



HAL
open science

Multiscale molecular simulations to investigate adenylyl cyclase-based signaling in the brain

Siri van Keulen, Juliette Martin, Francesco Colizzi, Elisa Frezza, Daniel Trpevski, Nuria Cirauqui Diaz, Pietro Vidossich, Ursula Rothlisberger, Jeanette Hellgren Kotaleski, Rebecca Wade, et al.

► To cite this version:

Siri van Keulen, Juliette Martin, Francesco Colizzi, Elisa Frezza, Daniel Trpevski, et al.. Multiscale molecular simulations to investigate adenylyl cyclase-based signaling in the brain. *WIREs Computational Molecular Science*, 2022, 10.1002/wcms.1623 . hal-03695888

HAL Id: hal-03695888












<https://hal.science/hal-03695888>

Submitted on 10 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multiscale molecular simulations to investigate adenylyl cyclase-based signaling in the brain

Siri C. van Keulen¹  | Juliette Martin²  | Francesco Colizzi³  |
 Elisa Frezza⁴  | Daniel Trpevski⁵  | Nuria Cirauqui Diaz²  |
 Pietro Vidossich⁶  | Ursula Rothlisberger⁷  | Jeanette Hellgren Kotaleski^{5,8}  |
 Rebecca C. Wade^{9,10}  | Paolo Carloni^{11,12} 

¹Computational Structural Biology Group, Bijvoet Center for Biomolecular Research, Science for Life, Faculty of Science – Chemistry, Utrecht University, Utrecht, The Netherlands

²CNRS, UMR 5086 Molecular Microbiology and Structural Biochemistry, University of Lyon, Lyon, France

³Molecular Ocean Laboratory, Department of Marine Biology and Oceanography, Institute of Marine Sciences, ICM-CSIC, Barcelona, Spain

⁴Université Paris Cité, CiTCoM, CNRS, Paris, France

⁵Science for Life Laboratory, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm

⁶Molecular Modeling and Drug Discovery Lab, Istituto Italiano di Tecnologia, Genoa, Italy

⁷Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne

⁸Department of Neuroscience, Karolinska Institute, Stockholm

⁹Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), Heidelberg, Germany

¹⁰Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, and Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Heidelberg, Germany

¹¹Institute for Neuroscience and Medicine (INM-9) and Institute for Advanced Simulations (IAS-5) “Computational biomedicine”, Forschungszentrum Jülich, Jülich, Germany

¹²INM-11 JARA-Institute: Molecular Neuroscience and Neuroimaging, Forschungszentrum Jülich, Jülich, Germany

Correspondence

Ursula Rothlisberger, Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne.

Email: ursula.rothlisberger@epfl.ch

Jeanette Hellgren Kotaleski, Science for Life Laboratory, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE-10044 Stockholm, Sweden; and Department of Neuroscience, Karolinska Institute, SE-17165 Stockholm.

Email: jeanette@kth.se

Abstract

Adenylyl cyclases (ACs) play a key role in many signaling cascades. ACs catalyze the production of cyclic AMP from ATP and this function is stimulated or inhibited by the binding of their cognate stimulatory or inhibitory G α subunits, respectively. Here we used simulation tools to uncover the molecular and sub-cellular mechanisms of AC function, with a focus on the AC5 isoform, extensively studied experimentally. First, quantum mechanical/molecular mechanical free energy simulations were used to investigate the enzymatic reaction and its changes upon point mutations. Next, molecular dynamics simulations were employed to assess the catalytic state in the presence or absence of G α subunits. This led to the identification of an inactive state of the enzyme

Siri C. van Keulen, Juliette Martin, Francesco Colizzi, Elisa Frezza, Daniel Trpevski, Nuria Cirauqui Diaz, and Pietro Vidossich contributed equally to the study.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *WIREs Computational Molecular Science* published by Wiley Periodicals LLC.

Rebecca C. Wade, Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), Schloss-Wolfsbrunnengasse 35, 69118 Heidelberg, Germany, and Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, and Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany.
Email: rebecca.wade@h-its.org;

Paolo Carloni, Institute for Neuroscience and Medicine (INM-9) and Institute for Advanced Simulations (IAS-5) "Computational biomedicine," Forschungszentrum Jülich, 52425 Jülich, Germany.
Email: p.carloni@fz-juelich.de

Present address

Nuria Cirauqui Diaz, European Molecular Biology Laboratory, Grenoble, France

Funding information

Horizon 2020 Framework Programme, Grant/Award Numbers: 945539, HBP SGA3; Klaus Tschira Foundation; Ramón y Cajal Fellow, Grant/Award Number: RYC2019-026768-I; Swedish Research Council, Grant/Award Numbers: VR-M-2020-01652, VR-M-2017-02806; European Union's Horizon 2020 research and innovation program through the ICEI project, Grant/Award Number: 800858; Spanish Ministry of Science and Innovation, Grant/Award Number: CEX2019-000928-S; Swiss National Computing Center CSCS; GENCI for computing on the CINES supercomputer OCCIGEN, Grant/Award Numbers: 2018-A0040710357, 2017-A0020707585, 2016-c201607758

Edited by: Modesto Orozco, Associate Editor

1 | INTRODUCTION

To perform their functions, cells need to sense and process various signals representing the state of their external and internal environment. Examples are signals representing the availability of nutrients, the level of cellular damage, and, particularly important for multicellular organisms, various communication signals that serve to coordinate cellular activity among tissues and/or organs. The required signal processing is achieved through specialized molecular circuits, called signal transduction cascades, that have evolved to elicit suitable responses to different stimuli. Examples include the activation of cellular motors to propel a bacterium toward nutrients, the start of cellular repair mechanisms, and the strengthening and weakening of synapses in the nervous system. For extracellular signals, the cascades typically start with cell-surface receptors with a high specificity for a particular molecule (an agonist), which upon binding triggers a chain of chemical reactions inside the cell, thus relaying the signal across the membrane. Very often, the beginning of a cascade involves activating an enzyme that amplifies the extracellular signal by producing large amounts of an intracellular second messenger. Such is

that is present whenever an inhibitory $G\alpha$ is associated, independent of the presence of a stimulatory $G\alpha$. In addition, the use of coevolution-guided multi-scale simulations revealed that the binding of $G\alpha$ subunits reshapes the free-energy landscape of the AC5 enzyme by following the classical population-shift paradigm. Finally, Brownian dynamics simulations provided forward rate constants for the binding of $G\alpha$ subunits to AC5, consistent with the ability of the protein to perform coincidence detection effectively. Our calculations also pointed to strong similarities between AC5 and other AC isoforms, including AC1 and AC6. Findings from the molecular simulations were used along with experimental data as constraints for systems biology modeling of a specific AC5-triggered neuronal cascade to investigate how the dynamics of downstream signaling depend on initial receptor activation.

This article is categorized under:

Structure and Mechanism > Computational Biochemistry and Biophysics
Molecular and Statistical Mechanics > Molecular Dynamics and Monte-Carlo Methods
Software > Molecular Modeling

KEYWORDS

adenylyl cyclases, molecular simulation, systems biology modeling

the case with the family of mammalian adenylyl cyclase enzymes (ACs). These catalyze the conversion of adenosine triphosphate (ATP) to cyclic adenosine monophosphate (cAMP)—one of the main cellular second messenger signaling molecules. The main regulators of AC activity are G proteins, the first proteins to be activated when the cell-surface receptors (the so-called G-protein coupled receptors or GPCRs) bind their extracellular ligand. The regulation of AC activity can be bi-directional, with some G proteins stimulating the enzymes and others inhibiting them. Having in mind that ACs are also regulated by many other molecules, such an arrangement allows not only for signal transduction, but also for the integration of multiple signals (relayed by the regulators) at the start of the signaling cascade. Studying ACs and their regulation is therefore of central importance for understanding how various cellular processes are activated and controlled.

In the past few years, a consortium within the Human Brain Project (HBP) EU flagship has systematically investigated structure, function, and dynamics of AC enzymes and their complexation with activated G proteins by multiscale molecular simulations, from quantum to coarse grained levels. Relevant findings were incorporated in systems biology network models of an AC-based signal transduction network that controls synaptic plasticity in striatal projection neurons. Taken together, these studies provide an unprecedented view on AC enzyme regulation and provide insight into the implications of the latter for complex brain processes involving synaptic plasticity and learning, such as reinforcement learning.¹ This focus article shows the predictive power and the synergies of different techniques when applied to the same system, and suggests that a similar approach could be successfully applied to other enzymatic systems.

2 | THE AC ENZYME AND ITS COMPLEXES WITH ITS COGNATE G PROTEINS

The enzymes of the AC family act as central integrators of extracellular neurochemical signals from first messengers, such as hormones and neurotransmitters.² The AC family consists of nine transmembrane isoforms (AC 1–9) and one

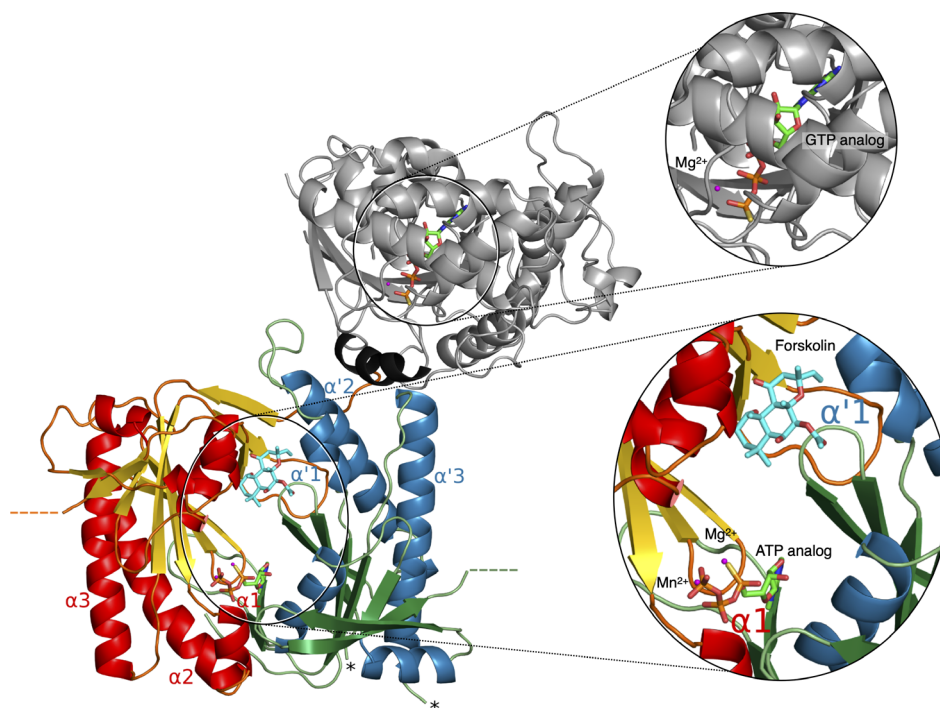


FIGURE 1 Representation of the X-ray structure of one of the two cytoplasmic domains of the adenylyl cyclases (AC) enzymes, bound to the α subunit of the stimulatory G protein.¹¹ The domain consists of the C1a subdomain from canine AC5 (red/yellow/orange) and the C2a subdomain (blue/green/light green) from rat AC2 from the catalytic unit. Dashed lines indicate the location of the C termini of the C1a and C2a subdomains to which the C1b and C2b subdomains, which were not resolved in this structure, are attached. Asterisks indicate the position of an unresolved flexible loop. The $\alpha'1$, $\alpha'2$, and $\alpha'3$ helices of the C2a subdomain form a groove to which bovine G_s binds (gray) via its switch II helix (dark gray). The $\alpha1$, $\alpha2$, and $\alpha3$ helices of the C1a subdomain form a similar groove in a pseudosymmetric position. The protein is shown in cartoon representation, three metal ions (Mn^{2+} and Mg^{2+}) are shown as magenta spheres, and ligands are shown in stick representation colored according to atom type. A sulfur derivative of GTP (green carbons) is bound to $G_{\alpha s}$ and a sulfur derivative of AC's substrate ATP (green carbons) and forskolin (cyan carbons), a promoter of AC activity, bind at the interface of the C1a and C2a subdomains. Figure adapted from Tong et al.¹⁸

soluble protein (AC 10 or sAC). Although abundantly available, AC isoforms differ in their tissue distribution. They are all expressed in the central nervous system (CNS), where they play a key role in synaptic plasticity, long-term potentiation, sensory perception and many other processes.^{3–5}

ACs catalyze the cyclization of ATP to the second messenger cAMP, thereby changing intracellular cAMP concentrations. cAMP, by binding to specific cytoplasmic proteins, activates a multitude of downstream events. AC function is modulated by a variety of cognate proteins. The main regulators are G proteins, heterotrimers consisting of α , β and γ subunits.¹ GTP-bound stimulatory $G\alpha_s$ and its structurally similar^{8,9} olfactory $G\alpha_{olf}$ subunits activate all AC isoforms, while the myristoylated inhibitory $G\alpha_i$ modulates the activity of ACs 1, 5, 6.^{7,10} The $G\beta\gamma$ subunits can also stimulate or inhibit ACs except AC 9.²

ACs consist of an N-terminal domain (N), two membrane-spanning domains (M1, M2), and two cytoplasmic domains (C1, C2). Several X-ray crystal structures of the stimulated catalytic domains, C1 and C2, complexed to $G\alpha_s$ and in an isolated form have been resolved^{11–15} as well as an activated AC9 structure including the membrane domains.¹⁶ Each of these structures includes a pseudosymmetric heterodimer (C1 and C2) with the catalytic site at the interface between the C1a and C2a subdomains¹⁷ (Figure 1). $G\alpha_s$ associates to the C2a subdomain and stabilizes the C1a/C2a heterodimer. This harbors the AC's catalytic site at its interface. ATP is stabilized by magnesium ions and residues in the pocket. In contrast, experimental structural information is lacking for the AC: G_i complex, hampering a complete understanding of the mechanism of inhibition of the enzyme by this cognate protein. Structural information is available for G_i in isolation¹⁹ and in absence of a key posttranslational modification present across all mammalian proteins, myristoylation of the N-terminus.^{20–22} Finally, it is unclear if trimeric $G\alpha_s$ -AC- $G\alpha_i$ complexes are formed in the cell and if so, if they are catalytically active or not.

3 | MULTISCALE SIMULATIONS OF AC5

We have applied a multiscale simulation approach to address these questions, and used these results to model a specific signaling pathway involving AC enzymes in memory rewarding processes in the brain. The main results of this joint effort are reported here, along with technical details. We refer the reader to References 23–26 for more comprehensive information about the simulation methodology employed. We focus mostly on the AC5 isoform from different mammals, including mouse, rat and human (Sections 3 and 4). Insights into other members of the superfamily were also deduced (Section 5). At the time this work began, the only available structural information was that of holo and apo ACs' catalytic domains in complex with $G\alpha_s$.^{11,12} All calculations presented here are based on this information.² We first studied AC on the quantum mechanical (QM) level via mixed QM/molecular mechanical (MM) simulations to investigate the free-energy landscape of its catalytic mechanism, converting ATP to cAMP and pyrophosphate (Section 4). Then we moved to an atomistic description of the model for which we evaluated different AC5 states in classical all-atom molecular dynamics (MD) simulations, such as AC in the presence and absence of $G\alpha_i$ and $G\alpha_{olf}$ to assess AC's stability and catalytic propensity (Section 5). Section 6 describes a normal mode analysis of the classical MD results, which leads to a deeper insight into AC5's conformational flexibility. Section 7 shows how coevolution data can be used in combination with metadynamics simulations to investigate the free-energy landscape of AC5 in the absence and presence of G proteins. Section 8 applies Brownian dynamics (BD) simulations to predict the association rate constants for G-protein binding to AC5 using protein structures from the classical MD results. These rate constants, as well as the results on AC5 regulation from the classical MD simulations, were then used in our deterministic kinetic models described in Section 9 to study the functional implications of these findings on the AC5-based signal transduction cascade which regulates synaptic plasticity in striatal projection neurons. Finally, Section 10 extends our study to other members of the AC family by comparing the sequences, structures and electrostatic potentials of AC5 with those of the other eight transmembrane AC isoforms.

4 | QM/MM SIMULATIONS OF A CHIMERIC MODEL OF AC

Within the range of applicability of transition state theory, parameters of enzymatic kinetic equations may be related to the reaction free-energy profile, which is accessible through atomistic simulations, provided that chemical bonds are allowed to break and/or form, and high energy configurations are properly sampled during the simulation. Quantum chemical methods are required to model the reorganization of chemical bonds, and given that the reaction takes place

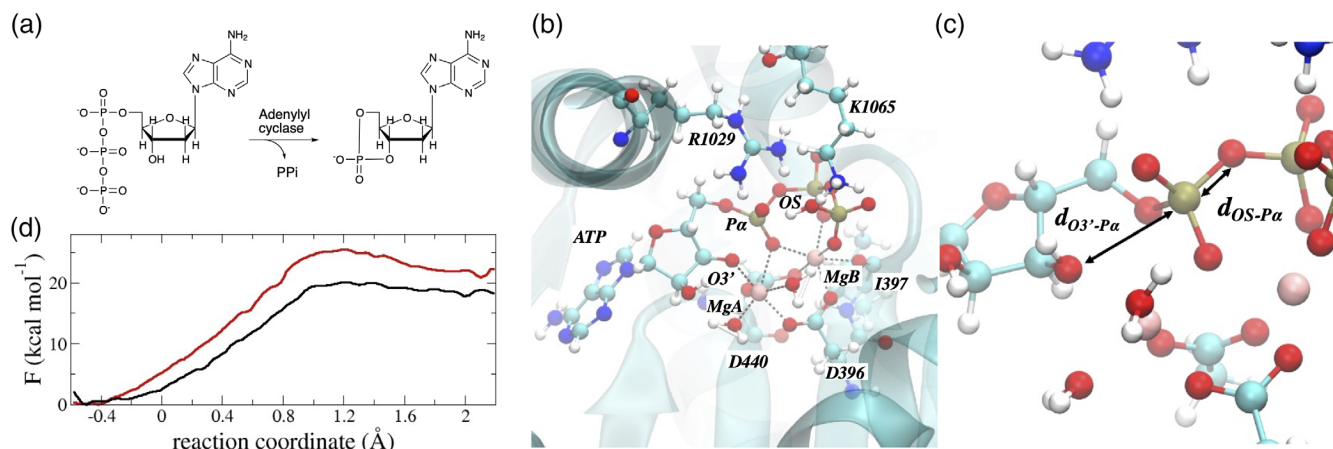


FIGURE 2 Quantum mechanical/molecular mechanical (QM/MM) modeling of the conversion of ATP to cAMP. (a) Chemical scheme of the reaction. (b) QM/MM setup of the adenylyl cyclase (AC):G α_s complex active site used to study the enzymatic reaction. The QM atoms are shown in ball and stick representation. All the other atoms in the system were treated with the classical force field. Bulk water is not shown for clarity. (c) Close up of the QM region indicating the O3'-P α and OS-P α distances used for the definition of the reaction coordinate. (d) Free energy profile along the reaction coordinate for the wild-type AC chimeric enzyme (black line) and the enzyme after quenching the Arg1029 side chain charges (red line) to mimic the effect of a point mutation to an uncharged residue

within the protein environment, the hybrid QM/MM approach is particularly suitable to account for steric and electrostatic effects. Employing enhanced sampling methods, high energy configurations may be visited and characterized during MD simulations. QM/MM simulations were used here to reveal relative changes in k_{cat} due to point mutations. Through the integrated approach outlined in this review, it may be possible to study how these changes are propagated at the system level.

The choice of the QM level is a key aspect to consider when performing mechanistic studies of chemical reactions. In this context, semi-empirical methods usually perform with lower accuracy (both in terms of activation energies and reaction pathways, thus possibly leading to wrong mechanistic conclusions) compared to more advanced QM methods (such as density functional theory, DFT), yet they offer the advantage of being computationally cheap and cover relatively long trajectories on the ns timescale. We should mention that, as exascale computer architectures become available, the usage of highly scalable codes, such as the one described by Bolnykh et al.^{27,28} allows DFT-based QM/MM simulations to reach timescales not vastly inferior to those obtained by semi-empirical methods²⁹

The conversion of ATP to form cAMP via a cyclization reaction involves the nucleophilic attack of the O3' oxygen on the α -phosphate of ATP, with an inversion of configuration and the dissociation of a pyrophosphate moiety (Figure 2a,b). This bond rearrangement follows a mechanism involving two Mg²⁺ ions, common to many nucleic acid processing enzymes.¹¹ The X-ray structure of a chimeric model of dog/rat AC domains in complex with ATP and the stimulatory G α subunit is discussed in Reference 11 (see Section S2 of Appendix S1 for method details). First, we used MD simulations to construct the Michaelis–Menten complex and achieve a catalytically competent reactant configuration, in which the O3' nucleophile is bound to a Mg²⁺ ion, a feature not seen in the experimental structure. Then, we predicted the free energy profile along the reaction coordinate shown in Figure 2c, using QM/MM-based umbrella sampling simulations (see Section S2 of Appendix S1 for details). We used the semi-empirical PM6 method³⁰ for the QM subsystem. The free energy barrier turned out to be approximately 20 kcal/mol, in fair agreement with reported estimates from experimental data that place the barrier between 15 and 19 kcal/mol.^{11,31,32} Furthermore, we found that the free energy barrier increases by roughly 5 kcal/mol upon quenching the R1029 side chain charge (Figure 2d). This suggests that the presence of this charged arginine is crucial for the efficiency of AC5, which is in line with the mutagenesis experiments on AC5.³²

5 | CLASSICAL MOLECULAR DYNAMICS SIMULATIONS OF AC5

Microsecond MD simulations were used to investigate the structural determinants of the AC5 protein in the presence or absence of cognate G α subunits, their impact on ATP conversion to cAMP, and the allosteric communication

between distant regulatory binding sites.^{33–37} The calculations were based on a modified Amber99SB force field^{38–41} and the GROMACS software package.^{42–45} The starting models were built based on available X-ray structures of AC5, $G\alpha_i$, and $G\alpha_{olf}$ and using homology modeling with Modeller⁴⁶ and in silico docking with Cluspro and HADDOCK2.2^{47–49} (Figure 3a,c). The AC5: $G\alpha_i$ complexes simulated in References 34,37,50 were compatible with the membrane location deduced from the recently determined full-length structure of AC9¹⁶ (Figure 3a).

Both Cluspro and HADDOCK allow to incorporate prior knowledge in the form of presumed binding site residues, which are used to guide the docking process toward these residues. In the HADDOCK runs, for example, mutagenesis information was included through position restraints to guide the modeling of the complex predictions of the catalytic AC5 domains and the $G\alpha_i$ subunit (Figure 3a,c). The recently developed machine-learning based AlphaFold (AF) method,⁵¹ which has revolutionized the field of protein structure prediction,⁵² could be interesting to apply here. Indeed, the multimer version of the algorithm⁵³ could be used to predict the structural determinants of some of the AC: $G\alpha$ complexes.³

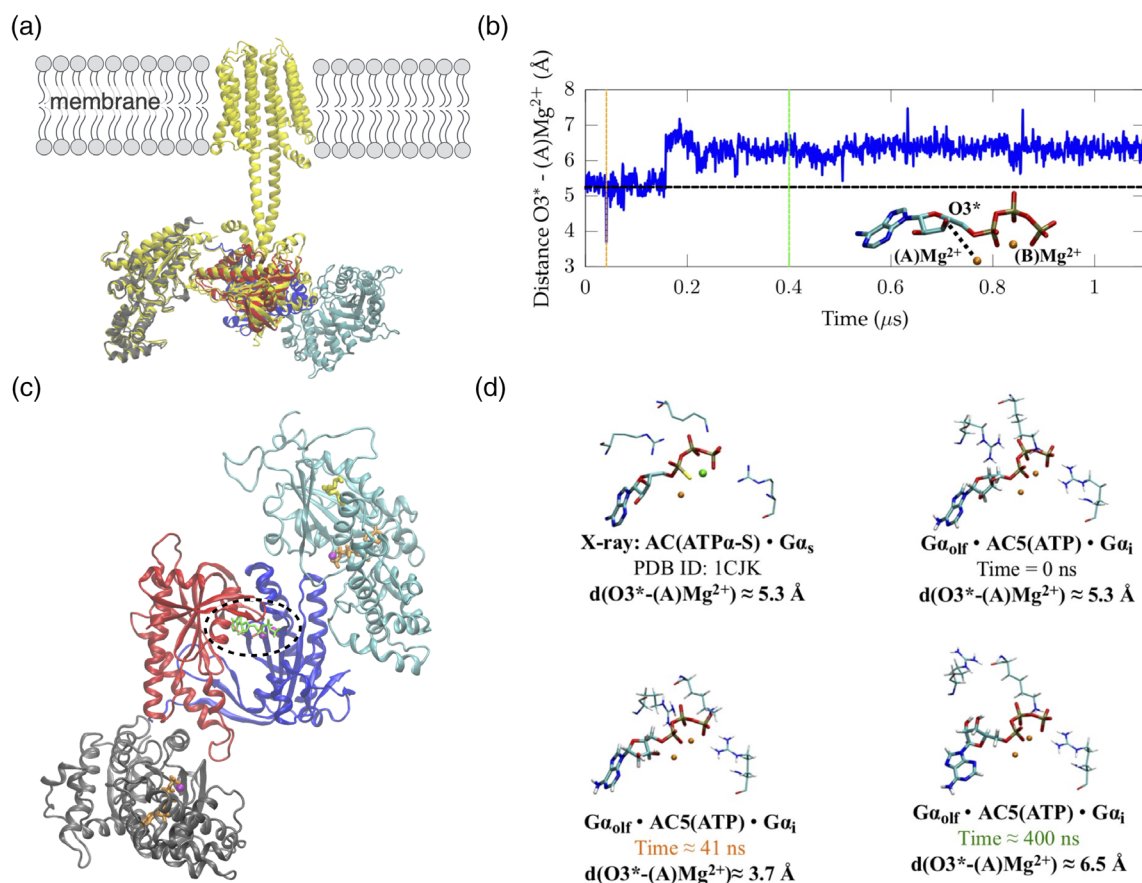


FIGURE 3 Structure and dynamics of the simulated $G\alpha_{olf}$:AC5: $G\alpha_i$ complex. (a) Orientation of the ternary complex with respect to the membrane, which is in line with the resolved full-length AC9 structure (PDB ID 6R3Q) in yellow.¹⁶ $G\alpha_{olf}$ is shown in gray, while $G\alpha_i$ and the catalytic domain of AC5 are depicted in cyan and red (C2a)/blue(C1b), respectively. (b) 1 μ s molecular dynamics (MD) trajectory of the ternary complex described in panel (a) and (c).¹ The distances between O3* and (a) Mg^{2+} during the dynamics is monitored: (i) short distances near 3 \AA represent the near-attack conformation of adenosine triphosphate (ATP) in the active site¹; (ii) long distances around 7 \AA represent inactive ATP conformations. In the trajectory, a conformational change in ATP (see panel d) is reflected by an increase in the O3*-(A) Mg^{2+} distance, leading to a state of the complex that is unable to sample the near-attack ATP conformation.¹ (c) Cytosol view on the $G\alpha_{olf}$:AC5: $G\alpha_i$ complex in which the color coding matches the one in panel (a). GTPs are highlighted in orange and the myristoyl moiety in yellow. ATP is encircled by a dashed black line and depicted in green. A detailed view of the active site is described in panel d. (d) Four conformations of ATP in the active site of the ternary complex ($G\alpha_{olf}$:AC5: $G\alpha_i$) are highlighted.¹ One representing the conformation found in PDB ID 1CJK, the active AC conformation of the binary $G\alpha_s$:AC complex, while the other three are snapshots from the MD trajectory shown in panel (b). The time at which the three snapshots are extracted are highlighted in orange (41 ns) and green (400 ns), which are represented by equivalently colored lines in panel b. At each frame, the O3*-(A) Mg^{2+} distance is shown in \AA .

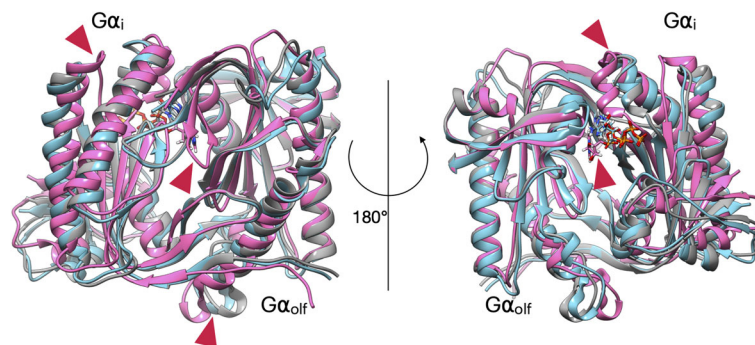


FIGURE 4 Conformations of AC5 observed in molecular dynamics simulations of binary complexes in presence of adenosine triphosphate. Conformations are averaged over the last 500 ns of 1.5 μ s trajectories.^{33,34} Gray: AC5(ATP), pink: AC5(ATP):G α_{olf} , cyan: AC5(ATP):G α_i , ATP molecules are shown in stick representation. G α subunits are not shown for clarity, but their binding sites are labeled. Red arrows point to main structural differences. Left: membrane side view. AC5 in interaction with G α_{olf} (pink) has a closed G α_i binding site and an open β 2 loop; AC5 in interaction with G α_i (blue) or alone (gray) has a closed G α_{olf} binding site. Right: cytosolic side view. The C1 and C2 domains are closer together in AC5:G α_{olf} (pink) than in other conformations, and the location of the ATP molecule differs.

Root-Mean-Square Deviations (RMSD) computed on separate domains in simulations of binary complexes^{33,34} revealed the existence of several long-lived substates in AC5 influenced by the presence or absence of G-protein subunits (Figure 4). These simulations also showed that the presence of regulatory G proteins and ATP affect the rigidity of AC5 in the isolated and the associated forms, as assessed by the Root-Mean-Square Fluctuation (RMSF) profiles. We analyzed the impact of G proteins and ATP on AC5 flexibility by comparing the difference in RMSF of AC5 residues in the different simulation setups. The conformations of the regulatory binding sites of AC5 were described using helix angle and distance descriptors. This allowed us to describe the different conformations of AC in interaction with regulatory proteins (Figure 4) and revealed the allosteric communication between the opposite binding sites of inhibitory and stimulatory G proteins,^{33–37} as observed by Internal Normal Mode analysis (see Section 6).

A structural analysis of AC5 alone and in complex with G proteins^{1,33,34,36,37} suggested that (i) AC5 is activated in the presence of G α_{olf} by locking ATP in a catalytic-compatible conformation; (ii) the protein is inhibited via increased conformational and positional fluctuations of ATP in the presence of G α_i (Figure 3c,d) together with a change of the shape of the active site that discourages ATP binding, as also observed in the coevolution-guided metadynamics simulations (see Section 7).^{33–37} (iii) a ternary complex consisting of AC5 simultaneously associated with active G α_{olf} and active myristoylated G α_i in its soluble form (not interacting with the membrane) may be formed in the absence or presence of ATP.^{36,37,50} However, the probability of ATP conversion of such a ternary complex (Figure 3) is significantly reduced relative to that of the binary AC5:G α_{olf} complex in Figure 1 (Figure 3b,d). This information^{36,37,50} was incorporated in the mathematical modeling of the reinforcement learning signaling network, which revolves around AC regulation via G-protein activation¹ (See Section 9).

6 | COARSE-GRAINED NORMAL MODE ANALYSIS IN INTERNAL COORDINATES OF AC5

Normal Mode Analysis (NMA) is one of the most common methods used to probe large-scale, shape-changing motions in biological molecules and it is a computationally inexpensive method. To obtain the frequencies ω_k and the eigenvectors A_{ik} , the following generalized eigenvector problem has to be solved:

$$\mathbf{HAW} = \mathbf{FA}, \quad (1)$$

where \mathbf{W} is a positive diagonal matrix whose elements are $W_{ik} = \delta_{ik}\omega_{ik}$ with δ_{ik} a Kronecker delta (i and k are two generic indices of the matrix), \mathbf{H} is the kinetic matrix and \mathbf{F} is the Hessian matrix. The kinetic matrix \mathbf{H} is defined by

$$H_{ij} = \sum_a m_a \frac{\partial \vec{r}_a}{\partial q_i} \frac{\partial \vec{r}_a}{\partial q_j}, \quad (2)$$

where q_i and q_j are the internal coordinates i and j , respectively, \vec{r}_a and m_a are the position vectors and the atomic masses of each atom a , respectively. The Hessian matrix \mathbf{F} is defined by $F_{ij} = \frac{\partial^2 E_p}{\partial q_i \partial q_j}$ with E_p the potential energy.

In this study of AC5, NMA was performed using a coarse-grained model to represent AC5. An anisotropic elastic network, in which pseudo-atoms are connected by springs if their distance is closer than a chosen cutoff, was used to compute the potential energy. The most common model contains only $C\alpha$ atoms.⁵⁴ However, representations that include several particles per residue better capture structural details along the movements.^{55–57} Among the latter, we used the coarse-grained model PaLaCe (named after the developers Pasi, Lavery, and Ceres)⁵⁸ to describe the protein(s) and their interactions with iNMA (Internal Normal Mode Analysis).⁵⁹ PaLaCe is characterized by a two-tier representation (one for bonded and another one for nonbonded interactions). Three backbone beads (N, $C\alpha$, and C') are used for backbone representation and one or two beads for the side chain. An implicit model took into account the solvent and the ions present in solution. PaLaCe was parameterized using a large body of structural information on soluble, globular proteins in their native state present in the PDB.⁶⁰ It is characterized by a residue-dependent level of coarse-graining. The two-tier representation, and in particular an atomic-scale backbone, with the torsional energies dependent on both the backbone ϕ, ψ , and side-chain torsions α, β, γ ($E_{\text{tor}} = \sum_{(\alpha, \beta, \gamma)} E_{\text{tor}}(\tau) + \sum_{(\alpha, \beta, \gamma), (\phi, \psi)} E_{\text{tor}}(\tau_1, \tau_2)$) allows avoiding any artificial restraints on secondary structure. PaLaCe can be used to properly maintain structures of folded proteins and to predict protein dynamic fluctuations and flexibility, as well as large-scale force-induced conformational changes.^{58, 59} However, it currently cannot be used for folding studies, and there is room for improvement regarding its performance and its ability to predict protein–protein interactions.

Here, iNMA, using simplified representations of the protein structures in the PaLaCe model,⁵⁸ was used to investigate the effect of regulatory proteins on AC5 conformational flexibility. Comparison was made with the results from MD simulations (see Section 5).^{33, 34} The calculations were performed on the backbone dihedral angles (ϕ and ψ) that capture proteins' global movements well.^{59, 61, 62} iNMA was used to investigate the global movements of AC5 alone and in complex with $G\alpha_{\text{olf}}$ and/or $G\alpha_i$ (i), the flexibility of the different molecular species (ii) and the impact on the flexibility of AC5 in the presence of either $G\alpha_{\text{olf}}$ and/or $G\alpha_i$ (iii), using similar strategies to the ones applied in the MD simulations. Our starting structures were based on the homology models and docking results used in the MD simulations.⁴ They are also compatible with new structures obtained by Korkhov et al.¹⁶ as described in Section 5. iNMA turned out to reproduce the RMSF of AC5 obtained by MD simulations (Figure S1). The first modes for AC5 show an allosteric communication between the two binding sites that is modulated by the presence of $G\alpha_{\text{olf}}$ and/or $G\alpha_i$ (Figure S2). They turned out to compare well with the ones obtained by Principal Component Analysis. The change in flexibility of AC5 upon binding of $G\alpha_{\text{olf}}$ and $G\alpha_i$ either in the $G\alpha_{\text{olf}}$ binding site or in its putative binding site was estimated by the difference in RMSF values for apo and the complex. iNMA calculations allowed us to rule out a model where $G\alpha_i$ would bind at the $G\alpha_{\text{olf}}$ binding site, where AC5 flexibility was barely affected. On the contrary, when $G\alpha_i$ is in its putative binding site, the intensities of the peaks are similar to those of AC5: $G\alpha_{\text{olf}}$. In addition, the presence of either $G\alpha_i$ in the putative binding site or $G\alpha_{\text{olf}}$ impacts the flexibility of both binding sites and the active site regions, as observed in MD simulations.

To better understand the impact on AC5 fluctuations upon $G\alpha$ binding, we also compared the modes of AC5 in two conditions: (i) by constraining the positions of either $G\alpha_i$ or $G\alpha_{\text{olf}}$ in the binary complex and (ii) by considering the two $G\alpha$ subunits and AC5 alone.⁵⁹ The presence of $G\alpha_i$ at the putative binding site strongly impacts the essential dynamics of AC5, since the collective movements (the first modes) differ from those of AC5 alone. Finally, we characterized the ternary complex by computing iNMA on AC5: $G\alpha_{\text{olf}}$ with $G\alpha_i$ at its putative binding site and $G\alpha_i$:AC5 with $G\alpha_{\text{olf}}$ at its binding site. $G\alpha_i$ or $G\alpha_{\text{olf}}$ turned out to impact the flexibility of AC5 in two opposite way: $G\alpha_i$ increased the flexibility of the $G\alpha_{\text{olf}}$ binding site and $G\alpha_{\text{olf}}$ rigidified the $G\alpha_i$ binding site (Figure 5). These results suggest that the formation of the ternary complex is possible only when $G\alpha_{\text{olf}}$ is already bound to AC5, in agreement with MD simulations. The latter sampled $G\alpha_i$ -compatible conformations of AC5 in the AC5: $G\alpha_{\text{olf}}$:ATP complex, and $G\alpha_{\text{olf}}$ -incompatible conformations in the AC5: $G\alpha_i$ complexes, with and without ATP.³⁴ We conclude that iNMA is a promising method to predict the flexibility of protein complexes and to understand the impact of the binding on a specific protein. Our results, coupled with the findings obtained by MD simulations, provide a crucial piece of information for the systems biology modeling.

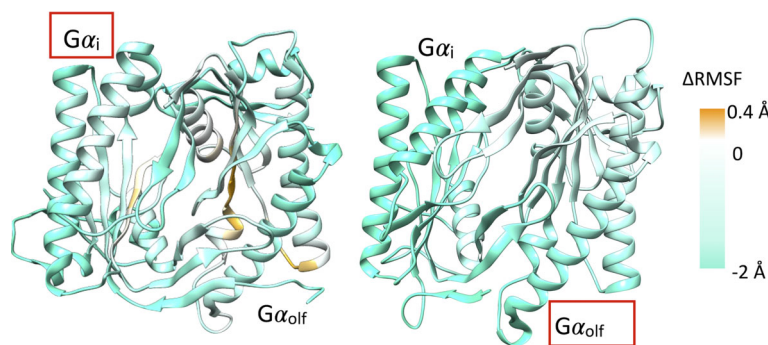


FIGURE 5 Changes in flexibility induced by the addition of each G protein on pre-formed binary complexes predicted by internal Normal mode analysis using PaLaCe. Left: Changes in flexibility when $G\alpha_i$ is added to the AC5: $G\alpha_{olf}$ complex on its putative binding site. Right: Changes in flexibility when $G\alpha_{olf}$ is added to the $G\alpha_i$:AC5 complex. In each case, the added protein is highlighted with a red box. Changes in flexibility are assessed by the difference in RMSF values (Δ RMSF), computed as RMSF in the ternary complex minus RMSF in the binary complex. RMSF for the AC5: $G\alpha_{olf}$ complex is reported in Figure S3. For more intense colors (orange for increased flexibility and cyan for decreased flexibility) correspond to differences with respect to the preceding structure (AC5: $G\alpha_{olf}$ and $G\alpha_i$:AC5, respectively) on a scale of -2 to $+0.4$ Å as also shown in the color bar. The addition of $G\alpha_i$ flexibilizes the $G\alpha_{olf}$ binding site (left).

7 | COEVOLUTION-GUIDED METADYNAMICS SIMULATIONS OF AC5

The extraction of coevolved residue–residue contacts from sequence information alone is a powerful tool to guide the search for native structures,^{64–66} as also demonstrated by AlphaFolds2's EvoFormer.⁵¹ Furthermore, coevolved contacts have been used to model conformational ensembles⁶⁷ and to identify functionally relevant transitions in proteins^{24,68} and nucleic acids.^{69,70} Indeed, pairwise coevolved amino acid positions correlate strongly with spatial proximity in three-dimensional space. The underlying basic principle is rather intuitive: if two protein residues come in contact with each other at some point in a protein's life, a destabilizing amino acid substitution at one position is expected to be compensated by a substitution with a complementary residue in the other position during evolution, in order for the protein to preserve its functional stability. We have recently shown that coevolutionary information can reduce the complexity of the configurational space in protein–protein allosteric regulations.²³ Homologues of human (h) AC5 were identified from the clustered UniProt database (uniclust30_2018_08)⁷¹ using HHblits⁷² to generate the corresponding multiple sequence alignment (MSA). Performing Direct Coupling Analysis (DCA)^{64,73} on an MSA of protein homologues, it is possible to select those coevolved pairs of residues that are not in contact in the native structure (Figure S4, red dots; Figure S4b, red lines; below), and which might thus be informative of alternative protein conformations.²⁴ DCA outputs a direct information score per pair of residues⁷³ that have been used as input for an automated protocol whereby coevolution contacts are filtered and introduced as ensemble restraints in coarse-grained discrete molecular dynamics simulations,^{24,74,75} which allowed us not only to detect functionally relevant conformations, but also to generate a trajectory connecting them—thus enabling its discretization⁷⁶ and its use as a reference for path-based simulations.⁷⁷ Orozco et al. have shown²⁴ that filtering of DCA contacts is not critical when abundant sequences ($>10,000$) are available, yielding strong evolutionary signals. However, when fewer homologues are aligned (on the order of a few thousand), conformational transitions could not be modeled without filtering DCA contacts due to noise in the coevolution map. A few high-quality coevolved pairs are thus necessary to robustly guide protein dynamics, making the detection of these constraints decisive.

A multiscale MD approach guided by coevolutionary data was used to reconstruct the free-energy landscape of a fully solvated atomistic model of hAC5 with and without the regulatory $G\alpha$ subunit bound²³; here $G\alpha$ was either the stimulatory $G\alpha_s$, or the inhibitory $G\alpha_i$. The multiscale consisted in coevolution-guided coarse-grained (alpha-carbons level) discrete MD simulations^{24,74,75} that were used to generate a reference functional conformational transition whose free-energy landscape was mapped with fully atomistic metadynamics simulations.^{78–81} We found AC5 populating two main conformational ensembles, an open and a closed one, with all the existing experimental structures falling just into the open ensemble (Figure 6). Specifically, the open ensemble included the experimental structures of activated AC5 bound to $G\alpha_s$ together with an ATP analog (P-site inhibitor) and the activator forskolin (PDB ID 1CJK),¹¹ a forskolin derivative alone (PDB ID 1AZS),¹² and a guanosine 5'-triphosphate (GTP)-based substrate analog and forskolin (PDB ID 6R4O) from cryo-EM experiments^{16,82} of the full-length membrane AC (Figure 6, green circles in the bottom panel).

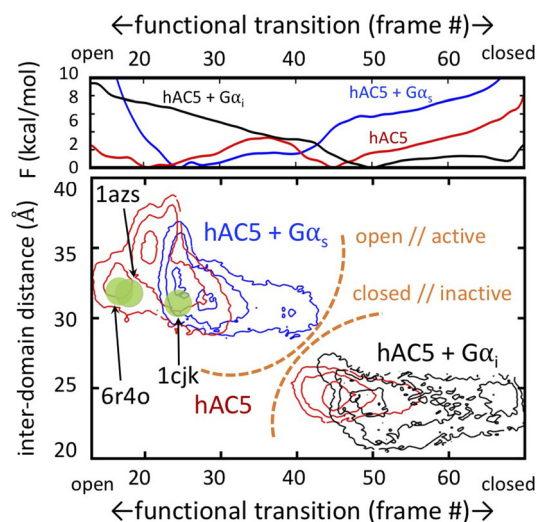


FIGURE 6 Modulation of hAC5 functional transition by the binding of stimulatory and inhibitory G proteins. Comparison of the free energy landscape populated by human AC5 with no regulatory protein bound (red plot), in complex with stimulatory $G\alpha_s$ (blue plot), and in complex with inhibitory $G\alpha_i$ (black plot); hAC5 alone can populate both closed and open states, the binding of $G\alpha_s$ stabilizes an open state, while the binding of $G\alpha_i$ favors the closed conformation of hAC5. The upper panel shows the overall depth of the free-energy profiles. In the bottom panel, contours are drawn at 1 kcal/mol intervals from 0 to 3 kcal/mol, after alignment of each minimum to zero. Gray transparent circles with corresponding PDB codes (PDB IDs 6R4O 1CJK, 1AZS) represent the positions of experimentally determined AC structures in the explored conformational space. Figure readapted from Reference 23.

Notably, AC5 was observed to shift from one ensemble to the other depending on which G protein it was bound to. In particular, when $G\alpha_s$ was bound to hAC5, the conformational ensemble of AC shifted to the open conformation that became approximately 6 kcal/mol more stable than the closed state (Figure 6, blue lines). In remarkable contrast, however, when hAC5 was simulated bound to the inhibitory $G\alpha_i$, the open/closed equilibrium shifted toward the closed conformation of AC that became approximately 8 kcal/mol more stable than the open state (Figure 6, black lines). The closure of the ATP-binding site is consistent with biochemical data indicating that P-site inhibitors (ATP analogs) bind with greatly reduced affinity to AC in the presence of $G\alpha_i$.¹¹ Qualitatively, this general behavior was also observed in the microsecond-long classical all-atom MD simulations described above.^{34,36} Overall, these multiscale simulations complemented structural and biochemical data (commented above), and extended our understanding of the mechanisms of protein-induced allostery in AC5. The model revealed quantitative details on the regulation mechanism and demonstrated that signal transduction in AC/G protein systems operates through the selective (de)stabilization of the particular state to which ATP preferentially binds following the classical “Monod-Wyman-Changeux” population-shift paradigm.^{83,84} Furthermore, the results provided a general framework for investigating and manipulating complex biomolecular regulations.

The procedure outlined above can be used to probe other allosteric regulations in uncharted conformational space for a wide range of complex systems. Advances in genomic sequencing make the procedure applicable to a large number^{23,24} of macromolecules for which functionally relevant transitions can be efficiently modeled and perturbed, with possible implications in parameter definition for system biology approaches.

8 | BD SIMULATIONS OF AC5

BD simulations were used to compute the bimolecular association rate constants for the binding of the G alpha protein domains to AC5.¹ BD provides a computationally efficient method to simulate the diffusional association of proteins that has been used to compute the bimolecular diffusional association rate constants of a range of protein-protein complexes.^{83–88} While the atomic-detail structures of the proteins are used in the simulations, the number of degrees of freedom is minimized by using an implicit solvent model and neglecting the internal flexibility of the proteins by representing each of them by one or a few predetermined conformations. In this study, the conformations were extracted from molecular dynamics simulations of the complexes for both apo and holo AC5 (see Section 2).^{36,37} Thus, for the

computation of rate constants the relative translational and the rotational diffusion of the two rigid-body proteins was simulated (neglecting hydrodynamic interactions) using the following equations⁸⁹:

$$\Delta r = \frac{D\Delta t}{k_b T} F + R, \quad (3)$$

where Δr is the translational displacement in timestep Δt , which is dependent on the intermolecular force, F , the relative diffusion coefficient D , and the random displacement R , for which

$$\langle R \rangle = 0 \text{ and } \langle R^2 \rangle = 6D\Delta t. \quad (4)$$

An analogous equation is used for the rotational displacement $\Delta \varphi$ of the proteins:

$$\Delta \varphi = \frac{D_{\text{Rot}}\Delta t}{k_b T} \Theta + R_{\text{rot}}. \quad (5)$$

Diffusion coefficients at infinite dilution were computed for each of the reaction species from their structures using HYDROPRO.^{90,91} For computational efficiency, the forces between the proteins were calculated from precomputed potentials on grids of points encompassing each of the proteins. The forces were the sum of and nonpolar desolvation terms⁹² and electrostatic interaction and electrostatic desolvation terms,⁹³ computed using the ECM effective charges to approximate the intermolecular forces calculated from numerical solution of the linearized Poisson–Boltzmann equation.⁹⁴ A variable timestep was used that depended on the distance between the proteins. The bimolecular diffusional association rate constant, k_{on} , was computed using the NAM method⁹⁵ as:

$$k_{\text{on}} = k_0(b) \frac{\beta}{1 - \frac{(1-\beta)k_0(b)}{k_0(c)}}, \quad (6)$$

where $k_0(b)$ and $k_0(c)$ are the rate constants for the association of the two molecules to separations of b and c , respectively, computed analytically using the Smoluchowski equation:

$$k_0(r) = 4\pi D r, \quad (7)$$

and β , the probability of binding after reaching the separation b , was computed from a large number (50,000 for each MD snapshot) of BD trajectories generated using the SDA 7 software⁹⁶ and starting with the proteins a distance b apart in random orientations and ending when the proteins reached a separation c . The satisfaction of reaction criteria for encounter complex formation was monitored to compute β .

The BD simulations were used here to provide an efficient prediction of the forward rate constants in the systems biology network model (Section 9). The computed rate constants were similar for the binding of $G\alpha_s$ and $G\alpha_i$ to AC5 with values of around $10^7 \text{ M}^{-1} \text{ s}^{-1}$. The values were not significantly affected by the binding of the other $G\alpha$ subunit or the presence of ATP in the active site of AC5. Clearly, the BD approach neglected the membrane anchoring of the proteins and any postdiffusional contributions to association rates, such as induced fit or conformational gating due to conformational changes of the binding groove of AC5 upon binding the $G\alpha$ subunit. Nevertheless, sensitivity analysis for the computed parameters showed that the values computed from the BD simulations were consistent with the ability of AC to perform coincidence detection effectively (see Section 9). Coordinate and input files for performing the BD calculations are available in the live paper at (<https://live-papers.brainsimulation.eu/#2019-bruce-et-al>).¹

9 | SIGNAL TRANSDUCTION NETWORK INVOLVING THE AC ENZYME

The molecular simulation methods described above are crucial to obtain an integrated view on how different G proteins regulate AC activity together and thus how cAMP is produced (Figure 7). The acquired results can be incorporated into kinetic models and simulations of the receptor induced signaling pathways/networks, where cAMP is one important

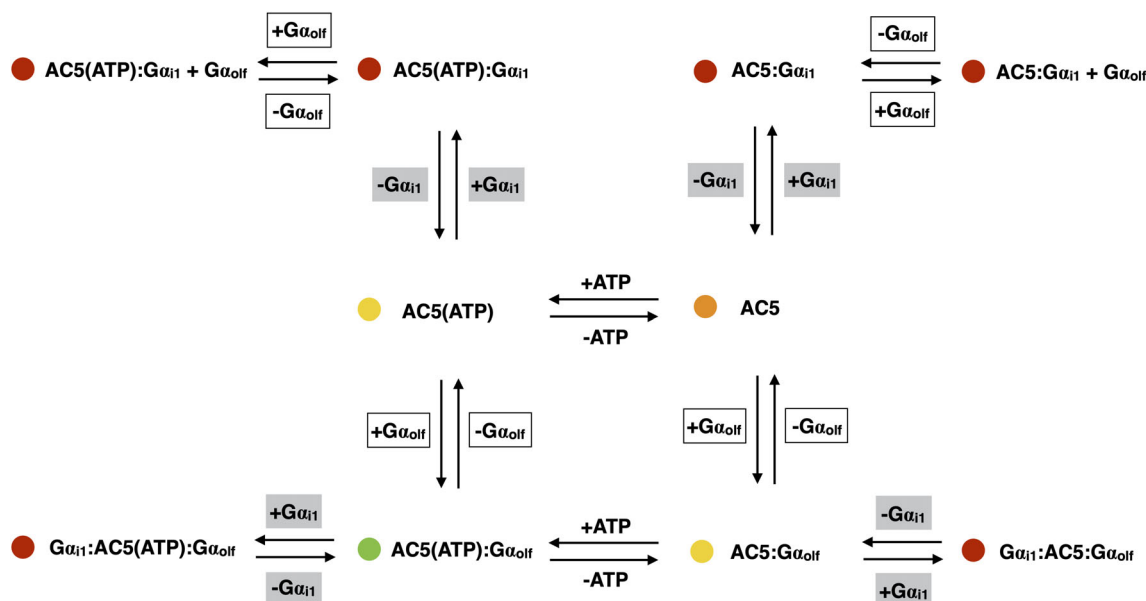


FIGURE 7 Scheme showing the catalytic probability of AC5 in the presence of different interaction partners. The catalytic activity is shown for a fully catalytically active conformation (green), AC5 states with reduced catalytic activity (yellow, orange) and the inactive form of AC5 (red). In this scheme, the results of all all-atom simulations supported by the NMA and coevolutionary-driven metadynamics studies are summarized. When AC5 is not associated to adenosine triphosphate (ATP) or to the olfactory $\text{G}\alpha$ subunit, the enzyme's conformation has a low probability to convert ATP to cAMP and is therefore depicted in orange. However, the active site is not perturbed as in all $\text{G}\alpha_i$ complexed states, which leaves the door open for $\text{G}\alpha$ -subunit and/or ATP association. When ATP binds, AC5 (ATP), the catalytic core is stabilized by the substrate's presence and increases the probability of ATP conversion, which leads to a more active conformation than AC5 in its isolated form (AC5). However, $\text{G}\alpha_{\text{olf}}$ association is required to form a fully functional catalytic pocket in AC5. Independent of the state or complex AC5 is in, whenever $\text{G}\alpha_i$ associates to AC5, the active site is perturbed due to the presence of the inhibitory $\text{G}\alpha_i$ subunit, leading to an AC5 catalytic pocket that is unfavorable for ATP conversion (depicted in red).

molecule, to determine whether and how the characteristics of the AC machinery affect the cAMP-dependent function of neurons and synapses.

Kinetic models are phenomenological models of a network of chemical reactions, which most often assume spatial homogeneity of the reactants, that is, macroscopically describe a well-mixed solution of freely diffusing reactants. They are formulated by translating the chemical reaction network into deterministic ordinary differential equations and then fitting the equation parameters to available experimental data. These models enable the study of reaction kinetics, that is, of how the concentration of all the compounds changes during the course of the reaction. A simple example of such a model is the one for a single reversible bimolecular reaction:



$$v = k_f[S_1][S_2] - k_r[P]^m, \quad (9)$$

$$\frac{dS_1}{dt} = \frac{dS_2}{dt}, \quad \frac{dP}{dt} = mv, \quad (10)$$

where S_1 and S_2 are two substrates, m is the number of molecules of the product of the reaction, k_f and k_r are the rate constants of the forward and reverse reaction, and v is the reaction rate.⁹⁷ Kinetic models are based on the law of mass action, which is the idea that the reaction rate is proportional to the probability of a collision of the reactants, which itself is proportional to the product of the concentrations of the interacting reactants. The rate constants describe how often a collision results in a transformation of the reactants, and are dependent on the temperature of the reaction.

Even though the cellular environment around signal transduction networks is highly structured, often involving precoupled molecular complexes with only some molecular species freely diffusing, kinetic models can nevertheless be tuned to phenomenologically match experimentally measured concentrations of each molecular species by

appropriately tuning the rate constants in the model. To study the functional consequences of the results from the MD and BD simulations (Sections 5 and 8), we built models of the AC signal transduction network where we compared the regulatory mechanism implied by those results to a null hypothesis. Namely, the results from the MD simulations suggested that the ternary complex could exist and that it is catalytically inactive,¹ and the null hypothesis used for comparison was one where the ternary complex cannot form or is only present in negligibly low concentration, that is, only binary complexes of AC5 with the regulatory G proteins occur.

A schematic of the signal transduction network in striatal projection neurons of the direct pathway (dSPNs) is given in Figure 8. It involves AC5. This network is the gate for synaptic potentiation, that is, significant amounts of cAMP need to be produced in order to enable the strengthening of corticostriatal synapses on dSPNs. Indeed, AC5 on its own has a low catalytic rate that produces a basal level of cAMP, not enough to allow synaptic potentiation.

The network starts with two GPCRs, one of which is a receptor for the neurotransmitter dopamine (Da) and the other for acetylcholine (ACh). Da allows for dopamine to bind the Da receptor and activate the $G\alpha_{olf}$ protein, which in turn can stimulate AC5, while ACh activates $G\alpha_i$ and hence inhibits AC5.

Both neurotransmitters have baseline tonic levels in the striatum. Da may not bind the Da receptor much at its baseline concentration, while ACh may do so for its target receptor.^{98,99} Hence, in the basal state of the signaling cascade the ACh/ $G\alpha_i$ branch is active, that is, with an active $G\alpha_i$ protein bound to AC5 and thus inhibiting it. The relevant transient neuromodulatory signals for this cascade are an elevation of dopamine (Da \uparrow , signaling a rewarding environmental stimulus for the organism) and a pause, or decrease of acetylcholine (ACh \downarrow). Da \uparrow stimulates AC5, while ACh \downarrow allows for freeing the ACh receptor of its ligand and thus stopping the activation of $G\alpha_i$ and the inhibition of AC5. The ability of the AC5 enzyme to perform coincidence detection of the two signals Da \uparrow and ACh \downarrow can be quantified in terms of the amount of cAMP produced when the two transient signals occur almost simultaneously at the cell-surface receptors.

Comparison of the regulatory mechanism with a ternary complex to the null hypothesis revealed that this regulatory mechanism indeed confers the ability for coincidence detection since a cAMP response is produced only when both signals occur. This is a consequence of the catalytic inactivity of the ternary complex as predicted by the atomistic simulation studies. Both an ACh \downarrow for relieving the inhibition of the AC5 enzyme and a Da \uparrow for significantly activating it

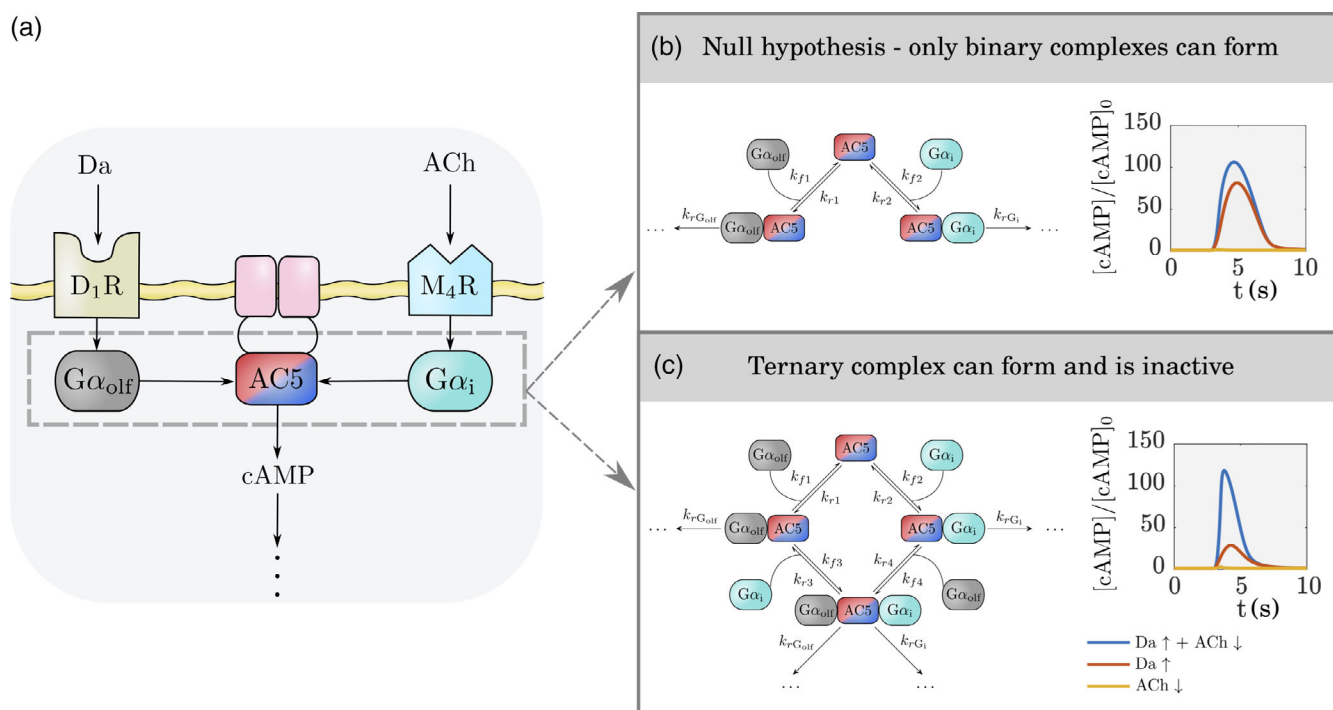


FIGURE 8 (a) Schematic of the AC5 signal transduction network in the direct pathway striatal projection neurons (dSPN). (b) Null hypothesis for the regulatory mechanism - only binary complexes can be formed between AC5 and the G proteins. (c) Regulatory mechanism where the ternary complex can form. The amount of cAMP predicted by each model is shown in (b) and (c). Illustration adapted from Bruce et al.¹

must occur in order to produce high levels of cAMP. The null hypothesis mechanism without a ternary complex can produce comparably high levels of cAMP for both $Da \uparrow$ alone and $Da \uparrow + ACh \downarrow$, not discriminating between these two situations. Additional simulations revealed that hindering $G\alpha_{olf}$ binding by the allosteric effects of $G\alpha_i$ binding (see Sections 5 and 6) to AC5 still allows for coincidence detection.¹ Additionally, the rate constants predicted by the BD simulations (see Section 8) fall in a parameter range that enables coincidence detection.

These results reveal a crucial requirement for synaptic potentiation in dSPNs, which is the coincidence of both a $Da \uparrow$ and an $ACh \downarrow$, and may inform on the nature of the environmental stimuli that are able to elicit potentiation and trigger learning. A $Da \uparrow$ is often associated with a rewarding stimulus, whereas the $ACh \downarrow$ has not been associated with any stimulus quality yet but occurs following salient stimuli.

Lastly, it should be mentioned that deterministic kinetic models often do not capture experimentally observed heterogeneity in cellular responses, and it has been shown that in these cases stochastic models are a more satisfactory description, since stochastic effects are often the source of the heterogeneity.¹⁰⁰ However, due to the highly structured synaptic environment with many molecules prebound in multimolecular complexes, it may be that in many cases an appropriate description of the molecular machinery is as logic circuits.

10 | ISOFORMS OTHER THAN AC5

AC5 is one of nine transmembrane AC isoforms. To consider whether the findings described above, which are specific for AC5 except for the QM/MM simulations, may also apply to other AC isoforms, we analyzed the sequences, structures, and electrostatic potentials of the AC isoforms. To identify the structural features influencing binding specificity and the binding sites for regulators on AC, we performed a comparative electrostatic analysis for the catalytic domains of mouse ACs 1–9¹⁸ (for sequence alignment (see SI1 in Reference 18). Comparative modeling techniques were first used to model the complete structures of all the AC enzyme catalytic domains, including loops not present in the template crystal structure (PDB ID 1AZS).¹² These structures were superimposed and their electrostatic potentials computed by solving the linearized Poisson–Boltzmann equation with APBS (Adaptive Poisson–Boltzmann Solver).¹⁰¹

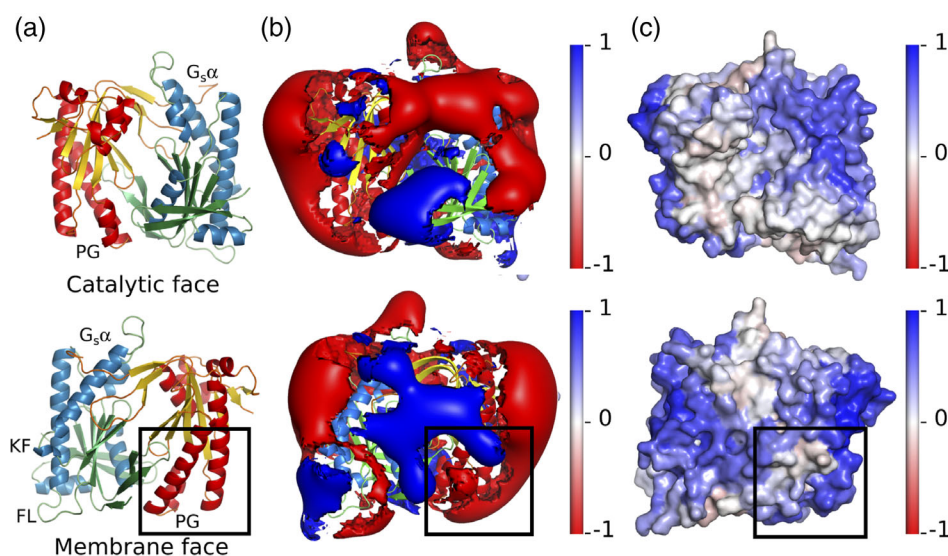


FIGURE 9 Identification of regions of conserved electrostatic potentials in $G\alpha_i$ -inhibited Adenylyl cyclases (AC) isoforms 1, 5, and 6 at the pseudo-symmetric groove (boxed). (a) AC viewed from the catalytic and membrane faces, (b) electrostatic potential of AC5 computed by solution of the linearized Poisson–Boltzmann equation and displayed as isopotential surfaces contoured at $1 kT/e$ (blue) and $-1 kT/e$ (red), (c) compactness score, C_S , mapped onto the surface of AC5.¹⁸ C_S is based on a silhouette coefficient that evaluates the compactness and degree of separation between clusters formed by grouping AC isoforms according to their electrostatic potential distributions in spheres of radius 10 angstroms centered on each solvent accessible Calpha atom of AC5. A score of -1 indicates bad clustering and a score of 1 denotes dense clusters. The region with high C_S score at the pseudo-symmetric groove corresponds to a negatively charged patch in AC1, 5 and 6 that is not found in the other AC isoforms, supporting the binding of G_i at this position (adapted and reproduced with permission from Tong et al.¹⁸).

The PIPSA (Protein Interaction Property Similarity Analysis) method^{102,103} was then employed to scan over the protein surface and compare the global and local electrostatic potential distributions across the isoforms by computing pairwise similarity indices. The PIPSA results supported the binding of $G\alpha_i$ at the pseudo-symmetric position with respect to the known $G\alpha_s$ binding site for AC isoforms 1, 5, and 6, which have a region of conserved negative potential at this position, indicating that the results described above for AC5 and its network might be similar for AC1 and AC6 (Figure 9). Other insights into the protein binding specificity of AC groups from the PIPSA analysis included the interaction of ACs 2, 4, and 7 with $G\beta\gamma$, the prediction of the binding site of RGS2, as well as of isoform-specific features of the small molecule binding sites on AC that could be exploited in the design of isoform-specific compounds.

11 | CONCLUSIONS

A wide range of molecular simulation methods was used to investigate AC structure, dynamics, function and binding to its cognate proteins $G\alpha_s/G\alpha_{olf}$ and $G\alpha_i$, with a focus on the AC5 isoform (Box 1). Quantum mechanical modeling provided a description of the free energy landscape associated with the AC enzymatic reaction as well as the catalytic rate. All-atom molecular dynamics simulations aided in explaining the mechanisms of allosteric regulation of AC5 by activating and inhibitory G proteins, the impact of ATP and the possible coupling between substrate binding and regulatory protein binding. Moreover, the classical MD results suggested that a ternary $G\alpha_s$ -AC5- $G\alpha_i$ complex can form but is catalytically inactive. NMA in internal coordinates were used to understand the global conformational changes of the AC5/ $G\alpha$ protein complexes. Metadynamics simulations along with coevolution analysis allowed us to suggest that the binding of $G\alpha$ -proteins to AC5 modulates its function via conformational selection mechanisms. BD simulations were used to estimate the rate constants for association of AC5 with activating and inhibitory $G\alpha$ proteins. The regulatory mechanism in which a catalytically inactive ternary complex can form was incorporated, together with the predicted rate constants for diffusional association, in a kinetic model of a signaling network important for brain plasticity and learning. The results showed that, with this regulatory mechanism, AC5 acts as a coincidence detector for transient

BOX 1 Section Highlights

3.1. QM/MM simulations were used to reveal relative changes in k_{cat} due to point mutations. Through the integrated approach outlined in this review, it may be possible to study how these changes are propagated at the system level.

3.2. According to classical MD simulations, ATP and regulatory G proteins impact flexibility and conformation of AC, providing structural basis for enzymatic regulation. A ternary complex consisting of AC5 simultaneously associated with active $G\alpha_{olf}$ and active myristoylated $G\alpha_i$ in its soluble form (not interacting with the membrane) may be formed in the absence or presence of ATP but is inactive.

3.3. NMA in internal coordinates, using simplified representations of the protein structures in the PaLaCe model, was used to investigate the effect of regulatory proteins on AC5 conformational flexibility.

3.4. A multiscale MD approach guided by coevolutionary data was used to reconstruct the free-energy landscape of a fully solvated atomistic model of AC5 with and without the regulatory $G\alpha$ subunit bound; here $G\alpha$ was either the stimulatory $G\alpha_s$, or the inhibitory $G\alpha_i$

3.5. BD simulations were used to provide an efficient prediction of the forward rate constants in the systems biology network model.

4.0. These results reveal a crucial requirement for synaptic potentiation in dSPNs, which is the coincidence of both a $Da \uparrow$ and an $ACh \downarrow$, and may inform on the nature of the environmental stimuli that are able to elicit potentiation and trigger learning. A $Da \uparrow$ is often associated with a rewarding stimulus, whereas the $ACh \downarrow$ has been associated with salient stimuli but not necessarily with the quality of such an event.

5.0. The PIPSA results supported the binding of $G\alpha_i$ at the pseudo-symmetric position with respect to the known $G\alpha_s$ binding site for AC isoforms 1, 5, and 6, which have a region of conserved negative potential at this position, indicating that the results described above for AC5 and its network might be similar for AC1 and AC6.

neurotransmitter signals, and suggested the kind of neurotransmitter signals required to trigger synaptic potentiation. Scale-bridging combinations of molecular modeling and simulation tools have been applied to other biochemical networks,¹⁰⁴ and the combination of methods applied to AC5 could readily be transferred to the study of protein complexes¹⁰⁵ as well as to other enzymes and their biochemical and signaling networks. The collaboration within the HBP consortium pointed to the advantage of using diverse and complementary methods—at different scales and resolutions—on the same system (the AC5 enzyme) to understand salient aspects of its function and regulation. It also shows the importance of sharing preliminary data among the partners and the synergy that can be achieved thereby. The effort has provided a basis for developing the molecular and subcellular simulation software infrastructure for the EBRAINS (<https://ebrains.eu/>) and FENIX (<https://fenix-ri.eu/>) services for multiscale brain simulation.

Taking these results together, we have shown that applying different molecular simulation techniques to a single enzyme (AC5) and its regulators provides a complementary and consistent picture of the enzyme's function (difficult or impossible to obtain by experiment), along with providing insight on its role in a specific neuronal cascade. The calculations also pointed out similarities between AC5 and some members of the AC family as regards regulatory interactions. We close by noticing that the simulation tools used here are fast improving. Particularly important is the arrival of exascale computers in the US, China, Japan and the EU, which will allow for unprecedented performance of many highly scalable codes, including MD and QM/MM. Also, the revolution in structural biology caused by advances in AI protein structure prediction, such as Alpha-fold, is paving the way toward the investigation of virtually any signaling cascade, by providing the structures of all human proteins involved in them. As a result, we expect that investigations like that presented here (which required several years) may be carried out much more efficiently (in terms of data and of codes) in the not-too-distant future.

AUTHOR CONTRIBUTIONS

Siri C. van Keulen: Writing – original draft (equal); writing – review and editing (equal). **Juliette Martin:** Writing – original draft (equal); writing – review and editing (equal). **Francesco Colizzi:** Writing – original draft (equal); writing – review and editing (equal). **Elisa Frezza:** Writing – original draft (equal); writing – review and editing (equal). **Daniel Trpevski:** Writing – original draft; writing – review and editing. **Nuria Cirauqui Diaz:** Writing – original draft (equal); writing – review and editing (equal). **Pietro Vidossich:** Writing – original draft (equal). **Ursula Rothlisberger:** Writing – original draft; writing – review and editing (equal). **Jeanette Hellgren Kotaleski:** Writing – original draft (equal); writing – review and editing (equal). **Rebecca C. Wade:** Writing – original draft; writing – review and editing (equal). **Paolo Carloni:** Writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

The simulations were performed on resources provided by the Swedish National Infrastructure for Computing at PDC (Center for Parallel Computing) and by GENCI for computing on the CINES supercomputer OCCIGEN (Grants 2016-c201607758, 2017-A0020707585, and 2018-A0040710357) and on the Swiss National Computing Center CSCS. The ICM-CSIC is recipient of the Severo Ochoa Centre of Excellence accreditation (CEX2019-000928-S) from the Spanish Ministry of Science and Innovation. We acknowledge the use of Fenix Research Infrastructure resources, which are partially funded from the European Union's Horizon 2020 research and innovation program through the ICEI project under the grant agreement No. 800858. We thank the Human Brain Project for supporting all the research projects presented here. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

Jeanette Hellgren Kotaleski, Paolo Carloni, Rebecca C. Wade: Horizon 2020 Framework Programme (945,539, HBP SGA3); Jeanette Hellgren Kotaleski: Swedish Research Council (VR-M-2017-02806, VR-M-2020-01652); Swedish e-science Research Center (SeRC). Francesco Colizzi is a Ramón y Cajal Fellow (RYC2019-026768-I). Rebecca C. Wade: Klaus Tschira Foundation.











CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Siri C. van Keulen  <https://orcid.org/0000-0001-6995-8389>
 Juliette Martin  <https://orcid.org/0000-0002-4787-0885>
 Francesco Colizzi  <https://orcid.org/0000-0001-5601-1452>
 Elisa Frezza  <https://orcid.org/0000-0003-0122-7859>
 Daniel Trpevski  <https://orcid.org/0000-0001-9068-6744>
 Nuria Cirauqui Diaz  <https://orcid.org/0000-0003-4752-9916>
 Ursula Rothlisberger  <https://orcid.org/0000-0002-1704-8591>
 Jeanette Hellgren Kotalleski  <https://orcid.org/0000-0002-0550-0739>
 Rebecca C. Wade  <https://orcid.org/0000-0001-5951-8670>
 Paolo Carloni  <https://orcid.org/0000-0002-9010-0149>

RELATED WIREs ARTICLES

[Biomolecular simulations: From dynamics and mechanisms to computational assays of biological activity](#)

ENDNOTES

- Other protein regulators include the regulator of G protein signaling protein 2 (RGS2),⁶ Ca²⁺/calmodulin-dependent protein kinase II (CaMKII), the protein kinases A and C (PKA and PKC) and calmodulin.⁷
- The cryo-EM structure of full-length AC9 enzyme, structurally different from all the other transmembrane members of the AC family, bound to G_{oIR} appeared after most of the presented research on AC5 was carried out.¹⁶ It confirmed the architecture of the catalytic domain of AC and the binding mode of G_{αs} observed in the absence of the transmembrane domain (See Section S1 in Appendix S1 for details).
- Predicting the structure of the ATP-bound state is currently not feasible as small molecule ligands cannot yet be included in this method.
- Indeed, iNMA can be successfully applied to homology models and it is insensitive to modeling errors (65). In particular, we studied the ability of iNMA to describe protein motions and flexibility in protein models taken from the CASP assessment, comparing them with their native structure.⁶³ NMA turned out not to be much affected by modeling errors, and it is accurate on protein models when the C α atoms' RMSD to the native structure is lower than 2–3 Å. This is the case for most homology models with target/template sequence identity higher than 60%.⁶³

REFERENCES

- Bruce NJ, Narzi D, Trpevski D, van Keulen SC, Nair AG, Röhrlisberger U, et al. Regulation of adenylyl cyclase 5 in striatal neurons confers the ability to detect coincident neuromodulatory signals. *PLoS Comput Biol*. 2019;15(10):e1007382.
- Hanoune J, Defer N. Regulation and role of adenylyl cyclase isoforms. *Annu Rev Pharmacol Toxicol*. 2001;41(1):145–74.
- Wu ZL, Thomas SA, Villacres EC, Xia Z, Simmons ML, Chavkin C, et al. Altered behavior and long-term potentiation in type I adenylyl cyclase mutant mice. *Proc Natl Acad Sci*. 1995;92(1):220–4.
- Iwamoto T, Okumura S, Iwatsubo K, Kawabe JI, Ohtsu K, Sakai I, et al. Motor dysfunction in type 5 adenylyl cyclase-null mice. *J Biol Chem*. 2003;278(19):16936–40.
- Bakalyar HA, Reed RR. Identification of a specialized adenylyl cyclase that may mediate odorant detection. *Science*. 1990 Dec 7; 250(4986):1403–6.
- Kehrl J. RGS2: a multifunctional regulator of G-protein signaling. *Int J Biochem Cell Biol*. 2002;34(5):432–8.
- Sadana R, Dessauer CW. Physiological roles for G protein-regulated adenylyl cyclase isoforms: insights from knockout and over-expression studies. *Neurosignals*. 2009;17(1):5–22.
- Simon MI, Strathmann MP, Gautam N. Diversity of G proteins in signal transduction. *Science*. 1991;252(5007):802–8.
- Oldham WM, Hamm HE. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol*. 2008;9(1):60–71.
- Sunahara RK, Dessauer CW, Gilman AG. Complexity and diversity of mammalian adenylyl cyclases. *Annu Rev Pharmacol Toxicol*. 1996;36(1):461–80.
- Tesmer JJJ, Sunahara RK, Johnson RA, Gosselin G, Gilman AG, Sprang SR. Two-metal-ion catalysis in adenylyl cyclase. *Science*. 1999;285(5428):756–60.
- Tesmer JJJ, Sunahara RK, Gilman AG, Sprang SR. Crystal structure of the catalytic domains of adenylyl cyclase in a complex with G_{src}·GTP γ S. *Science*. 1997;278(5345):1907–16.
- Zhang G, Liu Y, Ruoho AE, Hurley JH. Structure of the adenylyl cyclase catalytic core. *Nature*. 1997;386(6622):247–53.
- Tesmer JJJ, Dessauer CW, Sunahara RK, Murray LD, Johnson RA, Gilman Alfred G, et al. Molecular basis for P-site inhibition of adenylyl cyclase. *Biochemistry*. 2000;39(47):14464–71.

15. Mou TC, Masada N, Cooper DM, Sprang SR. Structural basis for inhibition of mammalian adenylyl cyclase by calcium. *Biochemistry*. 2009;48(15):3387–97.
16. Qi C, Sorrentino S, Medalia O, Korkhov VM. The structure of a membrane adenylyl cyclase bound to an activated stimulatory G protein. *Science*. 2019;364(6438):389–94.
17. Linder JU. Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation. *Cell Mol Life Sci*. 2006;63(15):1736–51.
18. Tong R, Wade RC, Bruce NJ. Comparative electrostatic analysis of adenylyl cyclase for isoform dependent regulation properties: electrostatic analysis of adenylyl Cyclases. *Proteins Struct Funct Bioinforma*. 2016;84(12):1844–58.
19. Coleman DE, Berghuis AM, Lee E, Linder ME, Gilman AG. Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science*. 1994;265(5177):1405–12.
20. Preininger AM, Van Eps N, Yu NJ, Medkova M, Hubbell WL, Hamm HE. The myristoylated amino terminus of Gαi1 plays a critical role in the structure and function of Gαi1 subunits in solution. *Biochemistry*. 2003;42(26):7931–41.
21. Dessauer CW, Tesmer JGG, Sprang SR, Gilman AG. Identification of a Giα binding site on type V adenylyl cyclase. *J Biol Chem*. 1998;273(40):25831–9.
22. Taussig R, Iniguez-Lluhi JA, Gilman AG. Inhibition of adenylyl cyclase by Gi alpha. *Science*. 1993;261(5118):218–21.
23. Colizzi F, Orozco M. Probing allosteric regulations with coevolution-driven molecular simulations. *Sci Adv*. 2021;7(37):eabj0786.
24. Sfriso P, Duran-Frigola M, Mosca R, Emperador A, Aloy P, Orozco M. Residues coevolution guides the systematic identification of alternative functional conformations in proteins. *Structure*. 2016;24(1):116–26.
25. Warshel A, Levitt M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol*. 1976;103(2):227–49.
26. Carloni P, Rothlisberger U, Parrinello M. The role and perspective of ab initio molecular dynamics in the study of biological systems. *Acc Chem Res*. 2002;35(6):455–64.
27. Bolnykh V, Olsen JMH, Meloni S, Bircher MP, Ippoliti E, Carloni P, et al. Extreme scalability of DFT-based QM/MM MD simulations using MiMiC. *J Chem Theory Comput*. 2019;15(10):5601–13.
28. Bolnykh V, Rossetti G, Rothlisberger U, Carloni P. Expanding the boundaries of ligand–target modeling by exascale calculations. *WIREs Comput Mol Sci*. 2021;11(4):e1535. <https://doi.org/10.1002/wcms.1535>
29. Chiariello MG, Bolnykh V, Ippoliti E, Meloni S, Olsen JMH, Beck T, et al. Molecular basis of CLC antiporter inhibition by fluoride. *J Am Chem Soc*. 2020;142(16):7254–8.
30. Stewart JJP. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J Mol Model*. 2007;13(12):1173–213.
31. Dessauer CW, Gilman AG. The catalytic mechanism of mammalian adenylyl cyclase. Equilibrium binding and kinetic analysis of P-site inhibition. *J Biol Chem*. 1997;272(44):27787–95.
32. Yan SZ, Huang ZH, Shaw RS, Tang WJ. The conserved asparagine and arginine are essential for catalysis of mammalian adenylyl cyclase. *J Biol Chem*. 1997;272(19):12342–9.
33. Frezza E, Martin J, Lavery R. A molecular dynamics study of adenylyl cyclase: the impact of ATP and G-protein binding. *PLoS One*. 2018;13(4):e0196207.
34. Frezza E, Amans TM, Martin J. Allosteric inhibition of adenylyl cyclase type 5 by G-protein: a molecular dynamics study. *Biomolecules*. 2020;10(9):1330.
35. van Keulen SC, Rothlisberger U. Effect of N-terminal myristoylation on the active conformation of Gαi1-GTP. *Biochemistry*. 2017;56(1):271–80.
36. Narzi D, van Keulen SC, Røthlisberger U. Gαi1 inhibition mechanism of ATP-bound adenylyl cyclase type 5. *PLoS One*. 2021;(1), 16, e0245197.
37. van Keulen SC, Narzi D, Rothlisberger U. Association of both inhibitory and stimulatory Gα subunits implies adenylyl cyclase 5 deactivation. *Biochemistry*. 2019;58:4317–24. <https://doi.org/10.1021/acs.biochem.9b00662>
38. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct Funct Bioinforma*. 2010;78(8):1950–8.
39. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem*. 2000;21(12):1049–74.
40. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct Funct Bioinforma*. 2006;65(3):712–25.
41. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic validation of protein force fields against experimental data. *PLoS One*. 2012;7(2):e32131.
42. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun*. 1995;91(1–3):43–56.
43. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4(3):435–47.
44. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013;29(7):845–54.
45. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701–18.

46. Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993 Dec;234(3):779–815.
47. Alekseenko A, Ignatov M, Jones G, Sabitova M, Kozakov D. Protein–protein and protein–peptide docking with ClusPro server. In: Kihara D, editor. Protein structure prediction [internet]. *Methods in molecular biology.* Volume 2165. New York: Springer US; 2020. p. 157–74. https://doi.org/10.1007/978-1-0716-0708-4_9
48. de Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc.* 2010;5(5):883–97.
49. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastiris PL, Karaca E, et al. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol.* 2016;428(4):720–5.
50. van Keulen SC, Rothlisberger U. Exploring the inhibition mechanism of adenylyl cyclase type 5 by N-terminal Myristoylated G_{i1}:GTP. *PLoS Comp Biol.* 2017;13(9):e1005673.
51. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–9.
52. Marx V. Method of the year: protein structure prediction. *Nat Methods.* 2022;19(1):5–10.
53. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer [internet]. *Bioinformatics.* 2021. <https://doi.org/10.1101/2021.10.04.463034>
54. Bahar I, Lezon TR, Bakan A, Shrivastava IH. Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem Rev.* 2010;110(3):1463–97.
55. Micheletti C, Carloni P, Maritan A. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. *Proteins.* 2004;55(3):635–45.
56. Hoffmann A, Grudinin S. NOLB: nonlinear rigid block normal-mode analysis method. *J Chem Theory Comput.* 2017;13(5):2123–34.
57. Kmiecik S, Gront D, Kolinski M, Wieteska L, Dawid AE, Kolinski A. Coarse-grained protein models and their applications. *Chem Rev.* 2016;116(14):7898–936.
58. Pasi M, Lavery R, Ceres N. PaLaCe: a coarse-grain protein model for studying mechanical properties. *J Chem Theory Comput.* 2013;9(1):785–93.
59. Frezza E, Lavery R. Internal Normal mode analysis (iNMA) applied to protein conformational flexibility. *J Chem Theory Comput.* 2015;11(11):5503–12.
60. Berman HM. The protein data bank. *Nucleic Acids Res.* 2000;28(1):235–42.
61. Frezza E, Lavery R. Internal coordinate normal mode analysis: a strategy to predict protein conformational transitions. *J Phys Chem B.* 2019;123(6):1294–301.
62. Cirauqui Diaz N, Frezza E, Martin J. Using normal mode analysis on protein structural models. How far can we go on our predictions? *Proteins Struct Funct Bioinforma.* 2021;89(5):531–43.
63. Schwede T, Diemand A, Guex N, Peitsch MC. Protein structure computing in the genomic era. *Res Microbiol.* 2000;151(2):107–12.
64. de Juan D, Pazos F, Valencia A. Emerging methods in protein co-evolution. *Nat Rev Genet.* 2013;14(4):249–61.
65. Sulkowska JJ, Morcos F, Weigt M, Hwa T, Onuchic JN. Genomics-aided structure prediction. *Proc Natl Acad Sci.* 2012;109(26):10340–5.
66. Morcos F, Pagnani A, Lunt B, Bertolino A, Marks DS, Sander C, et al. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc Natl Acad Sci.* 2011;108(49):E1293–301.
67. Sutto L, Marsili S, Valencia A, Gervasio FL. From residue coevolution to protein conformational ensembles and functional dynamics. *Proc Natl Acad Sci.* 2015;112(44):13567–72.
68. Morcos F, Jana B, Hwa T, Onuchic JN. Coevolutionary signals across protein lineages help capture multiple protein conformations. *Proc Natl Acad Sci.* 2013;110(51):20533–8.
69. Cutarello F, Tiana G, Bussi G. Assessing the accuracy of direct-coupling analysis for RNA contact prediction. *RNA.* 2020;26(5):637–47.
70. Weinreb C, Riesselman AJ, Ingraham JB, Gross T, Sander C, Marks DS. 3D RNA and functional interactions from evolutionary couplings. *Cell.* 2016;165(4):963–75.
71. Mirdita M, von den Driesch L, Galiez C, Martin MJ, Söding J, Steinegger M. Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* 2017;45(D1):D170–6.
72. Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods.* 2012;9(2):173–5.
73. Weigt M, White RA, Szurmant H, Hoch JA, Hwa T. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc Natl Acad Sci.* 2009;106(1):67–72.
74. Proctor EA, Ding F, Dokholyan NV. Discrete molecular dynamics. *WIREs Comput Mol Sci.* 2011;1(1):80–92.
75. Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Discrete molecular dynamics studies of the folding of a protein-like model. *Fold Des.* 1998;3(6):577–87.
76. Giorgino T. PLUMED-GUI: an environment for the interactive development of molecular dynamics analysis and biasing scripts. *Comput Phys Commun.* 2014;185(3):1109–14.
77. Branduardi D, Gervasio FL, Parrinello M. From A to B in free energy space. *J Chem Phys.* 2007;126(5):054103.
78. Bussi G, Laio A. Using metadynamics to explore complex free-energy landscapes. *Nat Rev Phys.* 2020;2(4):200–12.
79. Laio A, Parrinello M. Escaping free-energy minima. *Proc Natl Acad Sci.* 2002;99(20):12562–6.

80. Barducci A, Bussi G, Parrinello M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett*. 2008;100(2):020603.
81. The PLUMED Consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods*. 2019;16(8):670–3.
82. Khannpnavar B, Mehta V, Qi C, Korkhov V. Structure and function of adenylyl cyclases, key enzymes in cellular signaling. *Curr Opin Struct Biol*. 2020;63:34–41.
83. Northrup SH, Luton JA, Boles JO, Reynolds JCL. Brownian dynamics simulation of protein association. *J Comput Aided Mol Des*. 1988;1(4):291–311.
84. Gabdoulline RR, Wade RC. Protein-protein association: investigation of factors influencing association rates by Brownian dynamics simulations. *J Mol Biol*. 2001;306(5):1139–55.
85. Spaar A, Dammer C, Gabdoulline RR, Wade RC, Helms V. Diffusional encounter of Barnase and Barstar. *Biophys J*. 2006;90(6):1913–24.
86. Gabdoulline RR, Wade RC. Biomolecular diffusional association. *Curr Opin Struct Biol*. 2002;12(2):204–13.
87. Zhou HX, Bates PA. Modeling protein association mechanisms and kinetics. *Curr Opin Struct Biol*. 2013;23(6):887–93.
88. Öztürk MA, Wade RC. Computation of FRAP recovery times for linker histone – chromatin binding on the basis of Brownian dynamics simulations. *Biochim Biophys Acta BBA - Gen Subj*. 2020;1864(10):129653.
89. Ermak DL, McCammon JA. Brownian dynamics with hydrodynamic interactions. *J Chem Phys*. 1978;69(4):1352–60.
90. García de la Torre J, Huertas ML, Carrasco B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J*. 2000;78(2):719–30.
91. Ortega A, Amorós D, García de la Torre J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J*. 2011;101(4):892–8.
92. Gabdoulline RR, Wade RC. On the contributions of diffusion and thermal activation to electron transfer between *Phormidium laminosum* Plastocyanin and cytochrome *f*: Brownian dynamics simulations with explicit modeling of nonpolar Desolvation interactions and electron transfer events. *J Am Chem Soc*. 2009;131(26):9230–8.
93. Elcock AH, Gabdoulline RR, Wade RC, McCammon JA. Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. *J Mol Biol*. 1999;291(1):149–62.
94. Gabdoulline RR, Wade RC. Effective charges for macromolecules in solvent. *J Phys Chem*. 1996;100(9):3868–78.
95. Northrup SH, Allison SA, McCammon JA. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J Chem Phys*. 1984;80(4):1517–24.
96. Martinez M, Bruce NJ, Romanowska J, Kokh DB, Ozboyaci M, Yu X, et al. SDA 7: a modular and parallel implementation of the simulation of diffusional association software. *J Comput Chem*. 2015;36(21):1631–45.
97. Klipp E, Liebermeister W, Wierling C, Kowald A. *Systems biology: a textbook*. 2nd ed. Weinheim: Wiley-VCH; 2016. 488 pp.
98. Yapo C, Nair AG, Clement L, Castro LR, Hellgren Kotaleski J, Vincent P. Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. *J Physiol*. 2017;595(24):7451–75.
99. Nair AG, Castro LRV, El Khoury M, Gorgievski V, Giros B, Tzavara ET, et al. The high efficacy of muscarinic M4 receptor in D1 medium spiny neurons reverses striatal hyperdopaminergia. *Neuropharmacology*. 2019;1(146):74–83.
100. Wilkinson DJ. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat Rev Genet*. 2009;10(2):122–33.
101. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci*. 2001;98(18):10037–41.
102. Blomberg N, Gabdoulline RR, Nilges M, Wade RC. Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins Struct Funct Genet*. 1999;37(3):379–87.
103. Wade RC, Gabdoulline RR, De Rienzo F. Protein interaction property similarity analysis. *Int J Quantum Chem*. 2001;83(3–4):122–7.
104. Stein M, Gabdoulline RR, Wade RC. Bridging from molecular simulation to biochemical networks. *Curr Opin Struct Biol*. 2007;17(2):166–72.
105. Swanson JMJ. Multiscale kinetic analysis of proteins. *Curr Opin Struct Biol*. 2022;72:169–75.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: van Keulen SC, Martin J, Colizzi F, Frezza E, Trpevski D, Diaz NC, et al. Multiscale molecular simulations to investigate adenylyl cyclase-based signaling in the brain. *WIREs Comput Mol Sci*. 2022. e1623. <https://doi.org/10.1002/wcms.1623>