

Une solution pour les liquides : la statistique moléculaire

Daniel Borgis, Anne Boutin et Rodolphe Vuilleumier

Résumé La mécanique statistique moléculaire permet de faire le lien entre les interactions microscopiques entre molécules dans un liquide et les propriétés structurales, thermodynamiques ou spectroscopiques observées. Cet article montre comment les méthodes théoriques et numériques de l'état liquide qui se sont développées depuis les années 1960 ont permis d'aller de la caractérisation des liquides simples, des liquides moléculaires et des solutions à celle des liquides complexes, des milieux biologiques, et de la réactivité chimique dans ces différents milieux. L'accent est mis sur la détermination de l'énergie libre plutôt que de l'énergie, et sur l'élucidation des mécanismes réactionnels en solution. La notion de mécanique statistique multi-échelle pour aborder les milieux complexes est aussi développée.

Mots-clés **Liquides moléculaires, mécanique statistique, énergie libre, mécanismes réactionnels, modélisation multi-échelle.**

Abstract **A solution for liquids: molecular statistics**
Molecular statistical mechanics makes it possible to draw a link between microscopic interactions between molecules in a liquid and the observed structural, thermodynamic, or spectroscopic properties. This article shows how the theoretical and numerical liquid-state methods that have been developed since the 1960's have enabled to go from the characterization of simple and molecular liquids and solutions to that of complex liquids, biological media, and chemical reactivity in those different media. Focus is made on the determination of free energies rather than energies and on the elucidation of reaction mechanisms in solution. The notion of multi-scale statistical mechanics to tackle complex media is also developed.

Keywords **Molecular liquids, statistical mechanics, free energy, reaction mechanism, multi-scale modelling.**

Liquides et théorie

Le liquide est un état de la matière particulièrement important en chimie, les réactions chimiques et biologiques à la surface de la Terre se faisant pour l'essentiel en phase liquide.

Au mot *liquide*, on trouve dans le Larousse : « Corps qui prend la forme du récipient qui le contient mais dont le volume est déterminé. Dans l'état liquide, les molécules sont séparées par de faibles distances et sont soumises à une agitation thermique intense et aléatoire. »

Cette simple définition traduit bien la spécificité de cet état de la matière. Les phases liquides sont des milieux denses et désordonnés à grande échelle. Il existe cependant un certain ordre à courte portée. L'agitation thermique, nécessairement présente (une phase ne peut être liquide à température absolue nulle⁽¹⁾), conduit à un mouvement « aléatoire » des particules. Ainsi, dans ce milieu condensé, les molécules qui composent le liquide se rencontrent, interagissent (éventuellement réagissent) puis s'éloignent. Un liquide est donc un milieu désordonné où les interactions entre particules sont dominantes. Ces spécificités ont un fort impact sur les méthodes de chimie théorique pertinentes pour étudier cet état de la matière. Les outils théoriques pour décrire les liquides à partir de l'échelle moléculaire doivent prendre nécessairement en compte l'aspect désordonné, dynamique et les interactions intermoléculaires.

La description d'un liquide à l'échelle atomique ou moléculaire nécessite de prendre en compte un nombre important

de degrés de liberté et de réaliser des moyennes à température donnée. Ainsi, il est crucial de considérer l'entropie du système. La grandeur thermodynamique de base est donc l'énergie libre (plutôt que l'énergie potentielle d'interaction seulement) pour étudier aussi bien la solvation d'un soluté qu'une réaction chimique en phase liquide par exemple.

Les méthodes théoriques s'appuient donc sur la statistique moléculaire, théorie qui fait le lien entre la description du système à l'échelle microscopique et ses propriétés macroscopiques. L'obtention de moyennes statistiques ne peut se faire de manière analytique qu'au prix d'un certain nombre d'approximations (champ moyen...). L'avènement des simulations moléculaires a permis de pallier ce problème en résolvant les équations numériquement (voir [1]).

Les concepts de base de la mécanique statistique des liquides moléculaires

Le but que l'on se donne en mécanique statistique moléculaire est simple : comment peut-on retrouver et/ou prédire les propriétés thermodynamiques, spectroscopiques, de transport... d'un système moléculaire, c'est-à-dire toute observable que l'on peut mesurer expérimentalement pour des conditions thermodynamiques données, à partir de la donnée du potentiel d'interaction microscopique entre les particules qui le composent. Ces particules peuvent être des noyaux nucléaires entourés de leur cortège électronique (modélisation dite *ab initio* ou de premiers principes), des atomes formant des molécules (modélisation tout-atome),

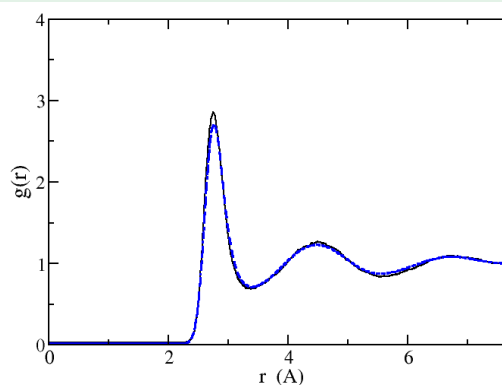
ou des grains de matière englobant un groupe plus ou moins grand d'atomes (modélisation gros grains). Puisque les forces entre particules dérivent directement du potentiel d'interaction, on parle de la définition du champ de forces. La modélisation *ab initio* [2], de par son coût numérique élevé, puisqu'il faut résoudre l'équation de Schrödinger des électrons pour chaque position donnée de tous les noyaux, est limitée à des systèmes moléculaires relativement petits, contenant au plus une centaine d'atomes. Une activité importante de la chimie théorique consiste donc à définir aussi rigoureusement que possible, pour un système donné, un champ de forces tout-atome à partir du potentiel *ab initio* sous-jacent, ou un champ de forces gros grains à partir du champ de forces tout-atome (voir l'article de I. Demachy et J.-P. Piquemal, p. 37). Le jeu de la mécanique statistique peut démarrer dès lors que ce champ de forces est connu.

Le calcul des propriétés se ramène en général à des moyennes des grandeurs microscopiques correspondantes sur toutes les positions – et éventuellement orientations – possibles de toutes les particules (on parle de configurations microscopiques). L'énergie interne, par exemple, s'obtient comme la moyenne sur toutes les configurations de l'énergie totale du système. Une autre grandeur caractéristique importante dans un liquide est la fonction de distribution de paire $g(r)$: la densité moyenne de particules à une distance r d'une particule donnée. Elle décrit la structuration microscopique au sein du liquide et contient la même information que le facteur de structure obtenu en diffusion de neutron ou diffraction de rayons X (voir encadré 1). Elle permet aussi de remonter, par des moyennes convenables, à certaines propriétés thermodynamiques comme la compressibilité, l'énergie interne, l'équation d'état. Mais comment effectuer une intégration sur toutes les configurations microscopiques possibles d'un système à N particules, avec N très grand, pour obtenir, par exemple, la fonction $g(r)$? Pour quelques systèmes modèles, tels qu'un fluide composé de particules sphériques impénétrables (des sphères dures), éventuellement chargées pour modéliser un sel fondu ou une solution ionique, on peut obtenir un certain nombre de résultats analytiques ou introduire des méthodes de résolution numérique rapides, telles que celles découlant de la théorie des équations intégrales [3-4]. Pour des modèles plus réalistes de liquides moléculaires, on doit recourir à des méthodes de simulations numériques, dans lesquelles on simule sur ordinateur le comportement microscopique du liquide et on fait en sorte que les particules qui le composent explorent convenablement leur espace de configuration, que ce soit en temps réel (dynamique moléculaire) ou dans un temps fictif où l'on génère des configurations au hasard, en s'assurant que le poids statistique des états explorés obéit aux lois de la mécanique statistique (méthode de Monte-Carlo). Il faut ensuite être capable d'exploiter ces trajectoires à N particules et de relier des grandeurs microscopiques calculées à des observables expérimentales. Pour la plupart des quantités thermodynamiques ou structurales, cela se fait par l'intermédiaire de valeurs moyennes calculées sur les trajectoires. Pour les propriétés de transport (coefficient de diffusion, conductivité thermique...) ou les propriétés spectroscopiques (spectres infrarouge ou Raman...), il est nécessaire de calculer par dynamique moléculaire des moyennes dépendant du temps, que ce soit pour un ensemble de trajectoires partant de conditions initiales différentes ou par l'intermédiaire de fonctions de corrélation dépendant du temps qui mesurent la corrélation que conserve une grandeur après un temps donné. La dynamique moléculaire peut aussi être utilisée comme un outil d'observation

Encadré 1

La fonction de distribution de paires $g(r)$

Cette fonction est un objet statistique fondamental en théorie des liquides. Elle mesure le nombre moyen de molécules situées dans une calotte sphérique entre r et $r + dr$ autour d'une molécule donnée, divisé par la même quantité dans le cas d'une distribution homogène (elle vaut donc 1 aux grandes distances). Elle est représentée sur la figure dans le cas de la fonction de distribution de paires oxygène-oxygène dans l'eau liquide. Elle est caractéristique d'une structure moyenne tétraédrique. Le premier pic à 2,8 Å donne la position moyenne de quatre molécules voisines liées directement par liaison hydrogène, tandis que le pic suivant vers 4,5 Å est caractéristique de la position des deuxièmes voisins dans un réseau tétraédrique. La courbe présentée correspond à une simulation de dynamique moléculaire *ab initio* de 128 molécules d'eau à $T = 25$ °C et 1 atm en conditions aux limites périodiques. Elle a été calculée soit par une méthode standard d'histogrammes pour un échantillon de 10 000 configurations (calcul « convergé », trait noir), soit par un estimateur amélioré possédant une propriété de « variance finie » pour un échantillon de seulement cent configurations (trait bleu pointillé).



* Borgis D., Assaraf R., Rotenberg B., Vuilleumier R., Computation of pair distribution functions and three-dimensional densities with a reduced variance principle, *Mol. Phys.*, **2013**, 22-23, p. 3486.

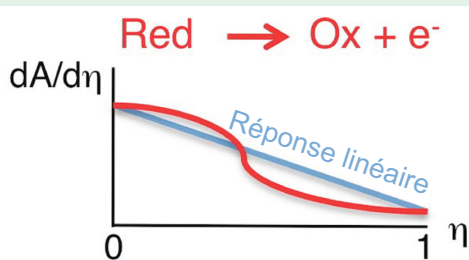
pour caractériser l'évolution moyenne d'un système hors d'équilibre, sa recherche d'un état d'équilibre et sa stabilisation.

La nécessité de recourir à des simulations numériques, au-delà de quelques systèmes modèles, fait que la mécanique statistique des liquides est une science récente qui ne s'est développée pleinement qu'avec l'avènement des ordinateurs modernes, d'abord à Los Alamos avec les travaux précurseurs de Metropolis *et coll.* [5] et Alder et Wainwright [6] sur le fluide de sphères dures, puis dans le monde entier avec le développement des centres nationaux de calcul, et en particulier du CIRCE à Orsay autour duquel une école française forte s'est développée dans les années 1960 [3, 7]. Aux liquides simples des débuts – fluides de Lennard-Jones, sels fondus, plasmas – a succédé l'étude des liquides moléculaires et des solutions [8], puis de systèmes moléculaires de plus en plus complexes [9] pour aller aujourd'hui par exemple – et pour ne citer que des travaux entrepris dans notre équipe – à l'étude par dynamique moléculaire *ab initio* de matériaux en conditions extrêmes comme les magmas du manteau terrestre, de la thermodynamique de réactions de transfert d'électrons dans l'eau (voir encadré 2), des propriétés infrarouges de l'eau en milieu confiné (voir encadré 3), ou à l'élucidation par simulation moléculaire tout-atome de la danse des molécules d'eau autour des protéines, ou de la déformation de milieux poreux après adsorption de molécules (voir encadrés 4, 5 et 6). Voir aussi l'application aux

Encadré 2

Détermination des potentiels d'oxydoréduction par simulation : théorie de Marcus et au-delà

Marcus, dans son célèbre travail sur les réactions de transfert d'électrons et les demi-réactions d'oxydoréduction, a proposé qu'un bon paramètre d'ordre pour décrire la réaction est la différence d'énergie entre les deux états réduits et oxydés du système. En remarquant que cette quantité est généralement très proche du potentiel électrostatique dû au solvant sur l'ion métallique, puis que ce potentiel dans un milieu diélectrique a une distribution de probabilité gaussienne, il a développé une formulation analytique du profil d'énergie libre en fonction du paramètre d'ordre, qui ne dépend que de deux paramètres, l'énergie libre de la réaction et l'énergie de réorganisation, qui est le travail dissipé lors d'une transformation idéale infiniment rapide. Suivant la théorie de Marcus, la force moyenne généralisée le long du paramètre d'ordre est linéaire en fonction de celui-ci, droite dont l'ordonnée à l'origine et la pente sont directement reliées à l'énergie libre de la demi-réaction et à l'énergie de réorganisation. En étudiant divers cations, nous avons identifié des situations où cette force généralisée n'est pas linéaire. C'est par exemple le cas pour les couples $\text{Cu}^+/\text{Cu}_2^+$ ou Ag^0/Ag^+ , dont la particularité est que la solvatation des espèces réduites et des espèces oxydées sont très différentes. Nous avons alors proposé deux extensions analytiques complémentaires à la théorie de Marcus qui reproduisent correctement le comportement non linéaire observé.



* Vuilleumier R., Tay K.A., Jeanmairet G., Borgis D., Boutin A., Extension of Marcus picture for electron transfer reactions with large solvation changes, *J. Am. Chem. Soc.*, **2012**, 134, p. 2067.

propriétés des liquides ioniques à température ambiante (*cf.* article d'A. Pádua, p. 63). Les records actuels en la matière concernent la dynamique moléculaire *ab initio* de systèmes impliquant plusieurs centaines d'atomes pendant des temps de l'ordre de la nanoseconde ou la dynamique moléculaire tout-atome d'un virus contenant plus d'un million d'atomes pendant quelques dizaines de nanosecondes [10]. Cette

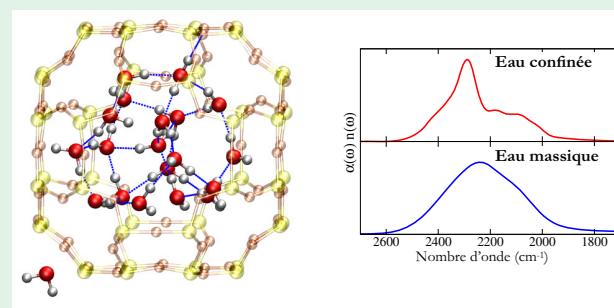
Encadré 3

Propriétés d'une gouttelette d'eau confinée

Propriétés électroniques, infrarouges

En utilisant différentes techniques de simulation moléculaire, nous avons étudié une large gamme de propriétés chimiques et physiques de l'eau confinée dans des matériaux zéolithiques hydrophobes et hydrophiles : les propriétés électroniques des molécules d'eau (par dynamique moléculaire *ab initio* de type Car-Parrinello), la structure et la dynamique du réseau de liaison hydrogène (par dynamique classique et *ab initio*) et la thermodynamique du phénomène d'adsorption et d'intrusion (par Monte-Carlo grand canonique et calcul d'énergies libres).

Nous avons montré que la dynamique des nanogouttes d'eau est proche de celle de l'eau massique. Nous avons caractérisé le réseau de liaisons hydrogène et observé que l'eau ne forme pas de liaisons hydrogène avec les oxygènes du matériau lorsqu'il est purement silicé. Le calcul direct du spectre infrarouge a révélé une bimodalité de la bande de vibration OH qui est la signature de la présence de deux types de liaison OH (impliquée dans une liaison hydrogène avec une autre molécule d'eau, ou pointant vers la zéolithe).

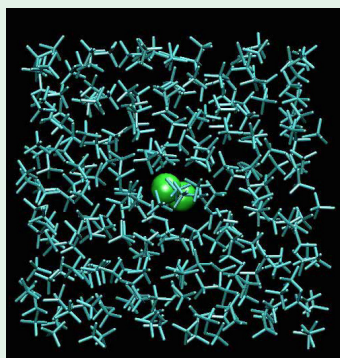


* Coudert F.-X., Cailliez F., Vuilleumier R., Fuchs A.H., Boutin A., Water nanodroplets confined in zeolite pores, *Faraday Discuss.*, **2009**, 141, p. 377.

capacité à augmenter les temps de simulation et à complexifier les systèmes étudiés est bien sûr le fruit de la performance sans cesse accrue des processeurs et de la fameuse loi de Moore⁽²⁾, mais pas seulement. La communauté a aussi été capable d'améliorer les algorithmes de simulation et de proposer des méthodes astucieuses d'exploration de l'espace des configurations, que ce soit en dynamique moléculaire ou Monte-Carlo, qui sont extrêmement efficaces et dites d'ordre N , c'est-à-dire avec un coût numérique croissant linéairement avec la taille du système. La stratégie

Encadré 4

Film du mouvement atomique lors d'une réaction chimique en phase liquide



La méthode de diffraction X résolue en temps appliquée à la phase liquide est une méthode récente et très prometteuse qui a pour but de mesurer en temps réel les déplacements atomiques lors de réactions chimiques photo-induites. Il est néanmoins nécessaire d'avoir recours à la simulation pour interpréter les signaux mesurés. Une des raisons principales pour ceci est la présence dans le

signal d'une contribution importante due au solvant lui-même, et non aux réactifs. Cette réponse du solvant est la trace des transferts

de bilan d'énergie au cours de la réaction photo-induite, qui conduit à un échauffement de l'environnement.

La dynamique du solvant contient alors des informations sur la cinétique de la réaction. Les modèles théoriques prédisent que la réponse du solvant correspond à une augmentation de température à volume constant, puis, de façon retardée, à une expansion thermique. Cette première étape, variation de structure due à un échauffement à volume constant, qui n'est pas couramment observée, peut être intéressante en soi pour comprendre la structure des liquides. La cinétique de l'expansion thermique dépend aussi beaucoup de la géométrie du dispositif expérimental que l'on peut modéliser à l'aide des équations de l'hydrodynamique.

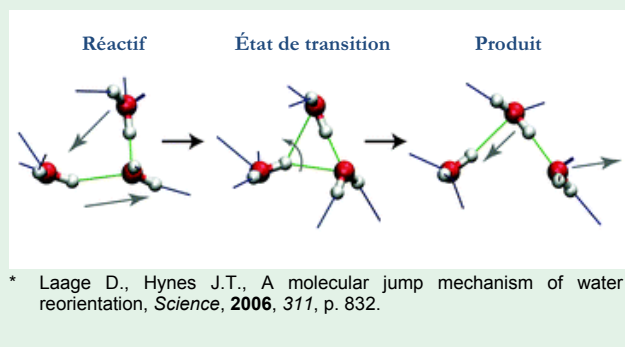
* Plecht A., Wulff M., Bratos S., Mirloup F., Vuilleumier R., Schotte F., Anfirrud P.A., Visualizing chemical reactions in solution by picosecond X-ray diffraction, *Phys. Rev. Lett.*, **2004**, 92, p. 125505.

Encadré 5

Découverte d'un nouveau mécanisme de réorientation de l'eau liquide

Dans l'eau pure, une façon de suivre les réorganisations du réseau de liaisons hydrogène est de mesurer la réorientation des molécules d'eau. Il était communément admis que cette réorientation se produisait selon un mécanisme de diffusion de Debye, correspondant à un mouvement brownien angulaire impliquant des pas angulaires infinitésimaux.

À partir de simulations de dynamique moléculaire, Damien Laage et Casey Hynes ont suggéré en 2006 un mécanisme radicalement différent où les molécules d'eau effectuent de grands sauts angulaires (60-70°) lors de l'échange d'accepteurs de liaison hydrogène. Ce mécanisme de saut peut être interprété comme une réaction chimique de type SN₂, sauf qu'ici il ne s'agit pas de liaisons covalentes mais de liaisons hydrogène dont les ruptures et formations se produisent de façon concertée. Le saut d'un accepteur à l'autre ne se produit qu'une fois qu'un nouvel accepteur s'est approché. Ce mécanisme contraste donc très fortement avec l'image antérieure où une molécule d'eau devait d'abord rompre totalement une ou plusieurs liaisons hydrogène pour devenir libre de se réorienter.



actuelle est d'utiliser des calculateurs massivement parallèles et de pouvoir distribuer les calculs sur des centaines, voire des milliers de processeurs.

Au-delà des progrès des algorithmes et des architectures informatiques qui permettent maintenant de simuler les fluides complexes avec des statistiques satisfaisantes, deux quantités thermodynamiques qui jouent un rôle essentiel en physique comme en chimie ne peuvent se réduire à de simples calculs de valeurs moyennes : l'entropie et l'énergie (ou enthalpie) libre.

Calcul de l'énergie libre

Déterminer l'énergie libre d'un système, ou son entropie, en échantillonnant, par dynamique moléculaire ou simulation Monte-Carlo, les configurations accessibles de ce système, revient à essayer de mesurer la surface d'un lac avec un fil à plomb. Cette image attribuable à Daan Frenkel illustre bien le cas particulier du calcul de l'énergie libre : le fil à plomb est parfait pour mesurer la profondeur moyenne du lac, mais ne nous renseigne pas sur son étendue. La situation n'est toutefois pas si désespérée. Nous sommes en fait le plus souvent intéressés à estimer la différence d'énergie libre entre deux états pour estimer la constante d'équilibre d'une réaction, une énergie libre de solvation, etc. Revenons alors à notre lac et à notre fil à plomb, et imaginons que nous cherchions à mesurer le rapport entre la superficie du lac et la surface recouverte d'algues, rapport dont le logarithme est l'analogie d'une différence d'énergie libre. Il suffit alors de se promener tranquillement en barque sur le lac et de temps en temps, de manière aléatoire par exemple, de laisser

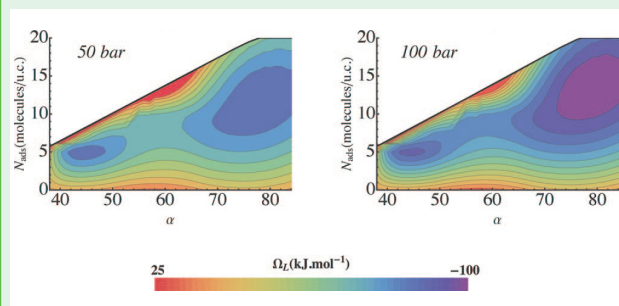
Encadré 6

Surface d'énergie libre pour la compréhension des changements de phases d'un matériau poreux induits par adsorption

Les solides hybrides organique-inorganique appelés « metal organic frameworks » (MOF), qui sont réalisés par auto-assemblage de centres métalliques reliés entre eux par des ligands organiques, présentent une gamme extrêmement variée de structures cristallines et de propriétés physico-chimiques.

L'une de leurs principales caractéristiques réside dans leur flexibilité, conduisant à d'importants changements de volume. Le problème posé par le couplage adsorption-déformation dans les solides compliants nanoporeux (MOF, et au-delà les cavités biomoléculaires telles que les canaux membranaires ou les cavités intraprotéiques), nécessite une approche rigoureuse, à l'échelle moléculaire, de thermodynamique statistique analytique et numérique. Un modèle thermodynamique analytique a été développé permettant de rationaliser les différents comportements observés expérimentalement dans divers MOF (comme l'existence ou non d'une ou plusieurs transitions de phases selon les matériaux et les adsorbats).

Très récemment, le calcul complet de la surface d'énergie libre en fonction de la pression du gaz et du volume du matériau a pu être réalisé dans un modèle simple. Cela a permis de caractériser complètement la thermodynamique d'un matériau compliant. On peut ainsi prédire le comportement de ces matériaux au cours de l'adsorption des gaz et leurs sélectivités d'adsorption, ouvrant la voie à l'optimisation de procédés de stockage et de séparation notamment.



descendre le fil à plomb. En comptant le nombre de fois où le fil à plomb remonte des algues, divisé par le nombre total d'essais, on détermine alors le rapport entre les deux superficies.

Cette image illustre la méthode d'insertion de Widom qui permet le calcul d'énergies libres de solvation. Cette méthode consiste à échantillonner les configurations d'un liquide, puis pour chaque configuration à tenter d'insérer le soluté de façon aléatoire. Ici, l'énergie associée, plus précisément le poids de Boltzmann correspondant à cette énergie, traduit le caractère favorable, ou non, de chaque configuration pour accommoder le soluté. Cependant, si très peu de configurations sont favorables, le risque est grand de ne pas réussir à insérer le soluté, ce qui se traduit par un bruit important sur le résultat. L'insertion d'une particule n'est par exemple favorable que si une cavité suffisamment grande s'est formée spontanément dans le solvant. Pour reprendre l'image précédente, si la surface recouverte d'algues est faible, il est important de n'échantillonner des configurations que dans une région restreinte du lac où on pense avoir *a priori* de grandes chances de trouver des algues. On va

alors biaiser l'échantillonnage des configurations de façon à améliorer la statistique ; il convient de soustraire l'effet du biais sur les résultats. Ces biais dépendent du problème et du système étudié et de nombreux biais ont été développés, enrichissant ainsi la boîte à outils des simulations.

Les relations thermodynamiques nous apprennent aussi que la différence d'énergie libre entre deux états est égale au travail le long d'une transformation réversible. Pour déterminer ce travail, on se ramène alors au calcul de la valeur moyenne d'une force, ce qui à nouveau est accessible par un échantillonnage des configurations. On parle alors souvent de potentiel de force moyenne pour décrire, par exemple, l'énergie libre le long d'une réaction de dissociation. Cependant, comme ce travail ne dépend pas du chemin suivi, puisque c'est la différence entre deux fonctions d'état, on peut tout à fait suivre des chemins non physiques, en faisant varier un paramètre d'ordre fictif. Ce peut être, par exemple, l'insertion progressive d'une particule en faisant croître son rayon.

Cette méthode d'intégration thermodynamique offre alors la possibilité de déterminer la différence d'énergie libre entre deux situations avec des potentiels d'interactions différents. Ceci a le gros avantage d'autoriser la détermination d'énergie libre absolue si on peut se ramener à un système pour lequel l'énergie libre est connue analytiquement (potentiels harmoniques, cristal d'Einstein) ou numériquement avec une bonne précision (comme pour le fluide de Lennard-Jones). Dans ce dernier cas, l'énergie libre est obtenue par intégration de l'équation d'état à partir de la limite basse densité, où l'on peut supposer le gaz parfait.

Jusqu'ici, nous avons contraint le système à une valeur fixée d'un paramètre d'ordre, qui décrit la transformation d'un état à un autre. Si au contraire le système est laissé libre, libre par exemple d'effectuer des réactions d'association-dissociation, alors une dernière définition de l'énergie libre peut être exploitée : il s'agit de l'énergie libre de Landau qui définit l'énergie libre comme le logarithme de la probabilité de trouver le système dans un état donné. La constante de la réaction d'association sera alors simplement écrite comme le logarithme du rapport de la probabilité de l'état dissocié et de la probabilité de l'état associé. La difficulté ici provient du fait qu'il est nécessaire d'échantillonner suffisamment les deux états ; or une barrière importante peut exister entre ceux-ci, ce qui va rendre les transitions très peu probables. La méthode d'Umbrella Sampling permet de forcer le système à échantillonner ces différents états intermédiaires, mais cette technique permet difficilement de traiter le cas où les différents états sont repérés par plusieurs coordonnées, appelées paramètres d'ordre. Il a alors été proposé d'accélérer l'exploration de cette surface d'énergie libre en parcourant les régions importantes de l'espace des paramètres avec un temps caractéristique beaucoup plus petit que le temps physique des transitions entre états stables. Là aussi, de nombreuses techniques ont été développées, dont les principales – l'algorithme de Wang-Landau [11] pour les simulations Monte-Carlo et la métadynamique [12] pour la dynamique moléculaire – ont pour principe de forcer le système à quitter une région déjà explorée. D'autres méthodes consistent par exemple à échantillonner la surface d'énergie libre à une température beaucoup plus grande que la température physique servant à définir l'énergie libre [13].

Mécanismes réactionnels en solution

Une surface d'énergie libre ainsi déterminée en fonction d'un petit nombre de paramètres d'ordre tels que la distance

entre des réactifs, une torsion ou la position d'une molécule clé du solvant, la polarisation du solvant, ou tout autre paramètre pertinent, constitue le point de départ pour étudier les mécanismes réactionnels, permettant par exemple d'identifier des états de transition. Cependant, il est nécessaire pour aller plus loin de connaître la dynamique sur cette surface d'énergie libre, dynamique qui n'est pas la simple dynamique de Newton car on a considérablement réduit la description du système. Seule cette dynamique permet de remonter à la cinétique de la réaction, mais aussi de distinguer une coordonnée lente qui va moduler cette cinétique d'une coordonnée rapide qui lui est indifférente. Surtout, la dynamique permet de tester le choix des paramètres d'ordre : est-ce qu'une fois passé l'état de transition putatif, le système évolue bien vers les produits de la réaction.

Pour étudier alors les mécanismes réactionnels, une nouvelle approche consiste à étudier directement les propriétés des trajectoires réactives. Ces trajectoires sont les trajectoires microscopiques qui dans le passé partent de l'état réactif et évoluent vers l'état produit (sans repasser par l'état réactif) : on parle de chemins de transitions [14]. Ces chemins peuvent être naïvement extraits d'une dynamique très longue au cours de laquelle un grand nombre de réactions entre les états produits et réactifs ont lieu. On ne regarde alors que les portions de trajectoire où on observe une telle transition entre les réactifs et les produits. Ceci est peu efficace car le système passe beaucoup de temps dans l'état produit ou dans l'état réactif avant d'effectuer une transition entre ces deux états. Il est cependant possible d'échantillonner directement les chemins de transition comme nous avons précédemment échantillonné les configurations, et là encore, la méthode Monte-Carlo est une méthode de choix. Plus encore pour l'échantillonnage des chemins, il est important d'utiliser des biais, et là aussi, des méthodes de biais ont été développées, en amenant progressivement des chemins partant des réactifs à rejoindre les produits.

Une fois obtenus ces chemins, la détermination des mécanismes réactionnels est ramenée à l'analyse des chemins de transition afin de déterminer les caractéristiques communes à ceux-ci. On cherche alors à décrire le mécanisme réactionnel par une coordonnée de réaction qui décrit l'évolution du système, des réactifs aux produits. Cette coordonnée de réaction doit distinguer les configurations qui appartiennent aux états réactifs et produits, mais aussi celles appartenant à l'état de transition. Pour cela, une nouvelle quantité a été introduite : le *commiteur* [14] d'une configuration, qui est la probabilité qu'une trajectoire aboutisse à l'état produit. Le commiteur permet de tester si un paramètre ou un ensemble de paramètres choisis conduit à une bonne coordonnée de réaction, distinguant les configurations de commiteurs différents. On pense à l'heure actuelle que la meilleure coordonnée de réaction est le commiteur lui-même : par le commiteur, on cherche à prédire l'évolution future probable d'un système vers l'état produit ou vers l'état réactif [15].

Il y a cependant une spécificité propre à la modélisation des réactions chimiques. Au cours de ces réactions, les électrons sont mis en jeu et les champs de forces utilisés doivent donc pouvoir rendre compte de ces réarrangements électroniques : brisures et formations de liaisons chimiques. Notons à ce propos que l'état électronique du système peut permettre aussi de définir un paramètre d'ordre pour décrire une réaction chimique : position des doublets électroniques, énergie des orbitales frontières... La dynamique moléculaire *ab initio* [16] consiste alors à coupler la dynamique moléculaire à un calcul de structure électronique à chaque

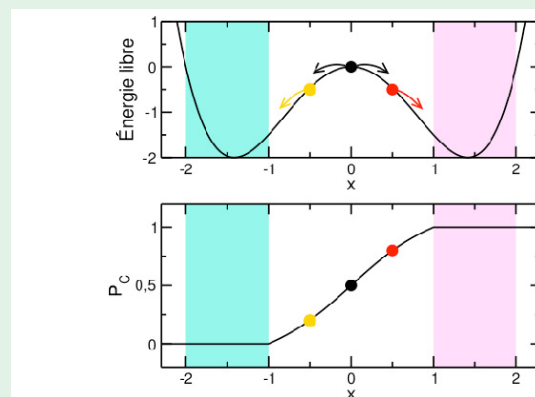
Encadré 7

Qu'est-ce que le commiteur ? Un exemple à une dimension

Considérons une particule brownienne à une dimension évoluant sur une surface d'énergie libre présentant un double puits (figure ci-dessous, panneau du haut). Cette surface d'énergie libre possède deux états stables, que nous désignerons comme l'état réactif (en bleu clair) et l'état produit (en rose clair) et on s'intéresse à la transition entre état réactif et état produit. Prenons trois configurations initiales du système : une (point orange) à gauche de la barrière, une (point noir) au sommet de la barrière, qui constitue l'état de transition du système, et une dernière (point rouge) à droite de la barrière. Des trajectoires partant du sommet de la barrière auront une chance sur deux, le système étant « indécis », d'aboutir dans l'état produit, et une chance sur deux de revenir dans l'état réactif (flèches noires). En revanche, des trajectoires partant des deux autres configurations initiales aboutiront avec une grande probabilité dans l'état produit ou dans l'état réactif respectivement (flèches rouge et orange).

Le commiteur, P_C , est alors la probabilité pour chaque configuration initiale qu'une trajectoire partant de ce point aboutisse dans l'état produit (panneau du bas). Il est proche de 0 pour le point orange et symétriquement proche de 1 pour le point rouge. Il vaut $\frac{1}{2}$ pour l'état de transition, point noir. De manière générale, le commiteur permet de distinguer les configurations qui appartiennent encore à l'état

réactif ($P_C < \frac{1}{2}$, point orange) des configurations appartenant déjà à l'état produit ($P_C > \frac{1}{2}$, point rouge). Dans des situations plus complexes que ce simple exemple, l'état de transition peut alors être repéré à partir de cette condition sur le commiteur, sans d'ailleurs nécessairement connaître la surface d'énergie libre.



configuration permettant d'obtenir les forces sur les atomes. Malgré de nombreux succès et un développement important, cette technique reste limitée à l'étude de quelques dizaines à centaines d'atomes. Il est possible d'augmenter significativement la taille des systèmes simulés, en particulier dans le but d'étudier la réactivité de molécules biologiques, par des méthodes mixtes classiques-quantiques. Dans ces méthodes, seule la région directement impliquée dans la réaction est traitée par des méthodes de chimie quantique, tandis que l'environnement (solvant, reste de la protéine) est représenté par des champs de forces classiques. Récemment, des champs de forces réactifs autorisant la réorganisation des liaisons chimiques avec un coût proche des potentiels empiriques ont été proposés.

La spectroscopie, dépendant du temps en particulier, permet d'étudier expérimentalement les mécanismes réactionnels. La modélisation, et là encore la dynamique moléculaire *ab initio* qui permet d'avoir accès à des grandeurs spectroscopiques, est complémentaire de ces techniques expérimentales en permettant l'interprétation des spectres expérimentaux et en validant les modèles développés (voir encadré 4, p. 73).

Mécanique statistique à différentes échelles

Même avec des méthodes astucieuses d'exploration de l'espace des configurations permettant de « forcer » certains événements rares (franchissements d'importantes barrières de potentiel), les systèmes moléculaires complexes impliquant des mouvements collectifs très lents peuvent difficilement être abordés avec des méthodes de simulation moléculaire tout-atome. C'est le cas par exemple des processus membranaires, des mécanismes de repliement des protéines ou d'auto-assemblage de biomolécules ou d'objets nanométriques. Il faut alors essayer de décrire la matière avec une résolution plus grossière que l'atome et définir des grains élémentaires regroupant plusieurs, voire un grand nombre d'atomes, ainsi que les interactions effectives entre ces grains : c'est le domaine de la modélisation gros grains. Une fois les grains et leurs interactions définis à l'échelle adéquate, les mêmes outils de mécanique statistique que ceux décrits précédemment et les mêmes outils de simulations

(Monte-Carlo, dynamique moléculaire) – ou dynamique brownienne si l'influence du solvant est modélisée par un écrantage électrostatique et par des forces de friction – peuvent être déroulés. Plusieurs modèles gros grains avec une échelle de description de l'ordre de la dizaine d'atomes ont été développés depuis une dizaine d'années pour l'étude des membranes lipidiques [17-18] et plus récemment pour l'étude des interactions protéine-protéine, protéine-acide nucléique et protéine-membrane [18-19]. Ce premier glissement d'échelle permet de gagner trois ordres de grandeur dans les temps de simulation accessibles et a permis d'éclaircir à un niveau quasi atomique nombre de processus membranaires comme la fusion, l'endo- ou exocytose, l'électroporation, la pénétration de molécules anesthésiques ou antibiotiques dans le milieu intracellulaire. L'extension de ces modèles à la prédiction systématique de la structure et de la stabilité des complexes de protéines est en cours. Pour continuer dans cette procédure de « zoom arrière » – en gardant autant que possible une approche systématique de chimiste théoricien –, les grains élémentaires peuvent être étendus à des structures secondaires de protéines, à des protéines entières, à des particules colloïdales ou à des nanoparticules, toutes ces entités pouvant être décrites par exemple par une frontière diélectrique et une distribution de charge. Dans cette veine, Marie Jardat et son équipe à Paris ont développé récemment des méthodes multi-échelles systématiques pour définir des potentiels effectifs rigoureux entre nano-objets chargés en présence de sel et pouvoir simuler, par exemple, des assemblées colloïdales ou des interactions protéines-ADN ou nanoparticules-ADN [20]. Aux États-Unis, A. Elcock et son équipe sont arrivés récemment à la simulation d'un modèle simplifié de cytoplasme bactérien incluant les 50 protéines les plus abondantes. Ils ont pu caractériser ainsi les phénomènes de diffusion et de rencontre dans ce type de milieu compact et engorgé [21]. Le défi à venir est de pouvoir conserver à ces échelles la notion de spécificité moléculaire, qui est essentielle pour la reconnaissance et la fonction, et de faire en sorte que des biomolécules puissent s'auto-assembler *in silico* pour former des structures biologiques opérantes.

Au bout du glissement d'échelle, la matière peut être assimilée comme un milieu continu et décrite par un ensemble de champs macroscopiques, grandeurs dont la valeur dépend

de la position dans l'espace (par exemple la densité locale de molécules, la densité locale de polarisation ou de charge ionique, le potentiel électrostatique local...) et éventuellement du temps. C'est le domaine de l'électrostatique des milieux continus, des équations cinétiques de réaction-diffusion et de l'hydrodynamique au sens large couplant réactivité et transport de matière et de chaleur. L'électrostatique des milieux continus en particulier est une approximation très répandue en chimie quantique pour incorporer à moindre coût les effets de solvant dans un calcul de structure électronique (méthode dite du milieu continu polarisé, PCM [22]), ou pour décrire l'environnement aqueux et ionique en modélisation biomoléculaire (méthode dite de Poisson-Boltzmann, PB). Bien que représentant l'alliance de la carpe et du lapin, le modèle microscopique le plus élaboré pour le soluté et une représentation très simplifiée du solvant environnant, ces méthodes peuvent être paramétrées pour produire des grandeurs énergétiques convenables à un coût numérique raisonnable, comparé par exemple à des méthodes hybrides QM/MM en chimie quantique (voir l'article de X. Assfeld et N. Ferré p. 43).

La notion de champ n'est cependant pas réservée au monde macroscopique, et on peut développer une mécanique statistique des champs, basée sur la notion de densité microscopique locale au sein du milieu, qui est valable à l'échelle moléculaire. C'est un défi actuel que de remplacer les modèles continus macroscopiques de solvation, type milieu continu polarisé ou Poisson-Boltzmann, par des méthodes continues de champs microscopiques, basées sur la théorie des équations intégrales [23] ou de la fonctionnelle de la densité moléculaire [24], et capables de reproduire/prédire de façon précise et extrêmement rapide par rapport aux simulations moléculaires la structure microscopique tridimensionnelle d'un solvant et de son cortège ionique au voisinage d'un soluté, ainsi que l'énergie libre de solvation correspondante.

Conclusion

Nous concluons en revenant sur le titre de notre article : les méthodes de mécanique statistique moléculaire, développées dans les années 1970 pour comprendre le comportement des liquides simples, se sont étendues progressivement dans les décennies suivantes pour comprendre les propriétés des liquides moléculaires, des mélanges et des solutions. L'orée du siècle a vu fleurir un bouquet de méthodes astucieuses d'exploration de l'espace des phases pour pouvoir caractériser des phénomènes aux temps longs – ou des événements rares –, en particulier les réactions chimiques en solution, avec des hypothèses minimales sur le mécanisme réactionnel attendu et les coordonnées de réaction pertinentes.

De façon concomitante, les algorithmes de simulation et les outils d'analyse statistique développés ont permis d'augmenter les échelles spatiales scrutées et d'aller vers l'étude, au niveau moléculaire, de liquides complexes (liquides ioniques à température ambiante, cristaux liquides, polymères fondus, solutions colloïdales...), d'interfaces liquide/solide, ou de la matière biologique.

Enfin, le développement de potentiels effectifs (ou gros grains) a permis d'appliquer les mêmes outils de mécanique statistique à des objets élémentaires de la taille du nanomètre, objets par ailleurs complexes vus à l'échelle atomique. L'étude des assemblées de tels objets et de leur possible auto-organisation (pour aller en biologie par exemple vers la modélisation supramoléculaire d'une cellule) constitue un des défis majeurs pour les années à venir.

Propriétés d'une gouttelette d'eau confinée

- (1) Les physiciens diront que l'hélium à 0 K est liquide à cause des fluctuations quantiques. Mais restons ici chimistes : gardons l'agitation thermique comme seule source du mouvement moléculaire !
- (2) Moore avait observé que la puissance des ordinateurs double environ tous les 18 mois.
- [1] Boutin A., Vuilleumier R., De Boltzmann aux expériences *in silico*, *L'Act. Chim.*, N° spécial « La chimie prépare notre avenir », vol. 2, **2011**, 353-354, p. 61.
- [2] Marx D., Hutter J., *Ab-initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, **2009**.
- [3] McQuarrie D., *Statistical Mechanics*, University Science Book, **2000**.
- [4] Hansen J.P., McDonald I.R., *The Physics of Simple Liquids*, Academic Press, 3rd ed., **2006**.
- [5] Metropolis N., Rosenbluth A., Rosenbluth M.N., Teller A.H., Teller E., Equation of state. Calculations by fast computing machines, *J. Chem. Phys.*, **1953**, **21**, p. 1087.
- [6] Alder B., Wainwright T.E., Phase transition for a hard sphere system, *J. Chem. Phys.*, **1957**, **27**, p. 1208.
- [7] Verlet L., Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.*, **1967**, **159**, p. 98.
- [8] Rahman A., Stillinger F.H., Molecular dynamics study of liquid water, *J. Chem. Phys.*, **1971**, **55**, p. 3336.
- [9] Levitt M., Warshel A., Computer-simulation of protein folding, *Nature*, **1975**, **253**, p. 694.
- [10] Freddolino P.L., Arkhipov A.S., Larson S.B., McPherson A., Schulten K., Molecular dynamics simulations of the complete satellite tobacco mosaic virus, *Structure*, **2006**, **14**, p. 437.
- [11] Wang F., Landau D.P., Efficient, multiple-range random walk algorithm to calculate the density of states, *Phys. Rev. Lett.*, **2001**, **86**, p. 2050.
- [12] Laio A., Parrinello M., Escaping free-energy minima, *PNAS*, **2002**, **99**, p. 12562.
- [13] Maragliano L., Vanden-Eijnden E., Single-sweep methods for free energy calculations, *J. Chem. Phys.*, **2008**, **128**, p. 184110.
- [14] Dellago C., Bolhuis P.G., Geissler P.L., Transition path sampling, *Adv. Chem. Phys.*, **2002**, **123**, p. 1.
- [15] Metzner P., Schütte C., Vanden-Eijnden E.J., Illustration of transition path theory on a collection of simple examples, *Chem. Phys.*, **2006**, **125**, p. 084110.
- [16] Car R., Parrinello M., Unified approach for molecular dynamics and density-functional theory, *Phys. Rev. Lett.*, **1985**, **55**, p. 2471.
- [17] Shelley J.C., Shelley M.Y., Reeder R.C., Bandyopadhyay S., Klein M.L., A coarse grain model for phospholipid simulations, *J. Phys. Chem. B*, **2001**, **105**, p. 4464.
- [18] Marrink S.J., Risselada H.J., Yefimov S., Tieleman D.P., de Vries A.H., The MARTINI force field: coarse grained model for biomolecular simulations, *J. Phys. Chem. B*, **2007**, **111**, p. 7812; Monticelli L., Kandasamy S.K., Periole X., Larson R.G., Tieleman D.P., Marrink S.J., The MARTINI coarse grained forcefield: extension to proteins, *JCTC*, **2010**, **4**, p. 819.
- [19] Basdevant N., Borgis D., Ha-Duong T. *et al.*, A coarse-grained protein potential derived from an all-atom force field, *J. Chem. Phys. B*, **2007**, **11**, p. 9390; Modelling protein-protein recognition in solution using the coarse-grained force-field SCORPION, *J. Chem. Theory Comp.*, **2013**, **9**, p. 803.
- [20] Dahirel V., Pailluson F., Jardat M. *et al.*, Non specific DNA-protein interaction: why proteins can diffuse along DNA, *Phys. Rev. Lett.*, **2009**, **102**, p. 228101; Effective interaction between charged nanoparticles and DNA, *PCCP*, **2011**, **13**, p. 12603.
- [21] McGuffee S.R., Elcock A.H., Diffusion, crowding and protein stability in a dynamic molecular model of the bacterial cytoplasm, *PLoS Comp. Biol.*, **2010**, **6**, p. e1000694.
- [22] Tomasi J., Menucci B., Cammi R., Quantum mechanical continuum solvation models, *Chem. Rev.*, **2005**, **105**, p. 2999.
- [23] Yoshida N., Imai T., Phongphanphanee S., Kovalenko A., Hirata F., Molecular recognition in biomolecules studied by statistical-mechanical integral-equation theory of liquids, *J. Phys. Chem. B*, **2009**, **113**, p. 873.
- [24] Borgis D., Gendre L., Ramirez R., Molecular density functional theory: application to solvation and electron-transfer thermodynamics in polar solvents, *J. Phys. Chem. B*, **2012**, **116**, p. 2504.



D. Borgis



A. Boutin



R. Vuilleumier

Daniel Borgis (auteur correspondant) et **Anne Boutin** sont directeurs de recherche au CNRS, et **Rodolphe Vuilleumier**, chargé de recherche, au Pôle de Physico-Chimie Théorique (UMR 8640 ENS/CNRS/UPMC), École Normale Supérieure, Paris*.

* Pôle de Physico-Chimie Théorique, UMR 8640 ENS/CNRS/UPMC, ENS, 24 rue Lhomond, F-75005 Paris.
Courriels : daniel.borgis@ens.fr ; anne.boutin@ens.fr ; rodolphe.vuilleumier@ens.fr