

Animaux et médicaments

Le concept de médicaments reste l'apanage de l'espèce humaine.

Ces dernières années, des primatologues ont observé des chimpanzés qui consommaient des produits naturels aux effets analogues à ceux de médicaments utilisés par les humains. D'où la question : les animaux pratiquent-ils une protopharmacie ? L'analyse du phénomène montre qu'il est abusif de qualifier ces comportements de « pratiques médicales » et que l'usage des médicaments, dans le sens où nous l'entendons généralement, reste l'apanage de notre espèce.

En 1989, Michael Huffman, un primatologue américain associé à l'Université de Kyoto, et son assistant tanzanien Mohamedi Seifu Kalunde publiaient une observation exceptionnelle d'utilisation d'une plante par un chimpanzé de Tanzanie : l'un des animaux du groupe observé, qui ne s'alimentait presque plus et souffrait de troubles intestinaux, cueillit des rameaux de *Vernonia amygdalina* – une plante au goût très amer, généralement évitée – pour les mâcher et en sucer le jus ; l'animal s'étant rapidement rétabli, les deux zoologistes avaient envisagé l'hypothèse d'une intentionnalité, pour un but curatif, dans le choix et l'usage de ce « médicament ».

La consommation d'argile, observée chez le chimpanzé et chez de nombreux autres singes, et celle de charbon de bois, tout récemment décrite chez le colobe rouge de Zanzibar par Thomas Stuhsaker et ses collègues des Universités de Caroline du Nord et du Wyoming, sont considérées comme des phénomènes de même nature, souvent comparées à des usages de médicaments par les animaux. En effet, les pharmacies vendent du charbon activé ou certaines formes d'argile pour traiter les troubles stomacaux et intestinaux.

En ce qui concerne les argiles, j'ai moi-même dosé, avec Léon Guéguen, de l'INRA de Jouy-en-Josas, les sels minéraux dans les échantillons qui étaient utilisés par les chimpanzés du Gabon, ainsi que dans d'autres terres argileuses ingérées par les entelles (des singes de la famille des Colobidae) du



C.M. Hladik

Peut-on dire qu'un chimpanzé recherche un effet « médicamenteux » lorsqu'il suce le jus des tiges d'*Hypselodelphis violacea* qui lui procure des acides aminés indispensables ? Ce comportement a un effet bénéfique comparable à celui de l'ingestion d'autres produits qui peuvent éliminer les tanins.

Sri Lanka. Nous avons comparé ces concentrations en sels minéraux à celles dans les aliments, principalement des fruits et des feuilles, consommés par les primates. À l'époque où nous abordions ce problème de la géophagie (consommation de terres), les zoologistes admettaient généralement que les animaux trouvaient dans le sol du chlorure de sodium ou d'autres sels minéraux. Nous avons écarté cette hypothèse : la concentration en sels assimilables était bien inférieure à celle que nous avons trouvée dans la plupart des aliments végétaux consommés.

Dans l'article que nous avons publié, en 1974 (*Comptes rendus de l'Académie des sciences*, Paris, t. 279), nous avons proposé une autre explication, qui supposait une interaction des argiles et des tanins, aujourd'hui bien établie pour le cas des singes. En effet, les tanins sont des molécules qui rendent indigestes une grande partie des protéines. Par exemple, lorsque des animaux consomment des feuillages riches en tanins, ceux-ci se lient aux protéines du bol ali-

mentaire, de la même façon que la peau d'un animal peut être transformée en cuir par des liaisons stables entre des tanins et des protéines animales, au cours de l'opération de tannage. L'ingestion d'argile, qui lie rapidement les tanins par adsorption, évite ces liaisons néfastes à la digestion.

Les animaux connaîtraient-ils alors l'usage des médicaments ? Certains zoologistes n'hésitent pas à utiliser le terme de médicament et de protopharmacie, mais il n'existe aucune différence entre la consommation de produits riches en sucres, en protéines, en sels minéraux, etc., indispensables à l'équilibre énergétique du régime alimentaire, et l'absorption de substances végétales ou minérales qui, de la même façon, contribuent à l'équilibre de l'alimentation, en éliminant les substances indigestes, lorsqu'elles sont mélangées aux autres aliments. Dirait-on, par exemple, que des rats carencés en vitamines (ou en certains sels minéraux) « se soignent » lorsqu'ils choisissent préférentiellement, quand on leur en propose, les aliments qui contiennent les vitamines ou les sels minéraux manquants ?

Le médicament, lui, est défini, dans le monde occidental, non seulement par son action spécifique (sur la base de leur action, de nombreuses substances ingérées sont comparables à des médicaments), mais aussi par le contexte où il est utilisé, en fonction du concept de maladie, que l'on oppose à celui de bonne santé.

Si le mode d'action de certaines substances utilisées par des primates non humains est très comparable à celui de certains de nos médicaments, il est abusif d'appliquer aux observations effectuées sur le monde animal cette notion de médicament, définie par rapport à un contexte social. Jusqu'à preuve du contraire, les notions de maladie et de non maladie n'existent que dans les cultures humaines.

Claude Marcel HLADIK
Laboratoire d'écologie générale
du Muséum national d'histoire
naturelle, Brunoy

La synthèse combinatoire

DANIEL MICHELET • CLAUDE HÉLÈNE

La chimie combinatoire donne rapidement toutes les combinaisons possibles d'un jeu de molécules. La diversité de ces molécules aide le chimiste à découvrir de nouvelles substances thérapeutiques.

Les pharmacologues sont sans cesse en quête de nouveaux médicaments. Leur tactique de recherche a évolué au fil du temps, les différentes stratégies étant souvent complémentaires : recherche systématique d'une éventuelle activité biologique parmi les molécules synthétisées dans un laboratoire ; collecte de substances naturelles ; conception assistée par ordinateur de molécules. Aujourd'hui, la chimie combinatoire autorise la synthèse de collections, ou banques, de molécules par combinaison d'un jeu de molécules de construction. Ce faisant, les chimistes ne font qu'imiter la nature qui, partant de quelques molécules élémentaires, a construit toute la variété des gènes et des protéines.

Au cours des années 1970, les progrès du génie génétique ont permis aux biologistes de préparer des quantités notables de protéines pures : pour ce faire, ils introduisent le gène codant la protéine recherchée dans le génome d'une bactérie, laquelle synthétise cette molécule comme s'il s'agissait d'une de ses propres protéines. Certaines, telles que l'insuline ou des facteurs de croissance, sont des médicaments. D'autres sont des cibles thérapeutiques : en inhibant, partiellement ou totalement, leur activité biologique, on peut maîtriser les mécanismes biologiques auxquels elles participent.

Les biochimistes ont ensuite mis au point des tests colorimétriques leur permettant d'évaluer l'affinité d'une molécule pour une protéine cible et de ne conserver ainsi, dans une collection, que les molécules présentant un intérêt pharmacologique potentiel. Ces tests permettent de « cribler » plusieurs dizaines de milliers de molécules par mois, mais les sociétés qui les avaient élaborés n'avaient pas de collections de molécules à tester !

Une nouvelle méthode de synthèse allait pallier cette insuffisance. Au cours des années 1980, les immuno-chimistes cherchaient à établir la structure des fragments des protéines virales (ou peptides viraux) qui déclenchent les réactions du système immunitaire d'un organisme infecté par un virus. Pour synthétiser de nombreux peptides de 5 à 12 acides aminés, ils ont mis au point une nouvelle méthode, nommée synthèse en parallèle. Les peptides produits par ces synthèses en parallèle ont permis aux immunologistes d'identifier de nombreux sites de liaison entre les anticorps et les protéines virales.

La méthode combinatoire est dérivée de cette synthèse de peptides en parallèle, mais elle est plus puissante, car elle produit un nombre de molécules croissant exponentiellement à chaque étape de la synthèse (dans la synthèse en parallèle, on obtient un nombre constant de peptides à chaque étape). Le principe de ce mode de synthèse divergent est illustré dans l'encadré de la page ci-contre.

De la synthèse à l'identification

Ainsi, au lieu de n'engager qu'un seul acide aminé à chaque stade de la synthèse, on fait réagir simultanément les 20 acides aminés naturels : on obtient alors, en n étapes, 20^n peptides représentant toutes les combinaisons possibles de ces 20 acides aminés en molécules ordonnées de n éléments. En pratique, la méthode de synthèse en parallèle divergente ne donne pas la diversité moléculaire escomptée, car les acides aminés en concurrence ne réagissent pas avec la même vitesse, de sorte que de nombreux peptides sont peu abondants, voire absents. D'autres méthodes donnant effectivement des

quantités équivalentes de toutes les combinaisons possibles de peptides ont été mises au point : la méthode de synthèse par répartition et mélange, et la synthèse contrôlée par la lumière.

Dans la méthode par répartition et mélange, la synthèse a lieu sur un support solide constitué de billes de polystyrène. Placées dans un solvant, elles gonflent, de sorte que l'intérieur devient accessible aux réactifs dissous. Chaque bille porte environ 10^{13} amorces qui se combinent avec les réactifs. Par réactions successives, on obtient des oligomères, des molécules constituées d'un enchaînement de plusieurs sous-unités. Dans le cas où les molécules de construction sont des acides aminés, on obtient des oligopeptides.

Quand on dispose de n molécules de construction, et que l'on réalise p étapes de condensation (voir l'encadré de la page ci-contre), on produit n^p molécules, chaque bille portant 10^{13} copies d'un seul type de molécules.

Comment identifie-t-on les billes qui portent des molécules actives ? En fin de réaction, les billes sont réparties dans des « puits » creusés dans une plaque, et une partie des peptides qu'elles portent sont libérés. Ces molécules sont alors solubilisées dans le liquide contenu dans le puits, et un test colorimétrique indique l'affinité d'une des molécules synthétisées pour une protéine cible. Quand un des puits se colore, cela signifie qu'une des billes porte une molécule intéressante. Les billes de ce puits sont réparties une par une sur une nouvelle plaque et le test colorimétrique indique quelle bille porte la molécule active. On la caractérise alors par diverses méthodes physico-chimiques.

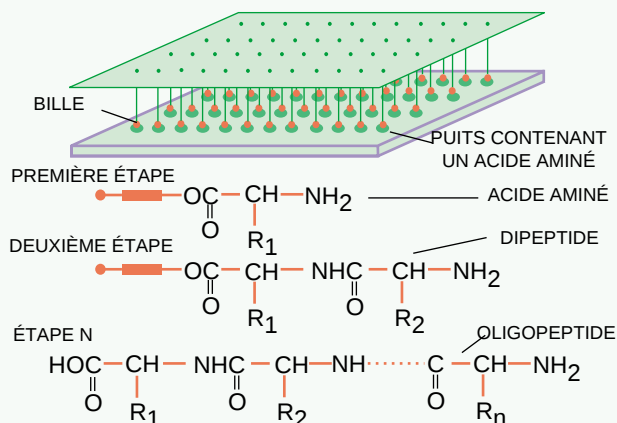
La deuxième méthode, la synthèse dirigée par la lumière, se déroule également sur support solide, sur un plan. Les acides aminés greffés sur la sur-

La synthèse en parallèle

Dans la synthèse en parallèle, une petite bille de polystyrène est fixée à une tige de polyéthylène ou «pin». Chacun des puits de la plaque contient un seul acide aminé.

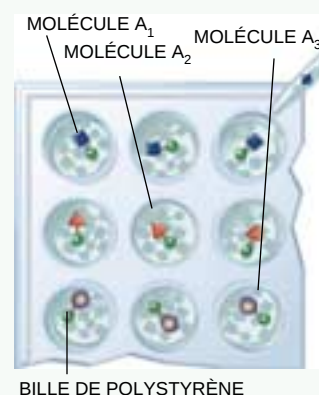
On plonge tous les pins dans les puits, et l'acide aminé présent dans chaque puits se greffe sur la bille immergée, par l'intermédiaire d'un «bras» (le cylindre rouge). Puis on change les acides aminés contenus dans les puits, et les billes sont à nouveau immergées. Le nouvel acide aminé se lie au premier, formant un dipeptide. Après n étapes, on obtient un oligopeptide qui est libéré dans le puits après rupture du bras qui le reliait à la bille de polystyrène.

Par cette méthode, on obtient autant de peptides différents qu'il y a de puits, et ces molécules sont formées par la succession des acides aminés qui ont été successivement introduits dans les puits.



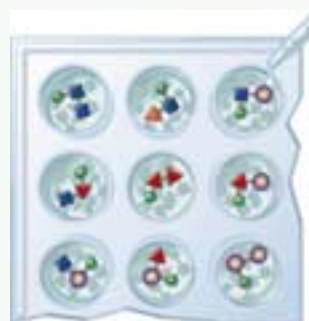
Les différentes méthodes de synthèse des banques combinatoires

La synthèse en parallèle divergente



Étape 1

On dispose d'un jeu de trois molécules A₁, A₂, A₃. Ces molécules sont fixées sur des billes de polystyrène différentes.



Étape 2

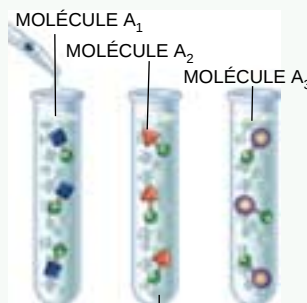
On ajoute dans chacun des puits le mélange des trois molécules A₁, A₂, A₃. On obtient les molécules A₁A₁, A₁A₂, A₁A₃; A₂A₁, A₂A₂, A₂A₃; A₃A₁, A₃A₂, A₃A₃ (pour simplifier, on n'a représenté qu'un type de molécules par puits, mais, en réalité, les trois types de molécules sont mélangés, sur chaque ligne).



Étapes suivantes

Au bout de deux étapes, on a synthétisé 3², soit neuf molécules différentes (en recommençant n fois la même opération, on obtient 3ⁿ molécules). Toutefois, les billes portent toutes un mélange de molécules, ce qui rend l'identification de la molécule active difficile.

La synthèse par répartition et mélange



Étape 1

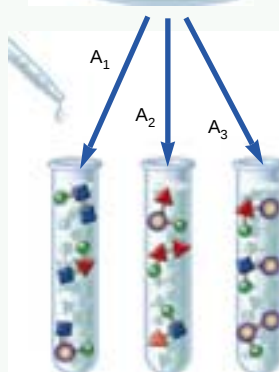
Des billes de polystyrène sont introduites dans les tubes à essai. Les molécules A₁ sont placées dans le premier tube, les molécules A₂ dans le deuxième, etc. On n'a représenté que trois tubes, mais on peut en utiliser des dizaines.

BILLE DE POLYSTYRÈNE



Étape 2

Le contenu des tubes est filtré et rincé, et toutes les billes sont regroupées et mélangées.



Étape 3

Le mélange est partagé en trois, et des molécules A₁ sont introduites dans le tube 1, des molécules A₂ dans le tube 2, etc.

Étapes suivantes

On recommence les opérations de mélange et de partage autant de fois qu'il est nécessaire pour ajouter toutes les sous-unités voulues. Une bille ne porte qu'un seul type de molécules, car, à chaque étape, les billes ne sont en contact qu'avec un seul réactif. Un test colorimétrique identifie la bille qui porte la molécule active.

MOLÉCULES

A ₁ A ₁	A ₁ A ₂	A ₁ A ₃
A ₂ A ₁	A ₂ A ₂	A ₂ A ₃
A ₃ A ₁	A ₃ A ₂	A ₃ A ₃

face plane portent un groupe protecteur sensible à la lumière qui les empêche de réagir. La lumière coupe les groupes protecteurs de sorte que seuls les peptides situés dans les zones éclairées se couplent à l'acide aminé présent dans la solution où est plongé le plan. En utilisant les stratégies de masquage de la photolithographie, on synthétise 2^n peptides avec n masques en n étapes de synthèse. La société californienne *Affymax*, qui a mis au point cette méthode, a obtenu 2^{16} (soit 65 536) peptides différents sur une surface de moins de deux centimètres carrés.

On localise les peptides potentiellement intéressants en détectant ceux qui se lient à la cible biologique voulue, par exemple grâce à des anticorps fluorescents. Comme on connaît la suite des acides aminés que l'on a ajoutés et que l'on a enregistré la position des masques, on déduit la composition des peptides de leur position sur le plan.

Les molécules de la synthèse combinatoire

La chimie combinatoire a d'abord servi à la synthèse de peptides à partir d'acides aminés naturels, puis, par le même procédé, des banques de peptides ont été préparées à partir d'autres acides aminés (artificiels), dont on espérait qu'ils résisteraient à l'action des protéases, les enzymes qui dégradent les protéines. Des pseudopeptides, nommés peptoides, ont ainsi été préparés. Les chimistes ont également synthétisé des banques combinatoires d'oligonucléotides. On synthétise un

mélange d'oligonucléotides (des fragments d'ADN) ; ces molécules sont composées de deux séquences fixes (les deux extrémités constituées d'une quinzaine de nucléotides) et d'une séquence centrale aléatoire de n nucléotides. Pour cette synthèse, on utilise les quatre nucléotides naturels (les constituants de l'ADN), de sorte que le mélange contient 4^n molécules.

Différentes techniques permettent de séparer, dans ce mélange, les molécules qui se fixent sur une cible choisie ; ensuite, on les amplifie par la méthode d'amplification des gènes, nommée PCR, qui a joué un rôle déterminant dans la sélection et l'identification d'oligonucléotides reconnaissant une cible biologique donnée. On recommence plusieurs fois le cycle de séparation et d'amplification, avant de séquencer les molécules intéressantes. C'est ainsi qu'ont été sélectionnés des oligonucléotides courts qui inhibent la thrombine, l'une des enzymes de la cascade de la coagulation du sang. Dans l'organisme, cette enzyme ne se fixe pourtant pas sur des acides nucléiques.

Les oligonucléotides ont la propriété de se replier en formant des «cages» qui ressemblent aux sites actifs des enzymes. Les molécules cibles s'y logent de façon très spécifique : par exemple, un oligonucléotide sélectionné pour fixer la théophylline (la principale substance stimulante contenue dans les feuilles de thé) reconnaît 10 000 fois moins bien la caféine, qui ne diffère de la théophylline que par un seul groupe méthyle (CH_3).

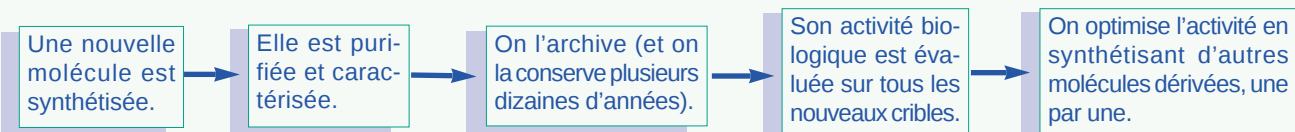
Dans tous les cas énoncés jusqu'à présent, la diversité moléculaire a été

obtenue par la synthèse de toutes les combinaisons possibles d'éléments de construction bifonctionnels (portant une fonction acide et une fonction amine dans le cas des peptides, et deux fonctions alcool reliées par des groupes phosphate dans le cas des oligonucléotides) dans un oligomère linéaire. Une autre façon de créer des banques de molécules intéressantes consiste à «décorer» un «répartiteur de fonctions chimiques» par les méthodes combinatoires.

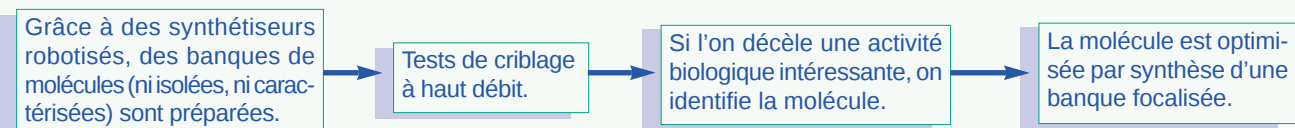
Considérons, par exemple, un cycle portant trois fonctions $-\text{COOH}$, protégées par des groupes qui les empêchent de réagir. Quand on élimine le groupe protecteur d'une fonction $-\text{COOH}$, on peut la combiner avec 20 amines différentes ; on obtient 20 amides $-\text{CO}-\text{NHR}_n$, où n prend les valeurs de 1 à 20. En éliminant ensuite les deux autres groupes protecteurs, on obtient 8 000 ($= 20^3$) triamides (des molécules portant trois fonctions amide, $-\text{CO}-\text{NH}-$). Sachant que quelque 5 000 acides et plus de 3 000 amines sont commercialisés, et que le nombre de fonctions portées par les répartiteurs varie quasi «à volonté», on conçoit la diversité des molécules ainsi obtenues. Outre leur diversité, ces molécules, en forme d'étoiles, sont de meilleurs produits pharmaceutiques potentiels que les peptides ou les oligonucléotides : elles ont, généralement, une meilleure biodisponibilité (leur concentration dans les organes cibles est supérieure à celle des peptides, car elles ne sont pas reconnues par les enzymes qui dégradent les protéines).

De la chimie classique à la chimie combinatoire

Chimie classique



Chimie combinatoire



Dans la méthode «classique», on consacre beaucoup de temps à synthétiser et à caractériser des molécules, sans savoir si elles auront les propriétés recher-

chées. Au contraire, dans la méthode combinatoire, on ne s'intéresse qu'aux molécules dont on sait déjà qu'elles ont les propriétés recherchées.

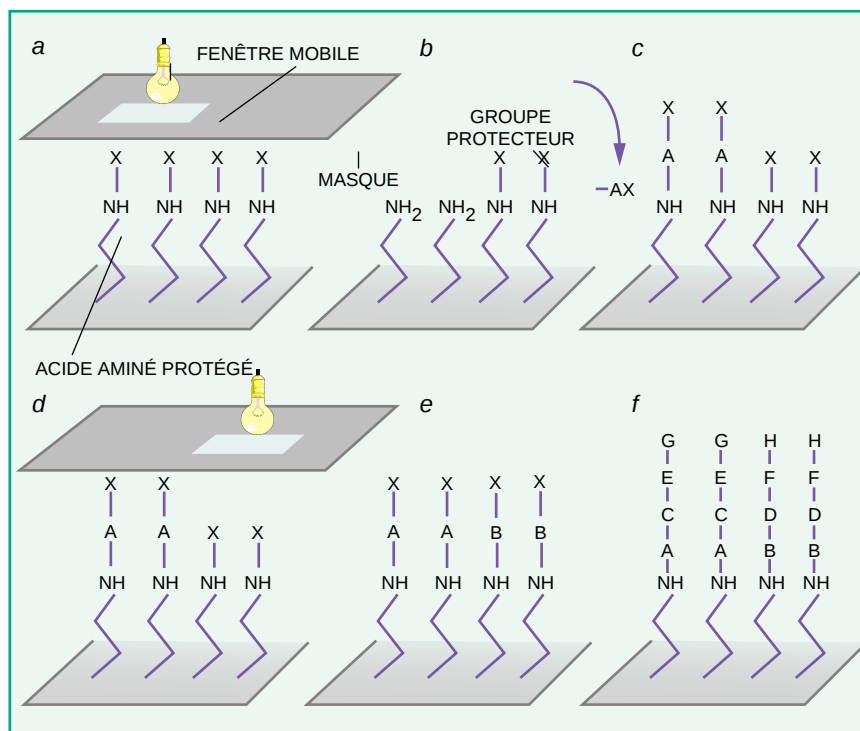
Toutefois, les chimistes rêvent de réaliser, grâce à la chimie combinatoire, des banques de molécules fondées sur l'établissement de liaisons carbone-carbone, et non pas seulement sur des liaisons amides ($-\text{CO}-\text{NH}-$). En fait, les chimistes ont synthétisé facilement des banques combinatoires de peptides et d'oligonucléotides parce qu'ils maîtrisaient si bien les couplages entre les éléments constitutifs qu'ils ont réussi à faire des synthèses multi-étapes sur support solide sans avoir à purifier les produits intermédiaires de synthèse. Les produits obtenus, en fin de réaction, sont même parfois utilisés sans être purifiés, ni caractérisés.

Le caractère univoque de ces synthèses a permis de s'abstraire de ce qui fait habituellement l'essentiel du travail du chimiste : la purification, l'isolement et la caractérisation, opérations artisanales qui doivent être réalisées molécule par molécule. Cette économie d'opérations fastidieuses confère à la chimie combinatoire son attrait. La synthèse se réduit à des opérations de mélange et d'agitation praticables en parallèle et à grande échelle par des robots. Toutefois, on ne pourra créer des banques combinatoires fondées sur d'autres réactions, et notamment sur la création de liaisons carbone-carbone, que si les réactions mises en jeu sont sélectives et si les rendements sont satisfaisants.

La généralisation de la synthèse combinatoire pose donc – enfin ! – des problèmes de chimie aux chimistes : trouver des conditions de synthèse telles que le rendement soit supérieur à 95 pour cent, quelles que soient les molécules qui réagissent.

Support solide ou solution homogène ?

Dans la synthèse sur support solide, la molécule en cours de synthèse est portée par une bille, ce qui permet de la séparer très simplement par filtration. Dans le cas des peptides et des oligonucléotides, 20 ans d'optimisation de la synthèse sur support solide ont précédé l'apparition de la chimie combinatoire : on dispose d'un arsenal de réactifs et de conditions opératoires qui permettent de réaliser des synthèses sur support solide avec un excellent rendement, quelle que soit la séquence synthétisée. Parce qu'ils espèrent synthétiser par la chimie combinatoire d'autres structures que des peptides ou



1. LA SYNTHÈSE COMBINATOIRE DIRIGÉE PAR LA LUMIÈRE est une autre méthode destinée à la préparation de banques combinatoires. Dans une première étape (a), tous les peptides greffés sur un support de silice sont protégés par un groupe X. L'ouverture est mise en place, et la lumière, qui pénètre uniquement par ce trou, libère les groupes protecteurs (b). On ajoute alors la brique de construction AX, qui ne peut réagir qu'avec les groupes non protégés (c). On déplace la fenêtre (d) et l'on ajoute le réactif BX. Seules quelques chaînes (dont on connaît la position) réagissent (e). En procédant de proche en proche, on fabrique des oligopeptides dont l'ordinateur a enregistré la séquence et la position sur la plaque (f).

des oligonucléotides, les chimistes tentent, aujourd'hui, sur support solide, des synthèses qui jusqu'à présent étaient réalisées dans des solvants. Cette «reconversion» n'est pas aisée.

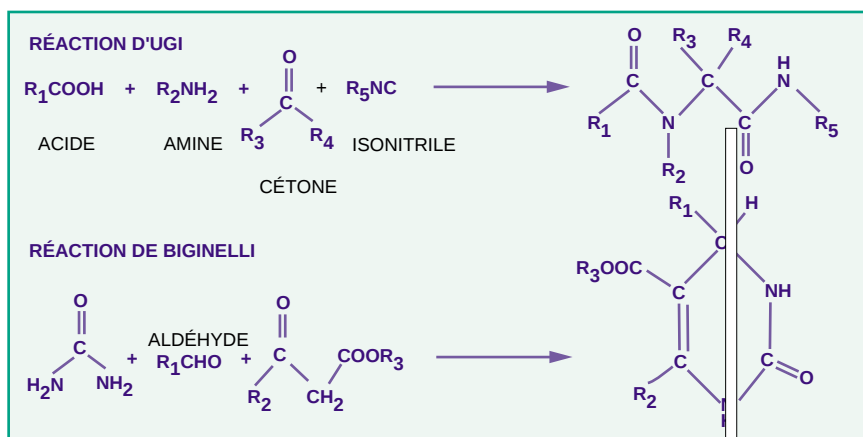
D'où viennent les difficultés ? Même si les billes de polystyrène faiblement réticulées gonflent dans le solvant pour donner une sorte de gel, ce qui permet aux réactifs de rencontrer les molécules attachées au polystyrène, le milieu n'est pas identique à une solution, et les rendements sont inférieurs.

Des progrès ont déjà été obtenus : les chimistes ont modifié les billes de polystyrène en y greffant notamment des chaînes de polyéthylène glycol, ce qui les rend compatibles avec un plus grand nombre de solvants. Ils mettront certainement au point des supports solides encore mieux adaptés, leur permettant de transposer plus facilement sur support solide les synthèses connues en solution homogène.

Une autre solution consiste à échapper au support solide. Ainsi, des supports à solubilité conditionnelle sont utilisés : le premier élément de la molécule à synthétiser est fixé à l'extré-

mité d'une molécule de polyéthylène glycol grâce à un bras clivable. Le polyéthylène glycol est soluble dans la plupart des solvants ; on peut ainsi réaliser la première étape de la synthèse dans le solvant utilisé normalement pour cette synthèse. En revanche, le polyéthylène glycol est quasi insoluble dans l'éther. On le fait précipiter avec la petite molécule qu'il porte, en ajoutant de l'éther, puis, après filtration, les autres étapes de la synthèse se poursuivent sur le même principe.

Une autre méthode encore consiste à faire des réactions à plusieurs composants, lesquelles permettent aussi d'échapper à la synthèse sur support solide : plusieurs molécules sont amenées à réagir simultanément et donnent une seule molécule (voir la figure 2). L'ensemble des combinaisons possibles à partir des molécules de départ, ou synthons, présente une diversité moléculaire notable. Toutefois, on connaît peu de réactions de ce type, mais étant donné leur potentiel en synthèse combinatoire, les chimistes recherchent de nouvelles réactions multicomposants.



2. LA SYNTHÈSE DE LIAISONS CARBONE-CARBONE par la méthode combinatoire est un objectif des chimistes. Certaines réactions, par exemple celles que l'on a schématisées, ont été utilisées pour créer des banques de molécules présentant de telles liaisons. En mélangeant des réactifs judicieusement choisis, on obtient une unique molécule contenant plusieurs substituants R_n . Grâce à la synthèse combinatoire, on peut réaliser toutes les combinaisons possibles à partir de divers acides, amines, cétones, etc.

La biochimie combinatoire

La biosynthèse par des micro-organismes génétiquement modifiés permet aussi d'obtenir des banques de séquences aléatoires de peptides. Pour ce faire, on fusionne une séquence aléatoire de nucléotides avec le gène codant une protéine de surface d'un bactériophage, un virus des bactéries. Une seule séquence aléatoire de nucléotides est introduite par phage, de sorte que l'on obtient une banque de phages recombinants : chacun présente à sa surface un peptide correspondant à l'oligonucléotide synthétique inséré dans son génome. Dès 1990, des banques de 10 à 100 millions de peptides ont été obtenues de cette façon. Ces banques de peptides présentées par des phages ont été criblées pour leur affinité vis-à-vis de divers récepteurs, et l'on a trouvé de nombreux peptides qui présentaient une affinité notable pour ces récepteurs.

De nombreux antibiotiques (les macrolides) sont des molécules cycliques (des macrolactones) élaborées à partir de différents acides par des enzymes bactériennes. La synthèse chimique totale est difficile, et l'on fait appel à des bactéries pour biosynthétiser ces macrolides. On obtient des macrolides différents selon la nature et l'ordre des acides incorporés. Différents gènes bactériens, regroupés en modules, codent les enzymes nécessaires à l'incorporation des différents acides dans le macrolide en construction. Si l'on dispose de n modules et que l'on biosynthétise des macrolides contenant p acides, le clo-

nage de toutes les combinaisons possibles de modules conduit à p^n produits différents. Il s'agit en quelque sorte d'une évolution artificielle où toutes les combinaisons sont réalisées en accéléré.

Des banques restreintes et ciblées

Au début des années 1990, des banques de peptides de plusieurs millions de molécules ont été synthétisées. Ces banques étaient constituées de mélanges où il était difficile d'identifier des molécules actives : plusieurs molécules de structure voisine et peu actives ajoutaient parfois leurs effets pour donner l'illusion que l'on avait une seule molécule très active. L'analyse de si grandes banques était fastidieuse et, aujourd'hui, les banques ne contiennent plus que 1 000 à 10 000 molécules au maximum ; dans la plupart des cas, il s'agit de molécules isolées ou de mélanges d'un petit nombre de molécules (entre 10 et 100), synthétisés en parallèle.

Ainsi, la taille des banques a été réduite, mais la diversité moléculaire préservée. Les molécules qui interviennent dans la synthèse des banques sont soigneusement sélectionnées. Des programmes tiennent compte de paramètres physico-chimiques et topologiques qui optimisent la diversité moléculaire en choisissant au mieux les molécules de construction (voir *Chimie combinatoire virtuelle*, par R. Lahana, page 56).

Aussi l'ambition des chimistes est-elle aujourd'hui de synthétiser des banques combinatoires plus restreintes,

mais aussi diverses que possible. De surcroît, quand on pressent qu'une structure chimique est essentielle pour obtenir l'activité recherchée, on peut synthétiser une banque de molécules qui, toutes, présentent la sous-structure souhaitée, mais substituée de façon combinatoire. On obtient alors une banque de molécules focalisée.

Plusieurs molécules pharmacologiques issues de la chimie combinatoire sont déjà en cours d'essais cliniques, et leur nombre va certainement augmenter au cours des prochaines années. On estime à 10^{200} le nombre de molécules dont la masse molaire est comparable à celle des médicaments usuels !

L'agrochimie, dont les cibles biologiques sont de même nature que celles de la pharmacie va vraisemblablement produire ses prochaines molécules par chimie combinatoire. Le champ d'application de la chimie combinatoire s'étend-il au-delà du domaine des sciences de la vie ? Oui, et Peter Schultz, à l'Université de Berkeley, a appliqué la méthode combinatoire à la synthèse d'oxydes mixtes, c'est-à-dire à la chimie minérale. Il a réalisé une banque combinatoire de films d'oxydes et a recherché, dans cette banque, les composés qui étaient supraconducteurs ou magnétorésistants. C'est ainsi qu'une nouvelle famille de matériaux de formule générale $\text{La}_x(\text{Ba}, \text{Sr}, \text{Ca})_y\text{CoO}_5$, présentant une magnétorésistance élevée, a été découverte.

La méthode combinatoire représente une façon, aujourd'hui incontournable, de rechercher des produits nouveaux, dont les applications devraient s'étendre bien au-delà de l'industrie pharmaceutique.

Daniel MICHELET a été membre de la direction scientifique du Groupe *Rhône-Poulenc*, de 1992 à 1997. Claude HÉLÈNE est directeur scientifique du Groupe *Rhône-Poulenc* et professeur au Muséum national d'histoire naturelle, où il dirige l'unité INSERM U 201.

L. BOUREL et al., *Synthèse combinatoire. Les autoroutes de la diversité*, in *L'actualité chimique*, pp. 33-40, décembre 1995.

F. BALKENHOHL et al., *Combinatorial Synthesis of Small Organic Molecules*, in *Angew. Chem. Int. Ed. Engl.*, vol. 35, pp. 2288-2337, 1996.

J. HOGAN, *Combinatorial Chemistry in Drug Discovery*, in *Nature Biotechnology*, vol. 15, n° 4, avril 1997.

Chimie combinatoire virtuelle

ROGER LAHANA

La modélisation de molécules obtenues par la méthode combinatoire et la prévision de leurs propriétés permettent aux chimistes de synthétiser seulement les substances qui semblent avoir l'activité biologique recherchée.

Concevoir de nouvelles molécules qui auront les propriétés thérapeutiques recherchées, tel est le Graal de tout laboratoire pharmaceutique. La chimie combinatoire (voir *La synthèse combinatoire*, par Daniel Michelet et Claude Hélène, page 50) est une puissante méthode de synthèse, offrant d'innombrables molécules originales. Toutefois, l'identification d'une molécule active parmi des centaines contenues dans le même mélange réactif reste souvent difficile. Les chimistes tentent de tourner la difficulté en modélisant les molécules obtenues par la méthode combinatoire avant de les synthétiser : ils évaluent, par des logiciels adaptés, leurs propriétés potentielles et seules les plus prometteuses de ces molécules virtuelles sont effectivement synthétisées.

Ainsi, la modélisation moléculaire facilite la sélection des molécules potentiellement intéressantes obtenues par la méthode combinatoire. L'ordinateur prévoit toutes les combinaisons que l'on pourrait obtenir en faisant réagir des molécules élémentaires judicieusement choisies, puis les chimistes-informaticiens doivent répondre à plusieurs questions : comment sélectionner les plus intéressantes ? Comment limiter le nombre de molécules étudiées sans écarter une molécule qui aurait pu, *in fine*, être intéressante ? Plusieurs de ces questions ouvertes peuvent être résumées en une seule : comment évaluer la diversité des molécules obtenues par la méthode combinatoire, c'est-à-dire déterminer en quoi elles diffèrent ?

Pendant longtemps, les algorithmes disponibles en modélisation ne visaient qu'à l'optimisation rationnelle du meilleur analogue d'une molécule déjà connue (un analogue est une molécule qui, tout en étant différente, se com-

porte comme la molécule mère). Puis, dans leur course aux nouvelles molécules, les laboratoires pharmaceutiques ont parfois abandonné cette approche rationnelle au profit d'une approche «aléatoire». On aurait, semble-t-il, davantage de chances de trouver une molécule intéressante nouvelle en testant plusieurs millions de molécules au hasard qu'en «construisant» une molécule d'après les caractéristiques idéales qu'elle devrait présenter.

En réalité, les deux approches sont complémentaires. Pour qu'une molécule nouvelle présente une activité biologique, il ne suffit pas qu'elle soit un analogue d'une molécule active connue : elle peut, par exemple, avoir une stabilité inadéquate ou interagir avec des cibles différentes de la cible visée, ce qui risque de déclencher des effets secondaires parfois désastreux.

De surcroît, les molécules des banques combinatoires – ou chimiothèques – ne sont pas conçues au hasard : elles sont fondées sur un motif moléculaire particulier, le synthon, sur lequel se feront des variations ; celles-ci sont de plus en plus souvent ciblées, et les chimistes tendent à restreindre la taille des chimiothèques (de plusieurs millions de molécules, il y a quelques années, à quelques milliers, pour faciliter l'identification des molécules actives). Toutefois, s'ils cherchent à réduire la taille des banques, les chimistes veulent conserver la biodiversité, c'est-à-dire les représentants de toutes les familles de molécules obtenues par la méthode combinatoire.

C'est alors qu'intervient la modélisation moléculaire. Fondant leur analyse sur divers paramètres, notamment sur les relations entre la structure d'une molécule et son activité, les chimistes modélisent les molécules des banques combinatoires et évaluent leur diver-

sité. Dès qu'ils ont détecté des molécules potentiellement intéressantes, ils réduisent la taille de la collection, en éliminant notamment les molécules redondantes. Ensuite, ils comparent les molécules des chimiothèques virtuelles et ils conservent uniquement les plus prometteuses. Ainsi, l'amélioration des techniques d'évaluation de la biodiversité moléculaire explique que l'on ait réussi à diminuer la taille des banques combinatoires, virtuelles, puis réelles.

L'évaluation de la diversité moléculaire

Comment évalue-t-on la diversité moléculaire ? S'il est facile d'affirmer que deux objets sont identiques, il est moins aisé de mesurer en quoi ils diffèrent. Supposons, par exemple, que l'on veuille décrire l'humanité en termes de diversité. On dira que les êtres humains appartiennent tous à la même espèce, qu'ils sont donc tous identiques, et que la diversité est nulle. Puis, si l'on différencie les êtres humains selon leur sexe, on crée deux groupes et, si l'on affine encore davantage leurs caractéristiques jusqu'à leur code génétique, on conclut que l'espèce humaine est constituée de plusieurs milliards d'individus tous différents.

Ainsi, selon les critères, ou descripteurs, une même population peut être considérée comme strictement monotone ou infiniment diverse ! Supposons maintenant que l'on classe l'ensemble des êtres vivants en fonction de leur sexe : un homme sera alors plus proche d'un cloporte mâle que d'une femme. La diversité dépend avant tout des critères que l'on choisit, elle ne correspond à aucune grandeur physique absolue.

Pour mesurer la diversité moléculaire d'une chimiothèque, on la décrit

sous forme informatique : il s'agit d'une chimiothèque virtuelle, qui n'existe que dans la mémoire d'un ordinateur, mais dont les caractéristiques sont bien définies. On peut alors optimiser les molécules dont on pense qu'elles auront une activité biologique intéressante, avant de les synthétiser.

Pour préciser la diversité moléculaire, on l'a d'abord définie comme l'opposé de la similitude : moins deux molécules se ressemblent, plus elles sont diverses. Cette quasi-lapalissade a l'avantage d'être simple, d'autant que l'on dispose d'outils informatiques de quantification de la similitude moléculaire, mais elle a plusieurs inconvénients.

Ainsi, comme dans l'exemple de la population humaine, le critère de «similitude moléculaire» est une information utile, mais partielle. En outre, pour évaluer ce critère, on doit comparer les molécules deux à deux, ce qui est rédhibitoire dès que les séries contiennent plus de quelques dizaines de molécules (le nombre de calculs nécessaires varie comme le carré du nombre de molécules) ; or, en chimie combinatoire, on doit traiter plusieurs dizaines ou centaines de milliers de molécules. Supposons que l'ordinateur mette une seconde à effectuer chaque calcul d'évaluation de l'indice de similitude : pour une chimiothèque de 100 000 molécules, il lui faudrait plus de trois siècles !

Pour pallier cette difficulté, les chimistes ont alors décrit les molécules d'une banque par des fonctions spécifiques, nommées pharmacophores. Un pharmacophore est une combinaison de groupes chimiques dont certaines caractéristiques, telles que la charge ou l'affinité pour l'eau, confèrent l'activité biologique à la molécule qui les porte. Les propriétés diffèrent selon la nature des groupes qui constituent les pharmacophores et selon la distance qui les sépare.

Supposons qu'un pharmacophore soit défini comme tous les triangles formés de trois types de groupes chimiques. Chaque triangle est décrit par ses trois sommets et par les longueurs de ses trois côtés. Si l'on considère que l'on a le choix entre six types de groupes différents, 216 combinaisons de sommets sont possibles ($6 \times 6 \times 6$) ; si la distance séparant deux sommets peut prendre dix valeurs différentes, on obtient 1 000 combinaisons de longueurs

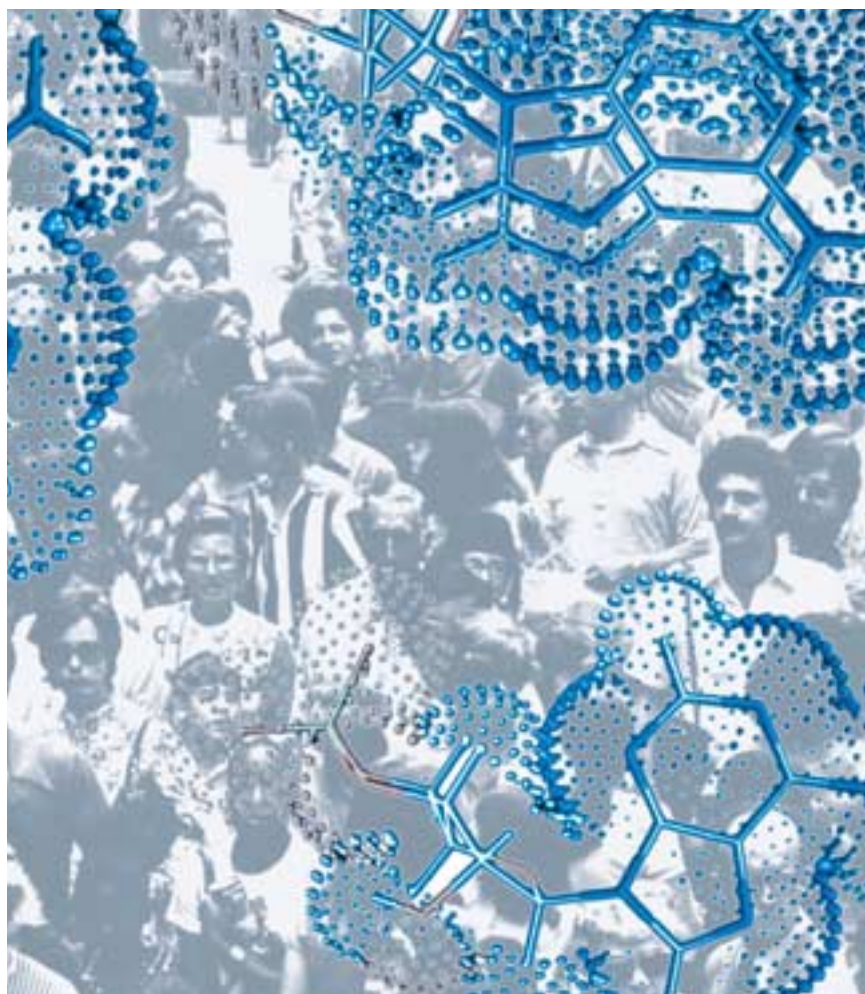
de côtés différentes ($10 \times 10 \times 10$). Toutefois, certaines combinaisons de longueurs sont géométriquement impossibles dans un triangle, et il ne reste que 500 possibilités environ. Pour coder sous forme informatique l'ensemble de ces combinaisons, c'est-à-dire 500 triangles portant les 216 pharmacophores possibles, on définit 108 000 valeurs possibles (500×216) par molécule, valeurs que l'on nomme des clés.

Cette méthode présente quelques inconvénients : un pharmacophore peut être défini par plus de trois centres, ce qui augmente exponentiellement la taille de la clé. Ainsi, si un pharmacophore est un tétraèdre défini par ses quatre sommets et par les six distances qui les séparent, la clé informatique pour une molécule est près de 10 000 fois plus lourde que celle du pharmacophore à trois sommets. Difficile de traiter des

séries de plusieurs centaines de milliers de molécules dans ces conditions (l'espace de stockage nécessaire serait égal à 1 000 gigaoctets pour 100 000 molécules, et le disque dur d'une station de travail est rarement supérieur à une dizaine de gigaoctets).

Une description complète de la diversité moléculaire

Et l'on est toujours confronté à la même question : comment décrire de la façon la plus exhaustive possible la diversité moléculaire ? Nous avons proposé une méthode qui décrit complètement une molécule par des centaines de paramètres : sa masse, son volume, le nombre de ses atomes, sa charge, sa solubilité, la présence ou l'absence de certains groupes chimiques, mais aussi sa topologie (par exemple, la ramification



DE LA DIVERSITÉ génétique à la diversité moléculaire : en combinant quelques briques de construction, les gènes façonnent des milliards d'hommes tous différents. De même, la chimie combinatoire crée d'innombrables molécules à partir d'éléments bien choisis. Avec la modélisation sur ordinateur des molécules des banques combinatoires virtuelles, on sélectionne les plus intéressantes, et l'on ne synthétise par voie chimique que les molécules potentiellement les plus intéressantes.

de ses liaisons), sa réactivité, ou encore le coût de sa synthèse ou la disponibilité commerciale des réactifs. Une analyse statistique préalable des variables permet de sélectionner les plus pertinentes. Par exemple, quand on explore une chimiothèque de peptides, molécules très flexibles, on utilise de préférence tous les descripteurs qui ne dépendent pas de la géométrie de la molécule, mais seulement de sa topologie, laquelle reste identique quelle que soit la géométrie.

L'ensemble des clés permet de définir la diversité des molécules d'une chimiothèque, de comparer deux chimiothèques, et de caractériser les «trous» de diversité, c'est-à-dire les plages de variations non décrites par certaines variables. On calcule aisément les distances séparant les molécules dans l'espace de leurs descripteurs : plus la distance qui sépare des molécules est grande, plus ces molécules sont diverses par rapport aux descripteurs utilisés.

Lors de la conception d'une chimiothèque virtuelle, on en limite la taille en évaluant la diversité apportée par chaque nouvelle molécule : soit l'enrichissement en diversité est suffisant, et la molécule est ajoutée à la banque combinatoire, soit il ne l'est pas, et la molécule est rejetée (il sera inutile de la synthétiser, puisque *a priori* elle n'apporte aucune information supplémentaire par rapport aux molécules déjà présentes dans la banque).

Une fois la chimiothèque virtuelle créée, une analyse statistique peut réduire encore la taille de la banque avec une perte limitée d'informations.

De nouveaux immunosuppresseurs

Nous avons appliqué notre méthode à la recherche d'immunosuppresseurs menée par une entreprise de biotechnologie américaine. Les immunosuppresseurs limitent, voire suppriment, les réactions du système immunitaire responsables notamment des rejets de greffe. On disposait d'une molécule qui semblait rendre des souris plus tolérantes aux greffes de cœur. Comment améliorer l'activité et la stabilité de cette molécule ?

La première molécule active était un peptide, c'est-à-dire un enchaînement d'acides aminés ; nous avons comparé sa séquence avec celle d'autres

peptides très proches dont l'activité avait été mesurée expérimentalement. Nous avons trouvé que l'on pouvait faire varier six acides aminés et que certains descripteurs permettaient de séparer les molécules actives des inactives.

Nous disposions pour cette étude de 35 types d'acides aminés différents (les 20 acides aminés naturels et 15 autres artificiels), la banque combinatoire potentielle contenait près de deux milliards de molécules, trop, même pour l'ensemble des ordinateurs de la planète ! Heureusement, les six positions variables étaient soumises à diverses contraintes physico-chimiques, et certaines combinaisons n'apportaient aucune diversité à la banque. De plus, nous savions que la synthèse de certaines combinaisons d'acides aminés ne serait pas réalisable, et nous les avons éliminées de la chimiothèque virtuelle.

Il ne restait alors plus «que» 279 936 molécules virtuelles à tester. Après avoir encore affiné les critères de sélection, il ne restait que cinq molécules présentant les caractéristiques nécessaires pour être biologiquement actives *in vivo*. Seules ces cinq molécules ont été réellement synthétisées, puis testées sur des souris greffées. Parmi ces cinq molécules, quatre ont été plus actives que la molécule testée initialement. L'une d'elles est environ 1 000 fois plus active : traitées par cet immunosuppresseur, les souris meurent de vieillesse sans avoir rejeté leur cœur étranger transplanté.

La chimie combinatoire est un outil puissant d'accès à la biodiversité. Grâce à cette stratégie, nous synthétisons seulement les molécules sélectionnées par l'ordinateur. La chimie combinatoire virtuelle optimise la sélection des molécules qui ont l'activité biologique potentielle la meilleure.

Roger LAHANA est Directeur des développements en modélisation moléculaire de la Société Synt:em.

R. LAHANA, *Les outils de la modélisation moléculaire*, in *Biofutur*, Technoscope, n° 144, pp.1-8, 1995.

R. LAHANA, *Similitude moléculaire et activité biologique*, in *Pour La Science*, pp. 54-60, juin 1995.

A. YASRI et al., *Rational Choice of Molecular Dynamics Simulation Parameters through the Use of Conformational Autocorrelation 3-D Method*, in *Protein Engineering*, vol. 9, n° 11, pp. 959-976, 1996.
