

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

Isabelle Demachy et Jean-Philip Piquemal

Résumé	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 forces est un subtil équilibre entre simplicité et représentativité physico-chimique. Des concepts de base aux enjeux actuels, cet article retrace les différents développements et spécificités des champs de forces.
Mots-clés	Énergie potentielle, champ de forces, conformations, structures, dynamique moléculaire, systèmes moléculaires complexes.
Abstract	Potential energy surface and force fields 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 fields development and properties.
Keywords	Potential energy, force field, conformation, structure, molecular dynamics, complexe molecular systems.

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 lui 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 de 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 relient suffit pour déterminer les propriétés. Pour autant, il faut savoir identifier ces régions et les explorer efficacement (voir dans ce numéro les articles traitant des

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 calculs 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 que l'on appelle un champ de forces, car de cette énergie potentielle dérivent 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 picosecondes (10^{-12} s) [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 domaine 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.

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 formation 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 encadré 1). C'est ainsi que Alder et Wainwright ont été les premiers à la fin des années 1950 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 [4]. 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 types 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 (encadré 1) dont les paramètres (k , l_0 , θ_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. À 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) (encadré 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

Encadré 1

Expression courante des champs de forces

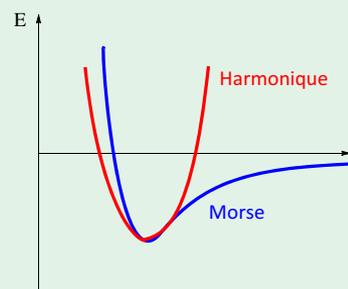
L'énergie potentielle est calculée avec un modèle empirique. Les interactions entre les atomes se décomposent en une composante intramoléculaire (énergies de déformation harmoniques des liaisons, des angles et dièdres) et une composante intermoléculaire décrivant les interactions entre atomes non liés,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

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

Déformation des liaisons

$$E_{liaisons} = \sum_i k_l (l_i - l_{i0})^2$$

où l est la longueur de liaisons instantanée, l_0 la longueur de référence, et k_l la constante de force.



Déformation des angles

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

où θ est la valeur instantanée de l'angle de valence, θ_0 l'angle de référence, et k_{angle_i} la constante de force.

Déformation 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 :

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

où r est la distance entre deux centres de 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 :

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

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 de 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 d'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'une 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 de 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 termes 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.

Les champs de force non polarisables de première génération : évolutions et limites

Il existe une variété de champs de forces additifs sans polarisation explicite, que l'on appelle champs de forces 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 physico-chimiques allant de la picoseconde à la milliseconde, voire à 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 l'article d'A. Pádua p. 63). La plupart des défis à relever concernent la modélisation de macromolécules et de 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 α et feuillets β , 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, constitue également un défi important. Leurs caractéristiques électroniques spécifiques, les nombreux centres chiraux, la grande variété de connexions pour former des polysaccharides linéaires ou branchés, le rôle des liaisons hydrogène leur confèrent des surfaces d'énergie potentielle complexes 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 et des interactions de type cation- π .

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 trois 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 (voir encadré 2). 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 base des valeurs des électronégativités et de dureté 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 quadripôles localisés sur les atomes et parfois sur les liaisons chimiques, permettant de retranscrire une plus grande anisotropie des interactions électrostatiques. Une méthode comme SIBFA va même beaucoup plus loin puisqu'elle permet une inclusion d'effets « non classiques » 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

Encadré 2

Champs de forces polarisables

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 :

$$M_i = M_i^0 + M_i^{\text{ind}} \quad (\text{A})$$

M représente la distribution de charge. Cette dernière peut être constituée de simples charges ponctuelles, d'un développement multipolaire comprenant charges, dipôles, quadripôles et/ou moments d'ordres 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 ce qui suit.

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

$$U_{\text{ele}} = \frac{1}{2} \sum_i \sum_{j \neq i} M_i^{\dagger} T_{ij} M_j \quad (\text{B})$$

avec la distribution de charge induite :

$$M_i^{\text{ind}} = \alpha_i^{-1} \sum_{j \neq i} T_{ij} (M_j^0 + M_j^{\text{ind}}) \quad (\text{C})$$

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 :

$$U_{\text{ele}} = \frac{1}{2} \sum_i \sum_{j \neq i} (M_i^0)^{\dagger} T_{ij} M_j^0 + \frac{1}{2} \sum_i \sum_{j \neq i} (M_i^{\text{ind}})^{\dagger} T_{ij} M_j^0 \quad (\text{D})$$

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

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 regard 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 leurs atomes porteurs, r_D et r_C leurs 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 (α) qui est fonction de la charge partielle de la particule de Drude et de la constante de force du ressort :

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

Modèle des charges fluctuantes

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

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

où Q_i est la charge partielle d'un site atomique i . χ 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 vers ou partant d'un atome donné. Ces paramètres sont optimisés afin de reproduire les moments dipolaires moléculaires et la réponse par polarisation moléculaire associée.

« 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.

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 (*encadré 3*) [24].

Pour les réactions chimiques, la réponse est donnée par plusieurs types d'approches. Les champs de forces et méthodes QM/MM (voir l'article de N. Ferré et X. Assfeld p. 43) 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 les contraintes géométriques de l'environnement. Les applications à la catalyse dans les enzymes et les complexes organométalliques sont nombreuses. La méthode EVB (« empirical valence bond ») [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.

Perspectives pour la simulation moléculaire

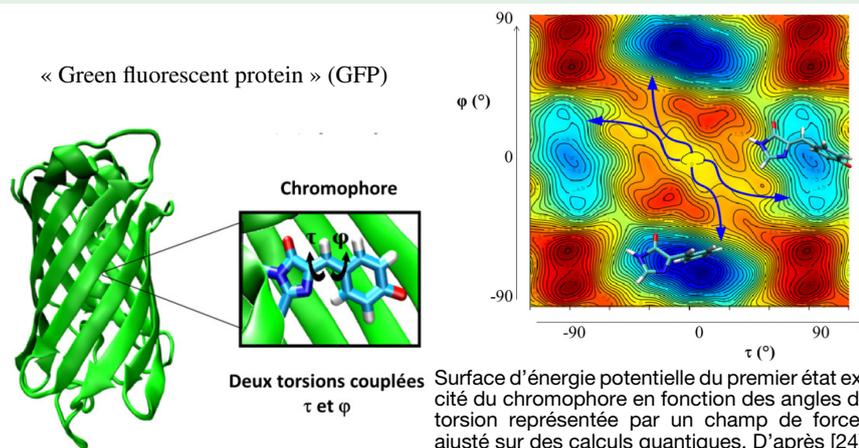
Le futur des simulations utilisant les champs de forces s'articule avec la mise à disposition prochaine pour les

Encadré 3

Champ de forces et état excité

Les protéines naturellement fluorescentes, comme la GFP (« green fluorescent protein »), 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 (τ et φ). 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 excluent les approches quantiques ou mixtes quantique-classique. Un champ de forces 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 τ et φ . Le champ de forces 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 forces 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 émergent 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.



communautés de machines de calculs massivement parallèles (calculs 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 permet 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]) permet d'ores et déjà d'anticiper que des simulations aux temps longs et donc d'intérêt biologique pourront devenir rapidement routinières.

Les auteurs remercient Gilberte Chambaud pour son aide lors des relectures de cet article.

Note et références

* Le petascale correspond à un quadrillion (10^{15}) d'opérations à virgule flottante par seconde et l'exascale à mille fois plus, soit un quintillion (10^{18}) d'opérations à virgule flottante par seconde.

- [1] Mc Cammon J.A., Gelin B.R., Karplus M., Dynamics of folded proteins, *Nature*, **1977**, 267, p. 585.
- [2] Chandler D.E., Hsin J., Harrison C.B., Gumbart J., Schulten K., Intrinsic curvature properties of photosynthetic proteins in chromatophores, *Biophys. J.*, **2008**, 95, p. 2822.
- [3] Halgren T.A., Potential energy functions, *Curr. Opin. Struct. Biol.*, **1995**, 5, p. 205.
- [4] Alder B.J., Wainwright T.E., Phase transition for a hard sphere system, *J. Chem. Phys.*, **1957**, 27, p. 1208.
- [5] Grundertofte K., Lijffors T., Norrby P.-O., Pettersson I., A comparison of conformational energies calculated by several molecular mechanics methods, *J. Comput. Chem.*, **1996**, 17, p. 429.
- [6] Lindorff-Larsen K., Maragakis P., Piana S., Eastwood M.P., Dror R.O., Shaw D.E., Systematic validation of protein force fields against experimental data, *PLOS One*, **2012**, 7, p. e32131.
- [7] Ponder J.W., Case D.A., Force fields for protein simulation, *Adv. Prot. Chem.*, **2003**, 66, p. 27.
- [8] Paton R.S., Goodman J.M., Hydrogen bonding and π -stacking: how reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions, *J. Chem. Inf. Model.*, **2009**, 49, p. 944.
- [9] Zhu X., Lopes P.E.M., MacKerell Jr A.D., Recent developments and applications of the CHARMM force fields, *WIREs Comput. Mol. Sci.*, **2012**, 2, p. 167.
- [10] Skjevik A.A., Madej B.D., Walker R.C., Teigen K., LIPID11: a modular framework for lipid simulations using Amber, *J. Phys. Chem. B*, **2012**, 116, p. 11124.
- [11] Foley B.L., Tessier M.B., Woods R.J., Carbohydrate force fields, *WIREs Comput. Mol. Sci.*, **2012**, 2, p. 652.
- [12] Vanommeslaeghe K. et al., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *Comput. Chem.*, **2010**, 31, p. 671.
- [13] Young D.C., Hoboken N.J., Molecular mechanics, chap. 10, *Computational Drug Design: a Guide for Computational and Medicinal Chemistry*, D.C. Young (ed.), Wiley, **2009**.
- [14] Wong C.F., McCammon J.A., Protein simulation and drug design, *Adv. Protein Chem.*, **2003**, 66, p. 87.
- [15] Jorgensen W.L., Challenges for academic drug discovery, *Angew. Chem. Int. Ed.*, **2012**, 51, p. 11680.
- [16] Jorgensen W.L. (ed.), Special issue on polarization, *J. Chem. Theory Comput.*, **2007**, 3, p. 1877.
- [17] Warshel A., Kato M., Pislakov A.V., Polarizable force fields: history test cases and prospects, *J. Chem. Theory Comput.*, **2007**, 3, p. 2034.
- [18] Piquemal J.P., Jordan K.D., From quantum mechanics to force fields: new methodologies for the classical simulation of complex systems, *Theor. Chem. Acc.*, **2012**, 131, p. 1207.
- [19] Gresh N., Cisneros G.A., Darden T.A., Piquemal J.P., Anisotropic, polarizable molecular mechanics studies of inter-, intra-molecular interactions, and ligand-macromolecule complexes. A bottom-up strategy, *J. Chem. Theory Comput.*, **2007**, 3, p. 1960.
- [20] Lopes P., Roux B., MacKerell A.D. Jr., Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability: Theory and applications, *Theor. Chem. Acc.*, **2009**, 124, p. 11.
- [21] Rick S.W., Stuart S.J., Berne B.J., Dynamic fluctuating charge force fields: Application to liquid water, *J. Chem. Phys.*, **1994**, 101, p. 6141.
- [22] Ponder J.W. et al., Current status of the AMOEBA polarizable force field, *J. Phys. Chem. B*, **2010**, 114, p. 2549.
- [23] Piquemal J.-P., Cisneros G.A., Reinhardt P., Gresh N., Darden T.A., Towards a force field based on density fitting, *J. Chem. Phys.*, **2006**, 124, p. 104101.
- [24] Jonasson G., Teuler J.-M., Vallverdu G., Mérola F., Ridard J., Lévy B., Demachy I., Excited state dynamics of the green fluorescent protein on the nanosecond time scale, *J. Chem. Theory Comp.*, **2011**, 7, p. 1990.
- [25] Warshel A., *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley, **1992**; Warshel A., Molecular dynamics simulations of biological reactions, *Acc. Chem. Res.*, **2002**, 35, p. 385.
- [26] van Duin AC.T., Dasgupta S., Lorant F., Goddard W.A., ReaxFF: a reactive force field for hydrocarbons, *J. Phys. Chem. A*, **2001**, 105, p. 9396.
- [27] Phillips J.C. et al., Scalable molecular dynamics with NAMD, *J. Comput. Chem.*, **2005**, 26, p. 1781.
- [28] Salomon-Ferrer R., Götz A.W., Poole D., Le Grand S., Walker R.C., Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle Mesh Ewald, *J. Chem. Theory Comput.*, **2013**, 9, p. 3878.
- [29] van der Spoel D., Hess B., GROMACS - the road ahead, *Comput. Mol. Sci.*, **2011**, 1, p. 710.
- [30] Lipparini F., Lagardere L., Stamm B., Cancès E., Schnieders M., Ren P.Y., Maday Y., Piquemal J.-P., *J. Chem. Theory Comput.*, **2014**, sous presse.
- [31] Stone J.E. et al., Accelerating molecular modeling applications with graphics processors, *J. Comput. Chem.*, **2007**, 28, p. 2618.
- [32] Shaw D.E. et al., Millisecond-scale molecular dynamics simulations on Anton, *Proceedings of the ACM/IEEE conference on supercomputing (SC09)*, Portland, Oregon: ACM, 1-11, **2009**.



I. Demachy

Isabelle Demachy

est enseignante-chercheur au Laboratoire de Chimie Physique d'Orsay*.

Jean-Philip Piquemal

est professeur au Laboratoire de Chimie Théorique, Sorbonne Université, UPMC**. Il est directeur de l'Institut Parisien de Chimie Physique et Théorique (IP2CT, FR 2622 UPMC/CNRS).



J.-P. Piquemal

* Laboratoire de Chimie Physique d'Orsay, Université Paris-Sud, Bât. 349, Campus d'Orsay, 15 avenue Jean Perrin, F-91405 Orsay Cedex.
Courriel : isabelle.demachy@u-psud.fr

** Laboratoire de Chimie Théorique, Sorbonne Universités, Université Pierre et Marie Curie, CC 137, F-75252 Paris.
Courriel : jpp@lct.jussieu.fr



102 avenue Georges Clemenceau - 94700 MAISONS ALFORT

Tél. : 01 43 53 64 00 - Fax : 01 43 53 48 00

edition@edif.fr - www.edif.fr