

# Chimie et innovation : les nouveaux outils du design moléculaire

Nicolas Froloff (86) et Hervé Plessix (86)\*,  
*L'Oréal Recherche, Aulnay-sous-Bois, France*

**L**A CHIMIE ne cesse de changer notre vie et notre quotidien. Chaque jour des équipes de recherche identifient de nouvelles molécules aux propriétés intéressantes : des molécules qui vont interagir avec le vivant et modifier son fonctionnement et qui deviendront des médicaments, des molécules qui vont s'assembler pour constituer des matériaux aux propriétés nouvelles (verres, céramiques, fibres de carbone, etc.), ou encore des molécules qui sont capables d'absorber la lumière visible ou les UV et qui deviendront des colorants ou des filtres solaires.

Aujourd'hui innover avec une vision industrielle dans les domaines de la chimie c'est l'art de savoir choisir les molécules utiles pour des applications bien définies. Cela implique, bien sûr, de maîtriser la synthèse de ces molécules, mais aussi et surtout de savoir identifier rapidement celles qui seront efficaces et optimales. Par exemple, dans l'industrie pharmaceutique, un médicament est une molécule optimisée satisfaisant un cahier des charges complexe, comportant de lourdes contraintes techniques et réglementaires : une bonne efficacité thérapeutique (interaction privilégiée avec une cible biologique), des effets secondaires minimisés (pas ou peu d'interactions avec d'autres cibles), une solubilité et une biodisponibilité optimales (la molécule atteint efficacement sa cible biologique dans l'organisme), une toxicité minimisée (bon rapport bénéfice/risque), un coût

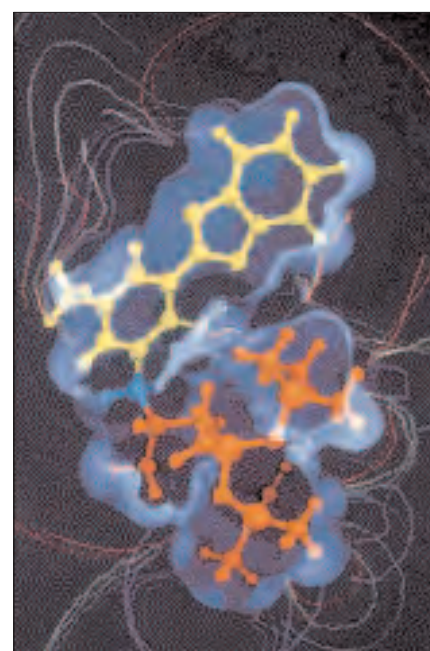
de synthèse acceptable, une protection industrielle solide (originalité et nouveauté de la molécule), etc.

Pour satisfaire un tel cahier des charges, le délai de développement d'un nouveau médicament depuis la première synthèse en laboratoire jusqu'à la mise sur le marché est aujourd'hui de l'ordre de dix ans, et le coût peut atteindre 1 milliard d'euros. Pour une molécule commercialisée, des dizaines de milliers auront été évaluées au cours des différentes étapes de recherche et de développement.

Pour trouver des molécules innovantes, la chimie a longtemps eu recours au hasard ou à l'analogie avec la nature.

Ainsi, la pénicilline (antibiotique) a été identifiée par hasard parce qu'elle tuait des colonies bactériennes dans une boîte de Pétri. La saccharine (édulcorant de synthèse) a été "goûtée" par le chimiste qui la synthétisa lorsqu'il porta à ses lèvres la cigarette qui en était imprégnée. Ou encore l'aspirine a été identifiée à partir de l'extrait de saule, arbre dont les décoctions étaient connues pour lutter contre la fièvre.

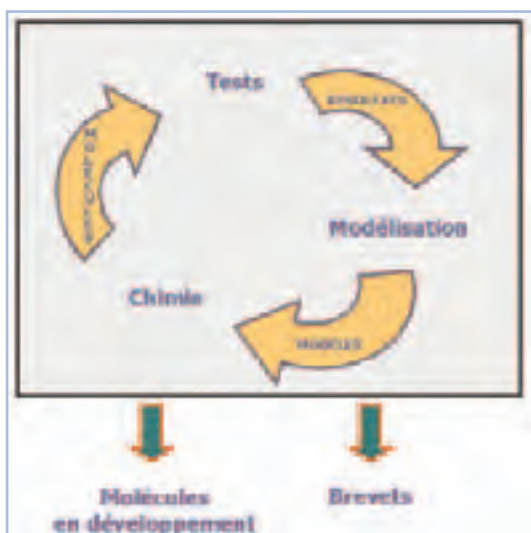
Mais aujourd'hui, les enjeux économiques liés à l'innovation ne permettent plus de s'en remettre exclusivement au hasard ou à l'intuition. En effet, "l'espace chimique" à explorer est infini, et la probabilité de trouver la bonne molécule du premier coup diminue au fur et à mesure de l'accroissement des contraintes de mise sur le marché.



Représentation en 3D du Mexoryl XL, filtre solaire breveté par L'Oréal, avec le volume occupé par la molécule et son champ électrostatique.

Ces enjeux ont conduit l'industrie chimique à rationaliser le processus d'innovation, tout en tenant compte du fait que les phénomènes chimiques complexes restent souvent abordés de manière empirique. Le secret d'un processus efficace d'innovation en chimie réside ainsi dans le bon fonctionnement de la boucle vertueuse qui relie les synthèses de nouvelles

\* Au sein de la Recherche avancée de L'ORÉAL, Nicolas FROLOFF dirige l'équipe de modélisation et d'analyse statistique et Hervé PLESSIX dirige le département de modélisation et d'évaluation cosmétique à haut débit.



Le cœur du processus d'innovation chimique est le cercle vertueux qui relie la synthèse de molécules, les tests expérimentaux et la modélisation des résultats permettant d'affiner le portrait-robot des molécules candidates au développement.

La tendance actuelle est à l'accélération : augmentation de la fréquence des cycles, du débit de molécules, de tests, d'informations et d'utilisation d'outils informatiques et de modélisation de plus en plus sophistiqués pour en extraire des connaissances pertinentes.

molécules, les tests des propriétés de ces molécules et la modélisation. Cette dernière permet, au fur et à mesure de l'expérimentation, de dégager des relations entre la structure chimique des molécules et les propriétés désirées. Ces modèles permettent alors d'affiner peu à peu le "portrait-robot" de la molécule optimale et d'orienter les nouvelles synthèses. Au plan humain, cela nécessite une relation étroite entre équipes pluridisciplinaires, dont chaque acteur maîtrise des techniques très pointues : chimistes, physiciens, biologistes, informaticiens, etc.

La réussite passe aussi par une accélération de ce processus d'innovation. La chimie à vocation industrielle a aujourd'hui recours à diverses technologies comme la synthèse à haut débit où les molécules ne sont plus synthétisées artisanalement une par une mais 100 par 100 (ou plus) à l'aide de robots, ou comme les tests miniaturisés à haut débit (*HTS* : "high throughput screening"), qui permettent d'évaluer rapidement les propriétés d'un grand nombre de molécules. Enfin, c'est l'extraction puis l'utilisation efficace de l'information qui dans ce domaine comme dans bien d'autres sont la clé de la réussite.

Ce sont ces techniques qui ont connu un grand développement ces dernières années dans d'autres secteurs de l'industrie chimique que nous nous employons à mettre en place au sein de la recherche de L'Oréal pour la mise au point de nouvelles matières premières cosmétiques.

## Conception assistée par ordinateur

Comme dans bien d'autres secteurs industriels, les chimistes ont cherché à faire "vivre" virtuellement leurs molécules afin de tester "in silico" (dans l'ordinateur) leurs propriétés physicochimiques ou biologiques avant de passer à la phase longue et coûteuse de la synthèse.

Ces techniques de Modélisation moléculaire ont connu leur plus grande application dans le domaine de la conception rationnelle de médicaments. Aujourd'hui, la conception d'un médicament se fait sur une cible biologique bien définie, cette cible étant souvent une protéine jouant un rôle clé dans la pathologie à traiter. Le "modélisateur" fait de la véritable conception assistée sur ordinateur en visualisant en trois dimensions la structure de la protéine cible sur son écran dans tous ses détails atomiques. La protéine possède généralement un "site récepteur", c'est-à-dire une poche de forme particulière dans laquelle il est possible de loger une petite molécule qui va altérer son fonctionnement (activation ou bien inhibition). Tout le travail du modélisateur consiste, avec le chimiste, à *designer* des molécules de forme optimale pour ce site récepteur qui pourront ainsi interagir de manière spécifique avec la protéine cible. Le chimiste synthétise alors une série de molécules selon cette première hypothèse, les molécules sont testées et les résultats des tests permettent d'affiner la conception d'une deuxième génération de molécules. Et ainsi de suite jusqu'aux premiers "candidats médicaments", qui sont alors soumis à des tests biologiques plus poussés.

Une réussite exemplaire dans le domaine de la conception assistée par ordinateur a été la mise au point récente de nouveaux médicaments pour traiter le SIDA, à partir de la structure 3D de certaines protéines cibles du virus HIV-1. Toute une génération de nouvelles molécules a ainsi vu le jour dans les laboratoires académiques et industriels américains dans les années quatre-vingt-dix, en particulier chez DuPont Merck qui a travaillé sur la protéase du virus HIV-1.

Contrairement aux autres médicaments, les inhibiteurs de cette protéase ont été conçus en grande partie sur ordinateur, en étudiant de manière théorique les interactions permettant à la molécule de se loger favorablement dans le site actif de l'enzyme pour l'inhiber durablement et empêcher ainsi le virus de se répliquer.

La structure 3D de cette enzyme a été déterminée à la résolution atomique, grâce à la technique de cristallographie par diffraction des rayons X. Là aussi, c'est une boucle d'optimisation vertueuse modélisation-expérimentation qui a permis d'aboutir à de bons médicaments-candidats : synthèse de nouvelles molécules, tests miniaturisés d'inhibition de l'enzyme, identification de "*hits*" (molécules répondant positivement dans les tests), cocrystallisation de l'enzyme avec le meilleur des *hits* et observation du complexe formé par diffraction des rayons X pour mieux comprendre comment la petite molécule se loge dans la cavité de l'enzyme, modélisation de molécules plus optimales sur la base de ces nouvelles informations, etc.

Des techniques analogues faisant appel à la chimie quantique sont utilisées au sein de la recherche de L'Oréal pour la modélisation des propriétés physicochimiques et la conception de nouveaux colorants et filtres solaires.

## Bases de données moléculaires

Gérer et exploiter toutes ces informations nécessite la construction de bases de données spécifiques aptes à manipuler les structures moléculaires dessinées en 2D que les chimistes uti-

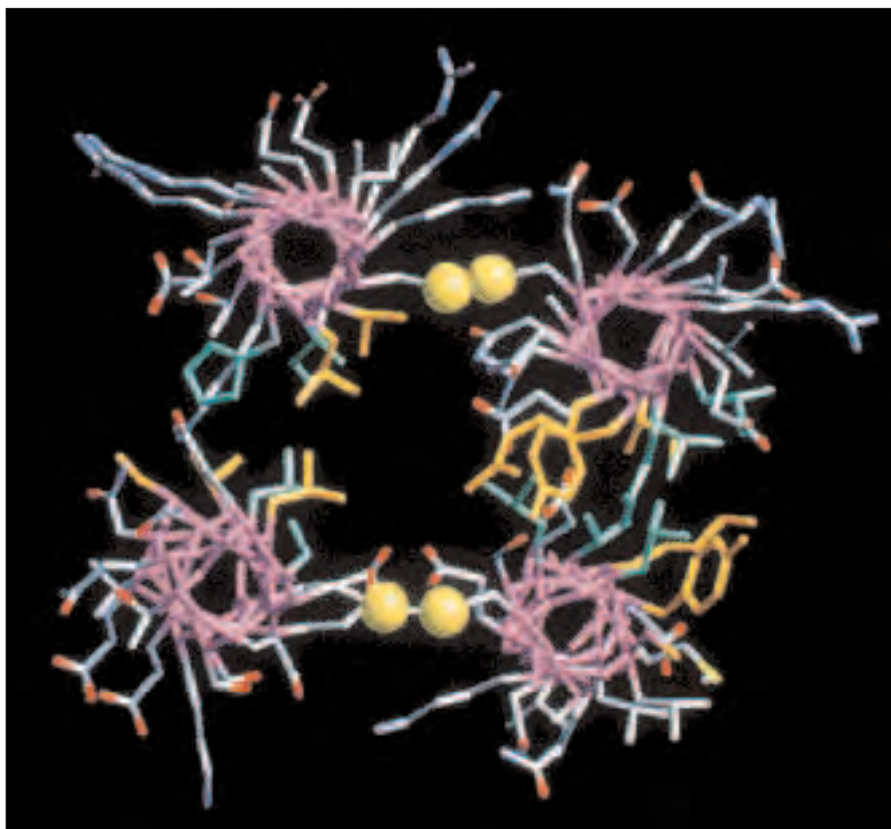
lisent pour représenter les molécules, ainsi que l'ensemble des données et propriétés associées.

Dans ces bases de données moléculaires, chaque molécule est stockée sous la forme d'un graphe en 2D ou 3D (formule chimique développée), et les recherches se font par structure chimique ou par sous-structure (des groupes chimiques fonctionnels par exemple). Les bases de données ont remplacé les catalogues des fournisseurs de produits chimiques. Elles permettent aussi de structurer toute l'information interne à une entreprise sur les molécules synthétisées et les données recueillies sur leurs activités biologiques et leurs propriétés physico-chimiques. En permettant un meilleur partage de l'information pour tous les partenaires d'un projet scientifique, ces bases de données accélèrent considérablement les temps de développement.

## Analyse de données

Toutes ces données accumulées n'ont de sens que si elles sont inlassablement exploitées par les chercheurs pour servir de moteur à leur créativité. Les outils d'analyse statistique ou de "data mining" qui permettent de construire des modèles mettant en relation la structure des molécules et leurs propriétés sont donc stratégiques dans ce processus, que ce soit pour optimiser une molécule, prédire une activité ou aider à la créativité en permettant des extrapolations fructueuses. Ce domaine est tout aussi passionnant techniquement car il se trouve au croisement de différentes techniques d'analyse de données : il s'agit par exemples de trouver des relations entre des graphes (les molécules) et des données chiffrées ou textuelles multivariées (les propriétés).

Comme d'autres domaines industriels, avec peut-être un peu de retard, la chimie s'ouvre à de nombreuses techniques de modélisation et d'analyse de données qui viennent enrichir et supporter le processus de synthèse. Des champs de compétence nouveaux s'ouvrent aux interfaces entre l'informatique, la chimie, les mathématiques, la biologie. Tout reste



Modèle moléculaire 3D de la kératine alpha, une protéine constitutive du cheveu possédant une structure en hélice alpha (ici 4 hélices vues du dessus, reliées par 2 ponts disulfures schématisés par les sphères jaunes).

encore à inventer : la méthodologie, les outils et les applications. Sans compter les changements de mentalité à opérer pour faire comprendre

que le chimiste de demain devra maîtriser à la fois les techniques de la synthèse organique et celles du traitement de l'information. ■



Représentation en 3D de la structure de la protéase du virus HIV-1 avec un inhibiteur. La protéine est schématisée par un ruban passant par sa chaîne principale, permettant de visualiser sa structure symétrique et les deux domaines mobiles (en haut) qui se referment sur le site catalytique (au centre) avec une molécule d'eau (sphère bleue). La structure de l'inhibiteur au sein du site est représentée sous forme de bâtons, à proximité des deux acides aspartiques catalytiques de la protéase (en rouge).  
Image réalisée par Teresa Larsen, Institut de Recherche Scripps, La Jolla, CA, USA.