



**HAL**  
open science

## La surface d'énergie potentielle vue par les champs de forces

Isabelle Demachy, Jean-Philip Piquemal

► **To cite this version:**

Isabelle Demachy, Jean-Philip Piquemal. La surface d'énergie potentielle vue par les champs de forces. L'Actualité Chimique, 2014, 388-389, pp.37-42. hal-00960389

**HAL Id: hal-00960389**

**<https://hal.science/hal-00960389v1>**

Submitted on 18 Mar 2014

**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.

## TITRE

### La surface d'énergie potentielle vue par les champs de forces

Isabelle Demachy, Laboratoire de Chimie Physique, Université Paris-Sud, 91405 Orsay

Jean-Philip Piquemal, Laboratoire de Chimie Théorique, Sorbonne Universités, Université Pierre et Marie Curie, CC 137, 75252 Paris

#### Abstract français

Dans le domaine de la chimie théorique, le champ de forces représente l'énergie potentielle d'une molécule déformable ou d'un ensemble de molécules en interaction, sans faire intervenir directement la mécanique quantique. Il décrit l'énergie potentielle selon une expression mathématique, fonction des coordonnées de déformation des molécules et des distances intermoléculaires et, dépendante de paramètres caractéristiques des forces d'interaction entre atomes. Expression approchée de l'énergie potentielle d'un système moléculaire, un champ de force est un subtil équilibre entre simplicité et représentativité physicochimique. Cet article retrace, des concepts de base aux enjeux actuels, les différents développements et spécificités des champs de forces.

#### Abstract anglais

In the field of theoretical chemistry, a force field represents the potential energy surface of a non-rigid molecule or of molecular fluids without any explicit intervention of quantum mechanics. A molecular mechanical force field describes the potential energy through an analytical expression of the deformation coordinates and the intermolecular distances. It depends on parameters representing molecular properties and interactions between atoms. This approximated expression results from a subtil balance between simplicity and realism. This article outlines both the basic concepts and the present challenges and stresses the major aspects of force field development and properties.

Mots clés :

**Energie potentielle / champ de forces/conformations/structure/dynamique moléculaire/  
systèmes moléculaires complexes**

## I- Introduction

Prédire l'enthalpie de vaporisation dans diverses conditions thermodynamiques, déduire la structure d'un polypeptide à partir de données RMN, déterminer la constante de vitesse de processus élémentaires, le chimiste peut le faire grâce à la modélisation moléculaire.

En effet, la modélisation moléculaire en chimie donne accès aux grandeurs structurales, thermodynamiques ou cinétiques observables à partir des interactions entre particules (noyaux et électrons) en explicitant les lois fondamentales de la physique à l'échelle atomique sous forme de modèles mathématiques et numériques. Une quantité essentielle dans cette approche est l'énergie du

système moléculaire. Considérant que la dynamique des noyaux est beaucoup plus lente que celle des électrons (approximation Born Oppenheimer), l'énergie s'écrit comme la somme de deux contributions, l'énergie potentielle dite totale et l'énergie cinétique des noyaux. L'énergie potentielle est définie pour une géométrie fixée des noyaux, elle inclut les interactions entre toutes les particules et s'écrit comme la somme de l'énergie électronique et de l'énergie coulombienne répulsive entre noyaux. Pour chaque état électronique, elle varie spécifiquement avec la géométrie du système et s'exprime donc comme une fonction des positions des noyaux, définissant l'hypersurface d'énergie potentielle.

De manière générale, il n'est pas nécessaire de connaître l'hypersurface complète et la connaissance des régions de la surface d'énergie potentielle au voisinage des minima et des chemins qui les relie suffisent pour déterminer les propriétés. Pour autant, il faut savoir identifier ces régions et les explorer efficacement (voir dans ce numéro les articles sur les méthodes de détermination de la structure électronique), ce qui pour un système chimique de grande taille (liquide, protéines en solution, interface solide-liquide...) va nécessiter un très grand nombre de calcul de l'énergie potentielle.

Le nombre de degrés de liberté c'est-à-dire le nombre de coordonnées dont dépend l'énergie potentielle est alors un paramètre clé pour le choix de la méthode de calcul de cette grandeur. Les méthodes dites *ab initio* ou de premiers principes, basées sur une représentation explicite des électrons et la résolution de l'équation de Schrödinger, ont un coût de calcul important et seront limitées à des systèmes d'une centaine d'atomes. Au-delà, se situe une activité importante de la chimie théorique qui consiste à définir aussi rigoureusement que possible une expression analytique de l'énergie potentielle. Elle est constituée d'un ensemble de termes mathématiques paramétrés décrivant en fonction des coordonnées des atomes ou groupements d'atomes, leurs interactions de liaisons ou non liées. C'est ce qu'on appelle un champ de forces car de cette énergie potentielle dérive les forces appliquées aux atomes ou groupements d'atomes, bases des algorithmes de dynamique moléculaire.

Les champs de forces sont devenus des outils essentiels aux études de phénomènes complexes dans de nombreux domaines dont la science des matériaux ou les simulations de macromolécules biologiques. La première simulation de protéines a été réalisée en 1970 sur le calculateur le plus puissant du moment pour un système de 458 atomes et pendant 9,2 ps [1]. En 2006, une dynamique moléculaire de plus d'un million d'atomes avec un champ de forces à l'échelle atomique a permis d'étudier la structure du virus mosaïque du tabac [2]. Ces avancées se font notamment par l'amélioration constante des champs de forces.

Les champs de forces doivent répondre à deux enjeux majeurs : permettre une bonne description des propriétés tant à l'échelle atomique (géométries, vibrations, conformations) qu'à l'échelle macroscopique et thermodynamique (enthalpies de vaporisation ou sublimation, énergie et entropie de solvation, moments dipolaires, paramètres spectroscopiques). C'est le choix de la forme mathématique et de la méthode de paramétrisation qui détermine alors la précision et la pertinence d'un champ de forces.

Le développement de champs de forces est un champ très actif en chimie théorique « qui relève autant de l'art que de la science » [3]. Mais avec l'accélération des simulations moléculaires elles-mêmes, la voie est ouverte pour une introspection réaliste des relations structure-dynamique-

propriétés de phases condensées complexes, là où les données expérimentales sont rares, inaccessibles ou ininterprétables sans modèle théorique.

## II- La philosophie des champs de forces classiques : concepts et hypothèses de base

Tant qu'un système moléculaire n'est le siège d'aucune réorganisation électronique (rupture ou formations de liaisons, transfert de charge), l'énergie potentielle varie selon des lois classiques lors d'une déformation intramoléculaire ou d'un réarrangement des molécules entre elles. Par exemple, l'élongation d'une liaison qui est décrite précisément par une courbe de Morse peut au voisinage de l'équilibre être remplacée par un terme harmonique (voir encart 1). C'est ainsi que Alder et Wainwright ont été les premiers à la fin des années 50 [4] à décrire une molécule comme un ensemble de boules et de ressorts, chaque boule correspondant à un atome et chaque ressort à une liaison covalente. Par ailleurs, en absence de réaction chimique, les longueurs de liaison C-C, C=C, C=O,..., ainsi que les angles de valence sont généralement très peu dépendants de la molécule considérée et de l'environnement. C'est la notion de groupement fonctionnel si essentielle en chimie qui va ici aussi simplifier et généraliser l'utilisation des champs de forces.

Les champs de forces les plus courants, dits 'tout atome', utilisent l'atome comme grain de base et définissent un ensemble de type d'atomes selon la nature des liaisons qu'ils forment. Ils décrivent l'énergie potentielle de déformation intramoléculaire par une somme de fonctions analytiques de type énergie d'élongation des liaisons, énergie de déformation des angles de valence et énergie de déformation des angles dièdres (encart 1) dont les paramètres ( $k$ ,  $l_0$ ,  $\theta_0$ ,  $n$ ...) sont déterminés pour un ensemble choisi de molécules de référence puis transférés à toute molécule contenant la même liaison, le même angle de valence. A cela, s'ajoute l'énergie d'interaction entre atomes non liés (séparés par plus de deux liaisons), c'est-à-dire les interactions entre leurs nuages électroniques. On les décompose par paires d'atomes sous la forme d'interactions électrostatiques, de répulsion, de dispersion (interaction attractive entre dipôles instantanés) (encart 1) et éventuellement de polarisation.

Dans les champs de forces dits de première génération, la polarisation très coûteuse en temps de calcul n'est pas représentée par un terme explicite.

Indépendamment de la forme fonctionnelle développée, la qualité et la précision d'un champ de forces résident dans celles de ses paramètres. La paramétrisation est en général réalisée sur des composés modèles de petite taille, elle combine des informations issues de calculs quantiques pour ajuster les charges ponctuelles et les constantes de forces et des informations expérimentales à reproduire via des simulations moléculaires pour ajuster les paramètres de répulsion et dispersion. De l'utilisation de données expérimentales vient le caractère empirique des champs de forces. C'est dans la stratégie de paramétrisation, souvent itérative, que se situe l'art de la détermination des champs de forces. Elle peut corriger certaines approximations introduites par la forme mathématique choisie, par exemple l'absence de polarisation par une correction des charges électrostatiques ou l'absence des effets à N corps dans la dispersion par l'ajustement des paramètres van der Waals sur des données thermodynamiques expérimentales.

Les champs de forces tout atome sont utilisés en mécanique moléculaire pour caractériser l'espace conformationnel de molécules comme les sucres ou les peptides et surtout en dynamique

moléculaire pour obtenir via la mécanique statistique des grandeurs moyennes structurales, thermodynamiques ou dynamiques de liquides, solutions, protéines, lipides...

La composition du système mais aussi le type de propriétés à modéliser guident le choix du champ de forces. C'est ainsi que d'autres modèles comme celui de l'atome unifié ou le modèle gros grain sont développés.

Le champ de forces 'atome unifié' inclut les atomes d'hydrogène aliphatiques et les carbones associés dans une même particule, définie par un type atome unifié et des paramètres. Il s'agit d'une approximation adaptée pour simuler des systèmes moléculaires dans lesquels les mouvements intermoléculaires sont beaucoup plus importants ou significatifs que les déformations intramoléculaires. C'est le cas par exemple des propriétés physiques et chimiques d'un matériau, d'un système hétérogène, du repliement d'une protéine ou d'interaction protéine-protéine.

La dynamique de grands agrégats macromoléculaires comme les macromolécules biologiques, les membranes phospholipidiques, les ribosomes ou les virus se passe à des échelles de temps de l'ordre de la milliseconde, voire la seconde, bien supérieures à celles accessibles dans les simulations tout atome. Dans ce cas, il est donc nécessaire de dépasser l'échelle atomique et d'envisager une description moins détaillée, en terme de grains élémentaires tels que les résidus, nucléotides, riboses, voire des entités moléculaires supérieures, c'est-à-dire une description mésoscopique ou 'gros grains'. Si « l'intégration » d'un grand nombre de degrés de liberté dans un modèle de « grain » plus grossier que le niveau atomique paraît naturelle, la paramétrisation de ces grains peut s'avérer ardue. Elle repose sur une description approchée et lissée de la surface d'énergie potentielle tout atome, mais aussi sur le mode utilisé pour l'exploration de la surface et se constitue d'un ensemble de paramètres interdépendants. Il s'agit d'un champ de recherche en fort développement.

De manière générale, les valeurs des paramètres d'un champ de forces sont dépendantes de la stratégie de paramétrisation choisie et sont interdépendantes. C'est pourquoi les paramètres, par exemple van der Waals ou électrostatiques, d'un champ de forces ne peuvent a priori pas être transférés dans un autre champ de forces, l'ensemble des paramètres devant rester cohérent. Cette limitation amène les groupes de recherche qui les développent mais aussi ceux qui les utilisent à enrichir régulièrement les champs de forces.

### **III- Les champs de force non polarisables de première génération : évolutions et limites**

Il existe une variété de champs de forces additifs sans polarisation explicite, qu'on appelle champs de force standard ou de première génération, selon le grain choisi pour définir les particules élémentaires du champ de forces, les molécules de référence utilisées pour établir les paramètres et le type de propriétés recherchées (MM2, CHARMM, AMBER, GROMOS, GROMACS, GLYCAM, OPLS, CVFF...).

De nombreuses études contribuent à comparer les performances des champs de forces les plus courants, par exemple [5-8]. Les termes électrostatiques vont influencer essentiellement la capacité à reproduire la réponse électronique du système à l'environnement, les termes van der Waals de répulsion et dispersion celle à reproduire la densité, la capacité calorifique, la chaleur de vaporisation, l'énergie libre d'hydratation. La corrélation inévitable entre tous ces paramètres peut facilement mettre à mal la qualité d'un champ de forces lorsqu'on l'applique à des systèmes trop

éloignés physiquement et chimiquement du jeu de molécules de référence. Ainsi, de nouvelles valeurs de paramètres d'un champ de forces donné sont fréquemment publiées par extension de l'ensemble de référence et/ou par des changements de méthodes de paramétrisation.

Avec l'augmentation de la puissance de calcul, la modélisation peut contribuer à des enjeux sociétaux par des simulations moléculaires de systèmes complexes réels simulant les mouvements des molécules et des processus physicochimiques allant de la picoseconde à la milliseconde voire la seconde. En bousculant les limites et la précision des champs de forces, ces recherches accélèrent leur développement. Ainsi, les liquides ioniques dont les applications en synthèse organique, en électrochimie ou électrodéposition sont des ouvertures majeures vers une chimie durable, sont activement étudiés (voir article A. Padua de ce numéro). La plupart des défis à relever concerne la modélisation de macromolécules et systèmes d'intérêt biologique. C'est le cas notamment de la modélisation du repliement des protéines et de la stabilité des structures secondaires de type hélice- $\alpha$  et feuillet  $\beta$ , qui se traduit par un travail sur les paramètres de torsion et leur couplage au sein de la chaîne d'acides aminés [9]. Un deuxième domaine concerne la modélisation des membranes des cellules donc de bicouches lipidiques [10], pour lesquelles différentes stratégies sont développées de façon à bien reproduire leur structure et leur dynamique et éviter la formation de gels. La modélisation des hydrates de carbone ou glucides dont les fonctions biologiques et interactions avec les protéines sont avérées mais mal connues, constituent également un défi important. Leurs caractéristiques électroniques spécifiques, les nombreux centres chiraux, la grande variété de connexion pour former des polysaccharides linéaires ou branchés, le rôle des liaisons hydrogène leur confèrent des surfaces d'énergie potentielle complexe à représenter. [11]. Face à la difficulté de trouver de nouveaux médicaments, la modélisation moléculaire peut contribuer au design de nouveaux agents thérapeutiques [9, 12-16]. Dans cette optique, l'interaction protéine-ligand doit être la plus réaliste possible, ce qui implique une amélioration de la description des cations, des interactions hydrophobes, des interactions de type cation  $-\pi$ .

#### **IV- Les champs de forces de nouvelle génération**

Ainsi, si l'on veut ordonner énergétiquement et de manière prédictive un certain nombre de ligands en interaction avec une protéine, i.e. afin de prévoir leur affinité relative réelle pour la cible, il devient parfois nécessaire de pouvoir gérer de très faibles différences d'énergie. Il devient alors critique de réduire les marges d'erreur sur le calcul de l'énergie potentielle et d'inclure des raffinements sur la fonction de calcul de l'énergie potentielle du champ de forces. Le perfectionnement le plus courant relève de l'inclusion explicite de la contribution de l'énergie de polarisation [15-19] qui correspond à la réorganisation de la distribution de charge d'une molécule soumise à une perturbation électrostatique externe étant ainsi la source principale de la non-additivité de l'énergie totale. Ainsi si 3 molécules A, B et C sont en interaction, formant par exemple un réseau de liaisons hydrogène, l'énergie totale du système  $E_{ABC}$ , dite à N-corps est, en réalité, plus importante que l'énergie totale à 2-corps correspondant à la somme des énergies de paires:  $E_{ABC} > E_{AB} + E_{BC} + E_{AC}$  (notons que l'égalité peut s'inverser dans certains cas impliquant des systèmes chargés !). Les champs de forces classiques étant par essence des potentiels de paires, ils ne peuvent au final qu'approximer l'énergie à 2-corps et ne permettent pas la prise en compte de la non-additivité de l'énergie. Pour cette raison, le développement de champs de forces de nouvelles générations, dites « polarisables » font l'objet d'intenses efforts avec des applications immédiates pour le traitement de systèmes très chargés ou contenant des métaux ou des solvants complexes.

Trois schémas de polarisation sont souvent introduits. Si toutes ces approches ont en commun d'inclure une polarisation explicite, diverses stratégies existent. Le premier, appelé modèle de Drude se base sur un modèle d'oscillateur (de Drude) qui introduit une particule auxiliaire attachée à chaque atome polarisable à travers un ressort harmonique mimant les redistributions électroniques.[20] Un second modèle propose de jouer sur la valeur de la charge d'un atome en la traitant comme une variable dépendant de l'environnement : c'est le modèle dit de charge fluctuante qui modélise le flot intramoléculaire d'électrons sur la bases des valeurs des électronégativités et de dureté (hardness) des atomes.[21] Enfin, le plus répandu reste le modèle de « dipôle ponctuel » [19] où chaque centre polarisable va porter une polarisabilité soit dérivée d'un modèle empirique ou de la chimie quantique, soit issue d'une expérience. Dans ce cas, le dipôle induit en un point est directement calculé comme le produit de la polarisabilité en ce point et de la valeur du champ électrique subit par le centre polarisable en interaction avec ses voisins. Bien évidemment, si l'on s'intéresse à la reproduction fine des énergies d'interaction intermoléculaire, il convient de s'intéresser aux autres contributions physiques possibles de l'énergie. Si les deux premières stratégies exposées (i.e. Drude et charges fluctuantes) visent surtout à maintenir une certaine simplicité des expressions analytiques et à conserver la représentation électrostatique de type « charge ponctuelle » d'une molécule des champs de forces classiques, les méthodes de dipôles ponctuels permettent d'inclure des raffinements supplémentaires. C'est le cas des approches AMOEBA [22] et SIBFA (Sum of Interactions Between Fragments Ab initio),[19] dites de seconde génération et qui comportent une représentation de la distribution de charge incluant des dipôles et des quadrupôles localisés sur les atomes et parfois sur les liaisons chimiques permettant de retranscrire une plus grande anisotropie des interactions électrostatiques. Une méthodes comme SIBFA va même beaucoup plus loin puisqu'elle permet une inclusion de effets –« non classique » d'électrostatique de courte distance, de champs de ligands et même de transfert de charge entre molécules. Récemment, une troisième génération de potentiels a été introduite. Ainsi le modèle GEM (Gaussian Electrostatic Model) [23] se base sur l'utilisation de densités électroniques « gelées » représentées sous la forme de gaussiennes dites de Hermite utilisées en lieu et place des multipôles distribués et permettant d'inclure de manière non empirique un certain nombre d'effets quantiques de courte distance dans la mécanique moléculaire.

## **V- Les champs de forces et la réactivité**

Les champs de forces décrits précédemment reposent sur l'hypothèse du maintien de la connectivité des atomes ou groupes d'atomes tout au long des simulations, ce qui exclut toute modélisation de réactions chimiques ou de processus impliquant un changement d'états électroniques comme les transferts d'électrons ou la fluorescence.

Pour ces derniers, la philosophie des champs de forces peut être maintenue moyennant un choix approprié des formes mathématiques et des paramètres comme nous l'avons développé dans l'étude de la dynamique de chromophore fluorescent à l'état excité (encart 3). [24]

Pour les réactions chimiques, la réponse est donnée par plusieurs types d'approches. Les champs de forces et méthodes QM/MM (voir dans ce numéro l'article de X. Assfeld et al) permettent d'étudier et de décrire un mécanisme réactionnel limité à un petit groupe d'atomes au niveau quantique tout en prenant en compte les effets électrostatiques et contraintes géométriques de l'environnement. Les applications à la catalyse dans les enzymes et les complexes organométalliques sont nombreuses. La méthode Empirical Valence Bond (EVB) [25] consiste à décrire la réaction chimique dans une

approche Valence Bond, comme la résonance entre les structures valence-bond les plus importantes dans les réactifs et les produits, au travers d'un hamiltonien effectif paramétré représentant les variations énergétiques de ces structures valence-bond et de leur couplage en fonction de coordonnées géométriques et grandeurs électrostatiques. Il n'y a pas de traitement explicite des électrons, ce qui rend la méthode beaucoup plus rapide que les approches QM/MM mais limite les réactions chimiques possibles à celles prédéfinies dans les formes valence-bond. Des champs de forces dits réactifs font également l'objet de développements et d'applications [26]. Ces champs de forces sont basés sur une description des ordres de liaison autour d'un atome. L'énergie de liaison n'est plus une fonction des distances mais une fonction de l'ordre de liaison entre deux atomes, de même les énergies de déformation angulaire dépendent des angles et des ordres de liaison. La conjugaison, la stabilité de paires libres sont décrites par d'autres termes spécifiques et les termes électrostatiques et van der Waals sont distribués entre toutes les paires d'atomes. Les champs de forces réactifs permettent de simuler près d'un demi-million d'atomes pendant quelques nanosecondes. Cette méthode trouve de nombreuses applications dans le domaine des hydrocarbures, des matériaux et des surfaces, des réactions de combustion, d'oxydation et des explosions.

## VI - Perspectives pour la simulation moléculaires

Le futur des simulations utilisant les champs de forces s'articule avec la mise à disposition prochaine pour les communautés de machines de calculs massivement parallèles (calcul petascale voir exascale). Ainsi grâce à l'évolution des logiciels, [27-29] des simulations de champs de forces classiques comprenant toujours plus d'atomes et pouvant utiliser des dizaines, voire des centaines de milliers de processeurs deviennent possibles. De même, concernant les champs de forces polarisables, leur optimisation en lien avec les mathématiques appliquées permettent aussi d'envisager une utilisation massivement parallèle [30] et donc l'accès à des systèmes toujours plus grands et à des temps de simulation en dynamique moléculaire toujours plus importants. Enfin, au-delà des supercalculateurs, l'avènement des processeurs dits « graphiques » de type GPU [31] ou dédiés aux simulations (ANTON [32]) permettent d'ores et déjà d'anticiper que des simulations aux temps longs et donc d'intérêt biologique pourront devenir rapidement routinières.

## Références

- [1] Mc Cammon JA, Gelin BR, Karplus M. *Nature*, 1977, 267 : 585-590.
- [2] Chandler DE, Hsin J, Harrison CB, Gumbart J, Schulten K. *Biophys J* 2008, 95, 2822-2836.
- [3] Halgren TA. *Curr Opin Struct Biol* 1995, 5, 205-210
- [4] Alder BJ, Wainwright TE. *J. Chem. Phys.* 1957, 27, 1208-1209.
- [5] Grundertofte, Lijrfors, Norrby, Petterson J. *J. Comput. Chem.* 1996, 17, 429-449
- [6] Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. *PLOS One*, 2012, 7,e32131
- [7] Ponder JW, Case DA. *Advances in Protein Chemistry*, 2003, 66, 27-78
- [8] Paton RS, Goodman JM. *J. Chem. Inf. Model* 2009, 49, 944-955
- [9] Zhu X, Lopes PEM, MacKerell Jr AD. *WIREs Comput Mol Sci* 2012, 2, 167 – 185
- [10] Skjevik AA, Madej BD, Walker RC, Teigen K. *J. Phys. Chem B*, 2012, 116, 11124-11136



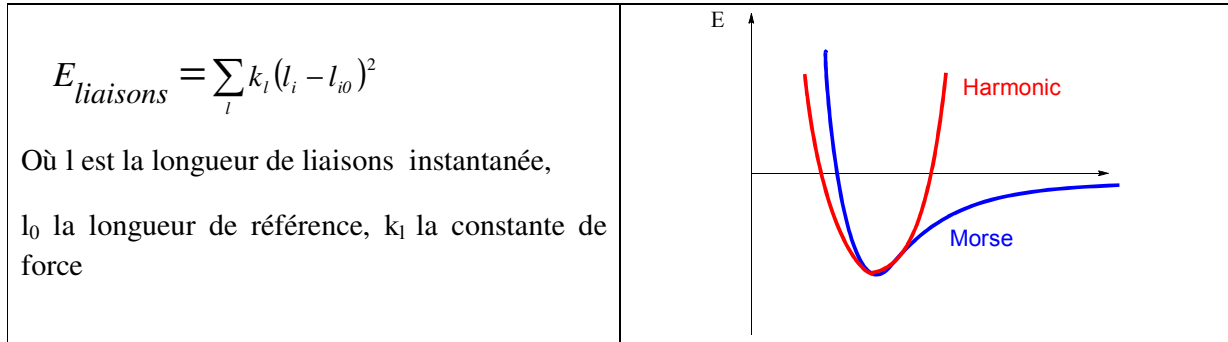
- [11] Foley BL, Tessier M B., Woods R J. WIREs Comput Mol Sci 2012, 2, 652-697
- [12] Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I et al Comput Chem 2010, 31, 671-690.
- [13] Young DC. Hoboken, NJ John Wiley and Sons, Inc 2009. DOI :10.1002/9780470451854.ch10.
- [14] Wong CF, Mac Cammon JA. Adv Protein Chem 2003, 66, 87-121.
- [15] Jorgensen, W. L. Angew. Chem. Int. Ed. 2012, 51, 11680–11684
- [16] Jorgensen, W. L. J. Chem. Theory Comput., 2007, 3, 1877–187
- [17] Warshel, M. Kato and A.V. Pisliakov, J. Chem. Theor. Comput., 2007, 3, 2034-2045.
- [18] Piquemal, JP ; Jordan K.D. Theor Chem Acc, 2012, 131, 1207
- [19] Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. p. J. Chem. Theory. Comput., 2007, 3, 1960
- [20] Lopes P, Roux B, MacKerell A., Theor Chem Acc, 2009, 124, 11-28.
- [21] Steven W. Rick, Steven J. Stuart and B. J. Berne J. Chem. Phys. 1994, 101, 6141
- [22] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. Di Stasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, J. Phys. Chem. B, 2010, 114, 2549
- [23] Piquemal, J.-P., Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A.; J. Chem. Phys., 2006, 124, 104101
- [24] Jonasson, G.; Teuler, J.-M.; Vallverdu, G.; Merola, F.; Ridard, J.; Lévy, B.; Demachy, I.; J. Chem. Theory Comp. 2011, 7, 1990
- [25] Warshel, A. Computer Modeling of Chemical Reactions in Enzymes and Solutions: New York. 1992. Warshel, A. Acc. Chem. Res. 2002, 35, 385-395
- [26] van Duin, A C. T.; Dasgupta S, Lorant F, Goddard W A. J Phys. Chem. A, 2001, 105, 9396–9409
- [27] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, Klaus Schulten. Journal of Computational Chemistry, 2005, 26, 1781-1802
- [28] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, J. Chem. Theory Comput. 2013, 9, 3878-3888
- [29] David van der Spoel and Berk Hess, Comput. Molec. Sci. 2011, 1, 710-715
- [30] Lipparini, F., Lagardere, L.; Stamm, B. ; Cances, E. ; Schnieders, M. ; Ren, P. Y.; Maday, Y; Piquemal, J.-P. J. Chem. Theory. Comput., 2014, sous presse, DOI: [10.1021/ct401096t](https://doi.org/10.1021/ct401096t)
- [31] John E. Stone, James C. Phillips, Peter L. Freddolino, David J. Hardy 1, Leonardo G. Trabuco, Klaus Schulten (2007). Journal of Computational Chemistry 2007, 28, 2618–2640.
- [32] David E. Shaw, Ron O. Dror, John K. Salmon, J.P. Grossman, Kenneth M. Mackenzie, Joseph A. Bank, Cliff Young, Martin M. Deneroff, Brannon Batson, Kevin J. Bowers, Edmond Chow, Michael P. Eastwood, Douglas J. Ierardi, John L. Klepeis, Jeffrey S. Kuskin, Richard H. Larson, Kresten Lindorff-Larsen, Paul Maragakis, Mark A. Moraes, Stefano Piana, Yibing Shan, and Brian Towles (2009). Proceedings of the ACM/IEEE Conference on Supercomputing (SC09) (Portland, Oregon: ACM): 1–11. DOI:10.1145/1654059.1654099

## ENCART 1

L'énergie potentielle est calculée avec un modèle empirique. Les interactions entre les atomes se décomposent en une composante intramoléculaire (énergies de déformation harmoniques des liaisons, des angles et dièdres) et une composante intermoléculaire décrivant les interactions entre atomes non liés, c'est-à-dire séparés par plus de deux liaisons.

$$E_{TOT} = E_{liaisons} + E_{angles} + E_{dièdres} + E_{vdW} + E_{electrostatique}$$

### Déformation des liaisons



### Déformation des angles

$$E_{angles} = \sum_l k_{anglei} (\theta_i - \theta_{i0})^2$$

Où  $\theta$  est la valeur instantanée de l'angle de valence,  $\theta_0$  l'angle de référence,  $k_{anglei}$  la constante de force

### Déformations des angles dièdres

Le terme énergétique de déformation des dièdres concerne trois liaisons et donc quatre atomes A,B, C et D consécutifs. L'angle dièdre correspond à l'angle entre les liaisons AB et CD, donc à la rotation autour de la liaison BC.

$$E_{\Phi} = \sum_{\Phi} [V_1(1 + \cos \Phi) + V_2(1 - \cos 2\Phi) + V_3(1 + \cos 3\Phi)]$$

### Interactions de van der Waals

Elles sont représentées par un potentiel en 6-12 :

$$E_{LJ}(r) = 4\epsilon \left[ \left( \frac{r_0}{r} \right)^{12} - \left( \frac{r_0}{r} \right)^6 \right]$$

$r$  est distance entre deux centre Lennard-Jones, en général deux atomes

### Interactions électrostatiques

La distribution des électrons considérés comme implicites est représentée par des charges partielles atomiques ponctuelles  $q_i$  et  $q_j$ . L'énergie électrostatique résulte de la somme des interactions coulombiennes par paires de charges :

$$E_{elec} = \sum_{i < j} \frac{q_i q_j}{\epsilon r_{ij}}$$

## ENCART2 : CHAMPS DE FORCES POLARISABLES.

### A. Modèle de dipôles ponctuels

Afin de décrire les interactions électrostatiques incluant les effets de polarisation, ce modèle considère une molécule comme une distribution de charges localisées sur des centres (atomiques ou non, i.e. paires libres par exemple). Ainsi la distribution de charge totale sur un site  $i$  est la somme des charges permanentes et induites.

$$\mathbf{M}_i = \mathbf{M}_i^0 + \mathbf{M}_i^{\text{ind}} \quad (\text{A})$$

$\mathbf{M}$  représente la distribution de charge. Cette dernière peut être constituée de simples charges ponctuelles, d'un développement multipolaire comprenant charges, dipôles, quadrupôles et/ou moments d'ordre supérieurs, ou une distribution de charge continue (i.e. de type densité électronique gelée, approche GEM). Ce modèle n'est pas limité par l'utilisation d'un type particulier de distribution de charge mais nous traiterons une distribution de type multipolaire pour plus de clarté dans l'exposé qui suit.

L'interaction électrostatique entre les charges sur les sites  $i$  et  $j$  est donnée par :

$$U_{\text{ele}} = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{M}_i^t \mathbf{T}_{ij} \mathbf{M}_j \quad (\text{B})$$

Avec la distribution de charge induite:

$$\mathbf{M}_i^{\text{ind}} = \alpha_i^{-1} \sum_{j \neq i} \mathbf{T}_{ij} (\mathbf{M}_j^0 + \mathbf{M}_j^{\text{ind}}) \quad (\text{C})$$

Il est possible en pratique de résoudre l'équation C de manière itérative afin d'obtenir les dipôles induits. On peut alors écrire l'énergie électrostatique finale :

$$U_{\text{ele}} = \frac{1}{2} \sum_i \sum_{j \neq i} (\mathbf{M}_i^0)^t \mathbf{T}_{ij} \mathbf{M}_j^0 + \frac{1}{2} \sum_i \sum_{j \neq i} (\mathbf{M}_i^{\text{ind}})^t \mathbf{T}_{ij} \mathbf{M}_j^0 \quad (\text{D})$$

Où le premier terme correspond à l'énergie électrostatique issue des moments "permanents" et le second terme correspond à l'énergie de polarisation générée par la présence de moments induits.

### B. Modèle de Drude

Dans le modèle de l'oscillateur de Drude, le phénomène de polarisation est décrit par une charge ponctuelle supplémentaire (l'oscillateur ou charge de Drude) attachée à tout atome (sauf les hydrogènes) par un ressort harmonique. Ainsi la charge ponctuelle peut bouger en regards de son site d'appartenance en réponse à une perturbation électrostatique issue de l'environnement. L'énergie électrostatique est la somme des interactions de paires entre les charges atomiques et les charges partielles des particules de Drude

$$E_{\text{ele}} = \sum_{A < B}^N \frac{q_C(A)q_C(B)}{|r_C(A)-r_C(B)|} + \sum_{A < B}^{N, N_D} \frac{q_D(A)q_C(B)}{|r_D(A)-r_C(B)|} + \sum_{A < B}^{N_D} \frac{q_D(A)q_D(B)}{|r_D(A)-r_D(B)|} + \frac{1}{2} \sum_A^{N_D} k_D (r_D(A)-r_C(B))^2 \quad (\text{H})$$

avec  $N_D$  et  $N$ , les indices des particules de Drude (sur tous les atomes sauf les hydrogènes),  $q_D$  et  $q_C$  respectivement les charges des particules de Drude et celles de leur atomes porteurs,  $r_D$  et  $r_C$  leur positions respectives, et  $k_D$  la constante de force harmonique entre l'oscillateur de Drude et l'atome porteur. Le dernier terme de l'équation H décrit le coût de la polarisation des particules de Drude. Ainsi, il est possible d'accéder à la polarisabilité atomique ( $\alpha$ ) qui est fonction de la charge partielle de la particule de Drude et de la constante de force du ressort :

$$\alpha = \frac{q_D^2(A)}{k_D} \quad (\text{I})$$

### C. Modèle des charges fluctuantes

Le formalisme dit des charges fluctuantes est basé sur la méthode de l'équilibration de charge modèle (CHEQ) dans laquelle le potentiel chimique est équilibré via une redistribution de la densité de charge. L'énergie d'un système apparait ainsi comme dépendant de la charge et s'exprime pour un système de  $M$  molécules contenant  $N_i$  atomes par molécule sous la forme:

$$E_{\text{CHEQ}}(R, Q) = \sum_{i=1}^M \sum_{\alpha=1}^{N_i} \chi_{i\alpha} Q_{i\alpha} + \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \sum_{\alpha=1}^{N_i} \sum_{\beta=1}^{N_j} J_{i\alpha i\beta} Q_{i\alpha} Q_{j\beta} + \frac{1}{2} \sum_{i=1}^{MN'} \sum_{j=1}^{MN'} \frac{Q_i Q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{j=1}^M \lambda_j (\sum_{i=1}^N Q_{ij} - Q_j^{\text{Total}}) \quad (\text{J})$$

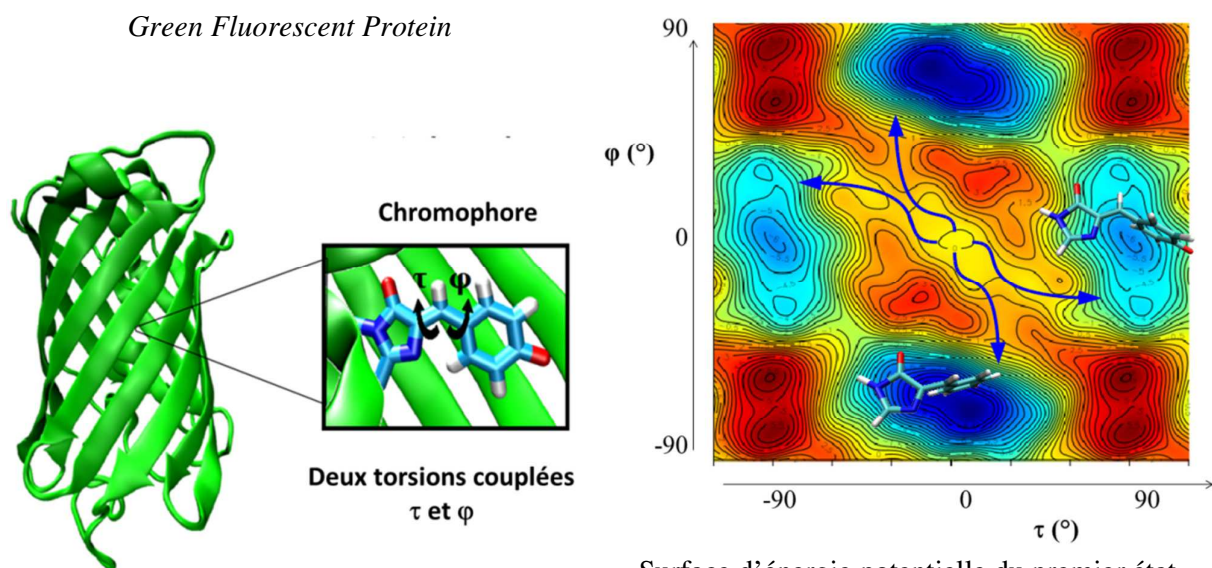
où  $Q_i$  est la charge partielle d'un site atomique  $i$ .  $\chi$  décrit l'électronégativité atomique contrôlant la directionnalité du flot de charge et  $J$  est la dureté (hardness) atomique qui représente la résistance du flot d'électrons allant ou partant d'un atome donné. Ces paramètres sont optimisés afin de reproduire les moments dipolaires moléculaires et la réponse par polarisation moléculaire associée.

### ENCART3 : CHAMP DE FORCE ET ETAT EXCITE

Les protéines fluorescentes, comme la GFP (green fluorescent protein), sont des protéines naturellement fluorescentes qui sont utilisées pour leur rôle de marqueur, permettant de localiser des protéines dans les cellules et d'en suivre les déplacements. De nombreuses études expérimentales et théoriques ont été menées ces dix dernières années sur les protéines fluorescentes. De là, se forge une compréhension essentiellement qualitative du rôle de la protéine vis-à-vis de l'obtention ou non d'une émission radiative : il apparaît que la protéine permet la fluorescence en bloquant les processus qui la désactivent ; ces processus de désactivation sont très rapides et efficaces (à l'échelle de la picoseconde) dans le cas du chromophore seul, et ils sont bien identifiés comme étant des torsions autour des liaisons intercycles ( $\tau$  et  $\varphi$ ). Dans la protéine, la sensibilité des temps de vie de fluorescence à des mutations proches ou non du chromophore, à des modifications de pH ou de température laisse supposer un contrôle de la dynamique du chromophore par différents paramètres, sans qu'ils soient pour autant identifiés et mis en relation.

La taille du système et la durée du processus à l'échelle nanoseconde exclut les approches quantiques ou mixtes quantique-classique. Un champ de force peut modéliser la surface d'énergie potentielle du chromophore à l'état excité moyennant des calculs quantiques de haute précision (SA2-CASPT2), et une interpolation des valeurs obtenues par une expression analytique appropriée en fonction des angles de torsion  $\tau$  et  $\varphi$ . Le champ de force ajusté sur les calculs quantiques est bâti comme une combinaison de produits de fonction cosinus et sinus de multiples des deux angles de torsion.

Ce champ de force donne accès à la modélisation de la désactivation non-radiative par conversion interne à l'échelle de la nanoseconde permettant d'identifier les mouvements moléculaires concertés de la protéine et du chromophore conduisant à cette extinction. De ces résultats, émerge une représentation plus complète du mécanisme qui libère la torsion du chromophore ou qui la déclenche et une interprétation des processus d'isomérisation du chromophore au sein des protéines photoactivables.



Surface d'énergie potentielle du premier état excité du chromophore en fonction des angles de torsion représentée par un champ de force ajusté sur des calculs quantiques

Jonasson G et al, JCTC, 2011, 7 : 1990