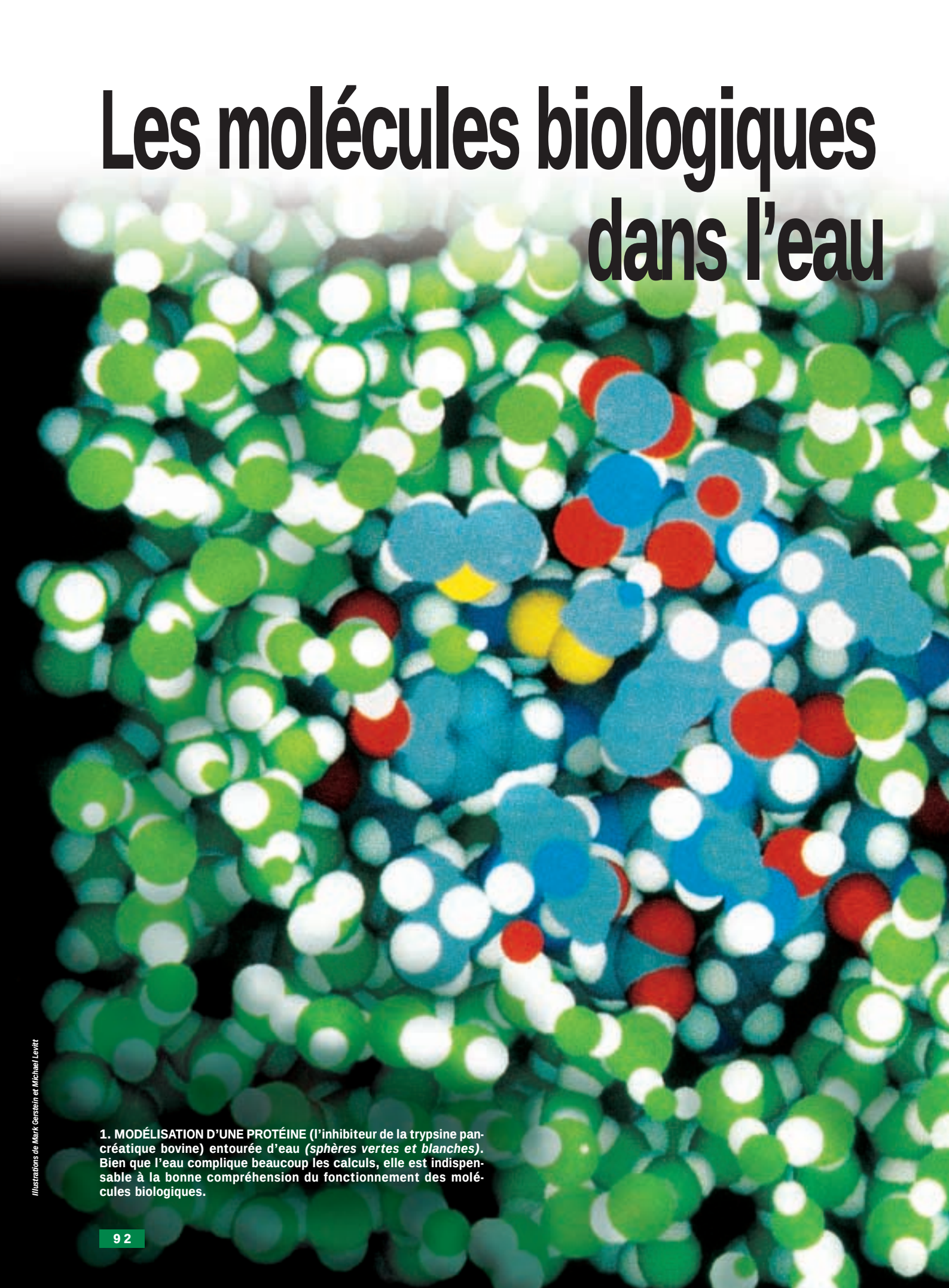


Les molécules biologiques dans l'eau



1. MODÉLISATION D'UNE PROTÉINE (l'inhibiteur de la trypsine pancréatique bovine) entourée d'eau (sphères vertes et blanches). Bien que l'eau complique beaucoup les calculs, elle est indispensable à la bonne compréhension du fonctionnement des molécules biologiques.

Les modélisations par ordinateur montrent l'influence de l'eau sur la structure et la dynamique des molécules biologiques.

MARK GERSTEIN • MICHAEL LEVITT

Au cours de l'été 1986, l'un de nous (M. Levitt) a dépensé plus de deux millions de francs pour une minuscule goutte d'eau. Il avait à peine de quoi mouiller une tête d'épingle, et l'eau n'était même pas réelle : à l'aide d'un puissant ordinateur, il calcula pendant deux semaines comment des molécules d'eau interagissent avec une protéine, modifiant sa structure et son mouvement.

La protéine étudiée était l'inhibiteur de la trypsine pancréatique bovine. Elle intéresse beaucoup les biologistes, car elle est petite, donc plus facile à étudier que la plupart des autres protéines. En 1977, Martin Karplus et ses collègues de l'Université Harvard avaient modélisé cette protéine, mais seulement dans le vide, sans interaction avec d'autres molécules. Personne ne savait comment l'inhibiteur de la trypsine pancréatique bovine se comporte quand la protéine est entourée de milliers de molécules d'eau, dans les cellules vivantes.

M. Levitt et sa collègue Ruth Sharon ont finalement découvert que le modèle dans le vide était inapproprié. Après eux, plusieurs équipes ont étudié la structure d'autres molécules biologiques dans leur milieu aqueux naturel.

Aujourd'hui, les ordinateurs de bureau mettent moins de deux heures pour modéliser des protéines telles que l'inhibiteur de la trypsine pancréatique bovine et ses molécules d'eau associées (l'expérience coûte alors moins de dix francs). Les biologistes ont simulé les structures dans l'eau de plus de 50 protéines et acides nucléiques tels que l'ADN.

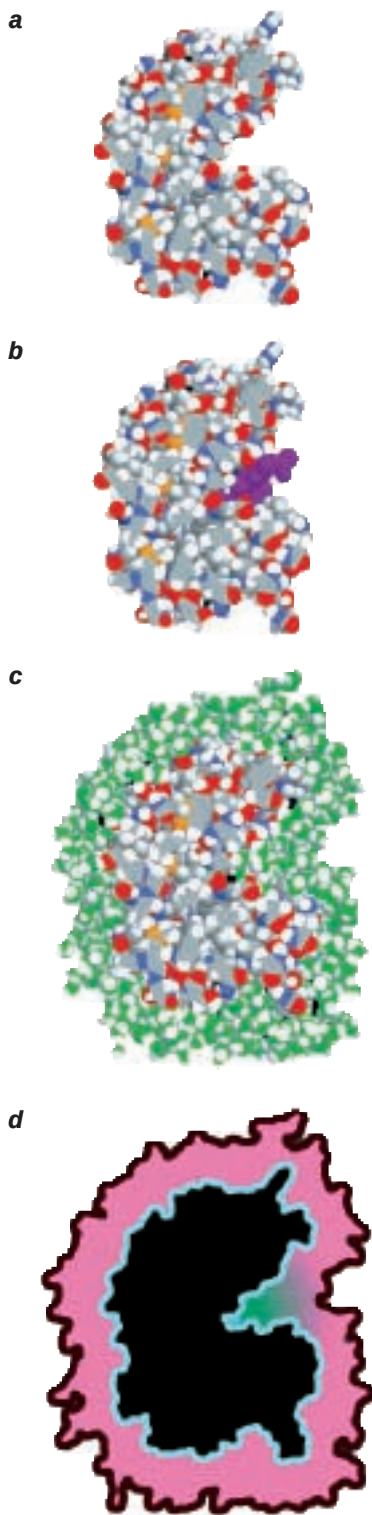
Pourquoi la compréhension des effets de l'eau sur la forme des molécules biologiques est-elle si importante? D'une part, les mécanismes moléculaires essentiels à la vie dépendent de la structure moléculaire et de ses interactions avec son milieu. D'autre part, la compréhension de la structure des molécules biologiques en milieu aqueux aidera les pharmaciens à concevoir de nouveaux médicaments capables de bloquer ou de faciliter diverses réactions biochimiques.

L'eau virtuelle

Une molécule d'eau (H_2O) est composée de deux atomes d'hydrogène et d'un atome d'oxygène. Chaque atome d'hydrogène apporte un électron sur une couche externe qui en contient jusqu'à deux, et l'atome d'oxygène apporte huit électrons, dont six électrons sur sa couche externe, qui en comporte huit au maximum. Les atomes se lient en mettant en commun leurs électrons, ce qui complète leur couche externe : l'atome d'oxygène met en commun deux électrons, chacun avec un atome d'hydrogène, de sorte que l'atome d'oxygène a alors huit électrons sur sa couche externe et que chaque atome d'hydrogène a alors deux électrons sur sa couche externe. Les quatre électrons restants de l'atome d'oxygène se regroupent par paires. Au total, la molécule d'eau a une forme de tétraèdre dont l'atome d'oxygène occupe le centre. Les deux atomes d'hydrogène et les deux paires d'électrons s'orientent chacun vers un des quatre sommets (voir la figure 3). Comme les électrons des atomes d'hydrogène sont plus souvent autour de l'atome d'oxygène qu'autour de leurs propres noyaux chargés, la molécule est polaire : elle a deux «nuages» de faible charge négative autour de l'atome d'oxygène, et ses deux atomes d'hydrogène restent avec de faibles charges positives. Au total, toutefois, ces deux types de charges s'équilibrent, et les molécules d'eau sont électriquement neutres.

Les chimistes représentent souvent les molécules d'eau sous forme d'un V (voir la figure 4). Chaque côté du V correspond à une liaison oxygène-hydrogène d'environ un dix-millionième de millimètre de longueur. L'angle formé par les deux cotés du V est proche de 105 degrés.

La polarité de l'eau favorise certaines interactions entre les molécules d'eau : chaque atome d'hydrogène, chargé positivement, d'une molécule d'eau est attiré par un atome d'oxygène, chargé négativement, d'une autre molécule. Ces interactions sont des liaisons



2. LE LYSOZYME est une enzyme qui tue les bactéries en décomposant les molécules de sucre de leur paroi. Son site actif se situe dans le sillon principal de la protéine (a). La forme du sillon s'adapte aux molécules (sphères violettes) qu'elle clive (b). La modélisation (c) de l'interaction de l'eau (sphères blanches et vertes) et du sillon permet aux chercheurs d'établir une carte du site actif (d). La zone verte indique les molécules d'eau facilement déplacées. De telles cartes sont fondamentales dans la conception de nouveaux médicaments inhibiteurs ou activateurs d'une enzyme.

hydrogène. Chaque molécule d'eau liquide forme quatre liaisons hydrogène. Cependant, le nombre de liaisons hydrogène autour d'une molécule d'eau liquide semble pouvoir varier entre trois et six, avec une moyenne de 4,5. Cette variabilité s'explique par la structure aléatoire et fluctuante de l'eau liquide, à la différence de la glace où les molécules d'eau se disposent selon un arrangement tétraédrique régulier. Le maintien d'une structure tétraédrique impose aux molécules d'eau une conformation plus lâche que celle de la plupart des autres liquides, tels les huiles ou l'azote liquide.

Pour simuler un volume d'eau, on doit représenter deux types de forces : intramoléculaires et intermoléculaires. On modélise les interactions des atomes, dans une molécule d'eau, par des forces à court rayon d'action, semblables à des ressorts : ces forces correspondent aux liaisons chimiques entre chaque atome d'hydrogène et l'atome d'oxygène de la molécule. Les forces intramoléculaires imposent des valeurs précises aux longueurs des liaisons entre l'oxygène et les atomes d'hydrogène, et à l'angle formé par ces liaisons. D'autre part, on modélise les interactions de plusieurs molécules d'eau par des forces électriques à plus longue portée.

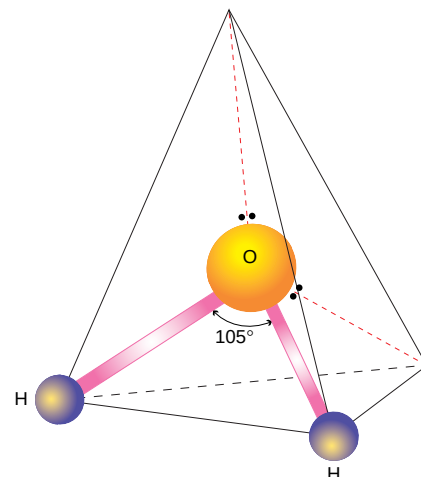
Les forces intermoléculaires ont un comportement différent : leur intensité diminue lorsque la distance augmente. Elles résultent de l'attraction des charges opposées et de la répulsion de charges identiques. Ces interactions engendrent les liaisons hydrogène, ainsi que des forces attractives plus faibles, nommées forces de Van der Waals.

Biologie simulée

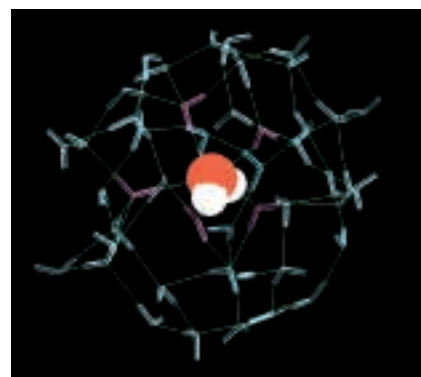
À la fin des années 1960, aux Laboratoires *Bell*, Aneesur Rahman et Frank Stillinger ont été les premiers à modéliser des molécules d'eau par ordinateur. Ils ont simulé le mouvement de 216 molécules d'eau dans une boîte rectangulaire pendant cinq milliardièmes de milliseconde (216 molécules s'agencent en un parallélépipède de six molécules de hauteur, de six molécules de largeur et de six molécules de profondeur). La simulation a reproduit quantitativement la plupart des propriétés de l'eau, telles que sa structure moyenne, sa vitesse de diffusion et sa température d'évaporation.

L'eau est importante dans les organismes vivants, parce qu'elle forme des

liaisons hydrogène avec d'autres molécules d'eau, et aussi parce qu'elle interagit avec divers types de molécules biologiques. Parce qu'elle est polaire, l'eau se lie à d'autres molécules polaires ou chargées, tels les acides, les sels, les sucres, et diverses régions des protéines et de l'ADN. Grâce à ces interactions, l'eau dissout les molécules polaires, ainsi qualifiées d'hydrophiles, ce qui signifie étymologiquement «qui aiment



3. LA MOLÉCULE D'EAU s'inscrit dans un tétraèdre dont le centre est occupé par l'atome d'oxygène (sphère orange). Deux des six électrons de la couche externe de l'oxygène sont mis en commun dans deux liaisons «covalentes» (barres roses) avec ceux des deux atomes d'hydrogène (sphères violettes). Ainsi les atomes ont leurs couches externes complètes. Les deux paires d'électrons restantes (points noirs) autour du noyau d'oxygène sont orientées vers les deux sommets.



4. LES LIAISONS HYDROGÈNE confèrent à l'eau ses propriétés particulières. Dans ce modèle d'eau liquide, la molécule centrale (sphères rouge et blanches) a formé des liaisons hydrogène (lignes vertes) avec cinq autres molécules d'eau (représentées par des V roses). Les atomes d'hydrogène de cette molécule (blancs) se lient à des atomes d'oxygène de deux autres molécules d'eau, et son atome d'oxygène (rouge) est lié à un atome d'hydrogène de trois autres molécules d'eau. Chaque molécule d'eau liquide crée en moyenne 4,5 liaisons hydrogène.

l'eau». En revanche, l'eau se lie très peu avec les molécules non polaires, telles les graisses, ce qui explique que l'eau et l'huile ne se mélangent pas à l'équilibre (les molécules non polaires sont hydrophobes, étymologiquement «qui n'aiment pas l'eau»).

Certaines molécules biologiques, par exemple les protéines et l'ADN, contiennent à la fois des parties hydrophiles et des parties hydrophobes, organisées en longues chaînes. La structure tridimensionnelle de ces molécules résulte du repliement des chaînes sur elles-mêmes, de sorte que les groupes hydrophiles sont en surface, où ils interagissent parfois avec l'eau, et les groupes hydrophobes sont enfouis à l'intérieur de la molécule, à l'abri de l'eau. Dès 1959, Walter Kauzmann indiqua que cet effet hydrophobe détermine le repliement de la protéine ; il est aujourd'hui encore très étudié (voir *Les mécanismes du repliement des protéines*, par Frederic Richards, *Pour la Science*, mars 1991).

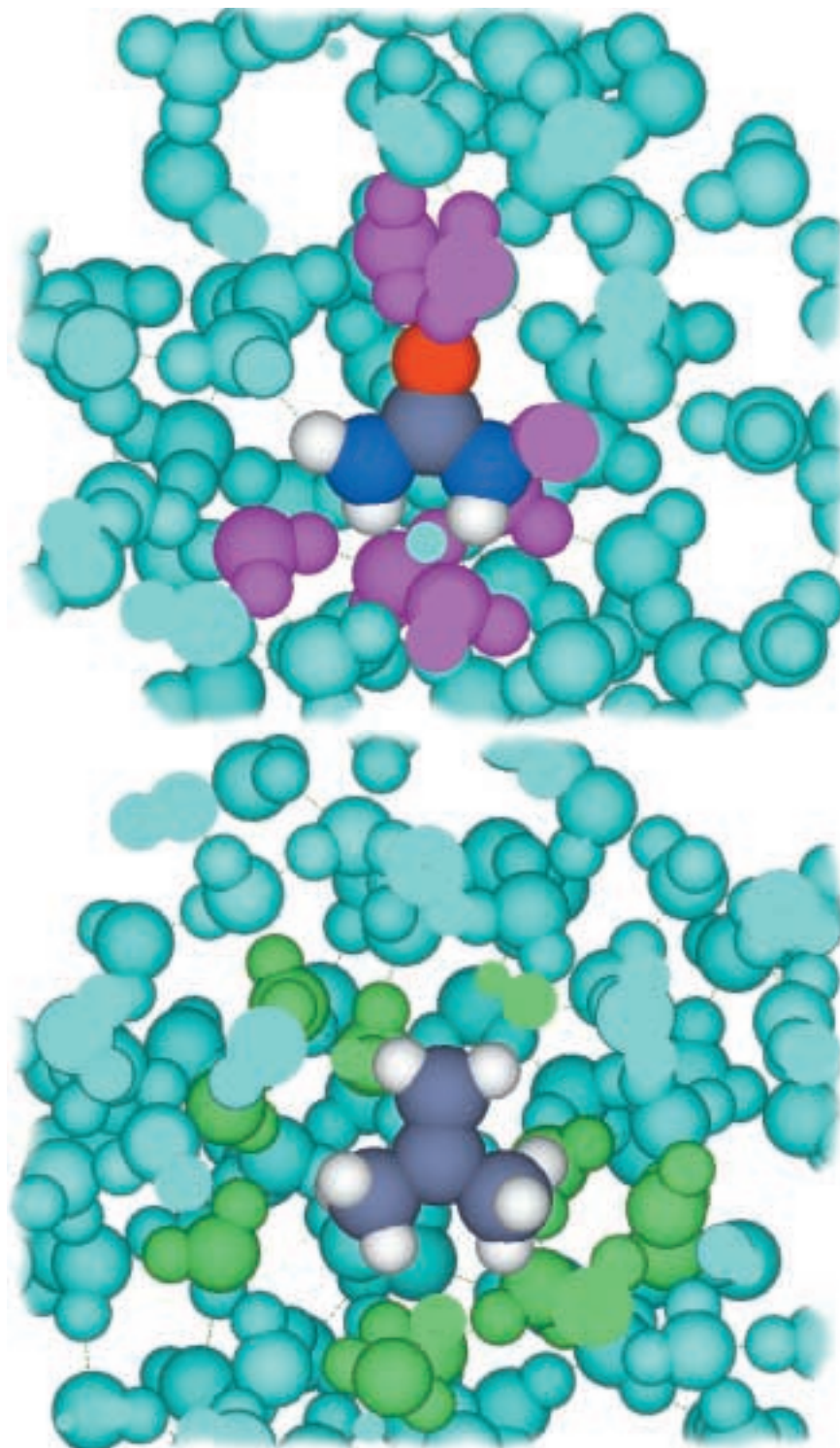
Les modélisations distinguent trois types d'eau : l'eau «ordonnée», qui entoure les molécules biologiques, interagissant fortement avec elles, l'eau «de masse», qui est au-delà et les molécules d'eau, qui sont enfouies dans les molécules biologiques. Une seule cellule contient des milliards de molécules d'eau, qui emplissent tous les espaces non occupés par des molécules biologiques. Au total, le corps humain contient environ 60 pour cent d'eau.

Comment modéliser les interactions des trois types d'eau avec les atomes d'une molécule biologique ? Nous spécifions tout d'abord les interactions élémentaires de tous les atomes, puis nous laissons le système évoluer selon les lois de la physique newtonienne. Une telle simulation requiert deux ingrédients fondamentaux : d'une part, la définition des interactions inter- et intramoléculaires de l'eau et des molécules biologiques ; d'autre part, la planification de leurs mouvements au cours du temps, c'est-à-dire la dynamique moléculaire.

Au cours de l'évolution d'un volume d'eau qui contient des molécules biologiques, chaque atome bouge. Sa position à chaque instant se détermine par rapport à sa position à l'instant précédent : on ajoute la distance parcourue entre les deux instants considérés. Quand aucune force n'agit sur l'atome, la distance qu'il parcourt est le produit du temps d'évolution par sa vitesse à la position précédente. Toutefois, les

forces exercées par les autres atomes ont généralement une résultante qui accélère l'atome. Quand les forces sont constantes pendant le laps de temps, nous calculons la nouvelle vitesse par l'application des lois de Newton (la

variation de la vitesse est proportionnelle à la force). Puis nous utilisons cette vitesse actualisée pour calculer la nouvelle position de l'atome. Les interactions, nombreuses dans un liquide, empêchent les atomes de se déplacer



5. DES MOLÉCULES aux formes semblables réagissent différemment avec l'eau selon qu'elles sont polaires ou non, c'est-à-dire selon qu'elles ont ou non une charge électrique partielle sur certains de leurs atomes. L'urée, une molécule polaire présente dans l'urine, forme des liaisons hydrogène avec les molécules d'eau (sphères violettes, en haut). En revanche, l'isobutène, apolaire, ne forme pas de telles liaisons ; les molécules d'eau créent des liaisons hydrogène entre elles et produisent une structure en forme de cage (sphères vertes, en bas).

notablement. Le temps utilisé est très court, de l'ordre d'une femtoseconde (un millionième de milliardième de seconde), pendant lequel une molécule d'eau ne bouge que de 1/500 de son diamètre.

Une longue simulation de tous les atomes d'une molécule biologique, avec son eau ordonnée, requiert de nombreux calculs à chaque laps de temps et produit une quantité de données considérable. Par exemple, pour simuler le mouvement d'une petite protéine dans l'eau, on doit calculer un demi-million de séries de coordonnées cartésiennes par nanoseconde (un milliardième de seconde); chacune décrit la position d'environ 10 000 atomes. La visualisation détaillée d'une telle simulation permet de voir chaque molécule tourner, se déplacer et vibrer pendant des millions d'images.

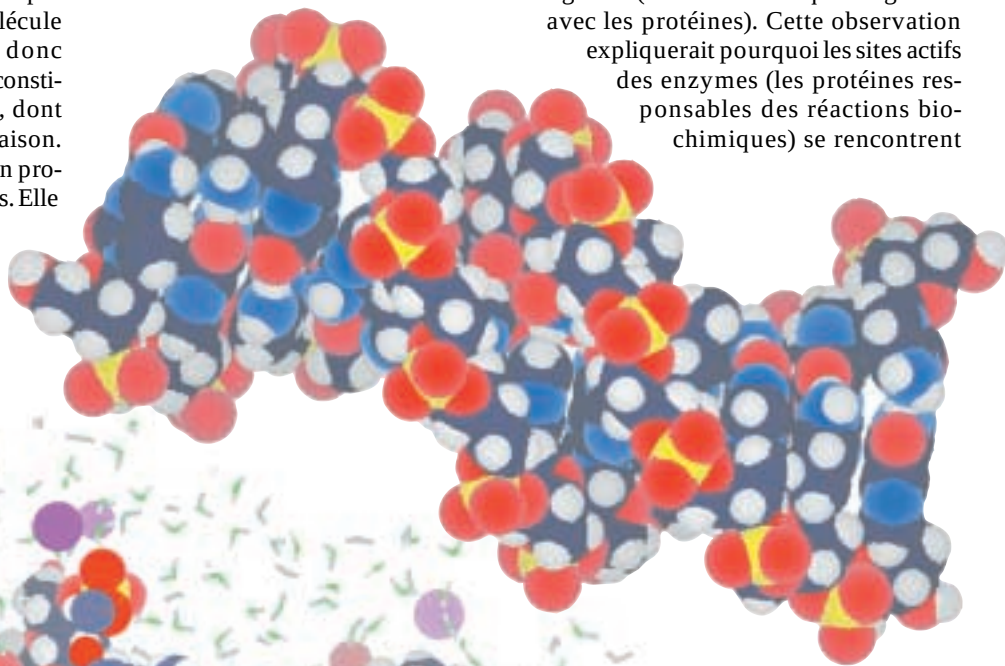
Comment l'eau détermine-t-elle la dynamique moléculaire? Pour répondre à cette question, examinons deux petites molécules organiques, l'isobutène et l'urée, qui ont des formes similaires, mais des propriétés chimiques différentes. L'isobutène, un combustible produit par raffinage du pétrole, est une molécule en forme de Y, non polaire (et donc hydrophobe), dont le squelette est constitué de quatre atomes de carbone, dont deux sont liés par une double liaison. L'urée, excrétée dans l'urine, est un produit du métabolisme des protéines. Elle a aussi une forme en Y, avec un groupe carbonyle (C=O) lié à deux groupes amine (NH₂). L'urée est une molécule polaire, c'est-à-dire hydrophile.

Les simulations de la dynamique moléculaire de l'isobutène et de l'urée montrent que l'eau se comporte de manière différente autour des deux molécules (voir la figure 5). Les molécules d'eau forment des liaisons hydrogène avec l'atome d'oxygène et les atomes d'hydrogène de l'urée. En revanche, les molécules d'eau autour de l'isobutène, hydrophobe, ne créent des liaisons hydrogène qu'entre elles. Elles forment ainsi une cage d'eau, qui entoure la molécule. Ce type d'études éclaire le comportement de l'eau autour des molécules biologiques plus complexes, telles que les protéines et les acides nucléiques. Par exemple, l'eau est indispensable au maintien de la structure de l'ADN. Les premières tentatives de modélisation de la dynamique moléculaire de l'ADN dans le vide ont échoué, car les forces répulsives entre les groupes phosphate, chargés négativement, qui constituent, avec des molécules de sucre, le squelette de la double hélice d'ADN, disloquent la molécule après seulement 50 picosecondes (un milliardième de milliseconde).

À la fin des années 1980, à Londres, M. Levitt et Miriam Hirshberg ont obtenu une modélisation de l'ADN de 500 picosecondes en «ajoutant» des molécules d'eau qui ont stabilisé la structure de la double hélice grâce à la formation de liaisons hydrogène avec les groupes phosphate (voir la figure 6). Les simulations ultérieures d'ADN en milieu aqueux ont montré que les molécules d'eau interagissent avec la plupart des parties de la double hélice d'ADN, y compris avec les paires de bases qui constituent le code génétique.

En revanche, l'eau ne pénètre pas jusqu'au cœur hydrophobe des protéines. Par conséquent, les simulations des protéines avec l'eau n'ont concerné que la surface de la protéine, qui est plus exposée que l'intérieur.

L'interaction des molécules d'eau et de la surface des protéines dépend de leur géométrie particulière et, notamment, des profonds sillons de la surface des protéines. Les molécules d'eau liées par les liaisons hydrogène se lient mal à ces sillons, de sorte qu'elles sont facilement remplacées par des ligands (les molécules qui réagissent avec les protéines). Cette observation expliquerait pourquoi les sites actifs des enzymes (les protéines responsables des réactions biochimiques) se rencontrent



6. LA DOUBLE HÉLICE DE L'ADN (en haut) est composée de deux montants torsadés autour des paires de bases (sphères grises, bleues, rouges et blanches); chaque montant est une alternance de sucres et de groupes phosphate (sphères rouges et jaunes). Une image extraite d'une simulation de l'ADN en milieu aqueux (en bas) montre que les molécules d'eau (V verts) s'immiscent jusqu'au centre de la double hélice de l'ADN et la stabilisent. Les sphères violettes représentent des ions sodium en solution.



7. LES HÉLICES ALPHA sont des motifs fréquents dans les protéines. Elles se déroulent plus facilement dans l'eau, car les molécules d'eau remplacent les liaisons hydrogène (lignes vertes) qui maintiennent normalement l'hélice. Une molécule d'eau (verte) relie ici un atome d'oxygène (rouge) à un groupe azote-hydrogène (bleu), qui habituellement forment à eux seuls une liaison hydrogène dans l'hélice repliée.

souvent dans des sillons (voir la figure 2). D'autre part, comme la disposition des molécules d'eau dans un site actif vide reproduit souvent la géométrie et la structure du substrat de l'enzyme, on utilise parfois les simulations pour concevoir des médicaments.

Biologie réelle

Faute de techniques expérimentales adaptées et précises, on apprécie mal la fidélité des simulations de molécules biologiques en milieu aqueux : on compare des résultats expérimentaux aux seules valeurs moyennes et globales obtenues par simulations.

La diffraction des neutrons et la diffraction des rayons X sont parmi les techniques le plus souvent utilisées pour la vérification des structures obtenues par simulation des molécules biologiques en milieu aqueux. Dans une expérience de diffraction des neutrons, un faisceau de neutrons est dirigé sur un petit échantillon cristallisé ; on enregistre la manière dont les neutrons sont dispersés par les molécules de l'échantillon. Chaque espace entre les molécules agit comme une petite fente qui produit un motif de diffraction caractéristique. L'analyse de ces motifs détermine l'espace-ment des molécules. Les résultats obtenus par la diffraction des neutrons sont proches des résultats moyens des simulations par ordinateur ; les distances coïncident.

On compare également la dynamique et le comportement prévus par les simulations moléculaires aux propriétés observées en laboratoire. Par exemple, la plupart des protéines contiennent au moins une «hélice alpha», une structure due à l'enroulement de la chaîne d'acides aminés en

raison de liaisons hydrogène entre des atomes de cette même chaîne. Grâce à des expériences, nous savons aujourd'hui que le chauffage des protéines déroule ces hélices alpha. Cependant, lors des premières tentatives de simulation du comportement d'une seule hélice alpha dans le vide à des températures élevées, l'hélice demeurerait intacte. En «ajoutant» de l'eau à la simulation, M. Levitt et Valerie Daggett, de l'Université de Washington, ont réussi à modéliser le comportement réel de l'hélice alpha (voir la figure 7).

Ces simulations par ordinateur fournissent de plus en plus d'informations sur la forme de diverses molécules biologiques et sur leur fonctionnement dans les organismes vivants. Cependant, les capacités de l'informatique et le coût d'utilisation des superordinateurs limitent en permanence les simulations de molécules biologiques en milieu aqueux. Quand les chimistes et les biologistes publient des modèles de molécules biologiques dans des revues, ils dessinent en général leurs modèles avec des couleurs vives sur un fond noir. Nous savons aujourd'hui que ce fond est aussi important que les molécules elles-mêmes et qu'il devrait avoir la couleur de l'eau.

Mark GERSTEIN est professeur adjoint à l'Université Yale. Michael LEVITT est professeur à l'Université de Stanford.

Michael LEVITT et Britt H. PARK, *Water: New You See It, New You Don't*, in *Structure*, vol. 1, n° 4, pp. 223-226, 15 décembre 1993.

Mark GERSTEIN et Cyrus CHOTHIA, *Packing at the Protein-Water Interface*, in *Proceedings of the National Academy of Sciences USA*, vol. 93, n° 19, pp. 10167-10172, 17 septembre 1996.

L'UNIVERS DES SCIENCES

L'UNIVERS DES SCIENCES

LES ÉTOILES

Vie et mort des soleils lointains

JAMES KALER



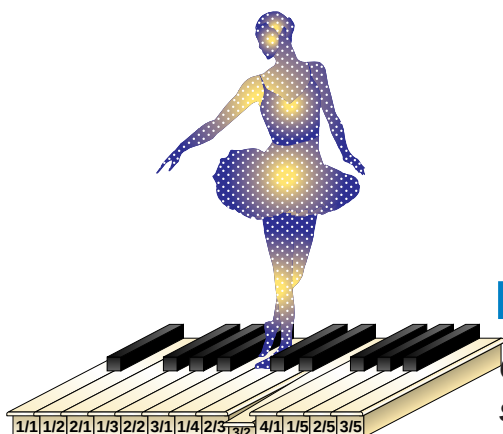
LES ÉTOILES

Ce livre expose l'origine et la nature de nos connaissances sur les étoiles, avant de les décrire dans leur dynamique et leur devenir. Les objets célestes les plus insolites, géantes rouges, supernovae, pulsars et trous noirs, s'insèrent alors naturellement dans le récit de l'évolution stellaire.

272 pages - 220 F
code 1892

**POUR LA
SCIENCE**
DIFFUSION BELIN

Bon de commande en p.98



Négligeable, mais troublant

JEAN-PAUL DELAHAYE

Certains ensembles infinis et partout denses sont négligeables... et ce n'est pas un paradoxe.

Le point n'a pas d'épaisseur : en enlevant successivement des points à une droite, nous ne la vidons jamais. Nous pouvons même en enlever 1 000 milliards sans y créer de trou véritable : un être se déplaçant dans ce qui reste, en faisant des pas aussi petits que l'on veut, pourra toujours aller de n'importe quel point à n'importe quel autre (voir la figure 1).

C'est seulement quand nous enlevons un intervalle complet à une droite que le déplacement des êtres aux trop petites jambes est impossible. En ôtant par exemple l'intervalle $[-1, +1]$ (tous les nombres compris entre -1 et $+1$), nous interdisons le passage de la partie positive à la partie négative à un marcheur dont les pas sont de longueur inférieure à 2, car il ne peut franchir le trou central, trop large.

Même la suppression d'un ensemble infini, comme l'ensemble des nombres rationnels (quotients de deux entiers), ne gêne pas la démarche à petits pas des êtres minuscules. Aussi n'est-ce pas la finitude ou l'infinitude qui permet d'élucider le problème : d'autres notions sont nécessaires pour étudier quels sont les trous négligeables de l'ensemble des nombres réels. Quelles sont ces notions ? Pouvons-nous examiner ce qui reste des ensembles « quand on leur a presque tout enlevé » ?

Le domaine des mathématiques, dénommé topologie de la droite réelle ou théorie du continu, s'est constitué pour traiter ce sujet. Cette théorie est troublante : alors que tout s'y développe de manière parfaitement rationnelle, le visiteur s'y trouve vite confronté à des situations presque paradoxales où l'intuition hésite et vacille.

La crainte de tomber sur des contradictions et l'idée préconçue que la compréhension de l'infini pris comme un tout (ce qu'on appelle l'infini actuel) était réservée à Dieu, ont longtemps paralysé les mathématiciens. Ce n'est qu'au XIX^e siècle que Bernhard Bolzano (1781-1848), Karl Weierstrass (1815-1897), Richard

Dedekind (1831-1916) et Georg Cantor (1845-1918) ont découvert que l'on peut classer et comparer les sous-ensembles infinis de l'ensemble des nombres réels sans rencontrer de contradictions. Ce fut

un émerveillement qui bouleversa les mathématiciens et en troubla gravement certains (Cantor est mort fou). Cela fait déjà un siècle que nous avons pénétré cet univers dangereux, et nous n'avons pas rencontré de paradoxes fatals. Après une période d'instabilité et de doute au début du XX^e siècle – période appelée « crise des fondements » –, les mathématiciens ont mis au point une théorie des ensembles qui apparaît inébranlable et sert de support à la topologie.

Est-il néanmoins si sûr que tout soit bien résolu ? Certains logiciens et philosophes soutiennent que nous sommes toujours au bord du précipice. L'heuristique de Borel que nous allons présenter et que les résultats récents de mathématiques expérimentales rendent vraisemblable est une des marques flagrantes de cette étrangeté du monde de la topologie.

LE NÉGLIGEABLE

L'heuristique de Borel se fonde sur la notion d'ensemble négligeable (ou ensemble de mesure nulle). Par définition, un sous-ensemble de l'ensemble des nombres réels est négligeable si on peut l'enfermer dans une suite – finie ou infinie – d'intervalles dont la longueur totale est aussi petite qu'on veut. Cette définition a été formulée pour la première fois en 1881 par A. Harnack dans un travail sur l'intégrale de Riemann, mais c'est Borel, en 1894, qui établira la théorie complète des ensembles mesurables.

Un point A à lui seul est un ensemble négligeable, car on peut l'enfermer dans l'intervalle $[A - 1/100, A + 1/100]$ qui a pour longueur $1/50$ (la longueur de l'intervalle $[a, b]$ des nombres compris entre a et b est $b - a$). En prenant $1/1\,000$ à la place de $1/100$, on enfermerait A dans un intervalle de longueur $1/500$, et finalement, quelle que soit la longueur L positive qu'on se donne, A peut être enfermé dans un intervalle de longueur L : $[A - L/2, A + L/2]$. Un ensemble fini de points, par exemple

