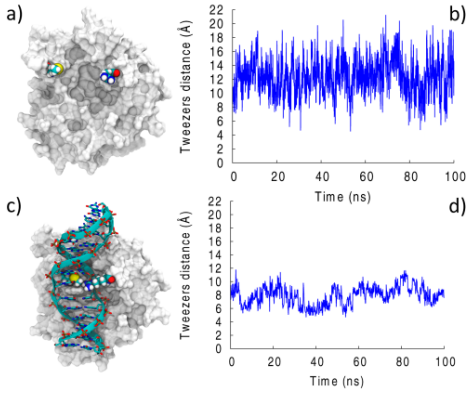


Figure 2. Representative snapshot and time evolution of the molecular tweezers width for the free APE1 (a,b) and for the complex between APE1 and the single AP containing DNA strand (c,d).

distance between the terminal amino acids fluctuates dramatically around 12 Å, hence leaving an easy access for the DNA strand to reach the active site. In stark contrast, when DNA is present, the molecular tweezers closes around the oligomer, the distance between the terminal edges drops down to 7 Å and the amplitude of the oscillations strongly diminishes. The closure of the tweezers upon complexation with DNA is once again driven by the interaction of positive arginine with the negative backbone; upon equilibrium, the tweezers is kept closed also thanks to the interaction between the electron rich sulfur of methionine and the arginine forming the tweezers' tips.

Interestingly, although the global DNA/APE1 structure is quite compact, we have observed three channels (see SI, Figure S4) allowing ingress of cations that may easily reach the active site, where the presence of Mg^{2+} is required for the cleavage reaction to occur. Note, however, that the simulations were only performed in the presence of Na^{+}; however, the fast and efficient ion entrance should not influence DNA recognition because it follows complexation with the DNA oligomers.

The behavior of APE1 enzyme free and in interaction with the single AP-containing DNA strand is depicted in Figure 2. The DNA/APE1 contact region is quite extended (see Figure 1) and is constituted chiefly of basic arginine and lysine residues. The positively charged amino acids provide an extensive electrostatic stabilization of the negatively charged DNA backbone. As a consequence the root-mean-square deviation (RMSD) of the DNA strand with respect to the crystallographic structure remains low throughout the trajectory. The stabilization of the DNA structure is mainly due to the electrostatic interactions between the negatively charged DNA backbone and the positively charged amino acids constituting the recognition region. Furthermore, DNA structure and position are also locked by the closure of the molecular tweezers that hence minimize structural deformations.

When introducing a second lesion to form clustered AP lesions, we have observed in all the models simulated in this study the formation of persistent complexes (representative conformations are reported in SI and Figure 3) spanning all

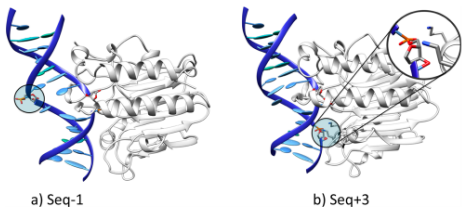


Figure 3. Representative snapshots for the seq-1 (a) and seq+3 (b) DNA/APE1 complex. DNA and protein are represented in carton. AP sites are evidenced in licorice representation. The interaction between the secondary AP site of seq+3 and the recognition region of APE1 is evidenced in the zoom of panel b.

100 ns. This is congruent with the relative high repair rate measured for human APE1. As already observed in all cases studied here, the APE1 enzyme induces strong geometrical constraints on the DNA oligomer, and indeed, we may see that its RMSD with respect to the reference crystalline form remains low and close to that found for the isolated AP-containing strand. In particular, the RMSD is much lower than that for the solvated strands, which are known to exhibit a strong structural deformation (Figure 4). As already evidenced in the case of isolated nucleotides,[39] the deviation in the RMSD is not only due to the AP and it involves the complementary base and the nearby residues, up to at least five base pairs distance. The case of seq0, that is, the sequence exhibiting the larger structural deformation in solution, is particularly striking. This fact may also explains why seq0 is highly repaired by APE1 thanks to the imposed geometrical constrains, whereas its larger structural deformation probably makes its recognition by Nfo particularly difficult.

The only exception to this general trend is observed for seq +3 for which the RMSD of the DNA/APE1 complex almost overlaps with that of the DNA strand in solution, hence showing a much larger deviation from the ideal crystal structure. Interestingly enough, seq+3 is the only sequence for which the processing rate is lower than 50% and for which repair resistance can be correlated to a worse recognition by APE1, resulting in a less stabilized complex. The molecular bases for poor recognition can be related to the fact that seq+3 is the only sequence in which the second AP site is in direct interaction with the APE1 recognition region (Figure 3b), whereas in all the other cases, the secondary lesions are exposed toward the bulk. Indeed, the AP site having a strong propensity to adopt extrahelical position, it strongly perturbs the DNA/APE1 complex when interacting with seq+3. Another evidence of this perturbation can be seen in the behavior of the molecular tweezers (Figure 5), as mirrored by the representative conformations and time evolution of the edge distance for seq+3. One may clearly observe a partial opening of the molecular tweezers in the course of the MD trajectory, which may once again safely conclude in favor of a less stable arrangement.

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

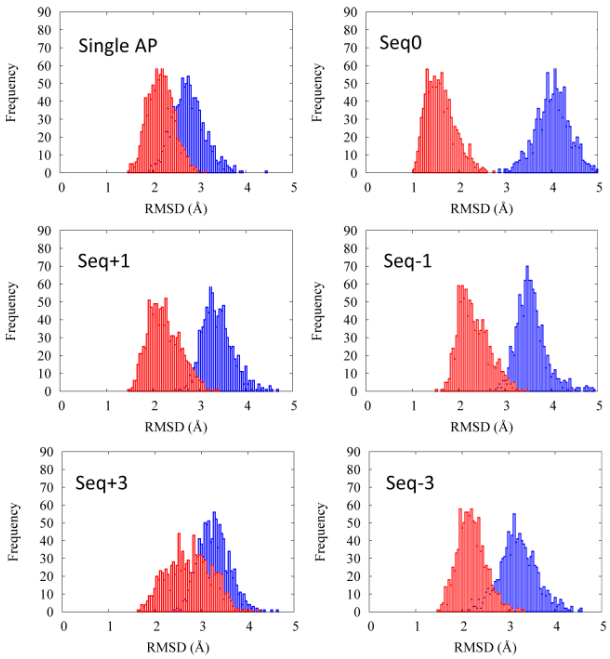


Figure 4. Distribution of the DNA root-mean-square deviation (RMSD) with respect to the crystal structure for the single containing and clustered AP lesions, inside the APE1 complex (red) and in solution (blue). The corresponding repair rates are given in Figure 1, whereas time series are reported in SI.

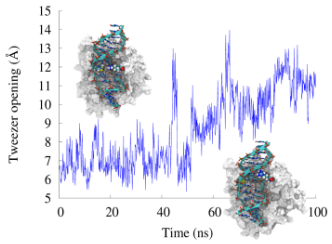


Figure 5. Representative conformations and time evolution of the molecular tweezers width for seq+3 showing the partial opening and the loss of stabilizing interactions.

The high efficiency of APE1 in processing clustered DNA lesions compared to bacterial endonucleases is clearly related to its capacity to induce strong mechanical constraints on the damaged DNA strands, and hence, to compensate for the large structural deviation engendered by these complex lesions. In particular, the presence of an extended positively charged region interacting with the DNA backbone and a tweezers closing around the DNA strand is crucial to ensure an adequate recognition and repair rate. Certain sequences (seq+3) exhibit a dramatically lower repair rate due to the perturbation induced by the secondary lesions on the recognition region, and in particular, on the stability of the closed molecular tweezers. The processing of AP clustered DNA is of high biological importance also because all DNA base lesions will be eventually converted to AP sites during reparation. The accumulation of complex AP DNA sites with increased processing efficiency by APE1 may in one hand reduce toxic AP lesions but on the other hand can lead to increased possibility of de novo DSB formation. For example, when two neighboring AP sites (<15−20 base-pair distance) are processed simultaneously by AP endonuclease, the occurrence of mutations and genomic instability may be increased through the generation of additional DSBs. Molecular modeling and simulations are complementary to the experimental studies to ultimately achieve the complete rationalization of the in vivo repair mechanisms of clustered DNA lesions, in particular focusing on the molecular basis of repair selectivity. Hence, the present investigation also paves the way not only to improve our understanding of DNA repair mechanisms but also to potentially design selective novel APE1 inhibitors, whose present efficiency is controversial,[51−53] ultimately being able to enhance radiotherapy efficiency.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.6b01692.

Description of the molecular dynamic simulation computational details and parameters for the AP site.

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

Chemical formulas of the AP sites, definition of the clustered DNA sequences nomenclature, definition of the tweezers and the tweezers' distance, representative snapshots of all the stable DNA/protein complexes, cations entrance paths, time series of the RMSD evolution for solvated and APE1 complexed oligomers. (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: antonio.monari@univ-lorraine.fr.
*E-mail: elise.dumont@ens-lyon.fr.
*E-mail: francois.dehez@univ-lorraine.fr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Calculations were performed using the local HPC resources of PSMN at ENS-Lyon and at Université de Lorraine—Nancy. This work was supported by the COST Action CM1201 "Biomimetic Radical Chemistry" is gratefully acknowledged. This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX0007) operated by the French National Research Agency (ANR). Dr. Georgakilas was supported by an EU grant MC−CIG-303514.

REFERENCES

(1) Fung, H.; Demple, B. Distinct Roles of Ape1 Protein in the Repair of DNA Damage Induced by Ionizing Radiation or Bleomycin. *J. Biol. Chem.* **2011**, *286* (7), 4968−4977.

(2) Gates, K. S. An Overview of Chemical Processes That Damage Cellular DNA: Spontaneous Hydrolysis, Alkylation, and Reactions with Radicals. *Chem. Res. Toxicol.* **2009**, *22* (11), 1747−1760.

(3) Clauson, C.; Schärer, O. D.; Niedernhofer, L. Advances in Understanding the Complex Mechanisms of DNA Interstrand Cross-Link Repair. *Cold Spring Harbor Perspect. Biol.* **2013**, *5* (10), a012732.

(4) Ayadi, L.; Coulombeau, C.; Lavery, R.; Arnott, S.; Chandrasekaran, R.; Birdsall, D. L.; Leslie, A. G. W.; Ratliff, R. L.; Ayadi, L.; Coulombeau, C.; et al. Abasic Sites in Duplex DNA: Molecular Modeling of Sequence-Dependent Effects on Conformation. *Biophys. J.* **1999**, *77* (6), 3218−3226.

(5) Hazel, R. D.; Tian, K.; de Los Santos, C. NMR Solution Structures of Bistranded Abasic Site Lesions in DNA. *Biochemistry* **2008**, *47* (46), 11909−11919.

(6) Wilson, D. M.; Barsky, D. The Major Human Abasic Endonuclease: Formation, Consequences and Repair of Abasic Lesions in DNA. *Mutat. Res., DNA Repair* **2001**, *485* (4), 283−307.

(7) Georgakilas, A. G.; Bennett, P. V.; Sutherland, B. M. High Efficiency Detection of Bi-Stranded Abasic Clusters in Gamma-Irradiated DNA by Putrescine. *Nucleic Acids Res.* **2002**, *30* (13), 2800−2808.

(8) Georgakilas, A. G.; Bennett, P. V.; Wilson, D. M.; Sutherland, B. M. Processing of Bistranded Abasic DNA Clusters in γ-Irradiated Human Hematopoietic Cells. *Nucleic Acids Res.* **2004**, *32* (18), 5609−5620.

(9) Freudenthal, B. D.; Beard, W. A.; Cuneo, M. J.; Dyrkheeva, N. S.; Wilson, S. H. Capturing Snapshots of APE1 Processing DNA Damage. *Nat. Struct. Mol. Biol.* **2015**, *22* (11), 924−931.

(10) Mol, C. D.; Izumi, T.; Mitra, S.; Tainer, J. A. DNA-Bound Structures and Mutants Reveal Abasic DNA Binding by APE1 and DNA Repair Coordination [Corrected]. *Nature* **2000**, *403* (6768), 451−456.

(11) Mazouzi, A.; Vigouroux, A.; Aikeshev, B.; Brooks, P. J.; Saparbaev, M. K.; Morera, S.; Ishchenko, A. A. Insight into Mechanisms of 3′-5′ Exonuclease Activity and Removal of Bulky 8,5′-cyclopurine Adducts by Apurinic/apyrimidinic Endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (33), E3071−E3080.

(12) Tsutakawa, S. E.; Shin, D. S.; Mol, C. D.; Izumi, T.; Arvai, A. S.; Mantha, A. K.; Szczesny, B.; Ivanov, I. N.; Hosfield, D. J.; Maiti, B.; et al. Conserved Structural Chemistry for Incision Activity in Structurally Non-Homologous Apurinic/apyrimidinic Endonuclease APE1 and Endonuclease IV DNA Repair Enzymes. *J. Biol. Chem.* **2013**, *288* (12), 8445−8455.

(13) Molina, R.; Stella, S.; Redondo, P.; Gomez, H.; Marcaida, M. J.; Orozco, M.; Prieto, J.; Montoya, G. Visualizing Phosphodiester-Bond Hydrolysis by an Endonuclease. *Nat. Struct. Mol. Biol.* **2014**, *22* (1), 65−72.

(14) Demple, B.; Sung, J.-S. Molecular and Biological Roles of Ape1 Protein in Mammalian Base Excision Repair. *DNA Repair* **2005**, *4* (12), 1442−1449.

(15) Kryston, T. B.; Georgiev, A. B.; Pissis, P.; Georgakilas, A. G. Role of Oxidative Stress and DNA Damage in Human Carcinogenesis. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **2011**, *711* (1), 193−201.

(16) Nikjoo, H.; O'Neill, P.; Terrissol, M.; Goodhead, D. T. Quantitative Modelling of DNA Damage Using Monte Carlo Track Structure Method. *Radiat. Environ. Biophys.* **1999**, *38* (1), 31−38.

(17) Watanabe, R.; Rahmanian, S.; Nikjoo, H. Spectrum of Radiation-Induced Clustered Non-DSB Damage - A Monte Carlo Track Structure Modeling and Calculations. *Radiat. Res.* **2015**, *183* (5), 525−540.

(18) Shikazono, N.; Noguchi, M.; Fujii, K.; Urushibara, A.; Yokoya, A. The Yield, Processing, and Biological Consequences of Clustered DNA Damage Induced by Ionizing Radiation. *J. Radiat. Res.* **2009**, *50* (1), 27−36.

(19) Nikitaki, Z.; Hellweg, C. E.; Georgakilas, A. G.; Ravanat, J.-L. Stress-Induced DNA Damage Biomarkers: Applications and Limitations. *Front. Chem.* **2015**, *3*, 35.

(20) Dumont, E.; Monari, A. Understanding DNA under Oxidative Stress and Sensitization: The Role of Molecular Modeling. *Front. Chem.* **2015**, *3*, 43.

(21) Singh, V.; Das, P. Condensation of DNA–a Putative Obstruction for Repair Process in Abasic Clustered DNA Damage. *DNA Repair* **2013**, *12* (6), 450−457.

(22) Georgakilas, A. G.; O'Neill, P.; Stewart, R. D. Induction and Repair of Clustered DNA Lesions: What Do We Know so Far? *Radiat. Res.* **2013**, *180* (1), 100−109.

(23) Lomax, M. E.; Folkes, L. K.; O'Neill, P. Biological Consequences of Radiation-Induced DNA Damage: Relevance to Radiotherapy. *Clin. Oncol. (R. Coll. Radiol).* **2013**, *25* (10), 578−585.

(24) Qi, Y.; Spong, M. C.; Nam, K.; Banerjee, A.; Jiralerspong, S.; Karplus, M.; Verdine, G. L. Encounter and Extrusion of an Intrahelical Lesion by a DNA Repair Enzyme. *Nature* **2009**, *462* (7274), 762−766.

(25) Ivanov, I.; Tainer, J. A.; McCammon, J. A. Unraveling the Three-Metal-Ion Catalytic Mechanism of the DNA Repair Enzyme Endonuclease IV. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (5), 1465−1470.

(26) Maiti, A.; Noon, M. S.; MacKerell, A. D.; Pozharski, E.; Drohat, A. C. Lesion Processing by a Repair Enzyme Is Severely Curtailed by Residues Needed to Prevent Aberrant Activity on Undamaged DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (21), 8091−8096.

(27) Cai, Y.; Geacintov, N. E.; Broyde, S. Nucleotide Excision Repair Efficiencies of Bulky Carcinogen-DNA Adducts Are Governed by a Balance between Stabilizing and Destabilizing Interactions. *Biochemistry* **2012**, *51* (7), 1486−1499.

(28) Radzimanowski, J.; Dehez, F.; Round, A.; Bidon-Chanal, A.; McSweeney, S.; Timmins, J. An "Open" Structure of the RecOR Complex Supports ssDNA Binding within the Core of the Complex. *Nucleic Acids Res.* **2013**, *41* (16), 7972−7986.

(29) Perez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. *Biophys. J.* **2007**, *92* (11), 3817−3829.

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

(30) Ayadi, L.; Coulombeau, C.; Lavery, R. The Impact of Abasic Sites on DNA Flexibility. *J. Biomol. Struct. Dyn.* **2000**, *17* (4), 645−653.
(31) Barsky, D.; Foloppe, N.; Ahmadia, S.; Wilson, D. M., III; MacKerell, A. D. MacKerell, a D. New Insights into the Structure of Abasic DNA from Molecular Dynamics Simulations. *Nucleic Acids Res.* **2000**, *28* (13), 2613−2626.
(32) Liu, Z.; Ding, S.; Kropachev, K.; Lei, J.; Amin, S.; Broyde, S.; Geacintov, N. E. Resistance to Nucleotide Excision Repair of Bulky Guanine Adducts Opposite Abasic Sites in DNA Duplexes and Relationships between Structure and Function. *PLoS One* **2015**, *10* (9), e0137124.
(33) Mu, H.; Kropachev, K.; Wang, L.; Zhang, L.; Kolbanovskiy, A.; Kolbanovskiy, M.; Geacintov, N. E.; Broyde, S. Nucleotide Excision Repair of 2-Acetylaminofluorene- and 2-Aminofluorene-(C8)-Guanine Adducts: Molecular Dynamics Simulations Elucidate How Lesion Structure and Base Sequence Context Impact Repair Efficiencies. *Nucleic Acids Res.* **2012**, *40* (19), 9675−9690.
(34) Churchill, C. D. M.; Eriksson, L. A.; Wetmore, S. D. DNA Distortion Caused by Uracil-Containing Intrastrand Cross-Links. *J. Phys. Chem. B* **2016**, *120* (7), 1195−1204.
(35) Bignon, E.; Gattuso, H.; Morell, C.; Dumont, E.; Monari, A. DNA Photosensitization by an Insider: Photophysics and Triplet Energy Transfer of 5-Methyl-2-Pyrimidone Deoxyribonucleoside. *Chem. - Eur. J.* **2015**, *21* (32), 11509−11516.
(36) Faraji, S.; Dreuw, A. Physicochemical Mechanism of Light-Driven DNA Repair by (6−4) Photolyases. *Annu. Rev. Phys. Chem.* **2014**, *65*, 275−292.
(37) Fujimoto, H.; Pinak, M.; Nemoto, T.; O'Neill, P.; Kume, E.; Saito, K.; Maekawa, H. Molecular Dynamics Simulation of Clustered DNA Damage Sites Containing 8-Oxoguanine and Abasic Site. *J. Comput. Chem.* **2005**, *26* (8), 788−798.
(38) Zálešák, J.; Lourdin, M.; Krejčí, L.; Constant, J.-F.; Jourdan, M. Structure and Dynamics of DNA Duplexes Containing a Cluster of Mutagenic 8-Oxoguanine and Abasic Site Lesions. *J. Mol. Biol.* **2014**, *426* (7), 1524−1538.
(39) Bignon, E.; Gattuso, H.; Morell, C.; Dehez, F.; Georgakilas, A. G.; Monari, A.; Dumont, E. Correlation of Bistranded Clustered Abasic DNA Lesion Processing with Structural and Dynamic DNA Helix Distortion. *Nucleic Acids Res.* **2016**, gkw773.
(40) Sutherland, B. M.; Bennett, P. V.; Sidorkina, O.; Laval, J. Clustered DNA Damages Induced in Isolated DNA and in Human Cells by Low Doses of Ionizing Radiation. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (1), 103−108.
(41) Akamatsu, K.; Shikazono, N.; Saito, T. Localization Estimation of Ionizing Radiation-Induced Abasic Sites in DNA in the Solid State Using Fluorescence Resonance Energy Transfer. *Radiat. Res.* **2015**, *183* (1), 105−113.
(42) Chastain, P. D.; Nakamura, J.; Rao, S.; Chu, H.; Ibrahim, J. G.; Swenberg, J. A.; Kaufman, D. G. Abasic Sites Preferentially Form at Regions Undergoing DNA Replication. *FASEB J.* **2010**, *24* (10), 3674−3680.
(43) Chastain, P. D.; Nakamura, J.; Swenberg, J.; Kaufman, D. Nonrandom AP Site Distribution in Highly Proliferative Cells. *FASEB J.* **2006**, *20* (14), 2612−2614.
(44) Lomax, M. E.; Cunniffe, S.; O'Neill, P. Efficiency of Repair of an Abasic Site within DNA Clustered Damage Sites by Mammalian Cell Nuclear Extracts †. *Biochemistry* **2004**, *43* (34), 11017−11026.
(45) Singh, V.; Kumari, B.; Das, P. Repair Efficiency of Clustered Abasic Sites by APE1 in Nucleosome Core Particles Is Sequence and Position Dependent. *RSC Adv.* **2015**, *5*, 23691−23698.
(46) Eccles, L. J.; Menoni, H.; Angelov, D.; Lomax, M. E.; O'Neill, P. Efficient Cleavage of Single and Clustered AP Site Lesions within Mono-Nucleosome Templates by CHO-K1 Nuclear Extract Contrasts with Retardation of Incision by Purified APE1. *DNA Repair* **2015**, *35*, 27−36.
(47) David-Cordonnier, M.-H.; Cunniffe, S. M. T.; Hickson, I. D.; O'Neill, P. Efficiency of Incision of an AP Site within Clustered DNA Damage by the Major Human AP Endonuclease. *Biochemistry* **2002**, *41* (2), 634−642.
(48) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(49) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; et al. Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* **2015**, *13* (1), 55−58.
(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781−1802.
(51) Al-Safi, R. I.; Odde, S.; Shabaik, Y.; Neamati, N. Small-Molecule Inhibitors of APE1 DNA Repair Function: An Overview. *Curr. Mol. Pharmacol.* **2012**, *5* (1), 14−35.
(52) Luo, M.; Kelley, M. R. Inhibition of the Human Apurinic/apyrimidinic Endonuclease (APE1) Repair Activity and Sensitization of Breast Cancer Cells to DNA Alkylating Agents with Lucanthone. *Anticancer Res.* **2004**, *24* (4), 2127−2134.
(53) Poletto, M.; Malfatti, M. C.; Dorjsuren, D.; Scognamiglio, P. L.; Marasco, D.; Vascotto, C.; Jadhav, A.; Maloney, D. J.; Wilson, D. M.; Simeonov, A.; et al. Inhibitors of the Apurinic/apyrimidinic Endonuclease 1 (APE1)/nucleophosmin (NPM1) Interaction That Display Anti-Tumor Properties. *Mol. Carcinog.* **2016**, *55* (5), 688−704.

DOI: 10.1021/acs.jpclett.6b01692
J. Phys. Chem. Lett. 2016, 7, 3760−3765

3654–3662 *Nucleic Acids Research, 2017, Vol. 45, No. 7* *Published online 28 February 2017*
doi: 10.1093/nar/gkx148

Conformational polymorphism or structural invariance in DNA photoinduced lesions: implications for repair rates

François Dehez[1,2,3,*], Hugo Gattuso[1,2], Emmanuelle Bignon[4,5], Christophe Morell[4], Elise Dumont[5,*] and Antonio Monari[1,2,*]

[1]CNRS, Theory-Modeling-Simulation, SRSMC F-54506 Vandoeuvre-lès-Nancy, France, [2]Université de Lorraine, Theory-Modeling-Simulation, SRSMC F-54506 Vandoeuvre-lès-Nancy, France, [3]Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana Champaign, [4]Institut des Sciences Analytiques, UMR 5280, Université de Lyon1 (UCBL) CNRS, ENS Lyon, Lyon, France and [5]Université de Lyon, ENS de Lyon, CNRS, Université Lyon 1, Laboratoire de Chimie UMR 5182, F69342 Lyon, France

Received December 20, 2016; Revised February 06, 2017; Editorial Decision February 20, 2017; Accepted February 21, 2017

ABSTRACT

DNA photolesions constitute a particularly deleterious class of molecular defects responsible for the insurgence of a vast majority of skin malignant tumors. Dimerization of two adjacent thymines or cytosines mostly gives rise to cyclobutane pyrimidine dimers (CPD) and pyrimidine(6-4)pyrimidone 64-PP as the most common defects. We perform all-atom classical simulations, up to 2 μs, of CPD and 64-PP embedded in a 16-bp duplex, which reveal the constrasted behavior of the two lesions. In particular we evidence a very limited structural deformation induced by CPD while 64-PP is characterized by a complex structural polymorphism. Our simulations also allow to unify the contrasting experimental structural results obtained by nuclear magnetic resonance or Förster Resonant Energy Transfer method, showing that both low and high bent structures are indeed accessible. These contrasting behaviors can also explain repair resistance or the different replication obstruction, and hence the genotoxicity of these two photolesions.

INTRODUCTION

DNA is well renown for its photostability, (1–3) which is a key feature to ensure genome stability and avoid mutations or carcinogenesis (4,5). Nucleobases and nucleic acids absorb in the UV spectrum, mostly in the UVB and slightly in the UVA region due to excitonic coupling. In spite of the existence of efficient photophysical channels (6) allowing non-reactive relaxation, UV absorption opens photochemical pathways leading to nucleobase modification (7). The most common DNA damages produced upon direct UV absorption are characterized by the dimerization of two adjacent, π-stacked pyrimidine rings in B-DNA (see Figure 1) (8–10). Ultimately, photoinduced DNA lesions may induce harmful effects at the cellular level that, in the case of eukaryotic organisms, may be translated into mutation, skin aging and carcinogenesis (11). Indeed, UV absorption, and unprotected sun exposure, are nowadays recognized as the main causes of malignant skin tumors. From a molecular point of view, tumor insurgence is mostly associated with the presence of cyclobutane pyrimidine dimers (CPD) and/or pyrimidine(6-4)pyrimidone 64-PP (8), both lesions being mutagenic and possibly inducing genome instability with differential biological effects (12,13).

UVB irradiation of cells (245 nm) generates CPD as the most abundant photolesions while 64-PP, resulting from the excitation to higher lying electronic states (14), accounts for ∼20% of the defects formed (15,16). Processing efficiency is especially low for CPD (17), and in particular way less efficient than for 64-PP (15). For the repair of both lesions the nucleotide excision repair (NER) mechanism is mobilized (18). However, NER is known to participate in case of strongly distorted DNA structures and in the presence of bulky lesions. Hence, since deformation plays a fundamental role in the lesion recognition (19) a detailed understanding of the structural and dynamical signatures of photolesions is necessary to allow for a rationalization of repair rates (20–23).

X-ray or nuclear magnetic resonance (NMR) structures are indeed available for 6-4PP complexed with repair enzymes (photolyase (17)), that lead to the understanding

*To whom correspondence should be addressed. Tel: +33 372745278; Fax: +33 372745271; Email: antonio.monari@univ-lorraine.fr
Correspondence may also be addressed to François Dehez. Tel: +33 372745076; Fax: +33 372745271; Email: francois.dehez@univ-lorraine.fr
Correspondence may also be addressed to Elise Dumont. Tel: +33 472728846; Fax: +33 4 72728080; Email: elise.dumont@ens-lyon.fr

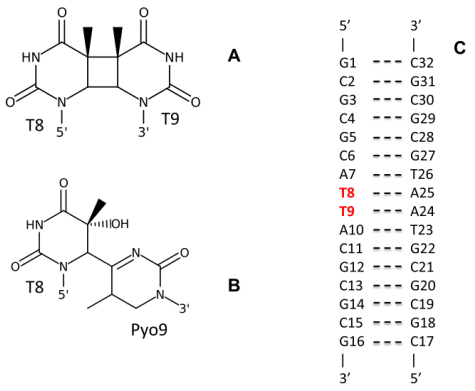


Figure 1. Molecular formula of CPD (**A**) and 64-PP (**B**). The 16 bp DNA double strand is also schematically reported with the bases' numbering (**C**). The two adjacent thymines at positions 8 and 9 (boldfaced in red) are the ones dimerizing to form the photolesions. In the case of 64-PP, the ninth nucleobase (T9) will be denoted Pyo9, being transformed into a pyrimidone unit.

of the enzymatic pathway toward repair at the molecular level (24), also thanks to molecular modeling and simulations (25,26). Yet structural resolution is far less often available for short oligonucleotides containing this photodamage. Two structures for solvated decamers containing 64-PP were reported (27–29), while most recently the structure of a 64-PP containing strand embedded within an histone (23) has been resolved. Other less direct measurements based on Förster resonant energy transfer (FRET) have also been published (30). Interestingly, the experimental determinations provided rather contrasting results: FRET measurements suggest a low DNA bending comparable to the one of the undamaged oligomers, as revealed by the large end-to-end distance of the strand, while X-ray and NMR agree on large bending going up to 44°. If the crystal packing or the interaction with histones can strongly modify the native structural properties of the DNA strands, still the experimental results in solution should be rationalized. Molecular dynamics (MD) simulations have been performed in the past on solvated oligonucleotides (31) or on oligonucleotides interacting with repair enzymes. (15,17,23). The pioneer simulation by Kollman's group (31) reported a very low bending for the 64-PP containing strand, however, the time-scales of the simulations were generally too small to capture their high structural flexibility. Also the structure of CPD containing oligomers has been resolved in the past concluding in favor of a general low bending of the DNA strand, (32) also in agreement with theoretical analysis (33). However, bending up to 30° have also been reported (34). While the repair mechanisms are now well understood owing to numerous and converging experimental (35) and modeling studies (36), insights are still scarce concerning the DNA-enzyme interaction leading to lesion recognition.

Structural deformation and strand flexibility have strong influence on the repair rates of DNA lesions. Indeed while bulky and important deformations are necessary to allow recognition by repair machinery (19), especially NER, flexibility has also been invoked as a factor enhancing the repair propensity of DNA lesions (37).

The use of well-calibrated (38,39), long-scale MD techniques allowing for an extended sampling of the conformational space of the DNA strand are needed to achieve a clear atomistic-scale characterization of the structural reorganization induced by photolesions. More importantly, sufficiently long MD will be able to capture the dynamical evolution of the structural deformation and hence will allow to properly assess the eventual polymorphisms exhibited by the damaged DNA oligonucleotides. MD simulations have been used to ultimately evidence the structural modification induced on DNA oligomers by oxidative lesions, intra- and inter-strand crosslinks, or via the interaction with exogenous and endogenous drugs (40–44). Recently, the structural deformations induced by clustered abasic sites have been correlated to the repair rate in human and bacterial organisms (21,22).

In this work we report an extensive all-atom MD simulation, up to the μs time scale, of DNA oligomers containing one CPD or one 64-PP damage, respectively; compared to the dynamics of the control undamaged B-DNA strand. The different structural behavior induced by the two lesions will be clearly assessed and compared to the evolution of the native B-DNA. The negligible deformations produced by CPD, and the extended polymorphism induced by 64-PP will be deeply analyzed and correlated with the low repair or the replication obstruction of the two photolesions.

MATERIALS AND METHODS

Computational protocol

For all the subsequent study we consider a 16-bp DNA strand (see Figure 1); the choice of a longer oligonucleotide instead of the 10-bp resolved by Lee *et al.*, has been done to avoid instability at the oligomer terminal bases. Starting from the same oligomer the lesion was produced by dimerization of the thymines at position 8 and 9, to obtain one CPD or a 64-PP, respectively. All atom MD simulations were performed with Amber16 (45) and NAMD (46) codes for the three systems, B-DNA, CPD and 64-PP. DNA nucleobases were represented by the parmbsc0 force field (47,48). The performance of the parmbsc0 force field has also been tested against the recently developed bsc1 (38) corrections and shown to perform analogously. In particular over the 150 ns dynamics we observed similar transitions between conformers as the ones described with parmbsc0 force field. The parameters for the CPD and 64-PP lesion were assigned according to the standard RESP protocol (49). We note that they are in very good agreement with the ones derived by Kollman *et al.* (31). Water molecules were represented by TIP3P model and 30 potassium cations were added to the simulation box to assure neutrality. A detailed description of the computational protocol and the force-field parameters for the damaged base can be found in Supplementary Data. After thermalization and equilibration production MD has been performed in the constant pressure and temperature (NPT) ensemble (300 K, 1 atm). In the case of 64-PP a 2 μs trajectory has been performed in order to properly

explore the conformational space evidencing the dynamical equilibrium between different conformers. In order to characterize the global DNA behavior the DNA bending angle was calculated along the three trajectories with Curves+ (50). The compactness of the structures was assessed by the calculation of the solvent accessible surface area (SASA) (51) obtained post-processing the MD trajectories with the measure command of VMD (52) and a probe radius of 1.4 Å. The local deformation was further explored by analyzing the extent of π-stacking between the nucleobases adjacent to the lesion as obtained by the distance between the aromatic rings.

RESULTS

In a first step we assess the structural distortion of the 16-bp duplex containing a photolesion through two global parameters namely the DNA helical bending, and the SASA. Helical bending in particular also allows for a straightforward comparison with experimental results. In Figure 2, we report the distributions of the bending angle as calculated by Curves+ (50) in the case of the native B-DNA and in presence of CPD or 64-PP. CPD bending distribution almost perfectly overlaps the one of the control double-stranded sequence, with maximum values at around 20° and 18°, respectively. Instead, the situation is stunningly different for 64-PP (Figure 2C), where the bending is much more dispersed. Indeed, one can evidence two peaks: a first one centered at around 20°, overlapping with the control B-DNA sequence and the CPD containing duplex, and a second one appearing at larger values (40°). We also point the presence of a less populated yet clear shoulder at even higher bending angles.

SASA values, presented in Figure 2 rightside, globally show similar tendencies as for the bending and clearly points in favor of marked deformations only in the case of 64-PP. Once again CPD and B-DNA distributions almost perfectly overlaps both in terms of the maximum value and of the band shape. On the other hand, three maxima, two well resolved peaks and a shoulder, can be evidenced for 64-PP. The solvent accessibility values indicate that 64-PP spans an ensemble of conformations, some of them being quite compact (low SASA) and other presenting a large surface area (high SASA).

In Figure 3 these tendencies are exemplified by the comparison of the representative structures of B-DNA and CPD containing oligomers (that indeed appears indistinguishable), with two highly distorted 64-PP conformations characterized by a low and high bending, and by different compactness due to the formation of hollow structures.

In order to assess the convergence of 64-PP simulation we also performed parallel tempered replica exchange (53) representing a total sampling time of about 5 μs. The distribution of bending and SASA obtained for 64-PP are reported in Figure 4. Globally the enhanced sampling procedure confirms the previous results and in particular the polymorphism exhibited by 64-PP containing strand, in particular both high and low bending values are once again represented. The SASA distribution is also extremely broad and point toward the coexistence of different structures exhibiting different compactness, however the peak for low SASA, i.e. compact structure, is much less well resolved, indicating a lower population of the corresponding conformers. Indeed, the relative population of the different structures appears difficult to reproduce and would necessitate a prohibitively computational effort, nonetheless the global consideration over the dynamic equilibrium leading to polymorphism still holds.

Hence, a first stunning difference between CPD and 64-PP can be unequivocally invoked, i.e. the structural invariance of the former compared to the important deformation, and the flexibility, induced by the latter. As indicated by the global descriptors 64-PP exhibits polymorphism and thus deserves a closer attention to clearly identify the complex behavior induced by this lesion. Figure 5 provides the time evolution of the distance between the nucleobases in the vicinity of the 64-PP lesion (see Supplementary Data for a proper definition). This parameter has been chosen since it allows to quantify the extent of the formation of π-stacking interactions. Along the 2 μs of the trajectory 64-PP explores a rather large conformational space with a complex equilibrium between different conformations, characterized by different π-stacking interaction patterns between the nucleobases. Quite importantly, the equilibrium is extremely dynamic, the stable conformations are separated by some transition regions, and the conformers can have multiple occurrences during the time series. In partial contrast to the global parameters (bending and SASA) the use of a more local descriptor clearly indicates the presence of at least five distinct conformations. However, the use of unconstrained MD does not allow to estimate the absolute population rates between the conformers nor to *a priori* exclude the existence of other possible stable structures.

Figure 6 provides representative snapshots of all the five stable equilibrium conformations together with schematic diagrams depicting the interactions taking place and stabilizing the specific conformations. It is important to stress out the considerable structural differences as well as the high variability of the coupling between the different nucleobases. Hence, one can safely conclude that 64-PP does not present a well-defined unique structure but rather a complex equilibrium between different conformations whose structure and properties strongly differ. Once again the same behavior is confirmed by the analysis of the distances distribution for the unbiased and parallel tempering simulation as reported in Supplementary Data, even though, as already evidenced for the SASA the relative population is slightly altered.

We seek to identify the rather complex reorganization pattern induced by 64-PP and the non-covalent interactions that come into play to stabilize the duplex after the induction of the covalent-linkage. Cartoon representations are given in Figure 6, together with schemes depicting the spatial arrangement of the nucleobases: hereafter the thymine at position 9 will be indicated as Pyo9 since it is the one forming the pyrimidone moiety (Figure 1). Indeed, the large conformational landscape explored at the μs range arises from the fact the duplex can rearrange to favor interstrand dispersive interactions (green boxes in Figure 6) and hydrogen bonding with offset nucleobases (dashed blue line), which in some cases can lead to the exclusion of vicinal nuclebases.

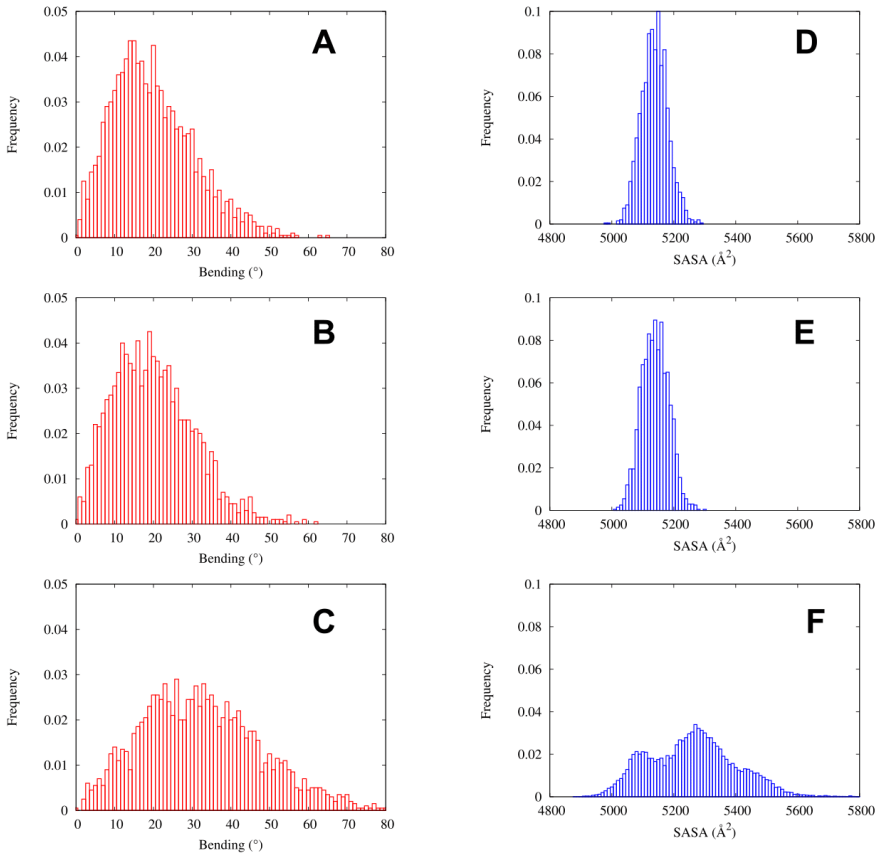


Figure 2. Distribution of the total DNA bending (in red) for B-DNA (**A**), CPD (**B**) and 64-PP (**C**); and of the SASA (in blue) for B-DNA (**D**), CPD (**E**) and 64-PP (**F**).

Conf-(1) maintains a spatial proximity between the damaged T8 and the originally complementary adenine A24, at a distance of ∼4Å which actually corresponds to the π-stacking. A24 can indeed no longer be Watson–Crick paired with Pyo9 and instead the helix readapts by a local shrinkage to afford a stacking between Pyo9 and A24, the latter behaving as an orphan nucleobasis.

The duplex then evolves toward Conf-(2) which is stabilized by conserving an extended stacking on the 64-PP complementary strand involving thymines and adenines (from T23 up to T26) while interstrand dispersive interactions are no more present. Interestingly, this conformation corresponds to a maximal curvature, as verified by the bend angle (54.3°).

After 1 μs, the duplex rearranges to adopt a more straight and compact conformation Conf-(3) with interstrand π-stacking between Pyo9 and T26, but also between A7 and A25. These interactions narrows the B-helix and contributes to decrease the bend (26.7°), whereas no new hydrogen bonds comes into play to stabilize the damaged duplex.

At around 1.8 μs a conformation rather similar to Conf-(2), that we will call Conf-(4), is found to be characterized by a proximity between the damaged Pyo9 and T26. Also this conformation presents an extended stacking on the lesion's complementary strand going from T23 to T26 and a high bending (49.6°). However, differently from Conf-(2), the stacking region is shifted downward.

Oppositely, Conf-(5) is the first structure exhibiting a gap between A24 and A25, which lies at about 7 Å, while T8 gets close to the A7–T26 base pair still locked in the Watson–Crick pairing. Ultimately, the alcohol function of T8 interacts with the oxygen O2 of T26 to form a stable hydrogen-

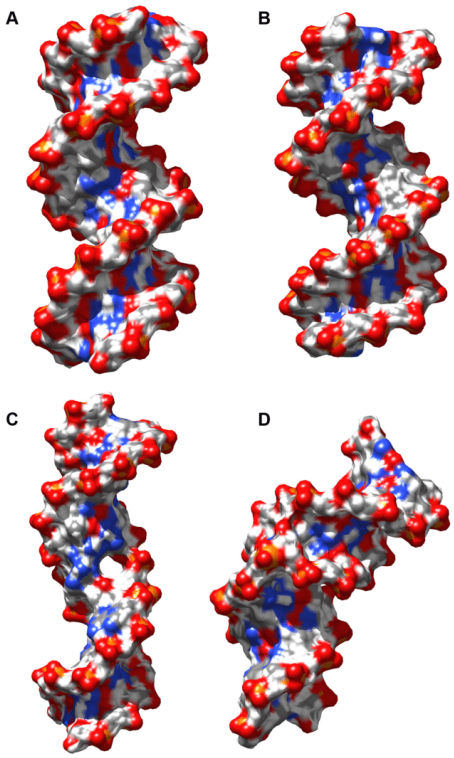


Figure 3. Representative snapshots rendered in surface mode for B-DNA (**A**), CPD (**B**) and 64-(PP) in a low bending high SASA conformation (**C**) and in a high bending low SASA one (**D**).

bond. Conf-(5) duplex presents a globally low bending just slightly higher than the average value for the control B-DNA sequence.

The tendency of A24 to seek opportunistic interactions to compensate for the lack of pairing with the complementary thymine is a constant all along our dynamics. The same behavior is also experienced by Pyo9 that in particular seeks to adopt conformations allowing dispersive interactions with A25. The coupled perturbation of the arrangements of the two strands propagates up to A10 and G22 which may develop non-conventional π-stackings interaction, while the Watson Crick pairing with T23 and C11 is maintained. Conf-(2) and Conf-(4) are both characterized by large bend angles of $\sim 50^{\circ}$, yet they differ markedly by the local structural rearrangement. Indeed, Conf-(4) is pushing A7 to an extrahelical position allowing to maintain C6…G23 and A10…T23 pairings. Note also that, Conf-(4) and Conf-(5) clearly exhibit a hole structure in correspondence of the lesions, leading to high values of SASA; the same behavior, even if less pronounced, is also be evidenced for Conf-(2). Correspondingly, the complex nucleobase rearrangement, necessary to accommodate the intrastrand cross-links created between Pyo9 and T8 leads to marked variations of the total bend angle. However, in all the cases, bulges and helical distortion in the region corresponding to the lesions are clearly identified. The existence of rather large structural deformations and a large reorganization of the helical parameters, in particular in the vicinity of the lesion, are also confirmed by the time-series and the distribution of some relevant DNA structural parameters extracted with Curves+ (50), such as the minor and major groove width and depth (see Supplementary Data). Indeed the base-pairs surrounding the lesions (from 7 to 10) appears as the most influenced and exhibit a strong variation correlating well with the distribution of the distances previously analyzed. Most notably, one can observe a general shrinking of the major groove especially evident for Conf-(2) and Conf-(3) while Conf-(1), Conf-(4), and especially Conf-(5) are characterized by a slight enlargement of the major groove width. Minor groove, on the contrary, appears less influenced and correlation extracted from this parameters less straightforward, even though the value of the groove depth in the correspondence of the lesion peaks to zero indicating an important shrinking of the strand.

DISCUSSION AND CONCLUSIONS

The behavior of the two most common DNA photoproducts, CPD and 64-PP, has been fully characterized by long timescale, all-atom simulations. In particular, it has been clearly evidenced that CPD induces a minimal structural deformation. The CPD containing duplex is indeed absolutely comparable to the native B-DNA, both in terms of bending and SASA distribution. In addition, only one stable conformation is observed and all the local and global structural parameters are much similar to the ones of the undamaged strand. CPD is one of the only lesions in B-DNA implying the formation of two covalent bonds tethering the adjacent thymines: rather logically, this confers more rigidity and locks the structure, decreasing the rotational freedom.

On the other hand, the situation is totally reversed in the case of 64-PP. Indeed, 64-PP is characterized by a dramatic structural deformation centered around the lesion and propagating to a large part of the duplex. At least three maxima can be evidenced in the distributions of both SASA and bending angle, pointing toward the coexistence of compact and surface exposed structures, as well as bent and straight arrangements.

Our simulations unify the complex and somewhat conflicting panorama of experimentally-based 64-PP, and CPD structural resolution. Indeed, while the general CPD low bending is recovered, as shown by the most probable angle, the distribution width clearly indicates that values close to 30° cannot be ruled out and can occur at room temperate. On the other hand, in the case of 64-PP our results show that both NMR (29) and FRET (30), i.e. high and low bending, are possible due to its polymorphisms and depending on the specific conformations. Indeed, we also report in Supplementary Data the distribution of the end-to-end distance of the DNA double strand we observe the presence of two peaks at distance of 55 or 60 Å that correlates well with the

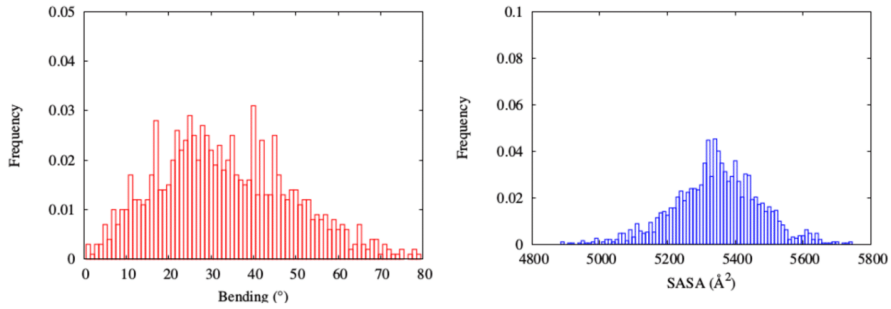


Figure 4. Distribution of the total DNA bending (red) and SASA (blue) obtained from replica exchange parallel tempering.

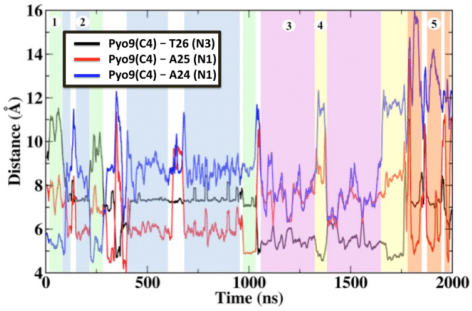


Figure 5. Time evolution of internucleobases distances measured between Pyo9 and A24, A25 and T26, characterizing the different π-stacking formations and hence the increased flexibility and polymorphism upon formation of 64-PP. The regions corresponding to different conformations are indicated by colored background.

values reported in (30) for strands having the same number of base pairs. The NMR predicted bending value of around 40° corresponds to one of the maximum of the bending distribution, on the other hand the FRET results are consistent with the second maximum appearing at lower angle and leading to bending comparable to the ones of native B-DNA.

Recently DNA polymorphism also in the case of non-damaged strands has been proved for instance by circular dichroism techniques (54) or by molecular simulation (55), it has also been correlated to the insurgence of different pathologies in particular when involving gene regulatory areas such as CpG islands (56) or telomeres (57,58).

As it is clearly shown by our simulations, 64-PP is actually experiencing structural polymorphism that ends up in a dynamic equilibrium between different conformations whose local structural parameters differ significantly. Such polymorphism is definitively striking also when compared to the more predictable structural rearrangements produced by oxidatively-induced intrastrand cross-link lesions (41,59). The former lesions rearrange to restore B-helicity and as a consequence they lead to less extreme bending values and a globally more rigid structure.

The observed contrasting structural behavior may indeed be correlated to the repair and replication rate of the two lesions, and hence their toxicity. Indeed, CPD extremely low repair rate may be seen as a direct consequence of the absence of relevant structural modifications compared to B-DNA. This fact indeed, hampers the lesion's recognition since a common recognition pattern is the structural deformation induced by the damage (37). In a simplified way, one may say that CPD actually masks itself and hides between the undamaged B-DNA, hence escaping recognition and repair. However, its rigidity also constitutes a blockage to the replication process that ultimately results in the limited mutagenic power especially as compared to 64-PP (60).

64-PP on the other hand adopts a totally different strategy, the large structural deformation experienced by this lesion is indeed consistent with the mobilization of the NER repair machinery and the lesions is processed much more efficiently than CPD (61). However, due to the observed polymorphism no single stable structure can be evidenced, and the lesion is constantly shifting back and forth between different conformations, and as a consequence the recognition by enzymes can be made more difficult. This is particular true since some of the conformers present very different local deformations and interaction patterns, as well as different global parameters such as bending and surface accessibility. Furthermore, the flexibility and polymorphisms of 64-PP also correlates with the absence of unyielding obstruction to the DNA replication process (62) leading to the accumulation of mutations at the photolesion spot. Indeed, the easy replication of 64-PP damaged strand is recognized as one of the reason of 64-PP high toxicity and mutagenicity (63), and the observed polymorphism and the different nucleobase interaction patterns may also explain the large spectrum of observed mutations.

It is however important to underline that we may anticipate a different behavior of 64-PP when complexed with histones in a nucleosomal environment, as surmised by Akihisa *et al.* (23). In particular the constraints due to interacting proteins and the compact nucleosomal environment may restrict the accessibility of the explored conformational

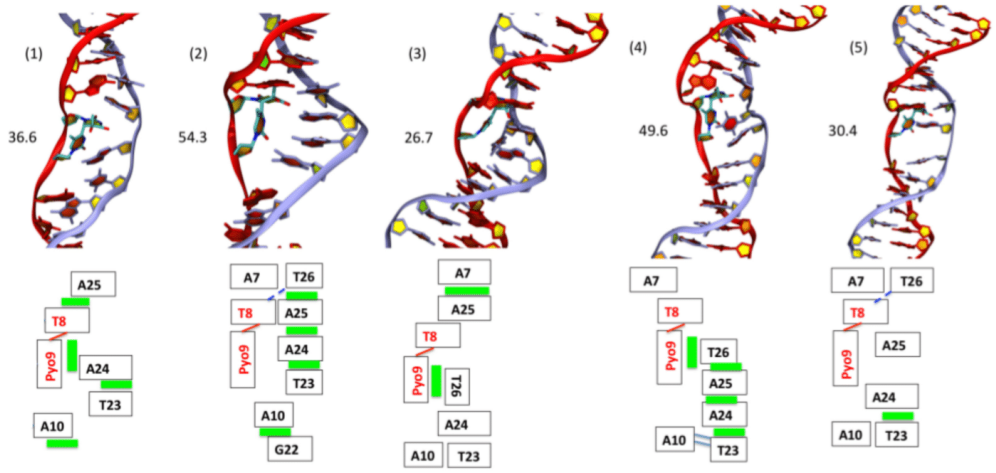


Figure 6. Cartoon representations of five representative conformation observed along the 2 μs trajectory. The schematic depiction of the interaction patterns is given in the bottom figures for each conformer, where green boxes represent π-stacking and blue lines hydrogen bonds. The value of the bending angle for each conformer is indicated in parenthesis.

landscape, still the influence of flexibility on the replication obstruction will hold.

Our simulations have allowed to shed a new light on the complex and dynamic structure of two very important lesions exploring complex free energy surfaces. They sketched two different and opposed behaviors: structural invariance (CPD) or conformational polymorphism (64-PP). In the following, and in order to draw a more precise correlation with repair efficiency, we preview to extend this study to the interaction between the different lesions' conformers and both nucleosomal environment and the repair enzymes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Supports from the University of Lorraine and ENS Lyon is gratefully acknowledged. AM thanks CNRS for support under the project 'Action pour l'interdisciplinarite 2016'. Support from the French Grand-Est region under the Idea project is acknowledged too. This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program 'Investissements d'Avenir' (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

FUNDING

University of Lorraine; ENS Lyon; CNRS under the project 'Action pour l'interdisciplinarite 2016' (to A.M.); French Grand-Est region under the Idea project; LABEX PRIMES of Université de Lyon [ANR-11-LABX-0063]; French National Research Agency (ANR) 'Investissements d'Avenir' program [ANR-11-IDEX-0007]. Funding for open access charge: University of Lorraine and Lyon.
Conflict of interest statement. None declared.

REFERENCES

1. Satzger,H., Townsend,D., Zgierski,M.Z., Patchkovskii,S., Ullrich,S. and Stolow,A. (2006) Primary processes underlying the photostability of isolated DNA bases: Adenine. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 10196–10201.
2. Barbatti,M., Aquino,A. J.A., Szymczak,J.J., Nachtigallova,D., Hozba,P. and Lischka,H. (2014) Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21453–21458.
3. Buchner,F., Nakayama,A., Yamazaki,S., Ritze,H.-H. and Loebcke,A. (2015) Excited-state relaxation of hydrated thymine and thymidine measured by liquid-jet photoelectron spectroscopy: experiment and simulation. *J. Am. Chem. Soc.*, **137**, 2931–2938.
4. Deman,J. and Van Larebeke,N. (2001) Carcinogenesis: mutations and mutagens. *Tumor Biol.*, 191–202.
5. Morley,A.A. and Turner,D.R. (1999) The contribution of exogenous and endogenous mutagens to in vivo mutations. *Mutat. Res.*, **428**, 11–15.
6. Gustavsson,T., Improta,R. and Markovitsi,D. (2010) DNA/RNA: building blocks of life under UV irradiation. *J. Phys. Chem. Lett.*, **1**, 2025–2030.
7. Cadet,J., Mouret,S., Ravanat,J.-L. and Douki,T. (2012) Photoinduced damage to cellular DNA: direct and photosensitized reactions'. *Photochem. Photobiol.*, **88**, 1048–1065.
8. Mitchell,D.L. and Nairn,R.S. (1989) The biology of the (6-4) photoproduct. *Photochem. Photobiol.*, **49**, 805–819.
9. Gonzalez-Ramirez,I., Roca-Sanjuán,D., Climent,T., Serrano-Perez,J.J., Merchan,M. and Serrano-Andres,L. (2011) On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers. *Theor. Chem. Acc.*, **128**, 705–711.
10. Rauer,C., Nogueira,J.J., Marquetand,P. and Gonzalez,L. (2016) Cyclobutane thymine photodimerization mechanism revealed by nonadiabatic molecular dynamics. *J. Am. Chem. Soc.*, **138**, 15911–15916.

11. D'Orazio,J., Jarrett,S., Amaro-Ortiz,A. and Scott,T. (2013) UV radiation and the skin. *Int. J. Mol. Sci.*, **14**, 12222–12248.
12. Lo,H.-L., Nakajima,S., Ma,L., Walter,B., Yasui,A., Ethell,D.W. and Owen,L.B. (2005) Differential biologic effects of CPD and 6-4PP UV-induced DNA damage on the induction of apoptosis and cell-cycle arrest. *BMC Cancer*, **5**, 135.
13. Pfeifer,G.P. and Besaratinia,A. (2012) UV wavelength-dependent DNA damage and human non-melanoma and melanoma skin cancer. *Photochem. Photobiol. Sci.*, **11**, 90–97.
14. Banyasz,A., Douki,T., Improta,R., Gustavsson,T., Onidas,D., Vayá,I., Perron,M. and Markovitsi,D. (2012) Electronic excited states responsible for dimer formation upon UV absorption directly by thymine strands: joint experimental and theoretical study. *J. Am. Chem. Soc.*, **134**, 14834–14845.
15. Balajee,A.S., May,A. and Bohr,V.A. (1999) DNA repair of pyrimidine dimers and 6-4 photoproducts in the ribosomal DNA. *Nucleic Acids Res.*, **27**, 2511–2520.
16. Perdiz,D., Grof,P., Mezzina,M., Nikaido,O., Moustacchi,E. and Sage,E. (2000) Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells: possible role of dewar photoproducts in solar mutagenesis. *J. Biol. Chem.*, **275**, 26732–26742.
17. Li,J., Liu,Z., Tan,C., Guo,X., Wang,L., Sancar,A. and Zhong,D. (2010) Dynamics and mechanism of repair of ultraviolet-induced (6-4) photoproduct by photolyase. *Nature*, **466**, 887–890.
18. Marteijn,J.A., Lans,H., Vermeulen,W. and Hoeijmakers,J. H.J. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Biol.*, **15**, 465–481.
19. Scharer,D.O. (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **5**, a012609.
20. Han,C., Srivastava,A.K., Cui,T., Wang,Q.-E. and Wani,A.A. (2016) Differential DNA lesion formation and repair in heterochromatin and euchromatin. *Carcinogenesis*, **37**, 129–138.
21. Bignon,E., Gattuso,H., Morell,C., Dehez,F., Georgakilas,A.G., Monari,A. and Dumont,E. (2016) Correlation of bistranded clustered abasic DNA lesion processing with structural and dynamic DNA helix distortion. *Nucleic Acids Res.*, **44**, 8588–8599.
22. Gattuso,H., Durand,E., Bignon,E., Morell,C., Georgakilas,A.G., Dumont,E., Chipot,C., Dehez,F. and Monari,A. (2016) Repair rate of clustered abasic DNA lesions by human endonuclease: molecular bases of sequence specificity. *J. Phys. Chem. Lett.*, **7**, 3760–3765.
23. Akihisa,O., Hiroaki,T., Wataru,K., Naoki,H., Syota,M., Mayu,H., Naoyuki,M., Tatsuya,T., Junpei,Y., Fumio,H. *et al.* (2015) Structural basis of pyrimidine-pyrimidone (64) photoproduct recognition by UV-DDB in the nucleosome. *Sci. Rep.*, **5**, 16330–16330.
24. Yamamoto,J., Martin,R., Iwai,S., Plaza,P. and Brettel,K. (2013) Repair of the (64) photoproduct by DNA photolyase requires two photons. *Angew. Chem. Int. Ed.*, **52**, 7432–7436.
25. Dokainish,H.M. and Kitao,A. (2016) Computational assignment of the histidine protonation state in (6-4) photolyase enzyme and its effect on the protonation step. *ACS Catal.*, **6**, 5500–5507.
26. Knips,A. and Zacharias,M. (2017) μABC: a systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA. *Sci. Rep.*, **7**, 41324.
27. Kim,J.-K., Patel,D. and Choi,B.-S. (1995) Contrasting structural impacts induced by cis-syn cyclobutane dimers and (6-4) adduct in DNA duplex decamers: Implication in mutagenesis and repair activity. *Photochem. Photobiol.*, **62**, 44–50.
28. Kim,J.-K. and Choi,B.-S. (1995) The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of Thymidylyl(3′−>5′)Thymidine by NMR and relaxation matrix refinement. *Euro. J. Biochem.*, **228**, 849–854.
29. Lee,J.-H., Hwang,G.-S. and Choi,B.-S. (1999) Solution structure of a DNA decamer duplex containing the stable 3′ T.G base pair of the pyrimidine(64)pyrimidone photoproduct [(64) adduct]: Implications for the highly specific 3′ T –> C transition of the (64) adduct. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 6632–6636.
30. Mizukoshi,T., Kodama,T.S., Fujiwara,Y., Furuno,T., Nakanishi,M. and Iwai,S. (2001) Structural study of DNA duplexes containing the (64) photoproduct by fluorescence resonance energy transfer. *Nucleic Acids Res.*, **29**, 4948–4954.
31. Spector,T.I., Cheatham,T.E. and Kollman,P.A. (1997) Unrestrained molecular dynamics of photodamaged DNA in aqueous solution. *J. Am. Chem. Soc.*, **119**, 7095–7104.
32. Wang,C.I. and Taylor,J.S. (1991) Site-specific effect of thymine dimer formation on dAn.dTn tract bending and its biological implications. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 9072–9076.
33. Cooney,M.G. and Miller,J.H. (1997) Calculated distortions of duplex DNA by a Cis, Syn cyclobutane thymine dimer are unaffected by a 3′ TpA step. *Nucleic Acids Res.*, **25**, 1432–1436.
34. Park,H., Zhang,K., Ren,Y., Nadji,S., Sinha,N., Taylor,J.-S. and Kang,C. (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15965–15970.
35. Ando,H., Fingerhut,B.P., Dorfman,K.E., Biggs,J.D. and Mukamel,S. (2014) Femtosecond stimulated raman spectroscopy of the cyclobutane thymine dimer repair mechanism: a computational study. *J. Am. Chem. Soc.*, **136**, 14801–14810.
36. Masson,F., Laino,T., Tavernelli,I., Rothlisberger,U. and Hutter,J. (2008) Computational study of thymine dimer radical anion splitting in the self-repair process of duplex DNA. *J. Am. Chem. Soc.*, **130**, 3443–3450.
37. Yang,W. (2011) Surviving the Sun: repair and bypass of DNA UV lesions. *Protein Sci.*, **20**, 1781–1789.
38. Ivani,I., Dans,P.D., Noy,A., Perez,A., Faustino,I., Hospital,A., Walther,J., Andrio,P., Goni,R., Balaceanu,A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, **38**, 55–58.
39. Perez,A., Luque,F.J. and Orozco,M. (2012) Frontiers in molecular dynamics simulations of DNA. *Acc. Chem. Res.*, **45**, 196–205.
40. Dumont,E. and Monari,A. (2015) Understanding DNA under oxidative stress and sensitization: the role of molecular modeling. *Front. Chem.*, **3**, 43.
41. Dumont,E., Drsata,T., Fonseca Guerra,C. and Lankas,F. (2015) Insights into the structure of intrastrand cross-link DNA lesion-containing oligonucleotides: G[85m]T and G[85]C from molecular dynamics simulations. *Biochemistry*, **54**, 1259–1267.
42. Dumont,E. and Monari,A. (2013) Benzophenone and DNA: Evidence for a double insertion mode and its spectral signature. *J. Phys. Chem. Lett.*, **4**, 4119–4124.
43. Nogueira,J.J. and Gonzalez,L. (2014) Molecular dynamics simulations of binding modes between methylene blue and DNA with alternating GC and AT sequences. *Biochemistry*, **53**, 2391–2412.
44. Spiegel,K., Rothlisberger,U. and Carloni,P. (2004) Cisplatin binding to DNA oligomers from hybrid car-parrinello/molecular dynamics simulations. *J. Phys. Chem. B*, **108**, 2699–2707.
45. Case,D., Betz,R., Botello-Smith,W., Cerutti,D., III,T.C., Darden,T., Duke,R., Giese,T., Gohlke,H., Goetz,A. *et al.* (2016) *AMBER 16*, University of California, San Francisco, CA.
46. Phillips,J.C., Braun,R., Wang,W., Gumbart,J., Tajkhorshid,E., Villa,E., Chipot,C., Skeel,R.D., Kale,L. and Schulten,K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
47. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
48. Perez,A., Marchan,I., Svozil,D., Sponer,J., III,T.E.C., Laughton,C.A. and Orozco,M.2007 Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma Conformers. *Biophysical.*, **92**, 3817–3829.
49. Wang,J., Wolf,R.M., Caldwell,J.W., Kollman,P.A. and Case,D.A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.*, **25**, 1157–1174.
50. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
51. Shrake,A. and Rupley,J.A. (1973) Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.*, **79**, 351–371.
52. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD—visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
53. Sugita,Y. and Okamoto,Y. (1999) Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, **314**, 141–151.
54. Kypr,J., Kejnovska,I., Renciuk,D. and Vorlickova,M. (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.*, **37**, 1713–1725.

55. Pasi,M., Maddocks,J.H., Beveridge,D., Bishop,T.C., Case,D.A., Cheatham,T. III, Dans,P.D., Jayaram,B., Lankas,F., Laughton,C. *et al.* (2014) μABC: a systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA. *Nucleic Acids Res.*, **42**, 12272–12283.
56. Dans,P.D., Faustino,I., Battistini,F., Zakrzewska,K., Lavery,R. and Orozco,M. (2014) Unraveling the sequence-dependent polymorphic behavior of d(CpG) steps in B-DNA. *Nucleic Acids Res.*, **42**, 11304–11320.
57. Cogoi,S., Paramasivam,M., Spolaore,B. and Xodo,L.E. (2008) Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res.*, **36**, 3765–3780.
58. Gaynutdinov,T.I., Neumann,R.D. and Panyutin,I.G. (2008) Structural polymorphism of intramolecular quadruplex of human telomeric DNA: effect of cations, quadruplex-binding drugs and flanking sequences. *Nucleic Acids Res.*, **36**, 4079–4087.
59. Churchill,C.D., Eriksson,L.A. and Wetmore,S.D. (2016) DNA distortion caused by uracil-containing intrastrand cross-links. *J Phys. Chem. B*, **25**, 1195–1204.
60. Sale,J.E., Lehmann,A.R. and Woodgate,R. (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat. Rev. Mol. Cell Biol.*, **13**, 141–152.
61. Courdavaultas,S., Baudouinb,C., Charveronb,M., Canguilhemb,B., Faviera,A., Cadeta,J. and Douki,T. (2005) Repair of the three main types of bipyrimidine DNA photoproducts in human keratinocytes exposed to UVB and UVA radiations. *DNA Repair*, **4**, 836–844.
62. Kuang,Y., Sun,H., Blain,J.C. and Peng,X. (2012) Hypoxia-selective DNA interstrand cross-link formation by two modified nucleosides. *Chemistry*, **18**, 12609–12613.
63. Gentil,A., Page,F.L., Margot,A., Lawrence,C.W., Borden,A. and Sarasin,A. (1996) Mutagenicity of a unique thymine-thymine dimer or thymine-thymine pyrimidine pyrimidone (64) photoproduct in mammalian cells. *Nucleic Acids Res.*, **24**, 1837–1840.

Chapitre 5 - Projets de recherche

I. Dynamique conformationnelle de la protéase ClpP

Les protéases caséinolytiques (ClpP) sont de grandes protéases cylindriques à sérines (environ 300 kDa) présentes dans les bactéries et la plupart des eucaryotes.[111] Bien que les ClpP soient capables de dégrader de petits peptides par elles-mêmes, une association avec leurs chaperones énergie dépendante (AAA+) est nécessaire pour dégrader des peptides et des protéines plus gros.[112] Les ClpP jouent un rôle actif dans la survie et la virulence des bactéries pathogènes. Par conséquent, le développement de médicaments ciblant les ClpPs constitue une stratégie prometteuse pour lutter contre les bactéries multirésistantes.[113–115] En outre, des indications selon lesquelles l'expression de ClpP mitochondriale est induite en réponse à un stress cellulaire, par exemple en conséquence de la genèse tumorale, souligne l'importance des protéases ClpP dans l'homéostasie des protéines eucaryotes.

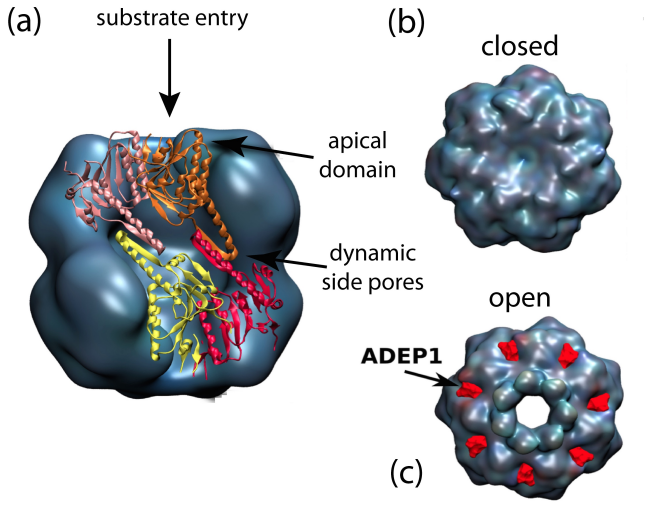


Figure 30. Vue schématique des données ClpP et de faisabilité. Vues de côté (a) et de dessus (b) de la structure ClpP à double anneau dans les états fermé (b) et ouvert lié à ADEP (c).

Dans son état fonctionnel, ClpP forme une structure complexe composée de deux anneaux heptamériques empilés face à face (Fig. 31).[116] La libération des produits de dégradation se produit par l'ouverture dynamique de pores latéraux (Fig. 1a). Deux principaux états structuraux de ClpP ont été observés dans les structures cristallographiques[117] : (i) un état comprimé dans lequel les deux cycles sont proches et la triade catalytique est dans un état de faible activité et (ii) une conformation active avec une triade catalytique idéalement positionnée. Cependant, plusieurs questions clés restent ouvertes concernant ces états et les transitions entre eux. (i) Les protéines non repliées entrent dans la chambre ClpP par la "région de déclenchement" N-terminale du domaine apical.

La liaison d'une chaperone régulatrice chargé de substrat (c'est-à-dire d'une protéine à dégrader) à ClpP entraîne un changement de conformation de la région de déclenchement en une conformation ouverte. Récemment, il a été démontré que de petites molécules qui déclenchent l'ouverture de ClpP éradiquent les bactéries résistantes à d'autres traitements antibactériens, ouvrant ainsi la voie à une nouvelle génération d'antibiotiques. Bien qu'il existe des informations structurelles sur la liaison de l'ADEP à ClpP, on ne sait pas comment la transduction du signal allostérique des sites de liaison à la région de porte N-terminale se propage. Ces sites de liaison correspondent à ceux des co-chaperones, de sorte que la compréhension des effets allostériques induits par le médicament fournira également un aperçu mécaniste du fonctionnement de la machinerie biologique chaperone / protéase. (ii) Un autre processus allostérique intéressant mais insaisissable concerne la communication entre différents sites actifs. Lorsque le petit médicament bortezomib,[118] qui se lie au site actif de ClpP, est ajouté en quantités sous-stœchiométriques, de sorte que seules quelques sous-unités de ClpP sont liées à un inhibiteur, les autres sites actifs de la structure du tétradécamère deviennent plus actifs, ce qui suggère un couplage allostérique les sous-unités voisines de la protéase, par un mécanisme encore insaisissable.
L'objectif est d'élucider les bases structurelles et dynamiques des divers états fonctionnels des protéases de ClpP, y compris le rôle de la transduction du signal allostérique lors de la liaison d'activateurs et d'inhibiteurs. Nous nous servirons des structures obtenues par Cryo-microscopie électronique de ClpP en complexe avec sa chaperone ClpX qui viennent juste d'être publiées.[119,120] Pour accéder aux transitions conformationnelles µs à ms, échelles de temps inaccessibles aux simulations standard tout-atome, nous utiliserons les algorithmes d'échantillonnage d'importance (par exemple, force de biais adaptative (ABF), échantillonnage général ou métadynamique).

Projet ANR 2019-2023

Targeting ClpP proteases: From dynamics and function to allosteric modulation by drugs

I. Proposal's context, positioning and objective(s)

a. Objectives and scientific hypotheses

Numerous processes in the cell depend on large molecular assemblies, which perform tasks as diverse as protein synthesis, protein degradation and chaperoning and enzymatic turnover of small metabolites. These "nano-machines" often rely on the orchestrated action of different subunits or different proteins forming a large entity. The long-range communications connecting distant parts to a functional unit are among the most intriguing and challenging phenomena of molecular-level biological systems. The central aim of this project is to provide insight into such long-range allosteric communications, combining state-of-the-art biophysical and structural approaches, with advanced *in silico* methods in a closely integrated manner.

This project deals with the ClpP machinery, a protein that is part of the cellular "protein quality control" system, which ensures protein homeostasis in the cell, and comprises chaperones, unfoldases, proteases and peptidases.[1] Molecular chaperones bind, protect and refold or unfold misfolded proteins, thus preventing their aggregation; proteins which need to be eliminated are processed by proteases and peptidases, often in concert with co-chaperones which unfold and deliver substrate proteins to the protease machineries.

The caseinolytic protease (Clp) system is a paradigmatic case of a protease machinery, present in bacteria and most eukaryotes. Clp plays an active role in survival and virulence of pathogenic bacteria. Therefore, the development of drugs targeting ClpPs has recently

emerged as a promising strategy to address multi-resistant bacteria.[2–7] Furthermore, indications that mitochondrial ClpP expression is induced in response to cellular stress, e.g. as a consequence of tumor genesis, stresses the importance of ClpP proteases in eukaryotic protein homeostasis.[8,9]

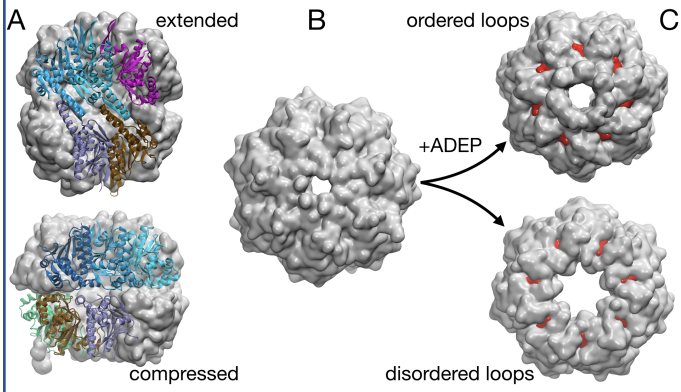


Figure 1: Schematic view of allosteric transitions in ClpP. Side view (A) of ClpP in its holo/active (top, 3MT6) and apo/inactive (bottom, 3QWD) conformations. Top view of the apo/inactive conformation (B). Addition of acydepsipeptide (ADEP, in red) resul-ting in conformations with ordered and disordered N-terminal loops, as observed in two independent crystal structures (C, 3MT6 and 3KTI, respectively).

The Clp system comprises a protease enzyme unit (ClpP) associated with an auxiliary AAA+ chaperone[10] (such as ClpX, ClpA and ClpC). The latter forms a hexameric ring-structure and is capable of unfolding proteins in an ATP-dependent manner by threading the polypeptide chain through its central pore.[11] ClpPs are large cylindrical serine proteases formed by two head-to-head stacked heptameric rings with a total molecular weight of ca. 300 kDa for the 14-mer (Figure 1). The tetradecameric structure forms a ca. 50 Å wide lumen with 14 active sites. The two adjacent rings are connected by the central α-helix E and a β-strand, forming inter-ring interactions. The Ser-His-Asp catalytic triad active sites are harbored in the barrel's inner lumen, in the vicinity of the central α-helix E. Different conformations of ClpP have been observed in crystal structures, including an extended conformation, in which the active site catalytic triad is optimally positioned for enzymatic activity, and compressed states, with misaligned catalytic triads. Thus, control of the activity can be achieved by a shift of the conformational state (extended *vs.* compressed).[12] Proteins to be degraded can gain access to the catalytic chamber only through the axial pores, located at the poles of the tetradecameric barrel structure, but in the resting state these pores are occluded by the N-terminal loops of the protease (Figure 1B,C). By itself, ClpP has, thus, only limited peptidase activity, and is capable of digesting only small peptides which enter the secluded proteolytic chamber via residual openings in the axial pores. This self-compartmentalized structure is a feature often encountered in the protein quality control machinery, providing the systems essentially "molecular sieve" properties, protecting the cell from uncontrolled protease activity. For processing larger polypeptides, the AAA+ chaperones are strictly required. Although no structural data at atomic resolution are available to date, a mechanism of this complex has been deduced from biochemical data. Furthermore, EM initially showed that hexameric ClpX rings stack coaxially onto one or both ends of the ClpP barrel to form singly- or doubly-capped structures.[13–15] The physical interaction between the double-heptameric ClpP barrel with the hexameric chaperone is thought to result in an opening of the axial pores. Furthermore, binding is thought to trigger the rearrangement of ClpP from an inactive compressed state to an active extended state, which has a longer axial dimension and places the catalytic triad in an ideal geometry for catalysis. Thus, polypeptides are transferred through the unfoldase and the open axial pores into the lumen of the active protease for digestion. Transiently opening lateral pores have been proposed as exit routes of the produced amino acid degradation products.[16] Thus, chaperone binding causes major conformational changes both in the axial pore and in the interface between the rings, transforming ClpP from its resting state to an active machinery. Intriguingly, binding of the chaperone at the distal axial binding sites appears to have a direct impact on the active site, and is capable of leading to the displacement of inhibitors bound to the active site.[17] The mechanistic basis of these long-range communications, connecting

chaperone-binding sites to the protease's entry gate and active sites over distances of >20 Å remains unresolved.

Dysregulation of the Clp system can cause major physiological defects in bacteria. In particular, acyldepsipeptide (ADEP), a natural peptide-based compound, has been shown to activate ClpP, converting it from a highly-regulated peptidase that can degrade proteins only with the aid of its partner AAA+ to an independent and unregulated protease. ADEP treatment was shown successful to be effective antibiotics for bacteria resistant to other treatments.[2,5] This finding placed ClpP among the promising new targets for antibiotic therapy and triggered intense research targeted towards deciphering the mode of action of these potential drugs. The binding sites of ADEP coincide with the putative interaction regions of the AAA+ chaperone, and drug binding thereby prevents the physiological interaction of the protease with its AAA+ chaperones. Furthermore, biochemical measurements show that binding induces an opening of the entry gates and increases the activity, thus mimicking the effects of chaperone-binding on the protease.[14] Conflicting views of the structural arrangement of the gate region have been reported. However, the structure of the gate in the ADEP bound state has not been resolved unambiguously, and it remains unclear how this long-range communication is achieved between distant parts of the protein. **Partner 2** has recently obtained further intriguing evidence for long-range communication in the ClpP protease. Binding of an inhibitor (bortezomib) to the active site of the protease at non-saturating concentrations leads to activation of the protease through (i) opening of the axial pores and (ii) enhanced catalytic efficiency at the unbound catalytic sites of neighboring ClpP subunits through promotion of the active extended conformation of the barrel (further outlined in section I.b).

The main aim of this project is to understand the structural, dynamical and mechanistic basis of long-range communications in ClpP proteases. To this end, we integrate a wide panel of data, ranging from biochemical activity measurements through crystallographic and electron-microscopy (EM) structures to atomic-resolution NMR data of dynamics and interactions, overall long-range structural views in solution provided by small-angle X-ray scattering (SAXS), and advanced in silico modelling through MD simulations. We will address the severe challenges associated with the study of such a large and dynamic protein assembly by novel developments of both experimental and computational approaches. We anticipate that this project will generate key new insights into long-range communications in large protein machineries. Through the mechanistic understanding of the action of ClpP activators and inhibitors, the foundations laid out by this project may enable the development of future antibiotic drugs that integrate the structural and dynamic properties of this machinery.

The aims of the **ProteaseInAction** are:

- To gain new insight into allosteric transitions in proteases, integrating structural biology with computational biophysics.
- To characterize the modulation of such allosteric transitions by drugs.
- To unravel the three-dimensional organization of proteases with their chaperones.
- To push beyond the current state of the art experimental and theoretical methods for the investigation of large, highly dynamic protein complexes.

II. Dynamique fonctionnelle des canaux pentameriques

II.1. Les récepteurs 5-HT3

L'intérêt pour la pharmacologie des récepteurs 5-HT3 est apparu conjointement à l'utilisation des chimiothérapies (par exemple, le cisplatine approuvé en 1978 pour les cancers de l'ovaire et du testicule). La dissection des caractéristiques du récepteur 5-HT3 et la découverte de ligands

spécifiques ont ensuite progressé parallèlement. Au moment de leur découverte dans les années 1980, les premiers antagonistes compétitifs étaient décrits comme «les médicaments les plus puissants de toutes les classes pharmacologiques décrites à ce jour» et moins de dix ans plus tard, la recherche antiémétique[121,122] sur 5-HT3 était aboutie pour l'industrie pharmaceutique avec la mise sur le marché de nombreux médicaments à succès cliniques (famille des -sétrons).

Le mode d'action moléculaire des antiémétiques a été étudié de manière approfondie par des études de liaison et d'électrophysiologie, tandis que des études structurales sur des protéines modèles solubles ont proposé des poses approximatives pour les médicaments liés. Récemment, les structures Cryo-EM du récepteur 5-HT3 ont révélé les conformations de l'état apo, de l'inhibition liée au tropisétron et de deux états liés à la sérotonine.[84] Cependant, la résolution limitée de la structure du tropisétron a empêché une analyse approfondie du site de liaison. Sur la base de nouvelles structures obtenus par nos collaborateurs Grenoblois, nous avons étudié de manière systématique l'association des -sétrons au récepteur 5-HT3 par dynamique moléculaire pour établir les bases moléculaires de l'effet de ces médicaments et caractériser les différences dans les modes d'actions.[123] Par la suite, je souhaite poursuivre le travail entrepris sur les interactions médicaments/récepteur 5-HT3 en m'intéressant à une autre molécule, la vortioxetine , utilisée pour traitres les 5-HT3 épisodes dépressifs caractérisés.[124] Cet antidépresseur a été très récemment autorisé sur le marché (2018 pour les USA, 2020 pour la France). A l'inverse des inhibiteurs de la recapture de la sérotonine, la vortioxetine (VTX) une activité multimodale ciblant les récepteurs 5-HT1A, 5-HT1B, 5-HT1D, 5-HT3, 5-HT7 et le transporteur de la sérotonine. Le VTX est un ligand atypique avec des effets à la fois agonistes et antagonistes. Plus précisément, la liaison au VTX induit une activité initiale suivie d'une inhibition apparemment persistante et insurmontable de la stimulation agoniste ultérieure,[125–127] faisant ainsi agir le VTX comme un antagoniste fonctionnel dans des conditions d'équilibre. Cet effet inhibiteur a été supposé refléter que le VTX induit et stabilise un état désensibilisé non conducteur du récepteur humain. Un autre aspect énigmatique du VTX est que ses effets diffèrent nettement entre les récepteurs 5-HT3A d'espèces étroitement apparentées. Plus précisément, l'activité agoniste observée au niveau du récepteur 5-HT3A humain est absente pour le récepteur 5-HT3A du rat.

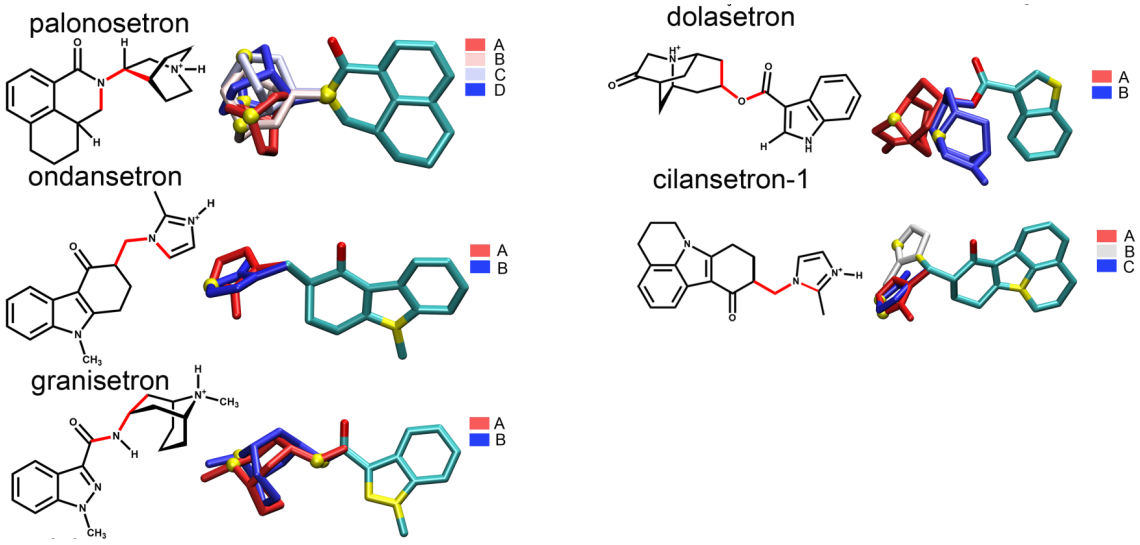


Figure 31. Structure des -sétrons, chaque conformation accessibles est représentée par son un code couleur différent.

La dynamique du récepteur 5-HT_3 est régie par les sous-unités qui le composent, la nature du ligand, le site de liaison considéré, ou encore la composition ionique de l'environnement, elle peut également être influencée par la composition lipidique de la membrane plasmique qui conditionne ses propriétés physiques, notamment sa fluidité. En effet, les proportions des différents lipides qui la composent varient suivant les tissus. Le cholestérol est un lipide connu pour avoir une influence sur la rigidité de la membrane plasmique, ainsi que pour sa capacité à se lier directement à certains récepteurs pour en moduler l'activité. En particulier des récepteurs comme le récepteur à l'acétylcholine ou au GABA, des récepteurs ionotropes qui possèdent une architecture pentamérique relativement proche du 5-HT3R, sont également affectés par les modifications du taux de cholestérol membranaire, ce qui suggère que cela pourrait également être le cas avec cet objet d'étude. Nous souhaitons caractériser l'association spécifique de molécules de cholestérol au canal 5-HT3 par le biais de simulation de dynamique moléculaire tout-atome mais également en utilisant des modèles gros-grains, qui grâce aux échelles de temps plus longues qu'ils permettent d'atteindre, devrait nous permettre d'identifier les éventuels acides aminés de la surface du récepteur capables de liés spécifiquement le cholestérol.

Enfin, fort de la connaissance des structures représentants différents états conformationnels le long du chemin d'activation de 5-HT3, nous envisageons de mettre en œuvre des algorithmes d'échantillonnage préférentiel pour décrire au niveau atomique les chemins de transition conformationnelle et révéler le réseau de communication allostérique sous-jacent à la fonction de cette protéine.

II.2. Les récepteurs nicotiniques de l'acétylcholine

Les récepteurs nicotiniques de l'acétylcholine (nAChR) assurent une communication synaptique rapide au niveau des synapses neuronales et des jonctions neuromusculaires. Ils sont impliqués dans de nombreux troubles neurologiques allant des syndromes myasthéniques à la maladie d'Alzheimer, et jouent un rôle central dans la dépendance à la nicotine. Le Torpedo nAChR est le membre prototypique de la superfamille de canaux ioniques pentamériques ligand-dépendants (pLGIC) qui comprend à la fois les nAChR musculaires et neuronaux, ainsi que les récepteurs de la sérotonine (5-HT3), du GABAA et de la glycine. Depuis 2015, les progrès rapides de la cryo-EM ont conduit à une croissance explosive de notre connaissance structurale des pLGICs révélant des détails fins concernant la liaison des médicaments [128], etc., ainsi que divers instantanés dans le paysage conformationnel.[84,129,130] Les nAChR neuronaux $\alpha 4\beta 2$ et $\alpha 3\beta 4$ ont été résolus dans des états liés aux agonistes.[131,132] Au cours de la dernière année, trois conformations du nAChR $\alpha 7$[133] et un état inhibé du nAChR Torpedo[134] ont également été publiés. L'explosion de nouvelles structures, cependant, n'a pas encore livré une image unificatrice de la façon dont la liaison d'un agoniste engendre la transition d'un pLGIC dans son état de repos vers des conformations ouvertes et finalement désensibilisées. Un facteur qui entrave notre compréhension du mécanisme de fonctionnement de pLGIC est la difficulté d'attribuer à une structure donnée un état physiologique. Par exemple, on ne sait toujours pas comment la plus faible constriction des pores au niveau d'un résidu chargé négativement observée dans le nAChR lié à agoniste désensibilisant, altère la perméation des ions.[135]

Pour aider à caractériser les différents états conformationnels des récepteurs nAChR, nous avons commencé à étudier par dynamique moléculaire l'ensemble des structures disponibles dans des environnements membranaires. Nous nous sommes concentrées sur les conformations présentées comme désensibilisées. Des simulations à l'équilibre ont dans un premier temps pu établir que des structures de divers récepteurs, pourtant extrêmement similaires, avaient un comportement très différent stables au cours du temps, certaines restant table, d'autres étant marquées par un collapse du pore hydrophobe. Nous avons réussi à montrer que ce collapse pouvait être prévenu par l'incorporations de lipides au niveau du pore, un modèle supporté par les traces de densité observées au niveau du pore par les cartes de densité électronique obtenues par cryo-EM (Figure 32).

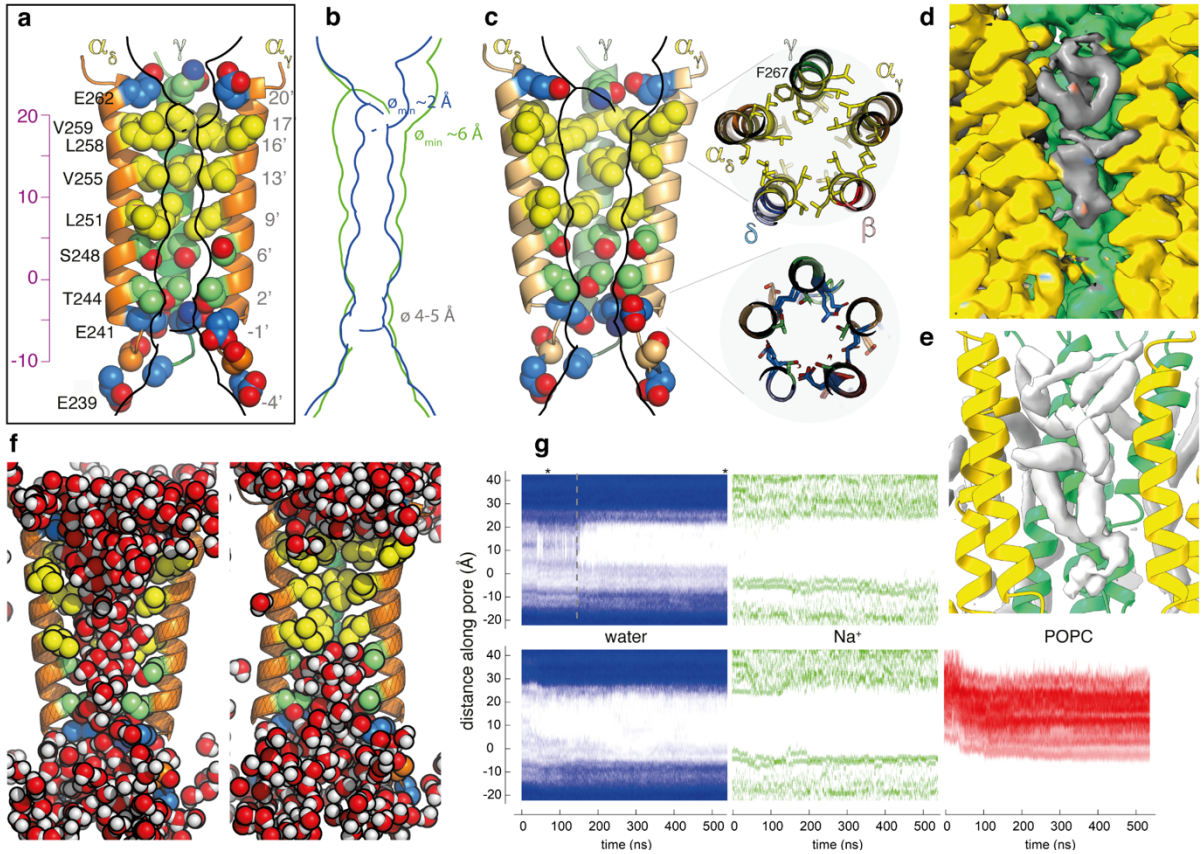


Figure 32. Conformations et hydrataion du pore. **a.** Vues latérales des hélices M2 de nAChRapo (les deux sous-unités de premier plan sont supprimées) avec des résidus face aux pores représentés sous forme de sphères (jaune hydrophobe, vert polaire, bleu chargé négativement). Le diamètre du pore accessible aux ions et aux molécules d'eau (calculée par CHAP) est indiquée par un contour noir. **b**. Superposition des diamètres du pore nAChRapo (bleu) et nAChRcarb (vert). La superposition montre la nette différence au niveau de la porte supérieure et la similitude au niveau -1'. **c.** Vues latérales des hélices M2 de nAChRcarb, avec vues de face tronquée au niveau de la constriction inférieure des pores. **d.** Vue latérale de la carte nAChRcarb (la densité autour des deux sous-unités de premier plan est supprimée) avec la densité observée à l'intérieur du pore en gris **e.** Instantané des hélices M2, pris lors d'une simulation MD de nAChRcarb avec 3 molécules POPC dans le pore. La densité moyenne des molécules POPC est indiquée. **f.** Vue latérale des hélices M2, prises lors d'une simulation MD de nAChRcarb, à 50 ns où les contraintes aux structures expérimentales sont toujours appliquées (à gauche) et à 600 ns après relaxation (à droite). Les molécules d'eau sont représentées par des sphères. **g.** Profils de densité d'eau, d'ions sodium et de POPC le long de l'axe du canal, lors d'une simulation MD de nAChRcarb seul (en haut, la ligne pointillée indique la libération de contraintes sur le pore) et lorsque 3 molécules de POPC sont ajoutées dans le pore (en bas).

Nous sommes en train de mettre en œuvre de simulations d'électrophysiologie computationnelle et des potentiels de force moyenne pour caractériser l'état de conductance des conformations dites désensibilisées, a priori non-conductrice.

Projet ANR 2021-2025

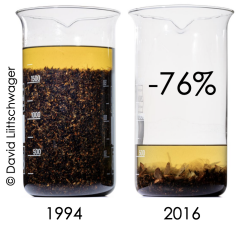
Molecular mechanism of insecticide action at pentameric receptors

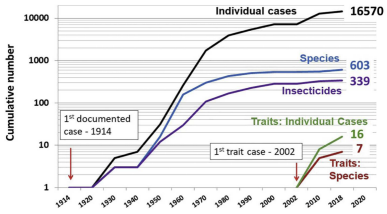
I. Context, positioning and objectives

a. Context and objectives

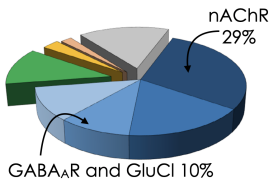
Insecticides in the anthropocene

Pest control is a central component of agriculture. Neonicotinoids (neonic), for instance, are hugely successful insecticides: they kill insects efficiently, have good systemic activity and relatively low toxicity in mammals[1]. Imidacloprid was the first neonic to be marketed, in 1990. And for a while, things appeared to go well. A 2008 review entitled "*Neonicotinoids – from zero to hero*"[2] , cited over 700 times, never mentioned the words 'bee' or 'pollinator' or 'beneficial', reflecting that environmental concerns were far from central a decade ago. The section on resistance conveyed an optimistic tone, as exemplified by "*Imidacloprid has proved remarkably resilient to resistance development, and cases that have been reported are still relatively manageable*". Things changed, however. In a recent review, the same author acknowledges that resistance to insecticides has become an "*extremely serious problem requiring a proactive approach*"[3] (see figure). The impact on ecosystems is also daunting: insecticides play a role in the decline in global insect populations[4]. Regulations were taken in Europe to ban pesticides deemed too poisonous for the environment, including some neonics. Yet, bans have been challenged—and in somes cases reversed, as for the sugar beet culture in France in 2020. In a time when agricultural yields should improve by 50% to feed the world population by 2050, the limited understanding of the molecular mechanism of operation of insecticides hampers the development of less toxic, more specific compounds.





Left: illustration of the decline of insects (National Geographic cover of the work of Hallmann et al.). Right cumulative increase in the number of: species resistant to insecticides (blue) insecticides (purple) or GMO traits (red) for which resistance has been reported [5].

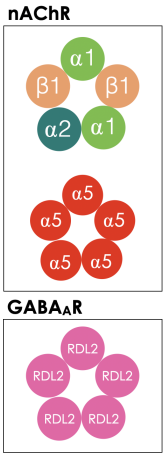


Few molecular targets correspond to a large share of insecticides

Clinical drugs targeting human ion channels represent only a low 10% of all drugs (and much less in value share). In stark contrast, insect ion channels, which are key players of the central nervous system and of muscle excitability, are the main molecular targets of neurotoxic insecticides, reaching up to ~70% of insecticide market value[1]. Moreover, only a restricted number of channels are targets. The worldwide insecticide sales market is currently dominated by products that target just five ion channels (green and blue on the diagram from[3]), two of them being at the core of the proposed research: the nicotinic acetylcholine receptors (nAChRs) and the type A ɣ-aminobutyric acid receptors ($GABA_A$Rs). Compounds targeting nAChRs are largely dominated by neonics, which are competitive modulators (market share of ~25% out of the 29% share for all compounds targeting nAChRs, which corresponds to ~5 billion USD in 2019), but also comprise newer fluorinated antagonists, allosteric modulators, and even the first commercial peptidic insecticide. $GABA_A$Rs also possess a rich pharmacology of older and newer insecticides involving a distinct molecular mechanism[6]: fipronil (banned for agricultural use, albeit authorized for veterinary use) is a non-competitive antagonist, whereas fluralaner and broflanilide (the latter authorized in Australia and Canada in 2020) are two allosteric modulators targeting different sites[1,3,6]). While resistance to older insecticides targeting $GABA_A$Rs has been linked to mutations in residues of the pore, resistance to neonicotinoids is more commonly due to overexpression of metabolic enzymes than to nAChRs mutants or variants.

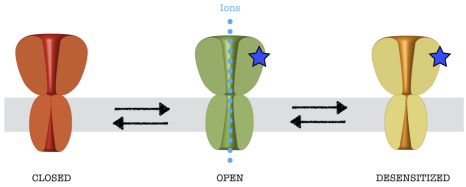
$GABA_A$Rs and nAChRs represent only a small sample of insect synaptic channels (e.g., the bee genome contains 103 genes encoding ion channel subunits) that also include channels known to be either targets for medicine in humans (e.g., CaV channels) or targets of venoms. However, those ion channels have been little studied in insects, and their genetic or pharmacological profiles have not been

thoroughly characterized (except in Drosophila). **We, therefore, focus the proposed research on $GABA_A$Rs and nAChRs both because they are important targets of insecticides, and because their genes and subunits are sufficiently characterized.**



Pentameric ligand-gated ion channels

GABA and acetylcholine are, respectively, the major inhibitory and excitatory neurotransmitters in the invertebrate nervous systems. Together they influence locomotion, olfactory learning, control of sleep, aggression and memory formation. $GABA_A$ and nACh receptors belong to the **pentameric ligand-gated ion channel family (pLGIC)**. The main RDL subunit of $GABA_A$Rs can form **homopentamers** and is generally encoded by one or a few genes (1 in the bee, 4 in the *Varroa destructor*, see example in the figure). nAChRs are most often **heteropentamers** containing 3 or more different subunits (there are 10 subunit genes in the Drosophila, 11 in the bee). pLGICs function as allosteric signal transducers[7] across the plasma membrane: upon binding of neurotransmitter molecules, the receptors undergo complex conformational transitions, which result in transient opening of an intrinsic transmembrane channel[8]. In mammals, the mechanistic details of these processes have been investigated over decades of pharmacology, electrophysiology[9], mutagenesis and recently in multi-state structural studies[10,11] (including work of P1 and P3). By contrast, no structure of any insect receptor is available, and insect nAChRs carried the reputation of being impossible to express only a year ago. Not even the subunit composition of native nAChR pentamers is known!

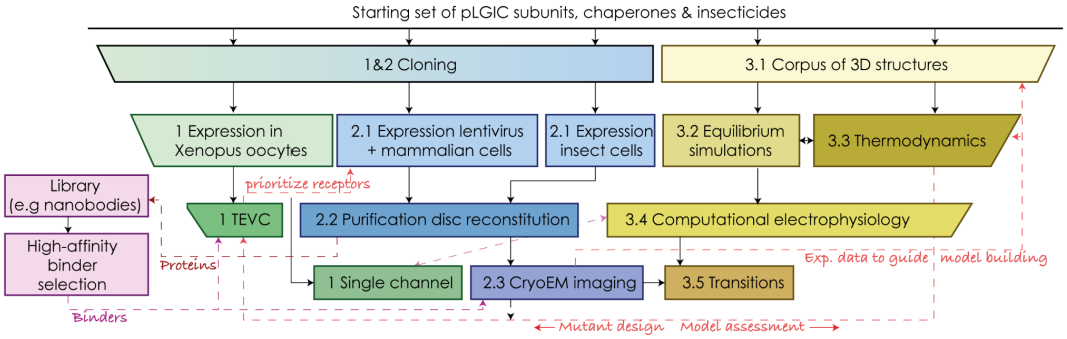


The Pesti-Penta project aims to understand, in molecular terms, how insect $GABA_A$ and nACh receptors (two major insecticide targets) operate and how they respond to insecticides. To establish this understanding, Pesti-Penta will combine cryoelectron microscopy (cryoEM) , molecular dynamics (MD) and electrophysiology to:

- Functionally characterize representative insecticides using electrophysiology, at the whole-cell and single-channel level, and including mutants in the binding pockets;
- Express, purify and image by cryoEM a few receptors in complex with representative insecticides;
- Characterize the dynamics of insecticide recognition using MD simulations and decipher the allosteric transition underlying receptor activation.

In order to understand how insect $GABA_A$ and nACh receptors operate at the molecular level, **we propose a work program consisting of four interdependent tasks.** These tasks will be undertaken for **several target receptors, with various levels of gain/risk ratios**, to maximise the chances of success. For a given receptor, we will:

- **Characterize the functional activity of insecticides**
- **Express, purify and image it by cryoEM in different conformations**
- **Model ligand binding dynamics and the allosteric transition underlying receptor activation**
- **Perform proof-of-concept screening of biologics on the purified receptor**



Task 3 - Modeling and Simulations

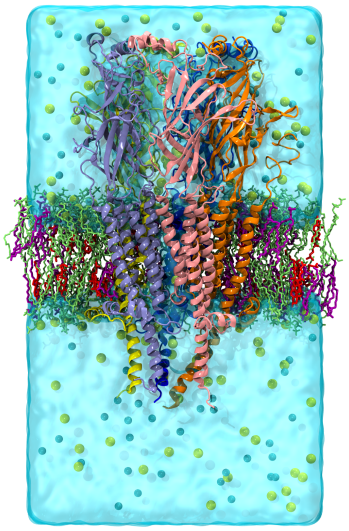
3.1 Corpus of 3D structures of GABAARs and nAChRs.

The main thrust of this work package is to establish a corpus of three-dimensional structures of both $GABA_A$Rs and nAChRs for the targets detailed in section *c.Methodology and risk management*. Our goal is to build for each receptor a set of structures for both beneficial and pest insects (*e.g., apis melifera*, *varroa destructor*, and *bumblebee*) representative of different conformational states along the activation pathways bound or not to insecticides (*e.g.,* Dielhdrin, Fipronil, Fluralaner, Broflanilide for $GABA_A$Rs and Imidaclopride, ω/κ-hexatoxin-Hv1a for nAChRs). We will resort to a hybrid strategy that, on the one hand, will incorporate to our corpus any 3D structure produced in the course of the project, and on the other hand, will rely on model structures built by homology for targets identified as good candidates for structure-dynamics-function studies in Tasks 1 and 2. It is clear that resolving 3D structures of insect receptors will bring our ensemble of structures to a level of precision way above modeling studies published so far, allowing us to study the allosteric transition underlying the activation of these channels at an unprecedented level of detail. The risk of lacking specific structural data for insect receptors is, however, counterbalanced by all known 3D structures of homologous receptors (GluCl, GlyR, human $GABA_A$R with sequence identity ~40-50 %), whose number has grown significantly over the past 2-3 years.

In practice, we will resort to a multiple-target homology approach as implemented in the MODELER package. We will first consider $GABA_A$ RDL receptors from beneficial and pest species from which P2 has preliminary functional data[11,12] and pursue with nAChR and for $GABA_A$R that will exhibit good expression, pentameric state after purification, reproducible currents, etc... along Tasks 1 and 2. The series of insecticides considered in our project will be systematically docked (through a flexible docking procedure as available in AutoDock) at the orthosteric binding sites (well-conserved among homologs) and at the allosteric binding sites (less-conserved and requiring probably structural information of Task 2.3). The ensemble of structures will evolve dynamically and incorporate any new structural data produced in the course of the project (improving homology models and/or replacing them whenever possible).

3.2 Multi-µs equilibrium MD simulations

The main objective pursued in this task is to provide at the molecular level an in-depth analysis of the dynamics of the different insect receptors considered herein in a physiologically meaningful environment. More specifically we will supplement the static structural corpus established in the previous task by means of a series of multi-µs all-atom MD simulations for all receptors embedded in a model lipid bilayer. All together, molecular-assays encompassing the pentameric channels, lipids, cholesterol, water, ions and neurotransmitters amount to ~240,00 atoms (see figure of a prototypical pentameric channel (here Torpedo fish nAChR in 3:1:1 POPC:POPA:Cholesterol bilayer with 150mM NaCl) for which we can produce up to 100 ns/day on one GPU (last generation) with the new version of NAMD[34]. It will allow establishing further the relevance of the corresponding structures in a proper environment at physiological temperature and pressure. MD trajectories of receptors in complex with inhibitors and allosteric modulators will provide a strong asset toward deciphering insecticide binding interfaces. Residues identified as crucial for the association will be systematically mutated and their effect on the receptor function characterized by means of electrophysiology experiments in Task 1. Such a synergetic theory-experimental work is anticipated to



Smoluchowski equation[43]. We will parametrize this equation by means of a potential of mean force (PMF) calculation, supplemented by the determination of the diffusivity and the boundary ion density value at a given voltage, which is required to evaluate the position-dependent current of ions in the steady state [44] . Computation of the PMF will benefit from recent developments in our group, whereby the adaptive biasing force (ABF) algorithm is combined with well-tempered metadynamics to map with utmost efficiency the rugged free-energy landscape that underlies ion permeation through the membrane protein. The resulting importance-sampling scheme, coined WT-meta-eABF[45], can be utilized in conjunction with a multiple-walker strategy to enhance ergodic sampling (MW-WT-meta-eABF)[46].

3.5 Conformational transition associated with agonist binding

The overarching goal of the theoretical part of this project is to provide a full atomic-picture of the allosteric transition underlying the activation of insect pentameric receptors upon binding of neurotransmitters. Such a strategy has been applied successfully to reveal the minimum free-energy gating pathways of the proton-activated bacterial GLIC channel, together with a number of intermediate states that could serve as a basis for identifying new drug prockets[47]. Armed with at least two structures of a given target receptor in both its resting/closed and active/open states, ideally resolved in the course of the project or built by homology and properly validated in previous tasks, we propose to model the reaction coordinate (RC) underlying the complex transitions that connect the open and closed conformations by means of a path-finding approach coined the string method with swarms of trajectories (SMwST) [48], which refines a trial transition pathway in multi-dimensional space coarse-grained to a set of collective variables (CVs) on the basis of their average dynamic drift. Machine-learning (ML) techniques will be utilized to efficiently explore the ensemble of high-dimensional functions of the positions of the atoms and find the best coarse-graining map. Among the available unsupervised ML schemes to achieve this task, auto-encoders are one of the most promising avenues[49]. They use a neural network, which, from the input layer, encodes through successive layers with fewer and fewer neurons up to a bottleneck with only a few neurons, and prior to decoding through successive layers with more and more neurons up to the last layer, which has exactly the same dimension as the input layer. Our objective is to approximate the identity map using the large amount of data generated by an MD trajectory of the conformational states of the protein. Finally, the free energy underlying the conformational transition will be determined along the path variable using the WT-meta-eABF approach [45]. This task, which is clearly high-risk, could possibly reveal new intermediate conformations along the activation pathways, providing a basis for identifying new druggable pockets for insect pentameric receptors.

III. Compaction et reconnaissance de l'ADN lésé

Maintenir l'intégrité du génome des organismes vivants est fondamental pour assurer le bon fonctionnement des cellules. Cependant, l'ADN est constamment exposé à des menaces externes, telles que la lumière UV, les rayonnements ionisants et le stress oxydatif, qui peuvent générer une variété de lésions complexes d'ADN, y compris la dimérisation de bases (figure 33), l'oxydation ou les ruptures de brin.[136] L'accumulation de lésions d'ADN dans les cellules peut in fine être responsable de la mort cellulaire et du développement de cancers ou encore de maladies associées au vieillissement.[44] Les cellules ont développé des systèmes de réparation spécifiques qui reconnaissent la présence d'une lésion dans un grand bassin de nucléobases intactes et ensuite catalysent le processus de réparation. Il existe divers mécanismes de réparation sophistiqués, [137]comme la voie de réparation de l'excision nucléotidique (NER), adaptée à l'élimination des lésions complexes et volumineuses. Comparé aux mécanismes de réparation de l'excision de base (BER), le NER qui implique des complexes multiprotéiques est beaucoup moins caractérisé. Chez les procaryotes, le NER repose sur les protéines Uvr.[138] UvrA (figure 34 d), agissant sous forme de dimère, avec UvrB, est responsable de la reconnaissance des dommages à l'ADN. Après la reconnaissance des dommages, UvrA se dissocie de l'ADN et UvrB forme un complexe pré-

incision stable avec l'ADN endommagé et recrute UvrC pour inciser l'ADN de chaque côté de la lésion. La spécificité et l'efficacité de la reconnaissance moléculaire est un processus très complexe, découlant de la formation de complexes macromoléculaires dynamiques aux sites de dommages à l'ADN. Les mécanismes moléculaires détaillés sous-jacents à la reconnaissance des dommages par UvrA et la nature et les caractéristiques communes des lésions d'ADN qui sont traitées par NER sont encore très mal compris.

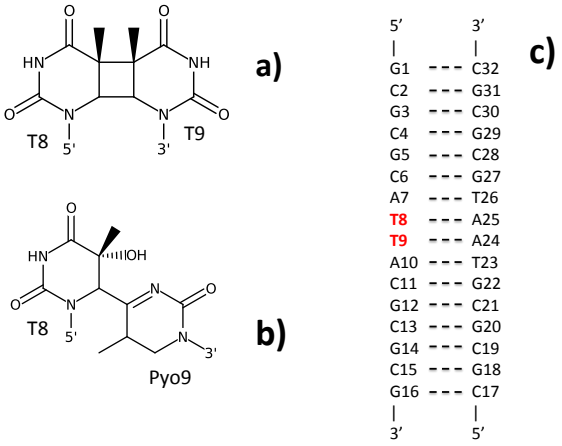


Figure 33. Formules moléculaires des deux photo-lésions CPD (a) et 634-PP (b). Sur le double brin (c) composé de 16 paires de bases, les deux thymines en position 8 et 9 sont celles qui dimérisent pour former la photo-lésion.

La présence de lésions clusters (plusieurs lésions qui se distribue consécutivement sur la double hélice d'ADN, figure 34 a) peut induire des effets coopératifs forts qui peuvent conduire à une structure et une évolution dynamique totalement différentes de qui est observé pour les lésions isolées, ce qui aura un impact sur l'efficacité de reconnaissance et de réparation. Par ailleurs dans la cellule, l'ADN n'est pas étendu mais plutôt condensé en structures compactes appelées chromatine[139] au moyen de protéines dédiées, les histones[140] (figure 34 b) chez les eucaryotes ou les protéines semblables à des histones telles que HU[141] (figure 34 c) chez les procaryotes. Il a été suggéré récemment que la réponse de l'ADN condensé à des lésions peut être, en grande partie, différente de ce qui a été appris jusqu'ici par des études in vitro d'ADN isolé. Par exemple, la présence d'histones peut faciliter la formation de ADN-protéine, ce qui peut à son tour affecter l'architecture nucléosomale. Comment l'ADN compacté se comporte sous l'effet de lésions reste une question largement inexplorée bien que fondamentale pour comprendre les mécanismes de réparation (et leurs éventuels) défauts in cellulo.

Je souhaite maintenant continuer le travail entrepris pour mieux comprendre les conséquences dynamique et structurale des lésions de l'ADN. A l'instar de ce qui a été fait pour les lésions clusters de sites abasiques, je m'intéresserais aux photo-lésions comme les dimères de thymines. J'étendrais les études de lésions simples et complexes réalisées sur des doubles brins étendues à des modèles ou ces lésions prennent place sur de l'ADN compacté en complexes avec des protéines histone ou histone-like (figure 34 a,b). Pour compléter la compréhension des mécanismes de réparation je modéliserais l'interaction de brins d'ADN lésés dans différentes conformations avec la protéine de reconnaissance du NER UvrA. L'objectif global de mon travail est de pouvoir, dans un premier, décrire précisément les effets d'un ensemble de lésions simples et complexes sur la structuration de l'ADN dans un contexte cellulaire, puis, dans un second temps, rationaliser les mécanismes de

reconnaissances de ces lésions par les protéines de réparation de la machinerie cellulaire. J'utiliserais ici les simulations de dynamique moléculaire non biaisés ou faisant appel à des techniques d'échantillonnages avancées (parallel tempering, Adaptive Biasing Force, …) ainsi qu'aux calculs d'énergie libre. J'envisage de faire appel à des champs de force atomique mais également à des champs de force de type gros-grains.

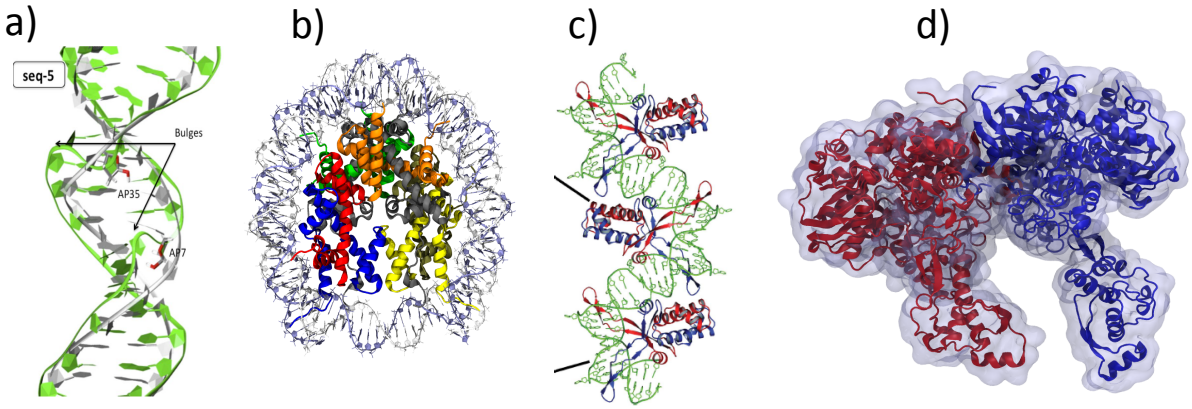


Figure 34. Structures tridimensionnelles de (a) un double brin d'ADN déformée par une lésion cluster formées de deux sites abasiques, (b) un nucléosome, entité de base de la chromatine, (c) une protéine bactérienne de type histone (protéine HU) liée à un double brin d'ADN et (d) un dimère d'UvrA, une protéine de la machinerie de réparation cellulaire du NER de Deinoccocus Radiodurans.

Références

1. Frenkel, D. & Smit, B. *Understanding Molecular Simulation*: F*rom algorithms to applications* (Elsevier, 2002).

2. Rupakheti, C. R., Roux, B., Dehez, F. & Chipot, C. Modeling induction phenomena in amino acid cation– $$\pi $$ π interactions. *Theor. Chem. Acc.* **137**, 174 (2018).

3. Lee, H., Dehez, F., Chipot, C., Lim, H.-K. & Kim, H. Enthalpy–Entropy Interplay in π-Stacking Interaction of Benzene Dimer in Water. *J. Chem. Theory Comput.* **15**, 1538–1545 (2019).

4. Le Sueur, C. R. & Stone, A. J. Practical schemes for distributed polarizabilities. *Mol. Phys.* **78**, 1267–1291 (1993).

5. Ponder, J. W. *et al.* Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **114**, 2549–2564 (2010).

6. Rappe, A. K. & Goddard, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95**, 3358–3363 (1991).

7. Rick, S. W., Stuart, S. J. & Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **101**, 6141–6156 (1994).

8. Lemkul, J. A., Huang, J., Roux, B. & MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **116**, 4983–5013 (2016).

9. Dehez, F., Soetens, J.-C., Chipot, C., Ángyán, J. G. & Millot, C. Determination of Distributed Polarizabilities from a Statistical Analysis of Induction Energies. *J. Phys. Chem. A* **104**, 1293–1303 (2000).

10. Celebi, N., Ángyán, J. G., Dehez, F., Millot, C. & Chipot, C. Distributed polarizabilities derived from induction energies: A finite perturbation approach. *J. Chem. Phys.* **112**, 2709–2717 (2000).

11. Chipot, C. *et al.* Alternative Approaches for the Calculation of Induction Energies: Characterization, Effectiveness, and Pitfalls. *J. Phys. Chem. A* **105**, 11505–11514 (2001).

12. Dehez, F., Chipot, C., Millot, C. & Ángyán, J. G. Fast and accurate determination of induction energies: reduction of topologically distributed polarizability models. *Chem. Phys. Lett.* **338**, 180–188 (2001).

13. Ángyán, J. G. *et al.* OPEP: A tool for the optimal partitioning of electric properties. *J. Comput. Chem.* **24**, 997–1008 (2003).

14. Soteras, I. *et al.* Derivation of Distributed Models of Atomic Polarizability for Molecular Simulations. *J. Chem. Theory Comput.* **3**, 1901–1913 (2007).

15. Dehez, F. *et al.* Modeling Induction Phenomena in Intermolecular Interactions with an Ab Initio Force Field. *J. Chem. Theory Comput.* **3**, 1914–1926 (2007).

16. Misquitta, A. J., Jeziorski, B. & Szalewicz, K. Dispersion Energy from Density-Functional Theory Description of Monomers. *Phys. Rev. Lett.* **91**, 033201 (2003).

17. Dehez, F., Archambault, F., Gutiérrez, I. S., Luque, F. J. & Chipot, C. An ab initio strategy for handling induction phenomena in metal ion complexes. *Mol. Phys.* **106**, 1685–1696 (2008).

18. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

19. Archambault, F. *et al.* Polarizable Intermolecular Potentials for Water and Benzene Interacting with Halide and Metal Ions. *J. Chem. Theory Comput.* **5**, 3022–3031 (2009).

20. Tang, K. T. & Toennies, J. P. New combining rules for well parameters and shapes of the van der Waals potential of mixed rare gas systems. *Z. F□ r Phys. At. Mol. Clust.* **1**, 91–101 (1986).

21. Luque, F. J., Dehez, F., Chipot, C. & Orozco, M. Polarization effects in molecular interactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 844–854 (2011).

22. Pohorille, A. & Chipot, C. Summary and Outlook. in *Free Energy Calculations: Theory and Applications in Chemistry and Biology* (eds. Chipot, C. & Pohorille, A.) 503–513 (Springer, 2007). doi:10.1007/978-3-540-38448-9_14.

23. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **3**, 198–210 (2013).

24. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

25. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).

26. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461 (2015).

27. Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **22**, 1420–1426 (1954).

28. Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23**, 187–199 (1977).

29. Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **78**, 2690–2693 (1997).

30. Andricioaei, I. Specialized Methods for Improving Ergodic Sampling Using Molecular Dynamics and Monte Carlo Simulations. in *Free Energy Calculations: Theory and Applications in Chemistry and Biology* (eds. Chipot, C. & Pohorille, A.) 277–322 (Springer, 2007). doi:10.1007/978-3-540-38448-9_8.

31. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).

32. Darve, E. & Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **115**, 9169–9183 (2001).

33. Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J.*

Comput. Phys. **22**, 245–268 (1976).

34. Liu, P., Dehez, F., Cai, W. & Chipot, C. A Toolkit for the Analysis of Free-Energy Perturbation Calculations. *J. Chem. Theory Comput.* **8**, 2606–2616 (2012).

35. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

36. Massova, I. & Kollman, P. A. Computational Alanine Scanning To Probe Protein−Protein Interactions: A Novel Approach To Evaluate Binding Free Energies. *J. Am. Chem. Soc.* **121**, 8133–8143 (1999).

37. Morrison, K. L. & Weiss, G. A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **5**, 302–307 (2001).

38. Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci.* **105**, 877–882 (2008).

39. Ramadoss, V., Dehez, F. & Chipot, C. AlaScan: A Graphical User Interface for Alanine Scanning Free-Energy Calculations. *J. Chem. Inf. Model.* **56**, 1122–1126 (2016).

40. Chebil, L. *et al.* Solubility of Flavonoids in Organic Solvents. *J. Chem. Eng. Data* **52**, 1552–1556 (2007).

41. Chebil, L. *et al.* Solubilities Inferred from the Combination of Experiment and Simulation. Case Study of Quercetin in a Variety of Solvents. *J. Phys. Chem. B* **114**, 12308–12313 (2010).

42. Bjelobrk, Z., Mendels, D., Karmakar, T., Parrinello, M. & Mazzotti, M. Solubility prediction of organic molecules with molecular dynamics simulations. *ArXiv210410792 Cond-Mat* (2021).

43. Epe, B., Pflaum, M. & Boiteux, S. DNA damage induced by photosensitizers in cellular and cell-free systems. *Mutat. Res. Toxicol.* **299**, 135–145 (1993).

44. Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N. Engl. J. Med.* **361**, 1475–1485 (2009).

45. Dougherty, T. J. *et al.* Photodynamic Therapy. *JNCI J. Natl. Cancer Inst.* **90**, 889–905 (1998).

46. Gattuso, H., Dumont, E., Chipot, C., Monari, A. & Dehez, F. Thermodynamics of DNA: sensitizer recognition. Characterizing binding motifs with all-atom simulations. *Phys. Chem. Chem. Phys.* **18**, 33180–33186 (2016).

47. Palmieri, F. Mitochondrial carrier proteins. *FEBS Lett.* **346**, 48–54 (1994).

48. Fiore, C. *et al.* The mitochondrial ADP/ATP carrier: Structural, physiological and pathological aspects. *Biochimie* **80**, 137–150 (1998).

49. Pebay-Peyroula, E. *et al.* Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426**, 39–44 (2003).

50. Nury, H. *et al.* Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Lett.* **579**, 6031–6036 (2005).

51. Ruprecht, J. J. *et al.* Structures of yeast mitochondrial ADP/ATP carriers support a domain-based alternating-access transport mechanism. *Proc. Natl. Acad. Sci.* **111**, E426–E434 (2014).

52. Ruprecht, J. J. *et al.* The Molecular Mechanism of Transport by the Mitochondrial ADP/ATP Carrier. *Cell* **176**, 435-447.e15 (2019).

53. Dehez, F., Pebay-Peyroula, E. & Chipot, C. Binding of ADP in the Mitochondrial ADP/ATP Carrier Is Driven by an Electrostatic Funnel. *J. Am. Chem. Soc.* **130**, 12725–12733 (2008).

54. Wang, Y. & Tajkhorshid, E. Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc. Natl. Acad. Sci.* **105**, 9598–9603 (2008).

55. Krämer, R. Influence of divalent cations on the reconstituted ADP, ATP exchange. *Biochim. Biophys. Acta* **592**, 615–620 (1980).

56. Krammer, E.-M. *et al.* High-Chloride Concentrations Abolish the Binding of Adenine Nucleotides in the Mitochondrial ADP/ATP Carrier Family. *Biophys. J.* **97**, L25–L27 (2009).

57. Dahout-Gonzalez, C. *et al.* Molecular, Functional, and Pathological Aspects of the Mitochondrial ADP/ATP Carrier. *Physiology* **21**, 242–249 (2006).

58. Kunji, E. R. S., King, M. S., Ruprecht, J. J. & Thangaratnarajah, C. The SLC25 Carrier Family: Important Transport Proteins in Mitochondrial Physiology and Pathology. *Physiology* **35**, 302–327 (2020).

59. Ravaud, S. *et al.* Impaired Transport of Nucleotides in a Mitochondrial Carrier Explains Severe Human Genetic Diseases. *ACS Chem. Biol.* **7**, 1164–1169 (2012).

60. Klingenberg, M. The ADP and ATP transport in mitochondria and its carrier. *Biochim. Biophys. Acta BBA - Biomembr.* **1778**, 1978–2021 (2008).

61. Mifsud, J. *et al.* The substrate specificity of the human ADP/ATP carrier AAC1. *Mol. Membr. Biol.* **30**, 160–168 (2013).

62. Bidon-Chanal, A. *et al.* How Do Membrane Transporters Sense pH? The Case of the Mitochondrial ADP–ATP Carrier. *J. Phys. Chem. Lett.* **4**, 3787–3791 (2013).

63. Chipot, C. *et al.* Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies. *Chem. Rev.* **118**, 3559–3607 (2018).

64. Berardi, M. J., Shih, W. M., Harrison, S. C. & Chou, J. J. Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* **476**, 109–113 (2011).

65. Zoonens, M. *et al.* Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2. *J. Am. Chem. Soc.* **135**, 15174–15182 (2013).

66. ACS Contributing Correspondents. Spotlights on Recent JACS Publications. *J. Am. Chem. Soc.* **135**, 14895–14895 (2013).

67. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2726990/.

68. Brüschweiler, S., Yang, Q., Run, C. & Chou, J. J. Substrate-modulated ADP/ATP-transporter dynamics revealed by NMR relaxation dispersion. *Nat. Struct. Mol. Biol.* **22**, 636–641 (2015).

69. Kurauskas, V. *et al.* How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine. *J. Phys. Chem. Lett.* **9**, 933–938 (2018).

70. Kurauskas, V. *et al.* Dynamics and interactions of AAC3 in DPC are not functionally relevant. *Nat. Struct. Mol. Biol.* **25**, 745–747 (2018).

71. Zhao, L., Wang, S., Run, C., OuYang, B. & Chou, J. J. Specific Lipid Binding of Membrane Proteins in Detergent Micelles Characterized by NMR and Molecular Dynamics. *Biochemistry* **55**, 5317–5320 (2016).

72. Steinmann, E. *et al.* Hepatitis C Virus p7 Protein Is Crucial for Assembly and Release of Infectious Virions. *PLOS Pathog.* **3**, e103 (2007).

73. Montserret, R. *et al.* NMR Structure and Ion Channel Activity of the p7 Protein from Hepatitis C Virus *. *J. Biol. Chem.* **285**, 31446–31461 (2010).

74. OuYang, B. *et al.* Unusual architecture of the p7 channel from hepatitis C virus. *Nature* **498**, 521–525 (2013).

75. Holzmann, N., Chipot, C., Penin, F. & Dehez, F. Assessing the physiological relevance of alternate architectures of the p7 protein of hepatitis C virus in different environments. *Bioorg. Med. Chem.* **24**, 4920–4927 (2016).

76. Oestringer, B. P. *et al.* Re-evaluating the p7 viroporin structure. *Nature* **562**, E8–E18 (2018).

77. Oestringer, B. P. *et al.* Hepatitis C virus sequence divergence preserves p7 viroporin structural and dynamic features. *Sci. Rep.* **9**, 8383 (2019).

78. Barnes, N. M., Hales, T. G., Lummis, S. C. R. & Peters, J. A. The 5-HT3 receptor – the relationship between structure and function. *Neuropharmacology* **56**, 273–284 (2009).

79. Unwin, N. Neurotransmitter action: Opening of ligand-gated ion channels. *Cell* **72**, 31–41 (1993).

80. Keramidas, A., Moorhouse, A. J., Schofield, P. R. & Barry, P. H. Ligand-gated ion channels: mechanisms underlying ion selectivity. *Prog. Biophys. Mol. Biol.* **86**, 161–204 (2004).

81. Nemecz, Á., Prevost, M. S., Menny, A. & Corringer, P.-J. Emerging Molecular Mechanisms of Signal Transduction in Pentameric Ligand-Gated Ion Channels. *Neuron* **90**, 452–470 (2016).

82. Thompson, A. J. & R. Lummis, S. C. 5-HT3 Receptors. *Curr. Pharm. Des.* **12**, 3615–3630 (2006).

83. Hannon, J. & Hoyer, D. Molecular biology of 5-HT receptors. *Behav. Brain Res.* **195**, 198–213 (2008).

84. Polovinkin, L. *et al.* Conformational transitions of the serotonin 5-HT 3 receptor. *Nature* **563**, 275–279 (2018).

85. Li, C. *et al.* Recent progress in drug delivery. *Acta Pharm. Sin. B* **9**, 1145–1162 (2019).

86. Jones, S. W. *et al.* Characterisation of cell-penetrating peptide-mediated peptide delivery. *Br. J. Pharmacol.* **145**, 1093–1102 (2005).

87. Bonhenry, D., Tarek, M. & Dehez, F. Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane. *J. Chem. Theory Comput.*

9, 5675–5684 (2013).

88. Bonhenry, D., Dehez, F. & Tarek, M. Effects of hydration on the protonation state of a lysine analog crossing a phospholipid bilayer – insights from molecular dynamics and free-energy calculations. *Phys. Chem. Chem. Phys.* **20**, 9101–9107 (2018).

89. Tieleman, D. P. The molecular basis of electroporation. *BMC Biochem.* **5**, 10 (2004).

90. Tarek, M. Membrane Electroporation: A Molecular Dynamics Simulation. *Biophys. J.* **88**, 4045–4053 (2005).

91. Gumbart, J., Khalili-Araghi, F., Sotomayor, M. & Roux, B. Constant electric field simulations of the membrane potential illustrated with simple systems. *Biochim. Biophys. Acta BBA - Biomembr.* **1818**, 294–302 (2012).

92. Gurtovenko, A. A. & Vattulainen, I. Pore Formation Coupled to Ion Transport through Lipid Membranes as Induced by Transmembrane Ionic Charge Imbalance: Atomistic Molecular Dynamics Study. *J. Am. Chem. Soc.* **127**, 17570–17571 (2005).

93. Delemotte, L., Dehez, F., Treptow, W. & Tarek, M. Modeling Membranes under a Transmembrane Potential. *J. Phys. Chem. B* **112**, 5547–5550 (2008).

94. Davalos, R. V., Mir, L. M. & Rubinsky, B. Tissue Ablation with Irreversible Electroporation. *Ann. Biomed. Eng.* **33**, 223 (2005).

95. Teissie, J., Golzio, M. & Rols, M. P. Mechanisms of cell membrane electropermeabilization: A minireview of our present (lack of ?) knowledge. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1724**, 270–280 (2005).

96. Dehez, F., Delemotte, L., Kramar, P., Miklavčič, D. & Tarek, M. Evidence of Conducting Hydrophobic Nanopores Across Membranes in Response to an Electric Field. *J. Phys. Chem. C* **118**, 6752–6757 (2014).

97. Chatterjee, N. & Walker, G. C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **58**, 235–263 (2017).

98. Li, X. & Heyer, W.-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* **18**, 99–113 (2008).

99. Spies, M. & Kowalczykowski, S. C. Homologous Recombination by the RecBCD and RecF Pathways. in *The Bacterial Chromosome* 389–403 (John Wiley & Sons, Ltd, 2004). doi:10.1128/9781555817640.ch21.

100. Sakai, A. & Cox, M. M. RecFOR and RecOR as Distinct RecA Loading Pathways *. *J. Biol. Chem.* **284**, 3264–3272 (2009).

101. Radzimanowski, J. *et al.* An 'open' structure of the RecOR complex supports ssDNA binding within the core of the complex. *Nucleic Acids Res.* **41**, 7972–7986 (2013).

102. Georgakilas, A. G., O'Neill, P. & Stewart, R. D. Induction and Repair of Clustered DNA Lesions: What Do We Know So Far? *Radiat. Res.* **180**, 100–109 (2013).

103. Georgakilas, A. G., Bennett, P. V. & Sutherland, B. M. High efficiency detection of bi-stranded abasic clusters in γ-irradiated DNA by putrescine. *Nucleic Acids Res.* **30**, 2800–2808 (2002).

104. Georgakilas, A. G., Bennett, P. V., Wilson, D. M., III & Sutherland, B. M. Processing of bistranded abasic DNA clusters in γ-irradiated human hematopoietic cells. *Nucleic Acids Res.* **32**, 5609–5620 (2004).

105. Gattuso, H. *et al.* Repair Rate of Clustered Abasic DNA Lesions by Human Endonuclease: Molecular Bases of Sequence Specificity. *J. Phys. Chem. Lett.* **7**, 3760–3765 (2016).

106. Schreier, W. J., Gilch, P. & Zinth, W. Early events of DNA photodamage. *Annu. Rev. Phys. Chem.* **66**, 497–519 (2015).

107. Lo, H.-L. *et al.* Differential biologic effects of CPD and 6-4PP UV-induced DNA damage on the induction of apoptosis and cell-cycle arrest. *BMC Cancer* **5**, 135 (2005).

108. Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151 (1999).

109. Dehez, F. *et al.* Conformational polymorphism or structural invariance in DNA photoinduced lesions: implications for repair rates. *Nucleic Acids Res.* **45**, 3654–3662 (2017).

110. Mizukoshi, T. *et al.* Structural study of DNA duplexes containing the (6–4) photoproduct by fluorescence resonance energy transfer. *Nucleic Acids Res.* **29**, 4948–4954 (2001).

111. Yu, A. Y. H. & Houry, W. A. ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett.* **581**, 3749–3757 (2007).

112. Alexopoulos, J. A., Guarné, A. & Ortega, J. ClpP: A structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* **179**, 202–210 (2012).

113. Kirstein, J. *et al.* The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol. Med.* **1**, 37–49 (2009).

114. Leung, E. *et al.* Activators of Cylindrical Proteases as Antimicrobials: Identification and Development of Small Molecule Activators of ClpP Protease. *Chem. Biol.* **18**, 1167–1178 (2011).

115. Moreno-Cinos, C. *et al.* ClpP Protease, a Promising Antimicrobial Target. *Int. J. Mol. Sci.* **20**, 2232 (2019).

116. Wang, J., Hartling, J. A. & Flanagan, J. M. The Structure of ClpP at 2.3 Å Resolution Suggests a Model for ATP-Dependent Proteolysis. *Cell* **91**, 447–456 (1997).

117. Lee, B.-G. *et al.* Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat. Struct. Mol. Biol.* **17**, 471–478 (2010).

118. Moreira, W. *et al.* Target Mechanism-Based Whole-Cell Screening Identifies Bortezomib as an Inhibitor of Caseinolytic Protease in Mycobacteria. *mBio* **6**, e00253-15.

119. Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S. A. & Raunser, S. Cryo-EM structure of the ClpXP protein degradation machinery. *Nat. Struct. Mol. Biol.* **26**, 946–954 (2019).

120. Fei, X., Bell, T. A., Barkow, S. R., Baker, T. A. & Sauer, R. T. Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *eLife* **9**, e61496 (2020).

121. Peroutka, S. & Snyder, S. Antiemetics: Neurotransmitter Receptor Binding Predicts Therapeutic Actions. *The Lancet* **319**, 658–659 (1982).

122. Basch, E. *et al.* Antiemetics: American Society of Clinical Oncology Clinical Practice

Guideline Update. *J. Clin. Oncol.* **29**, 4189–4198 (2011).

123. Zarkadas, E. *et al.* The Binding of Palonosetron and Other Antiemetic Drugs to the Serotonin 5-HT3 Receptor. *Structure* **28**, 1131-1140.e4 (2020).

124. Colle, R. & Corruble, E. La vortioxétine : un nouvel antidépresseur pour traiter les épisodes dépressifs caractérisés. *L'Encéphale* **42**, 48–58 (2016).

125. Andersen, J. *et al.* Molecular determinants for selective recognition of antidepressants in the human serotonin and norepinephrine transporters. *Proc. Natl. Acad. Sci.* **108**, 12137–12142 (2011).

126. Dale, E. *et al.* The multimodal antidepressant vortioxetine may facilitate pyramidal cell firing by inhibition of 5-HT3 receptor expressing interneurons: An in vitro study in rat hippocampus slices. *Brain Res.* **1689**, 1–11 (2018).

127. Ladefoged, L. K. *et al.* Modeling and Mutational Analysis of the Binding Mode for the Multimodal Antidepressant Drug Vortioxetine to the Human 5-HT3A Receptor. *Mol. Pharmacol.* **94**, 1421–1434 (2018).

128. Masiulis, S. *et al.* GABAA receptor signalling mechanisms revealed by structural pharmacology. *Nature* **565**, 454–459 (2019).

129. Kumar, P. *et al.* Cryo-EM structures of a lipid-sensitive pentameric ligand-gated ion channel embedded in a phosphatidylcholine-only bilayer. *Proc. Natl. Acad. Sci.* **117**, 1788–1798 (2020).

130. Rovšnik, U. et al. Characterization of the dynamic resting state of a pentameric ligand-gated ion channel by cryo-electron microscopy and simulations. (2020) https://www.biorxiv.org/content/10.1101/2020.06.19.161356v1 doi:10.1101/2020.06.19.161356.

131. Walsh, R. M. *et al.* Structural principles of distinct assemblies of the human α4β2 nicotinic receptor. *Nature* **557**, 261–265 (2018).

132. Gharpure, A. *et al.* Agonist Selectivity and Ion Permeation in the α3β4 Ganglionic Nicotinic Receptor. *Neuron* **104**, 501-511.e6 (2019).

133. Noviello, C. M. *et al.* Structure and gating mechanism of the α7 nicotinic acetylcholine receptor. *Cell* **184**, 2121-2134.e13 (2021).

134. Rahman, Md. M. *et al.* Structure of the Native Muscle-type Nicotinic Receptor and Inhibition by Snake Venom Toxins. *Neuron* **106**, 952-962.e5 (2020).

135. Gielen, M. & Corringer, P.-J. The dual-gate model for pentameric ligand-gated ion channels activation and desensitization. *J. Physiol.* **596**, 1873–1902 (2018).

136. Friedberg, E. C. DNA damage and repair. *Nature* **421**, 436–440 (2003).

137. Wood, R. D. Dna Repair in Eukaryotes. *Annu. Rev. Biochem.* **65**, 135–167 (1996).

138. Rouillon, C. & White, M. F. The evolution and mechanisms of nucleotide excision repair proteins. *Res. Microbiol.* **162**, 19–26 (2011).

139. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).

140. Ramakrishnan, V. Histone Structure and the Organization of the Nucleosome. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 83–112 (1997).

141. Grove, A. Functional Evolution of Bacterial Histone-Like HU Proteins. *Curr. Issues Mol. Biol.* **13**, 1–12 (2011).

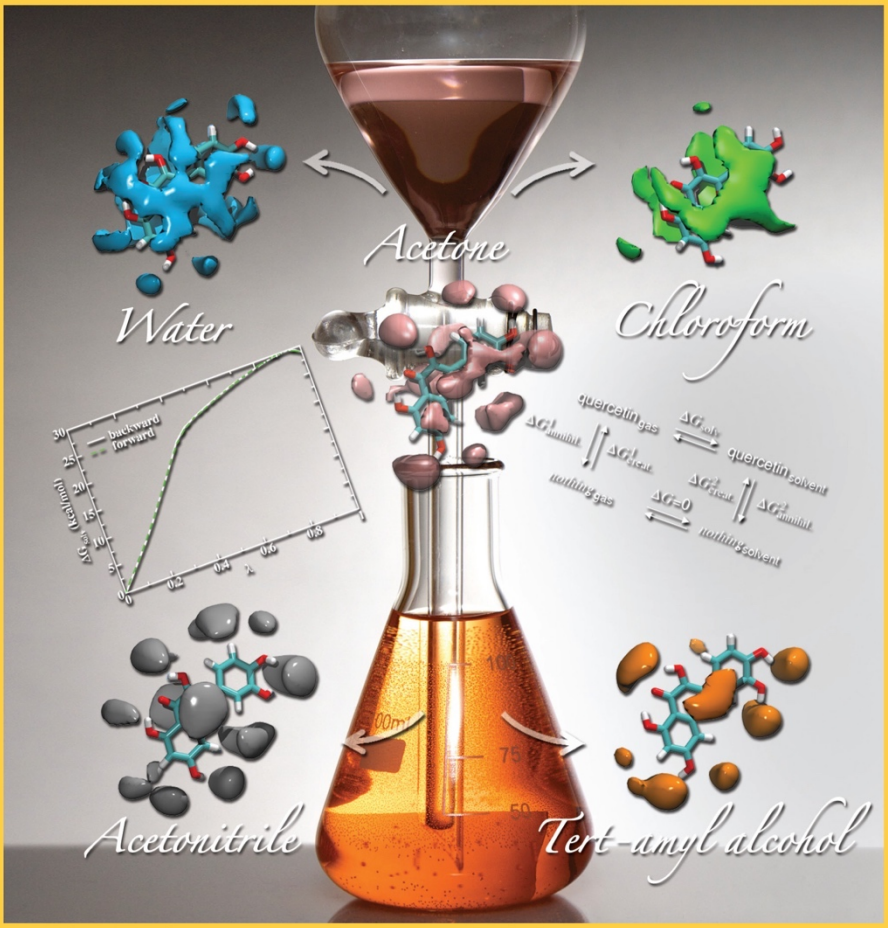
Covers


ACS
chemical
biology
JULY 2012
VOLUME 7, NUMBER 7
www.acschemicalbiology.org
The Community of Chemical Biologists
A New Way to Hit a Nerve
Assaying for Arginine Methylation Inhibitors
Tripling Up to Control Gene Transcription
ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.
www.acs.org

JPCBFK

SEPTEMBER 30, 2010
VOLUME 114
NUMBER 38
pubs.acs.org/JPCB

THE JOURNAL OF PHYSICAL CHEMISTRY B



Solubilities Inferred from the Combination of Experiment and Computed Free Energies of Solvation (see page 6A)

SOFT CONDENSED MATTER AND BIOPHYSICAL CHEMISTRY

ACS Publications
High quality. High impact.

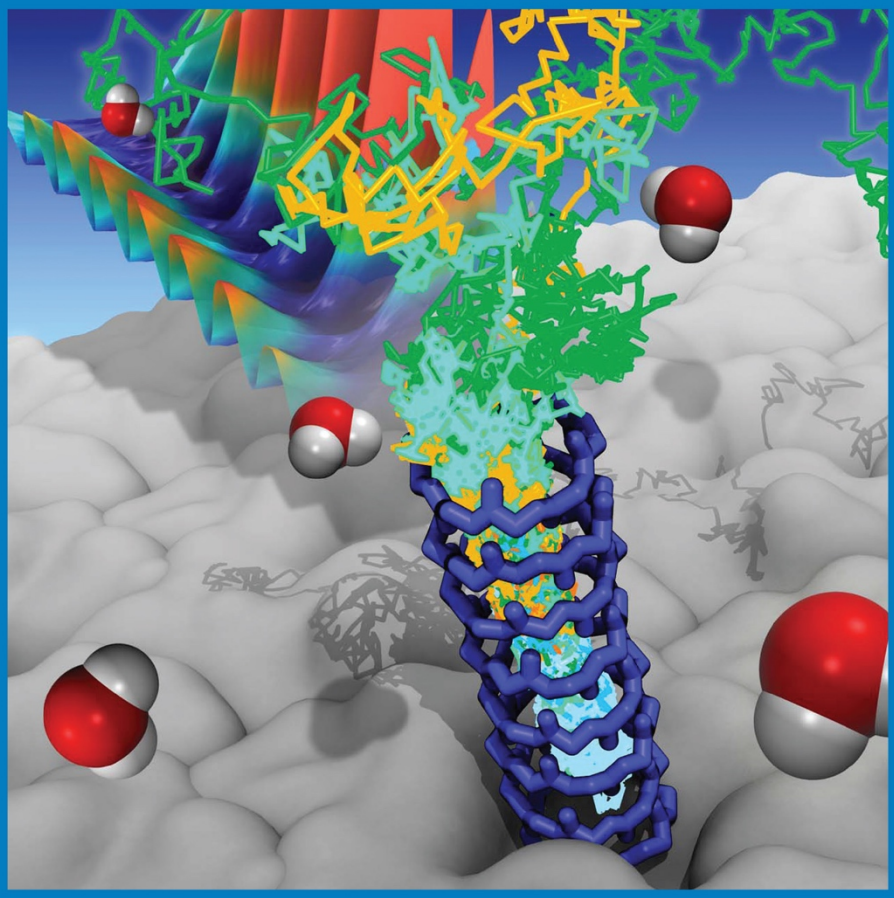
www.acs.org

JPCCCK

DECEMBER 19, 2013
VOLUME 117
NUMBER 50

pubs.acs.org/JPCC

THE JOURNAL OF PHYSICAL CHEMISTRY C

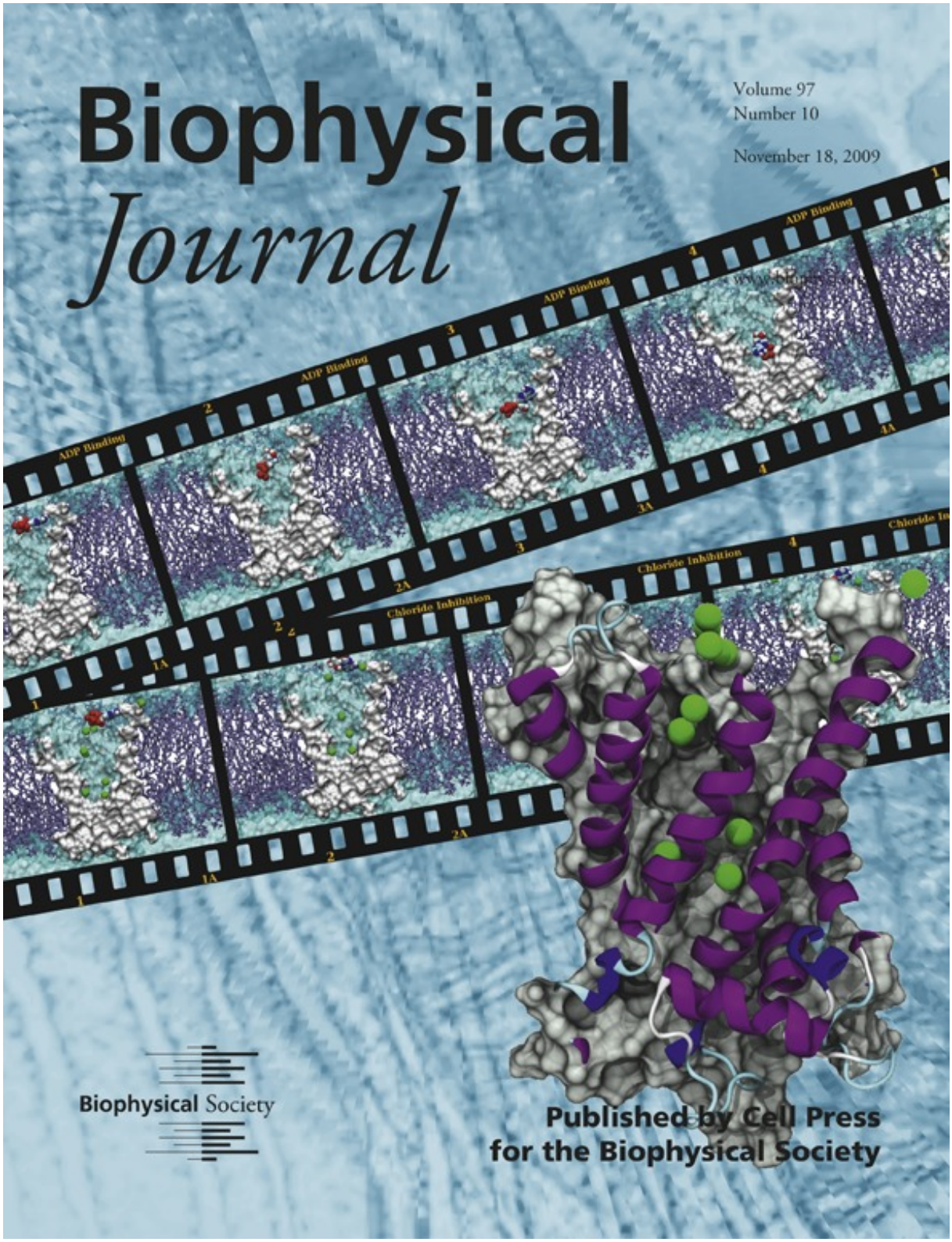


Reconciling Thermodynamics and Kinetics: Water Permeation through a Membrane-Spanning Peptide Nanotube (see page 5A)

ENERGY CONVERSION AND STORAGE, OPTICAL AND ELECTRONIC DEVICES, INTERFACES, NANOMATERIALS, AND HARD MATTER

ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.

www.acs.org


Biophysical
Journal
Volume 97
Number 10
November 18, 2009
ADP Binding
Chloride Inhibition
Biophysical Society
Published by Cell Press
for the Biophysical Society

BIOCHEMICAL SOCIETY

PORTLAND PRESS

BIOCHEMICAL JOURNAL

EXPLORING MOLECULAR MECHANISMS THAT UNDERPIN BIOLOGICAL PROCESSES

VOLUME 476 • ISSUE 11 • JUNE 2019

biochemj.org