

# l'actualité chimique

recherche - industrie - enseignement

Mensuel

Septembre  
2000

n° 9

## ■ Numéro spécial : la chimie combinatoire

Pharmacie

Criblage

Matériaux

Banque

Catalyse

Diversité



Expérimentation à haut débit

# SOMMAIRE



## CHIMIE COMBINATOIRE

### Éditorial

- La chimie combinatoire : nouvelle discipline ou stratégie de développement ?, par B. Sillion..... 3

### Banques de molécules

- La synthèse à haut débit (synthèse combinatoire), une discipline arrivée à maturité dans l'industrie pharmaceutique ?, par F. Bellamy..... 4
- Linkers et stratégies de clivage en synthèse sur support solide. Bilan et perspectives, par B. Carboni, F. Carreaux, J.-F. Pilard..... 9
- Réactions en cascade catalysées par des complexes du palladium. Nouvelles stratégies pour la chimie combinatoire, par G. Balme, D. Bouysi, N. Monteiro..... 15
- La RMN haute résolution à l'angle magique en chimie combinatoire, par G. Lippens, J.-M. Wieruszski, R. Warrass, P. Rousselot-Pailley, G. Montagne..... 21
- Contribution de la spectrométrie de masse au développement de la chimie combinatoire, par P.-H. Lambert, S. Bertin, J.-P. Volland..... 25
- Les condensations à composants multiples et leur utilisation en synthèse combinatoire, par H. Bienaymé, P. Schmitt..... 29

### Nouvelles thématiques

- L'approche combinatoire en catalyse hétérogène : quelles stratégies et perspectives pour la recherche publique ?, par C. Mirodatos..... 35
- Catalyse combinatoire. Nouveau concept et nouvelles techniques, par O. Lavastre..... 42
- Tails and steels versus heads and beads : the case for fluororous phase methodology in combinatorial and parallel syntheses, par C. Rocaboy, J. A. Gladysz..... 47
- Chimie combinatoire et chimie supramoléculaire, par I. Huc, J.-M. Lehn..... 51
- Synthèse et application de polymères fonctionnels à structure poreuse pour la chimie combinatoire, par A. Mercier, H. Deleuze, O. Mondain-Monval..... 56
- Le screening virtuel à haut débit (v-HTS), par J. R. Chrétien, M. Pintore, F. Ros..... 60
- Similarité moléculaire et criblage virtuel : méthodes de recherche *in silico* d'analogues actifs pour la découverte de composés thérapeutiques, par D. Horvath, C. Jeandenans..... 64

### En conclusion

- ... un outil complémentaire pour l'instant, par O. Lavastre..... 68

## MANIFESTATIONS ..... 69

## ACTIVITÉS DE LA SFC ..... 71

**sfc**  
société  
française  
de chimie

### L'Actualité Chimique

Revue de la Société Française de Chimie  
250, rue Saint-Jacques, 75005 Paris  
Tél. : 01.40.46.71.60 - Fax : 01.40.46.71.61.

## Rédaction

Rédacteur en chef : Bernard Sillion

Rédacteur en chef adjoint : Thérèse Chaudron

Secrétaire de rédaction, coordination, réalisation, mise en page : Évelyne Girard

Secrétariat : Roselyne Messal

Comité de rédaction : E. Bordes (div. Cat.), J. Buendia (SCI), J.-L. Butré, M. Carréga (div. Matér. polym.), G. Chambaud (com. interdiv. Enseignement), N. Cheymol (CPGE), J.-B. Donnet (ancien président de la SFC), R.-E. Eastes (ENS), J.-P. Foulon (UDP), J. Fournier (club Histoire), J.-F. Gruson (IFP), P.-C. Lacaze (div. Chim. phys.), J.-F. Lambert (div. Cat.), J.-C. Mendelsohn (Elf Atochem), G. Ourisson (ULP, Strasbourg), A. Picot (Prévention des risques chimiques), A. Proust (div. Chim. coord.), D. Rutledge (div. Chim. anal.), J.-M. Schmuck (club de jeunes), G. Schorsch (SFC), F. Sécheresse (div. Chim. coord.), P. Vermeulin (CNRS), J.-N. Verpeaux (div. Chim. orga.), C. Viel (club Histoire), D. Vivien (div. Chim. solide).

Publication analysée ou indexée par : Chemical Abstracts, la base de données PASCAL.

## Édition

Société Française de Chimie.  
Directeur de la publication : Marc Julia, ancien président de la Société Française de Chimie.

Imprimerie : SPEI, BP 26, 54425 Pulnoy.

ISSN 0151 9093

Commission paritaire n° 0402 G 75884

Publicité : EDIF, Le Clémenceau, 102, avenue Georges Clémenceau, 94700 Maisons-Alfort.  
Tél. : 01.43.53.64.00. Fax : 01.43.53.48.00.

Toute représentation ou reproduction, intégrale ou partielle, fait sans le consentement de l'auteur, ou de ses ayants droits, ou ayants cause, est illicite (loi du 11 mars 1957, alinéa 1<sup>er</sup> de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait une contrefaçon sanctionnée par les articles 425 et suivants du Code pénal. La loi du 11 mars 1957 n'autorise, aux termes des alinéas 2 et 3 de l'article 41, que les copies et les reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective d'une part, et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration.

## Tarifs 2000

L'Actualité Chimique (11 numéros par an)

• Particuliers/Institutions  
France ..... 1 100 FF  
Étranger ..... 1 325 FF

• Étudiants\*  
France ..... 420 FF  
Étranger ..... 660 FF

\* Sur présentation de la carte d'étudiant

• Membres de la SFC : service inclus dans la cotisation.

## Abonnement

Société Française de Chimie, 250, rue Saint-Jacques, 75005 Paris. Tél. : 01.40.46.71.60. Fax : 01.40.46.71.61.

Prix de vente au numéro : 120 FF

© SFC, 2000 - Tous droits réservés - Dépôt légal : septembre 2000

## Index des annonceurs

Agilent.....	p. 19	Eurobiochem.....	p. 6	Mimotopes.....	p. 2	Solvay.....	p. 14
Argonaut Technologies.....	p. 40-41	France Biochem.....	p. 70	Polymer.....	p. 13 et 46	Varian.....	p. 50
Aventis.....	1 <sup>re</sup> de couv.	Gilson.....	p. 34 et 11e de couv.	Sanofi.....	p. 63	Waters.....	p. 39 et 45
Bruker.....	p. 24	Laboratoire Fournier.....	p. 59	Sedac.....	p. 7 et 8	Wiley - VCH.....	p. 55
Chemspeed.....	p. 54	Mettler Toledo.....	11e de couv.	Sigma Aldrich.....	p. 20		

# *La chimie combinatoire : nouvelle discipline ou stratégie de développement ?*

**L**a synthèse rapide et contrôlée d'une famille d'individus chimiques a révolutionné la stratégie de recherche de l'industrie pharmaceutique. Cette approche à la fois moléculaire et statistique n'a pu se mettre en place que grâce à un développement des méthodes analytiques et à un renouveau de la chimie des supports.

De plus, en ce qui concerne l'industrie pharmaceutique, l'intérêt de la chimie combinatoire n'est validé que si il y a mise au point concomitante de méthodes spécifiques de tests. On voit par là que, dans ce domaine, la chimie combinatoire participe davantage d'une stratégie d'économie de développement que de l'émergence d'une nouvelle discipline.

Cependant, dès lors que la chimie combinatoire intègre la recherche de sélectivité, s'approprie les assemblages moléculaires, on voit apparaître une nouvelle discipline. D'une manière générale, la gestion de la synthèse et des propriétés des systèmes moléculaires complexes n'est-elle pas l'aboutissement du concept de chimie combinatoire ?

On pourra se rendre compte de l'évolution extrêmement rapide et de l'impact de ces recherches dans ce numéro spécial conçu par **Olivier Lavastre** que la rédaction tient à remercier vivement.

**Bernard Sillion**  
Rédacteur en chef

*PS : Le nombre d'annonceurs, présents dans ce numéro et que nous remercions, témoigne de l'intérêt de l'industrie pour la chimie combinatoire.*

# La synthèse à haut débit (synthèse combinatoire), une discipline arrivée à maturité dans la recherche pharmaceutique ?

François Bellamy\* docteur

**Summary :** *High throughput synthesis (combinatorial chemistry) in the pharmaceutical industry, a technique that has reached maturity*

*Combinatorial chemistry, which should more appropriately be termed high throughput synthesis, was the necessary offspring in the early 1990's of high throughput screening. Beginning with the synthesis of large non-oriented peptide library mixtures, the discipline has now matured whereby targeted libraries respond to standard pharmaceutical prerequisites pertaining to molecular weight, lipophilicity, chemical and metabolic stability and trans-membrane transport. In most cases, syntheses are run in parallel and not as a mixture, with each product being analysed and purified, if necessary. Beginning as a technical curiosity, this approach is now a cornerstone of drug discovery.*

**Mots clés :** *Synthèse combinatoire, drug discovery.*

**Key-words :** *Combinatorial synthesis, drug discovery.*

## La synthèse combinatoire, une réponse aux nouvelles techniques de screening à haut débit

Les deux principales étapes de la découverte d'une nouvelle molécule biologiquement active sont l'identification d'une cible biologique pertinente ainsi que l'identification de composés qui interagissent spécifiquement avec cette cible.

Beaucoup d'espoirs ont été mis dans la génomique et la protéomique pour générer de nouvelles cibles biologiques associées à des pathologies aujourd'hui encore mal soignées, mais il est encore un peu trop tôt pour juger de la pertinence de ces approches.

L'avènement de la biologie moléculaire, associée au fantastique développement de l'informatique et de la robotique, a permis la mise au point de tests

de screening à haut débit, le plus souvent sur des cibles humaines. Ainsi, le screening a énormément gagné en pertinence grâce à l'utilisation de cibles humaines et son débit a augmenté de façon considérable. En l'espace d'un peu plus d'une décennie, on est passé de quelques molécules étudiées chaque jour sur un modèle animal (ou de quelques dizaines en inhibition d'enzyme ou en organes isolés) à plusieurs milliers, dizaines de milliers sur une cible bien caractérisée. Le « ultra high throughput screening » permet aujourd'hui d'atteindre 100 000 molécules/j et au-delà.

Parallèlement à ce fantastique développement du HTS, les progrès de la synthèse chimique étaient certes réels mais quantitativement sans commune mesure et, très rapidement, l'approvisionnement du screening en molécules est devenu une étape limitante du processus de drug discovery. Or, ce n'est pas le nombre de molécules potentiellement disponibles qui constitue une limitation (on estime en effet qu'il y a  $10^{180}$  nouvelles entités chimiques possibles possédant un poids moléculaire  $\leq 750$ ), mais

bien la vitesse à laquelle il est possible de les synthétiser. Les raisons de cette « lenteur » de la synthèse organique résident essentiellement dans le fait que les chimistes avaient l'habitude de faire les molécules une à une et que, par ailleurs, ils passaient en moyenne 70 % de leur temps à traiter les mélanges réactionnels, à en extraire le produit désiré puis à le purifier, et tout cela en vue d'en obtenir des quantités non négligeables (souvent plusieurs centaines de mg).

Il est donc logique que la première tentative et les premières tentatives des chimistes aient été de préparer des produits en mélange et de les faire tester tels quels, c'est-à-dire en l'absence de toute purification, éliminant ainsi les limitations mentionnées ci-dessus. La chimie qui se prêtait le mieux à ces premiers essais de haut débit était incontestablement la synthèse peptidique qui, grâce à l'optimisation des réactions de couplage entre acides aminés, permettait d'obtenir de bons rendements et qui, grâce à l'utilisation des résines de type Merrifield, permettait un traitement aisé des milieux réactionnels.

\* Laboratoires Fournier, 50, rue de Dijon, 21121 Daix.  
Tél. : 03.80.44.75.50. Fax : 03.80.44.77.23.  
E-mail : f.bellamy@fournier.fr

## La course aux grands nombres : production massive de composés en mélange

Le début des années 90 a vu un véritable emballement de la synthèse de banques de peptides avec une course spectaculaire au gigantisme. Certaines équipes ont revendiqué avoir synthétisé des librairies de plusieurs dizaines de millions de composés. L'objectif d'approvisionner le screening avec un nombre suffisant de molécules était ainsi atteint mais en ayant quelque peu perdu de vue la finalité qui a toujours été de découvrir de nouveaux médicaments, c'est-à-dire, dans la grande majorité des cas, des produits présentant une stabilité chimique et métabolique raisonnable et actifs après administration par voie orale. Or les peptides ne répondent généralement pas à ces critères.

Par ailleurs, le screening de mélanges contenant un grand nombre de molécules différentes a vite montré ses limites :

- La nécessité de déconvoluer les mélanges trouvés actifs sur un test biologique, c'est-à-dire d'identifier le (ou les) molécule(s) active(s) contenues dans le mélange. Cette phase de déconvolution, souvent longue et fastidieuse, faisait perdre une grande partie du temps gagné auparavant.

- Il est bien connu que le fait de tester des mélanges (qu'ils soient issus de la synthèse combinatoire ou d'extraction de plantes ou de tissus/organes) conduit à de nombreux « faux positifs » : d'un mélange trouvé actif en screening, on ne peut identifier aucun composant actif, d'où une perte sèche de temps (et d'argent).

Des parades à ces divers inconvénients ont été imaginées puis mises en œuvre par différents groupes : les techniques de déconvolution se sont multipliées, simplifiées, ont gagné en rapidité mais cette étape reste quelque peu frustrante pour les chercheurs qui sont toujours avides de connaître la structure des « hits » (produits trouvés actifs sur un test de screening primaire) le plus rapidement possible.

Parallèlement, on a vu apparaître des procédés « d'étiquetage » des molécules à l'aide de différentes techniques,

telle que celle développée par C. Still ou l'utilisation d'une puce émettant une radiofréquence incluse dans la bille de support solide.

## Synthèse en parallèle et diversité moléculaire

D'autre part, la meilleure façon d'éviter la déconvolution et/ou l'étiquetage étant de tester les molécules une par une, il paraissait logique de les synthétiser également une par une, mais en tirant profit de l'expérience récemment acquise avec les mélanges, à savoir la miniaturisation, la robotisation et la synthèse sur phase solide qui facilite le traitement des expériences et permet une purification grossière.

En même temps, ce type de synthèse a été appliqué à des molécules de moins en moins peptidiques, voire « totalement organiques » afin de s'affranchir des inconvénients inhérents aux structures peptidiques. Ce changement de chimie requerrait, dans de nombreux cas, une synthèse en phase liquide et non plus en phase solide. L'ère de la synthèse en parallèle était née mais la course aux grands nombres toujours d'actualité. Toutefois, les chimistes médicaux ont commencé à se poser des questions quant à l'utilité de tester des banques contenant de nombreux produits structurellement très proches (le bien connu méthyle, éthyle, butyle, futile !). De là est née la polémique relative à la diversité moléculaire, polémique car d'un côté, il peut paraître peu efficace, mais par contre coûteux, de tester des molécules très proches en vue d'identifier un hit actif à  $10^{-5}$ M sur une cible donnée alors que l'espace chimique potentiel est si vaste ; d'un autre côté, chaque chimiste médicinal connaît au moins un cas où, dans une série chimique donnée, en remplaçant tel méthyle (ou tout autre groupe) par « autre chose », on perd toute l'activité biologique.

Alors, diversité ou pas diversité ? Tout dépend à quoi va servir la librairie en question : Hit Generation, Lead optimization...? Diverses ou pas, les librairies se sont multipliées aussi bien au sein de l'industrie pharmaceutique que dans une multitude de « start up » qui se sont créées à travers le monde (mais principalement aux États-Unis et en

Angleterre) et qui proposent, soit des librairies, soit leurs compétences en chimie à haut débit. Des millions de produits ont été testés sur des centaines de cibles générant des milliers de hits. Mais, même en tenant compte de la lenteur des processus de recherche et développement dans l'industrie pharmaceutique, il a fallu reconnaître que peu de ces hits ont conduit à des produits entrant en développement clinique. Ces merveilleuses techniques de screening à haut débit et de synthèse combinatoire n'auraient-elles abouti qu'à augmenter proportionnellement la taille de la meule de foin et la vitesse à laquelle on y cherche l'aiguille ? L'efficacité globale du processus de drug discovery n'en aurait pas été augmentée ?

## Objectif : synthétiser des molécules devant devenir des médicaments (drug like)

De cette réflexion est né un changement de philosophie et donc de stratégie en privilégiant la « qualité » aux dépens de la « quantité ».

Progressivement, les brochures de présentation des sociétés de service en chimie combinatoire se sont mises à vanter le fait que les molécules de leurs librairies sont toutes « drug like », c'est-à-dire qu'elles possèdent un certain nombre de propriétés physico-chimiques, communément regroupées sous le nom de « règles de Lipinsky » du nom d'un chimiste médicinal de chez Pfizer qui a observé que la plupart des médicaments actifs par voie orale ont :

- un poids moléculaire  $\leq 500$ ,
- un  $\log P \leq 5$ ,
- moins de 5 possibilités de donner une liaison hydrogène,
- moins de 5 possibilités d'accepter une liaison hydrogène,
- moins de 5 liaisons à libre rotation.

A peu près en même temps, on s'aperçoit que le fait de tester des bruts réactionnels, des produits non purifiés, coûtait cher car générait de nombreux faux positifs. Aussi, actuellement, la grande majorité des librairies qu'elles soient commerciales ou faites « in house » ne contiennent que des produits dont la pureté est supérieure ou égale à 80 %. Pour obtenir une telle pureté, dif-

férentes approches doivent être mises en œuvre :

- la chimie utilisée se doit d'être efficace en termes de sélectivité et de rendements,

- utilisation, à chaque étape ou presque, de techniques grossières de purification. Fréquemment, on élimine l'excès de réactif (qui permet d'augmenter le rendement) à l'aide de résines « scavengers » (par exemple, lors de la formation d'une liaison amide par réaction d'une amine sur un chlorure d'acide, on va piéger l'excès de chlorure d'acide à l'aide d'une résine basique) ;

- analyse des produits de la librairie finale, le plus souvent par LC-MS. Les produits ne présentant pas une pureté suffisante sont soit jetés à la poubelle, soit purifiés.

### La synthèse à haut débit s'intègre dans le processus de drug discovery

Ce critère de pureté a donc deux conséquences importantes, d'une part la

nécessité d'une analyse des produits finals et, d'autre part, la mise en place d'un système de purification à haut débit. On voit ainsi qu'il ne s'agit plus seulement de synthèse à haut débit pour répondre aux besoins du screening à haut débit, mais de tout un processus et donc de toute une organisation qui se met en place pour tenter d'augmenter l'efficacité du drug discovery et pour réduire les délais entre l'identification d'une nouvelle cible biologique et la première administration d'un produit à l'homme (phase I). En effet, il serait vain d'augmenter le débit d'une étape sans augmenter proportionnellement le débit des étapes en amont et en aval de celle-ci. En fait, la chimie combinatoire ou plus exactement la synthèse à haut débit n'est plus une « bête de cirque », mais constitue une discipline qui, tout en modifiant les habitudes des chimistes, mais aussi celles des analystes, s'intègre dans le processus plus global du drug discovery.

Cette intégration est d'ailleurs facilitée par le fait que les réactions chimiques utilisables à haut débit se

multiplient, puisqu'il est maintenant possible de travailler sous vide, sous pression, sous atmosphère contrôlée, à basse comme à haute température, etc., autorisant l'accès à des structures finales de plus en plus complexes. Aussi, la synthèse à haut débit est maintenant utilisée aux différents stades du drug discovery, à savoir identification de hits, passage de hit à lead, optimisation de lead, etc.

C'est en cela que l'on peut dire qu'elle a atteint un degré de maturité qui en fait un des outils indispensables du drug discovery.

### Perspectives

Le débit a augmenté tout au long du processus entraînant une augmentation des coûts, notamment au niveau du screening et plus particulièrement pour le poste « radioactivité », sachant que de nombreux tests biologiques font appel à des réactifs radiomarqués. D'où le retour, prôné par certains, au screening de produits en mélange. Il ne s'agit plus, cette fois, de screener des mélanges de milliers de produits analogues, mais plutôt de screener des mélanges de quelques (quelques dizaines) de produits, obtenus par synthèse parallèle, purs à plus de 80 % et sélectionnés pour être chimiquement très différents, ceci afin d'éviter, autant que faire se peut, les problèmes de faux positifs.

Une autre façon de baisser les coûts du screening est d'en pousser encore plus la miniaturisation. Toutefois, on atteint assez vite la limite de sensibilité des techniques de détection de la radioactivité, ce qui explique l'émergence des techniques de fluorescence. On parle maintenant de nanotechnologies. Un des systèmes de screening les plus puissants utilise des plaques à 3456 puits avec 1 à 2  $\mu\text{L}$  de solution de produit à tester dans chaque puits. Les chimistes vont-ils pouvoir suivre dans cette course à l'infiniment petit ? Le « lab on a chip » est-il une utopie ou une réalité de demain ? Si la synthèse combinatoire a plus été une révolution culturelle que technique, la « lab on a chip » ne pourra que résulter d'une vraie révolution technique sur laquelle plusieurs groupes travaillent déjà.

Mais d'ici là, la synthèse à haut débit devrait avoir fait ses preuves en ayant contribué à la mise en évaluation clinique de nombreux candidats médicaments et ce, dans des délais nettement plus courts que ceux auxquels nous avons été habitués, prouvant ainsi qu'elle a atteint sa pleine maturité.

# MODIFICATION RATIONNELLE D'ANTIGÈNE PEPTIDIQUE ET STRATÉGIE VACCINALE

Bertrand GEORGES, Sedac-therapeutics  
Bernard MAILLÈRE, CEA-Saclay

La caractérisation des épitopes T effecteurs capable d'induire une immunité protectrice représente l'un des enjeux actuels des nouvelles approches vaccinales. Cette stratégie prend tout son sens sur le plan immunologique car elle restreint l'élaboration du vaccin aux seuls éléments fonctionnels des protéines antigéniques<sup>(1)</sup>. En outre, ces épitopes peptidiques obtenus par synthèse répondent bien aux contraintes et exigences de l'industrie pharmaceutique en termes de caractérisation des produits et d'échelle de production.

Depuis une dizaine d'années, l'immunologie connaît une révolution dans le domaine de la présentation antigénique par le "décryptage" moléculaire des mécanismes de reconnaissance de l'antigène (par les récepteurs immunologiques notamment les molécules du CMH de classe I et II ou les récepteurs T). Ces connaissances nous laissent envisager la possibilité de modifier de manière rationnelle les épitopes T sur lesquels reposent l'induction de la réponse immunitaire afin d'en améliorer la réactivité immunologique. Cette stratégie a de vastes champs d'application dans le domaine biomédical notamment pour le développement d'immunothérapie contre certaines infections virales chroniques.

L'augmentation de l'immunogénicité par des "agonistes" d'épitopes T peut s'opérer de deux manières. D'une part, il est établi qu'un antigène comporte en général un nombre très restreint d'épitopes T induisant une réponse immune dominante au milieu d'épitopes dits "subdominants"<sup>(2)</sup> aux fonctionnalités trop faibles pour induire une réponse immunitaire significative. L'amélioration de la réactivité d'épitopes subdominants représente un intérêt considérable dans le cas d'infections virales chroniques<sup>(3)</sup>. Deux types de modifications peuvent être envisagées, l'une consistant à substituer les motifs d'ancrage pour les molécules de CMH<sup>(4,5)</sup>, l'autre à substituer certains résidus en contact avec le récepteur T tout en maintenant une reconnaissance croisée acceptable vis-à-vis de l'épitope natif<sup>(6)</sup>.

D'autre part, le polymorphisme fonctionnel des molécules HLA (Human Leucocyte Antigen) représente l'un des principaux écueils aux stratégies vaccinales basées sur l'utilisation de peptides. En effet, rares sont les épitopes T "promiscuous" c'est-à-dire capables d'interagir efficacement avec un large panel de molécules HLA. L'analyse au niveau moléculaire de l'interaction de peptides aux molécules HLA laisse toutefois entrevoir la possibilité de modifier les épitopes par

substitutions de positions définies afin d'augmenter l'affinité de liaison pour certains allèles.

L'ensemble de ces approches nécessite une analyse détaillée des motifs d'ancrage aux molécules HLA et ce pour chaque épitope T d'intérêt. En phase initiale, les tests de liaison sur molécule de CMH purifiées permettent une analyse fine et rapide des améliorations pouvant être apportées aux épitopes. Sur le plan fonctionnel, la validation des modifications peut être effectuée sur des modèles de souris transgéniques pour les molécules HLA ou plus directement par immunisation *in vitro* de cellules d'individus HLA typés.

La conception d'analogues d'épitopes T par substitutions dirigées sur les bases moléculaires de la reconnaissance immunologique constitue une approche originale et prometteuse de la vaccination, s'apparentant conceptuellement aux stratégies rationnelles de drug-design de l'industrie pharmaceutique.

- (1) Les épitopes peptidiques représentent les éléments centraux de la reconnaissance par les cellules T et sont nécessaires au développement et à la régulation de l'immunité à médiation cellulaire.
- (2) Les épitopes subdominants se caractérisent généralement par une faible réactivité pour les molécules de CMH de classe I ou de classe II limitant l'expansion des lymphocytes T précurseurs.
- (3) Aux cours d'infections virales chroniques notamment l'hépatite C, une perte de réactivité vis-à-vis d'épitopes dominants peut être observée et assimilée à un phénomène de tolérance.
- (4) Chaque molécule de CMH de classe I et de classe II présente une importante variabilité fonctionnelle liée à l'existence d'un polymorphisme allélique. Au niveau moléculaire, le site de liaison des molécules de CMH présente plusieurs poches d'ancrage au sein desquelles certaines chaînes latérales du peptide antigénique pourront interagir. Chacune de ces poches d'ancrage présente un certain nombre de résidus polymorphiques, événement ayant un impact direct sur la nature des peptides présentés par différents allèles.
- (5) AHLERS, J.D. *et al.* Enhanced immunogenicity of HIV-1 vaccine construct by modification of the native peptide sequence. *Proc.Natl.Acad.Sci. USA.* 1997, 94 : 10856-10861.
- (6) ROSENBERG S.A. *et al.* Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nature Med.* 1998, 4 : 321-327

## Bertrand GEORGES

Sedac-therapeutics - Institut de Biologie de Lille  
1, rue de Pr Calmette - 59011 Lille Cedex  
bertrand.georges@ibl.fr

## Bernard MAILLÈRE

DIEP - CEA-Saclay - 91191 Gif-sur-Yvette  
maillere @ cea.fr

# Linkers et stratégies de clivage en synthèse sur support solide

## Bilan et perspectives

**Bertrand Carboni\*** directeur de recherche CNRS, **François Carreaux\*** enseignant-chercheur, **Jean-François Pilard\*** enseignant-chercheur

**Summary :** *Linkers and cleavage strategies for solid phase synthesis. Overview and future prospects*

*Since the last decade, the explosion of interest in combinatorial and parallel synthesis has greatly stimulated developments in solid phase organic chemistry. Nowadays, their potential applications in such diverse areas as drug discovery, material science, cosmetics, catalysis or plant protection generates a tremendous demand for new linkers and cleavage processes. This brief review summarizes some of the recent advances involving these two crucial elements in solid phase synthesis.*

**Mots clés :** *Synthèse sur support solide, liens, clivage, chimie combinatoire.*

**Key-words :** *Solid phase synthesis, linkers, cleavage, combinatorial chemistry.*

Au cours de la dernière décennie, la chimie combinatoire, associée au criblage robotisé à haut débit, est devenue une des stratégies majeures pour la découverte de nouvelles molécules d'intérêt thérapeutique [1a]. Plus récemment, cette approche a permis également d'obtenir des résultats très prometteurs dans les domaines de la catalyse et des matériaux [1b]. La majorité des chimiothèques (banques ou collections de molécules) ont été préparées par synthèse sur support solide. Le succès de cette méthodologie repose sur un choix judicieux du polymère support, des « briques moléculaires » que l'on va réunir et sur l'efficacité des réactions chimiques mises en œuvre pour réaliser cet assemblage. Deux autres paramètres tout aussi importants, qui ne peuvent être dissociés l'un de l'autre, sont également à prendre en compte : le linker, c'est-à-dire l'élément qui assure le lien entre le support et le substrat initial, et la méthode de clivage qui permet de libérer le produit final [2].

Facilement accessible et, si possible bon marché, le linker doit permettre de réaliser le greffage du composé de départ avec un rendement élevé. Il doit également présenter une très bonne stabilité chimique durant toute la séquence réactionnelle. Enfin, le clivage final devra être réalisé dans des conditions douces et produire le moins possible d'impuretés. Cet article a pour objectif d'illustrer la diversité des stratégies d'accrochage et de clivage utilisés en synthèse sur support solide, en privilégiant les développements les plus

récents et les perspectives qui en découlent. Pour la suite de la discussion, nous avons choisi une classification basée sur la nature du linker, simple groupe protecteur ou source d'une nouvelle fonctionnalité.

### Linker passif : un groupe protecteur insoluble

Initialement développés pour les synthèses de peptides et d'oligonucléotides, ces linkers, et les résines auxquelles ils sont associés, jouent le rôle de groupes protecteurs insolubles. Le groupement fonctionnel (alcool, amine, acide...) impliqué dans le greffage est régénéré lors de la libération du produit final. Dans ce domaine, les exemples sont nombreux et on retrouve les stratégies couramment utilisées dans les séquences protection-déprotection. Les systèmes les plus employés restent basés sur des enchaînements de type benzyloxy (Wang **1a**, Sasrin **1b**) ou trityl **2**, mais il est apparu plus récemment de nouveaux linkers mettant en jeu des motifs dihydropyrane **3**, diol **4**, ou chlorosilane **5** (schéma 1). L'introduction de ces nouveaux systèmes visait principalement à améliorer l'efficacité et la sélectivité des réactions d'accrochage, la stabilité des linkers pendant les diverses étapes conduisant au produit recherché et la mise en œuvre de conditions les plus douces possibles pour réaliser le clivage.

En complément des méthodologies précédemment décrites s'est graduellement fait sentir le besoin de développer des procédés de clivage moins dénaturants pour la synthèse d'espèces sensibles (oligosaccharides, par exemple). Trois principales approches sont actuellement explorées. Elles font appel, soit à des organismes biologiques (enzymes), soit à des procédés photo- ou électrochimiques.

\* Synthèse et Électrosynthèse Organiques, UMR CNRS 6510, campus de Beaulieu, 35042 Cedex. Tél. : 02.99.28.67.47. Fax : 02.99.28.69.78. E-mail : bertrand.carboni@univ-rennes1.fr  
francois.carreaux@univ-rennes1.fr  
jfpilard@univ-rennes1.fr



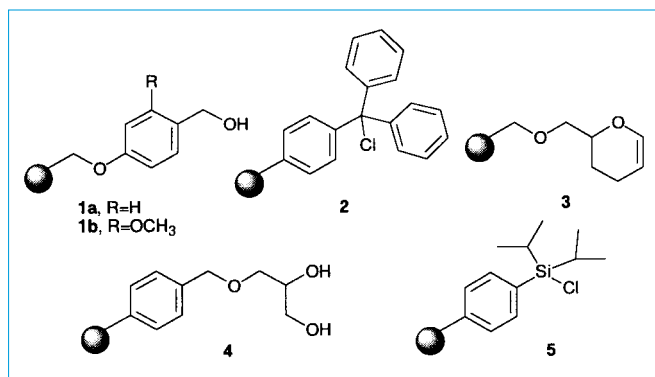


Schéma 1 - Linkers jouant le rôle de groupes protecteurs.

### Procédé de décrochage enzymatique

L'utilisation d'enzymes en tant que réactifs de coupure est actuellement en phase de maturation, particulièrement à l'interface chimie-biologie, afin d'accéder à des peptides, glycopeptides ou hydrates de carbone. Initié en 1992 par Elmore [3], le processus de dégreffage s'effectue par l'intermédiaire d'une reconnaissance spécifique du linker de type enzyme-substrat. Cette association intervient de manière intra-chaîne ou sur un fragment excroissant (extra-chaîne) en fonction de la nature du matériel biologique utilisé, respectivement une endo- ou une exo-enzyme (schéma 2).

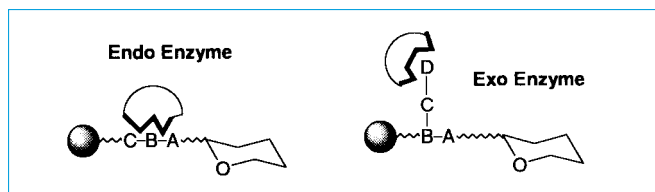


Schéma 2 - Différentes approches de coupure enzymatique.

Le clivage s'opère en milieu aqueux à pH et température contrôlés et présente l'avantage de limiter les rejets organiques. L'utilisation de l' $\alpha$ -chymotrypsine a été pour la première fois mise à profit pour la synthèse d'une glycoprotéine [4] (schéma 3).

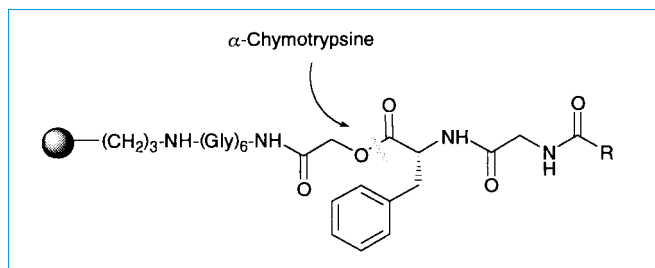


Schéma 3 - Coupure enzymatique par l' $\alpha$ -chymotrypsine.

L'utilisation d'une exo-enzyme, la pénicilline amidase a été décrite plus récemment [5]. Dans ce cas, le mécanisme de coupure fait apparaître, après hydrolyse enzymatique de la fonction amide, un hémiaminal qui se décompose spontanément à température ambiante en libérant la fonction alcool (schéma 4).

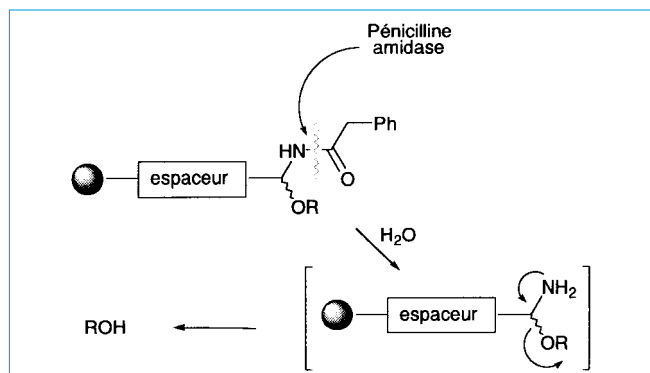


Schéma 4 - Mécanisme de clivage par l'intermédiaire d'une exo-enzyme (pénicilline amidase).

### Procédé de décrochage photochimique

En marge des développements précédents a été imaginé un procédé nécessitant uniquement une activation régiosélective du linker par exposition à une certaine longueur d'onde. Le concept de dégreffage photolytique a été décrit pour la première fois il y a plus d'une vingtaine d'années. Basés sur l'excitation de groupements nitro- ou céto-aromatiques photosensibles (schéma 5), les développements ultérieurs réalisés par Holmes [6] en synthèse peptidique supportée représentent une avancée très significative dans l'amélioration des conditions de coupure (sélectivité, durée d'irradiation).

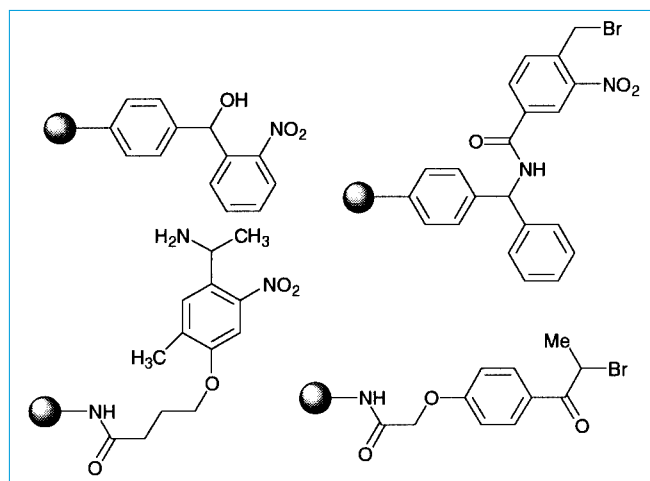


Schéma 5 - Quelques exemples de linkers supportés photolabiles.

Issues de ces recherches, les résines photosensibles actuellement privilégiées présentent une excellente stabilité chimique autorisant, durant les étapes d'élongation successives, l'utilisation de réactifs acides ou basiques (schéma 6) [7].

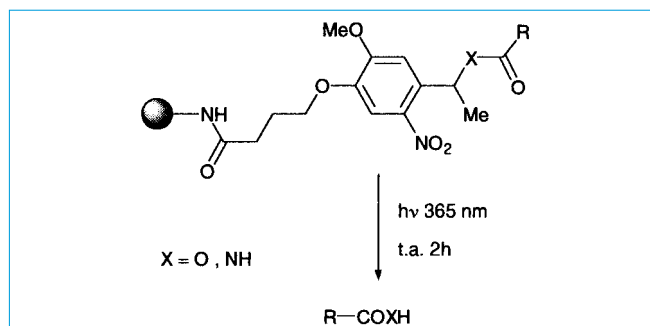


Schéma 6 - Décrochage de peptides par voie photochimique.

## Procédé de décrochage électrochimique

Plus récemment, le concept de linker électrochimiquement labile a conduit au développement d'une nouvelle approche basée sur des matrices conductrices de type polythiophènes 3-fonctionnalisés [8] (schéma 7).

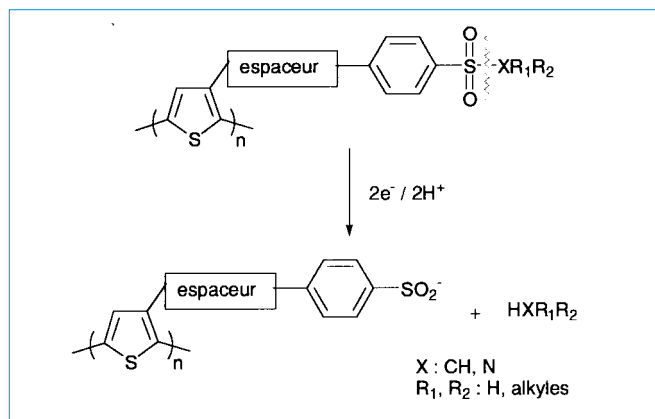


Schéma 7 - Coupure électrolytique de linkers phényl-sulfone ou sulfonamide sur polymères conducteurs.

Ces résines électrogénérées peuvent être déposées sur des supports de nature (platine, carbone) ou de forme (plaques, filaments, barreaux ou disques) variées et présentent une stabilité chimique comparable à leurs homologues polystyréniques. L'originalité de la méthode réside dans l'utilisation d'électrons comme agents de coupure agissant par scission cathodique de liaisons arylS-N ou arylS-C. Le mécanisme de dégreffage consomme deux électrons par mole et produit un sulfinate et une amine ou un dérivé carboné libre de tout résidu de coupure. En outre, le recyclage de la résine devient possible par transformation de l'entité sulfinate en chlorure de sulfonyle.

## Linker actif : un réactif supplémentaire

Pour ce type de linkers, le groupe fonctionnel libéré lors du dégreffage de la résine est différent de celui qui a permis d'accrocher le substrat initial sur le support. On distingue alors deux familles suivant le nombre d'étapes nécessaires pour réaliser la libération du produit final, celle pour qui le procédé de dégreffage se fait en une seule étape, et celle qui nécessite une activation préalable au décrochage.

## Procédé de décrochage direct

De très nombreuses résines commerciales peuvent être regroupées dans cette catégorie. La résine amide Rink est la plus couramment utilisée en chimie combinatoire pour libérer une fonction amide (schéma 8) [9]. Dans l'exemple décrit, une réaction de UGI à quatre composants a été mise en œuvre. La résine joue le rôle d'un équivalent d'ammoniac insoluble, le transfert de l'atome d'azote au produit final s'effectuant lors de l'étape de dégreffage.

Dans cette catégorie, on doit également inclure les linkers « traceless ». Ce terme est principalement utilisé pour

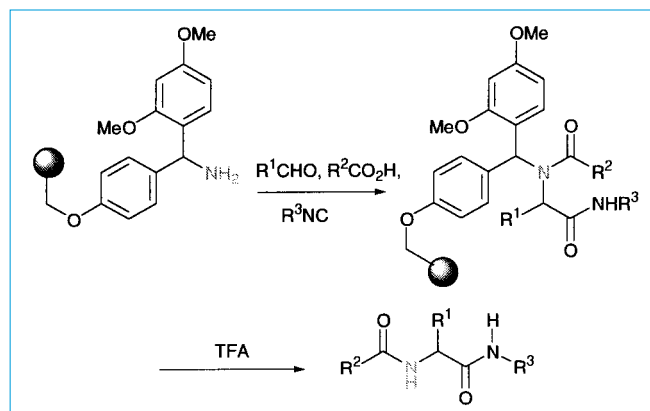


Schéma 8 - Utilisation de la résine Rink amide comme réactif.

décrire un système engendrant durant l'étape de dégreffage une nouvelle liaison carbone-carbone ou carbone-hydrogène à l'endroit où la molécule était liée à la résine [10]. La molécule obtenue après clivage ne présente plus aucune trace ou mémoire du support solide sur lequel elle était fixée. Les premières publications montrant l'intérêt d'une telle approche en chimie combinatoire sont apparues récemment en 1995. Ainsi, l'atome de silicium est remplacé par un hydrogène dans l'exemple représenté sur le schéma 9 [11].

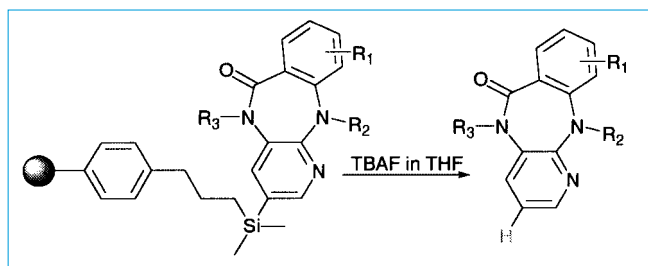


Schéma 9 - Exemple de linker « traceless ».

Dans la majorité des cas, les linkers « traceless » sont décrits comme pouvant générer lors du dégreffage des liaisons autres que C-C ou C-H. C'est notamment le cas des dérivés silylés qui peuvent donner naissance aux halogénures correspondants. Le groupement boronate, quant à lui, présente l'avantage de pouvoir être impliqué dans des réactions de protodéboronation, oxydation et des couplages au palladium [12].

Ce type de dégreffage multidirectionnel offre l'avantage principal de générer de la diversité lors de l'étape de coupure, augmentant ainsi le nombre de molécules finales. A titre d'exemple supplémentaire, on peut citer le cas des acides carboxyliques attachés à la résine Wang qui sont libérés sous l'effet de l'acide trifluoroacétique, alors que le traitement de la même résine avec une amine ou dans des conditions réductrices génère la fonction amide ou alcool correspondante [13] (schéma 10).

Une approche différente, mais également très efficace, est basée sur la mise en œuvre de réactions de cyclisation. La création d'une nouvelle liaison par voie intramoléculaire s'accompagne alors de la libération du produit. Cette stratégie présente un avantage déterminant. Seules les molécules présentant une structure appropriée seront cli-

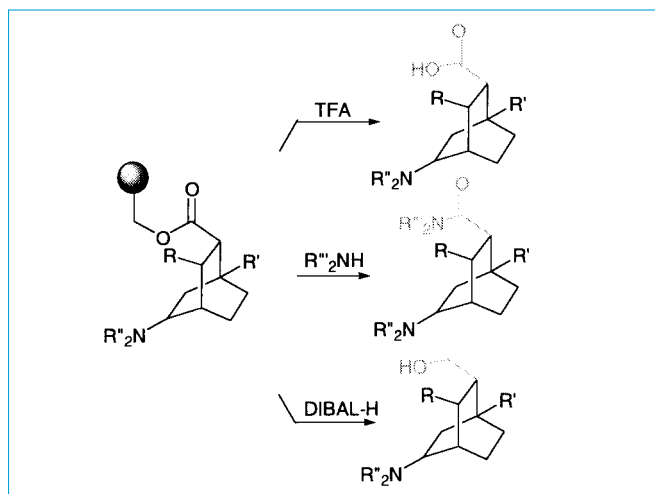


Schéma 10 - Exemples de clivage multidirectionnel.

vées. Par conséquent, même si l'une des réactions précèdent le clivage s'effectue de façon imparfaite, la pureté du produit final ne sera pas affectée.

Ce type de clivage a trouvé de nombreuses applications, principalement lors de la préparation d'hétérocycles par formation d'une liaison carbone-hétéroatome. Plus récemment, la réaction de métathèse a été utilisée avec succès pour accéder à des composés cycliques par création d'une liaison carbone-carbone [14] (schéma 11).

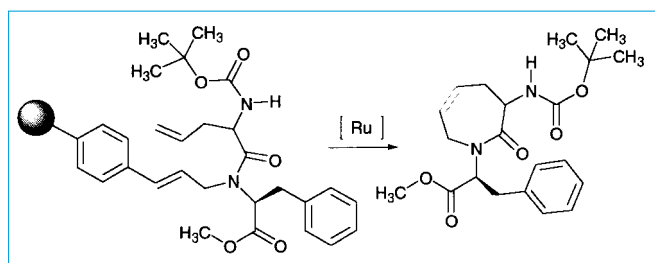


Schéma 11 - Exemple de clivage par cyclisation impliquant une réaction de métathèse.

## Procédé de décrochage en deux étapes

Parmi les stratégies développées spécifiquement pour la synthèse supportée de petites molécules organiques, le principe du « safety catch » est devenu de plus en plus populaire. En effet, dans ce cas, l'étape de dégreffage ne peut avoir lieu qu'après une étape préliminaire d'activation. Ceci permet d'avoir un linker stable vis-à-vis d'un grand éventail de conditions réactionnelles, incluant celles du dégreffage. Ce principe, décrit pour la première fois par Kenner, en 1974, dans le domaine de la chimie des peptides [15], était basé à l'origine sur la réactivité du groupe sulfonimide. Depuis, ce concept a été employé avec succès par plusieurs autres groupes, en particulier pour la synthèse de pyrimidines [16] (schéma 12).

Le décrochage par une amine (attaque nucléophile) ne peut être effectué qu'après oxydation. Cette stratégie offre une grande flexibilité, puisque cette étape peut être programmée à n'importe quel stade de la synthèse, à condition que les produits soient stables vis-à-vis des conditions d'acti-

tion et de libération. Dans le schéma 12, l'attaque nucléophile peut être réalisée par des amines, mais également, par l'azoture de sodium, des oxanions et des carbanions, ce qui augmente considérablement la taille et la diversité des chimiothèques.

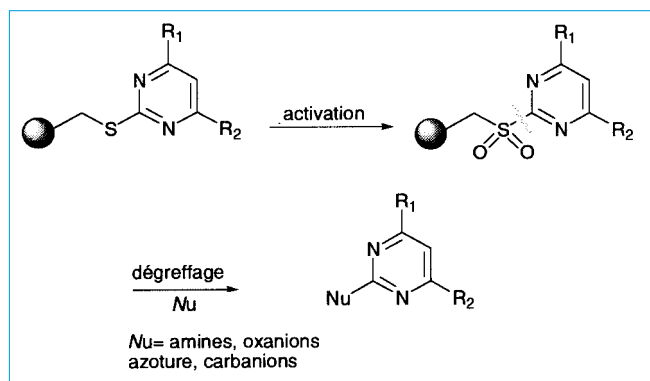


Schéma 12 - Principe et exemple d'un linker activable.

Enfin, citons également les résines de type REM où le greffage est réalisé grâce à l'addition de Michael d'une amine sur un motif acrylate (ou, plus récemment, vinylsulfone) [17]. La quaternisation par un agent électrophile introduit une diversité supplémentaire et simultanément active le linker en vue d'une réaction d'élimination en milieu basique (schéma 13).

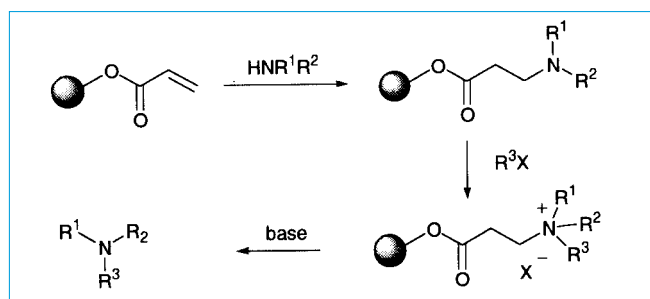


Schéma 13 - Principe d'une résine de type REM.

## Conclusion

Comme on a pu le constater à la lecture de ce document, la nature des linkers et l'efficacité des processus de clivage constituent deux éléments clés en synthèse sur support solide. Dans la recherche d'un système idéal, peu coûteux, produisant le moins d'impuretés possibles et permettant d'accéder à une grande diversité structurale, les efforts des chimistes se sont principalement portés sur la découverte ou l'optimisation de linkers « traceless » et photoactivables, ainsi que sur les stratégies de clivages multidirectionnels. La mise au point de conditions de dégreffage les plus douces possibles, en particulier compatibles avec des biomolécules, reste toujours d'actualité. Nul doute que, comme pour la chimie plus classique en solution, le respect de l'environnement deviendra dans le futur une condition *sine qua non* à la conception de nouvelles stratégies en chimie sur support solide.

## Références

- [1] (a) Thompson L.A., Ellman J.A., *Chem. Rev.*, **1996**, *96*, p. 555 ; Dolle R.E., *Molecular Diversity*, **1998**, *3*, p. 199 ; Dolle R.E., Nelson K.H., *J. Combinat. Chem.*, **1999**, *1*, p. 235.  
(b) Jandeleit B., Schaefer D.J., Power T.S., Turner H.W., Weinberg W.H., *Angew. Chem. Int. Ed.*, **1999**, *38*, p. 2494.
- [2] Backes B.J., Ellman J.A., *Current Opinion in Chemical Biology*, **1997**, *1*, p. 86 ; Morphy J.R., *Current Opinion in Drug Discovery and Development*, **1998**, *1*, p. 59 ; James I.W., *Tetrahedron*, **1999**, *55*, p. 4855. Divers sites Web consacrés à la chimie combinatoire sont désormais accessibles. Parmi eux, on peut citer celui ouvert par B.A. Bunin : <http://www.combinatorial.com>
- [3] Elmore D.T., Guthrie D.J.S., Wallace A.D., Bates S.R.E., *J. Chem. Soc. Chem. Commun.*, **1992**, p. 1033.
- [4] Schuster M., Wang P., Paulson J.C., Wong C.H., *J. Am. Chem. Soc.*, **1994**, *116*, p. 1135.
- [5] Böhm G., Dowden J., Rice D.C., Burgess I., Pilard J.-F., Guilbert B., Haxton A., Hunter R.H., Turner N.J., Flitsch S.L., *Tetrahedron Lett.*, **1998**, *39*, p. 3819.
- [6] Holmes C.P., Jones D.G., *J. Org. Chem.*, **1995**, *60*, p. 2318.
- [7] McKeown S.C., Watson S.P., Carr R.A.E., Marshall P., *Tetrahedron Lett.*, **1999**, *40*, p. 2407.
- [8] Marchand G., Pilard J.F., Simonet J., *Tetrahedron Lett.*, **2000**, *41*, p. 883.
- [9] Tempest P.A., Brown S.D., Armstrong R.W., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, p. 640.
- [10] Reitz B., *Current Opinion in Drug Discovery & Development*, **1999**, *2*, p. 358.
- [11] Woolard F.X., Paestch J., Ellmann J.A., *J. Org. Chem.*, **1997**, *62*, p. 6102.
- [12] Carboni B., Pourbaix C., Carreaux F., Deleuze H., Maillard B., *Tetrahedron Lett.*, **1999**, *40*, p. 7979 ; Pourbaix C., Carreaux F., Carboni B., Deleuze H., *Chem. Commun.*, **2000**, sous presse.
- [13] Ley S.V., Mynett D.M., Koot W.-J., *Synlett*, **1995**, p. 1017.
- [14] Piscopio A.D., Miller J.F., Koch K., *Tetrahedron Lett.*, **1997**, *38*, p. 7143.
- [15] Kenner G.W., Mc Dermott J.R., Sheppard R.C., *J. Chem. Soc., Chem. Commun.*, **1971**, p. 636.
- [16] Obrecht D., Abrecht C., Grieder, Villalgordo J.-M., *Helv. Chim. Acta*, **1997**, *80*, p. 65.
- [17] Quyan X., Armstrong R.W., Murphy M.M., *J. Org. Chem.*, **1998**, *63*, p. 1027.

# Réactions en cascade catalysées par des complexes du palladium

## Nouvelles stratégies pour la chimie combinatoire

**Geneviève Balme**\*<sup>1</sup> directrice de recherche CNRS, **Didier Bouyssi**\*<sup>2</sup> maître de conférences, **Nuno Monteiro**\*<sup>3</sup> chargé de recherche CNRS

**Summary :** *Palladium-mediated cascade reactions : new strategies for combinatorial chemistry*

*The urgent need for new bioactive molecules has prompted the search for new strategies that allow the efficient generation of large numbers of structurally diverse organic compounds targeted at the agrochemical and pharmaceutical industries.*

*Among them, palladium-mediated multicomponent cascade reactions offer major opportunities for the design of innovative multireaction « one pot » processes. The new methodologies enable a combination of complexity and diversity to be achieved that is not available by conventional chemistry.*

**Mots clés :** *Chimie combinatoire, catalyse, processus en cascade, réactions multi-composants.*

**Key-words :** *Combinatorial chemistry, catalysis, cascade process, multicomponent reactions.*

La chimie combinatoire occupe aujourd'hui une place importante dans la découverte de nouvelles molécules bio-actives. En effet, la recherche de nouvelles têtes de série ou « hits » par les sociétés pharmaceutiques et agrochimiques est actuellement basée sur l'évaluation de tests biologiques du plus grand nombre de molécules. Afin d'alimenter les tests de criblage, de nouvelles stratégies doivent être développées permettant de créer rapidement des banques ou bibliothèques de composés analogues, généralement de petites molécules de grande diversité chimique et présentant une certaine originalité vis-à-vis des molécules commerciales.

La mise au point de nouvelles voies d'accès rapides, simples et efficaces à des systèmes complexes et variés a tou-

jours été un des objectifs majeurs du chimiste organicien. Parmi celles-ci, les réactions en cascade [1], appelées également réactions tandem ou domino, constituent d'élégantes méthodes de synthèse permettant de créer plusieurs liaisons carbone-carbone ou carbone-hétéroatome en une seule opération. Compte tenu de leur potentiel en chimie combinatoire, la découverte de nouvelles réactions tandem ou domino, principalement pour la synthèse de composés hétérocycliques, présente donc un intérêt croissant. Dans ce contexte, les remarquables potentialités des réactions initiées par des métaux de transition avaient déjà été exploitées dès le début des années 80 [2]. Parmi les métaux de transition, le palladium

occupe une place de choix pour la formation de liaison carbone-carbone selon différentes réactions de couplage dont la plupart seraient impossibles par les voies de la chimie classique. La première étape, commune à la plupart de ces réactions, est une addition oxydante d'un halogénure ou triflate insaturé sur le palladium zérovalent (*figure 1*). L'intermédiaire ainsi formé peut alors réagir :

- soit avec des dérivés du bore, du zinc, du cuivre, du silicium, de l'étain ou de l'aluminium (réaction de Suzuki, Negishi, Sonogashira, Stille) [3]. L'étape de transmétallation est suivie d'une étape d'élimination réductrice conduisant au produit de couplage et régénérant le catalyseur.

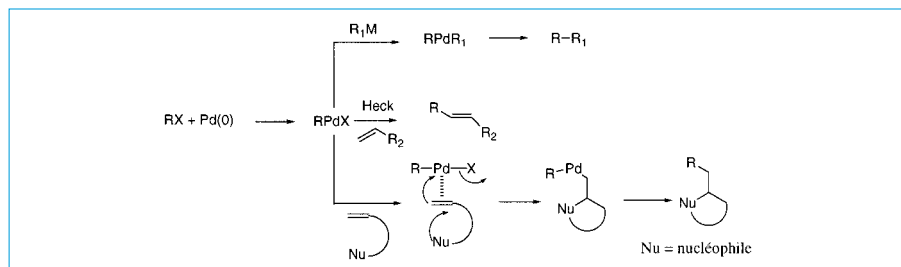


Figure 1.

\* Université Claude Bernard Lyon I, Laboratoire de chimie organique 1, UMR CNRS 5622, CPE, bât. 308, 43, bd du 11 Novembre 1918, 69622 Villeurbanne Cedex. Tél. : 04.72.43.14.16. Fax : 04.72.43.12.14.

<sup>1</sup> balme@univ-lyon1.fr

<sup>2</sup> bouyssi@univ-lyon1.fr

<sup>3</sup> monteiro@univ-lyon1.fr

- soit avec des alcènes (réaction de Heck) [4]. Dans ce cas, une *syn* insertion de la double liaison sur l'intermédiaire organopalladié est alors suivie d'une *syn*  $\beta$ -élimination d'hydrure de palladium.

- soit avec des dérivés éthyléniques ou acétyléniques comportant un carbo- ou hétéronucléophile [5]. Le complexe organopalladié joue ici le rôle d'agent d'activation de l'insaturation vis-à-vis de l'attaque du nucléophile. Cette cyclisation « tandem » se déroule ici selon un processus de *trans* addition de « type Wacker » et est suivie d'une élimination réductrice régénérant le catalyseur.

Ces réactions, permettant l'accès à de nombreux dérivés carbocycliques et hétérocycliques en une seule étape et avec un fort degré de régio- et stéréosélectivité, connaissent actuellement un grand essor. En effet, les contraintes récentes imposées par le respect de l'environnement (économie de matière, d'énergie, procédés propres...) s'ajoutant aux problèmes posés par la création de banques de molécules non peptidiques incitent au développement de nouvelles méthodologies de synthèse, et les réactions en cascade réalisées en présence d'un catalyseur répondent parfaitement à ces exigences.

### Réactions en cascade catalysées par le palladium

Le champ d'application de la réaction de Heck, connue depuis 1968, s'est considérablement élargi à la fin des années 1980, grâce aux travaux pionniers de plusieurs groupes [6]. Ces différents auteurs ont en effet montré que l'insertion d'un complexe organopalladié, réalisée sur une insaturation convenablement choisie, pouvait conduire à la formation d'un nouvel intermédiaire métallé ayant une durée de vie suffisamment longue pour être engagé dans de nouvelles réactions de couplage, créant ainsi une liaison carbone-carbone ou carbone hétéroatome supplémentaire (figure 2).

Ce principe a permis la réalisation de nouvelles combinaisons multiples de plusieurs réactifs plus communément appelées réactions multi-composants et est à la base de la plupart des réactions

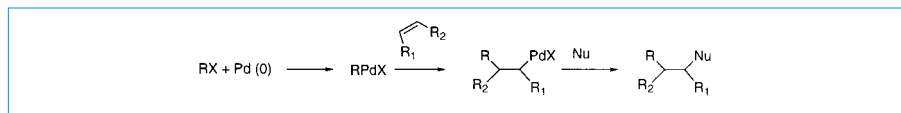


Figure 2.

en cascade palladocatalysées que l'on voit paraître dans la littérature. Parallèlement à ces travaux, une nouvelle méthode de cyclisation catalysée par le palladium s'est développée permettant la synthèse en une seule étape de composés de structures très variées. La grande diversité de nucléophiles et de précurseurs de complexes organopalladiés qui peuvent être mis en jeu dans cette réaction dite de « type Wacker » en fait un nouvel outil très performant particulièrement adapté à la synthèse rapide de composés analogues principalement hétérocycliques. La synthèse de nombreuses énolactones à partir de composés commerciaux montre le potentiel synthétique de cette nouvelle réaction [7] (figure 3).

S'ajoutant à toutes ces réactions domino pallado-catalysées, de nouvelles réactions multi-étapes en une seule opération ont été récemment développées, les différents partenaires mis en jeu réagissant dans des étapes réactionnelles dont une seule peut être catalysée par le palladium (par exemple réaction de Heck suivie *in situ* d'une réaction de Diels-Alder).

Le développement de ces réactions pour la production de bibliothèques de molécules nécessitent, cependant, que plusieurs critères soient respectés :

- les différents partenaires mis en jeu dans ces réactions doivent être commerciaux ou aisément accessibles et capables d'apporter une grande diversité au sein des composés synthétisés ;

- la réaction doit être générale, facile à mettre en œuvre et permettre une isolation rapide des produits d'arrivée en évitant les étapes fastidieuses de purification. L'idéal serait d'optimiser la

réaction en solution homogène de manière à limiter la formation de sous-produits et l'utilisation d'excès de réactifs. Ceci étant rarement réalisable, il est apparu judicieux de transposer la réaction sur phase solide permettant ainsi d'isoler les produits désirés par simple filtration [8]. Enfin, une approche complémentaire en phase liquide consiste à piéger sélectivement, en fin de réaction, le ou les réactifs en excès et éventuellement les sous-produits de la réaction par l'utilisation de résines fonctionnalisées ou « scavengers » [9].

Les quelques exemples présentés ci-dessous vont tenter d'illustrer les différentes possibilités qu'offrent ces réactions palladocatalysées pour la chimie combinatoire.

### Réactions de type Heck

Lorsque la réaction de Heck est réalisée sur des substrats ne permettant pas une terminaison classique du cycle catalytique par une *syn*  $\beta$ -élimination d'hydrure de palladium, une nouvelle liaison carbone-carbone ou carbone-hétéroatome peut être formée par réaction de couplage avec des alcènes, alcyne, des composés organométalliques ou intramoléculairement par des nucléophiles. Une telle situation apparaît, lorsque le substrat insaturé est un composé bicyclique rigide tel que le norbornène, un alcyne disubstitué ou en version intramoléculaire une double liaison 2,2-disubstituée. Différentes réactions de couplage, ayant été mises au point sur le norbornène en phase homogène par plusieurs groupes dans les années 80, ont été récemment adap-

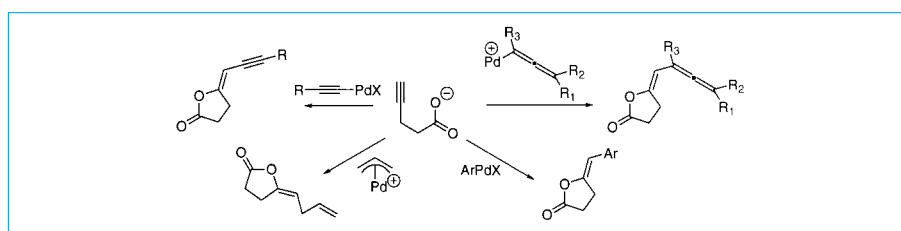


Figure 3.

tées en phase solide pour la synthèse de dérivés du tropane [10] (*figure 4*). Cet exemple montre que la transposition de réactions multi-composants peut poser d'énormes problèmes aux chimistes organiciens, les nouvelles conditions expérimentales laissant apparaître des réactions de couplage compétitives entre les différents partenaires. Le succès de cette réaction tandem est donc ici limité par le fait que l'intermédiaire organopalladié doit être isolé avant d'être remis en réaction avec le troisième composant.

De nombreux composés hétérocycliques ont pu cependant être obtenus sur phase solide lorsque la réaction de Heck est effectuée en présence de dérivés acétyléniques [11] ou de composés éthyléniques en version intramoléculaire [12] (*figure 4*).

Des réactions tandem peuvent également être réalisées lorsque la réaction de Heck est appliquée à des systèmes polyinsaturés. Cette réaction conduit à la formation d'intermédiaires  $\pi$ -allyliques qui peuvent ensuite être piégés intra- ou intermoléculairement par divers nucléophiles. De très nombreuses réactions multi-composants basées sur ce principe ont été réalisées en phase homogène ; certaines d'entre elles, transposées sur phase solide, ont récemment permis la création de diffé-

rentes bibliothèques d'hétérocycles azotés [13] (*figure 5*).

On peut également distinguer la remarquable séquence réactionnelle dans laquelle un halogénure d'aryle, un bicyclopropylidène et un diéophile réagissent sélectivement dans une réaction domino Heck-Diels-Alder. Trois liaisons carbone-carbone sont alors formées en une seule opération avec un haut degré de régiosélectivité. Son application en phase solide peut conduire, lors du dégreffage de la résine, à la formation d'une nouvelle liaison carbone-carbone [14] (*figure 5*).

### Réaction en cascade de « type Wacker »

De nombreux hétérocycles de structures très variées ont été récemment synthétisés par cette méthode en phase homogène et l'application en phase solide d'une synthèse d'indoles trisubstitués a été récemment développée [15] (*figure 6*).

Une série de composés tétrahydrofuraniques hautement fonctionnalisés a été également récemment atteinte par une nouvelle réaction tandem mettant en jeu trois composants commerciaux ou aisément accessibles. Cette nouvelle approche s'appuie sur la création en

une seule étape d'une liaison O-C et de deux liaisons C-C par l'addition de type Michael d'un alcoolate allylique sur un arylidène (ou alkylidène) malonate, suivie, *in situ*, d'une carbocyclisation palladocatalysée de type Wacker [16]. Cette réaction nécessitait cependant l'ajout au pousse-seringue de l'alcoolate sur le mélange des deux autres partenaires afin d'éviter des réactions secondaires. Cependant, lorsque l'alcool allylique est remplacé par son homologue acétylénique, l'alcoolate, l'accepteur de Michael et l'halogénure aryle peuvent être ajoutés ensemble. Cette réaction a été transposée sur phase solide mais elle n'a cependant pas un caractère général. En conséquence, une optimisation des conditions réactionnelles a été entreprise en phase liquide : les trois composants ont pu être ajoutés en quantité équimoléculaire et mis en réaction simultanément. Dans la plupart des cas étudiés, une conversion pratiquement totale de ces trois réactifs en le produit furanique ont simplifié les étapes de purification.

### Conclusion

La nécessité de créer de nouvelles bibliothèques de molécules pour le secteur biomédical et phytosanitaire est à

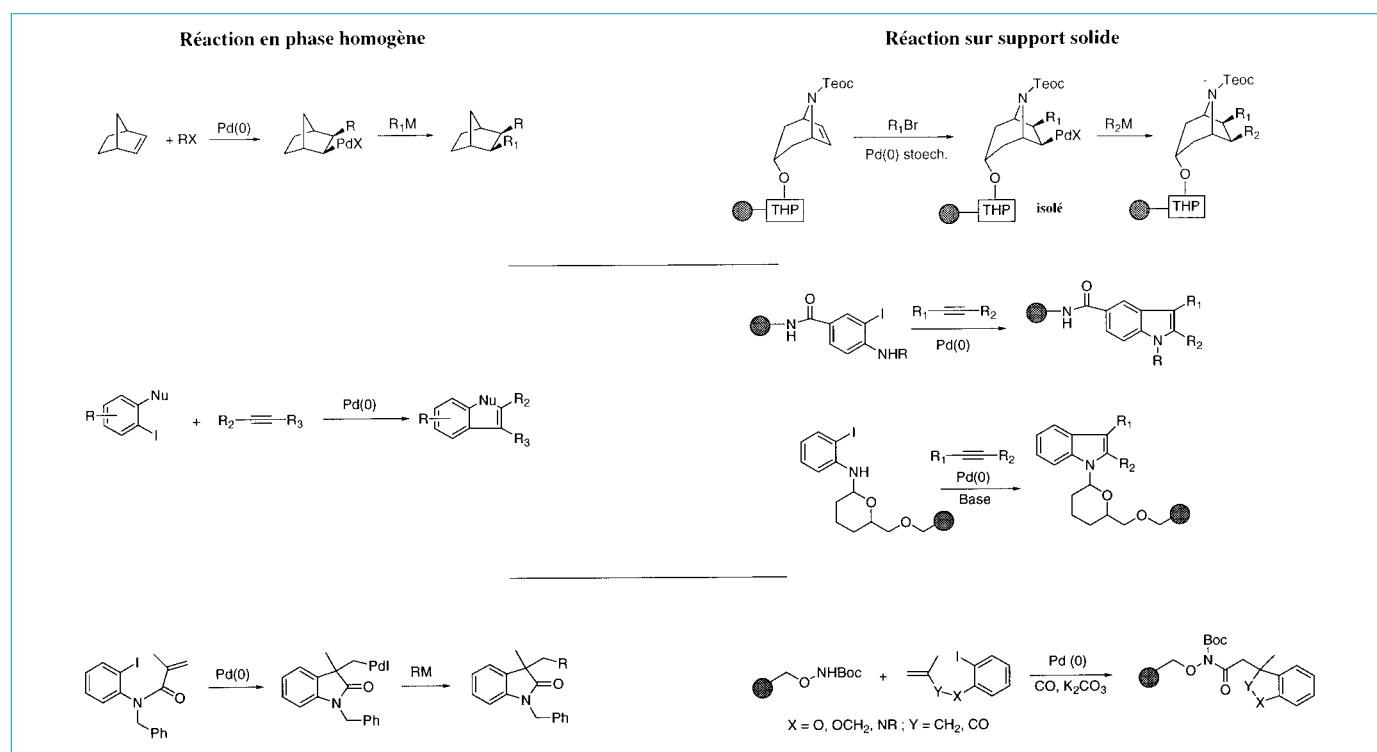


Figure 4.

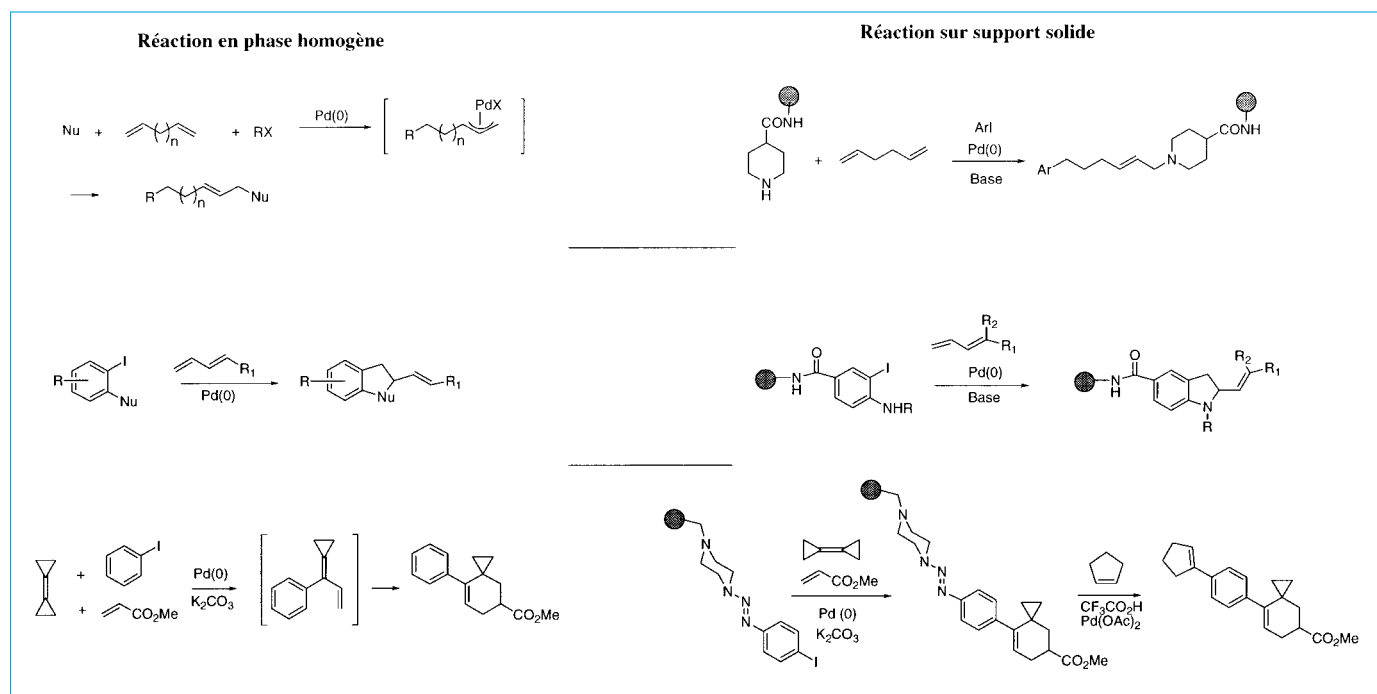


Figure 5.

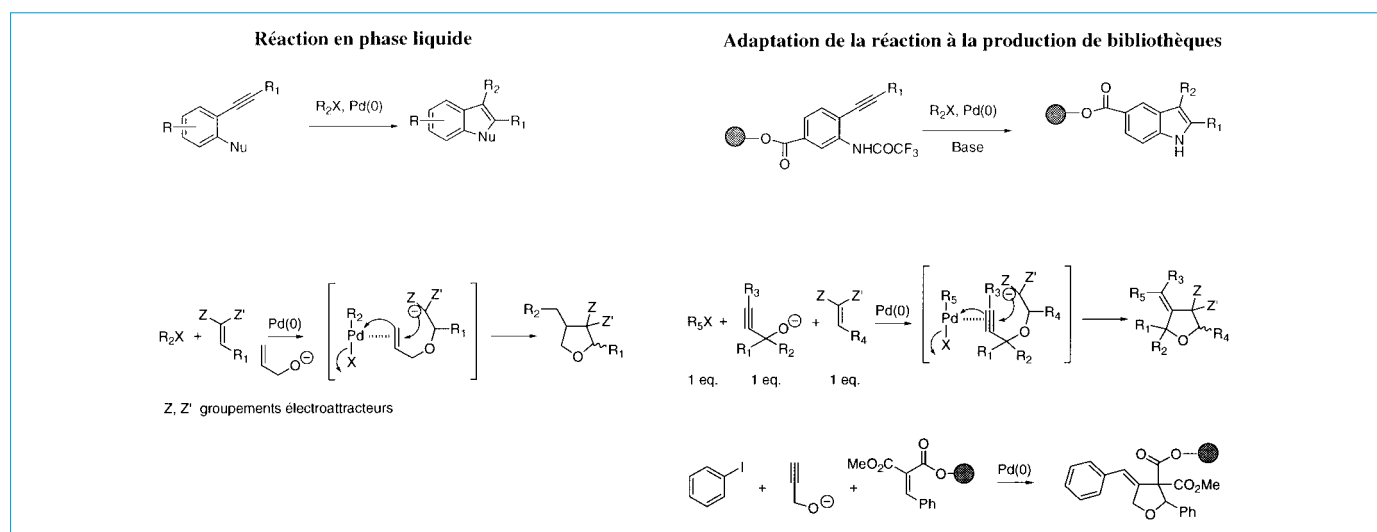


Figure 6.

l'origine du récent essor des réactions tandem et surtout multi-composants. De nouvelles voies d'accès élégantes et astucieuses à des molécules complexes et des synthons multifonctionnels ont ainsi été élaborées prouvant ainsi que la chimie combinatoire pouvait être une source de créativité pour les chimistes organiciens. Cependant, l'application de ces nouvelles méthodologies de synthèse à la production de bibliothèques de molécules est rarement réalisée par ces mêmes chimistes organiciens, peut-être en raison des difficultés souvent rencontrées lors de la transposition sur phase solide. De nouvelles techniques permettant une purification rapide des

produits synthétisés, sans modifier les conditions réactionnelles mises au point en phase homogène, devraient sans doute rapidement voir le jour.

## Références

- [1] Hoffmann H.M.R., *Angew. Chem. Int. Ed. Engl.*, **1992**, p. 1332-1334 ; Tietze L.F., *Chem. Rev.*, **1996**, p. 115-136.
- [2] Clinet J.-C., Dunach E., Vollhardt K.P.C., *J. Am. Chem. Soc.*, **1983**, p. 6710-6712 ; Schubert U., Fischer H., Hofmann P., Weiss K., Dötz K.H., Kreissl F.R., *Transition Metal Carbene Complexes*, Verlag Chemie, Weinheim, **1983**.
- [3] Tsuji J., *Palladium Reagents and Catalysts, Innovations in Organic Synthesis*, Wiley, Chichester, **1995**.
- [4] Heck R.F., *Palladium reagents in organic synthesis*, Academic Press, New York, **1985**.
- [5] Inoue Y., Itoh Y., Yen J.F., Imaizumi S., *Journal of Molecular Catalysis*, **1990**, *60*, L1-L3 ; Balme G., Bouyssi D., *Tetrahedron*, **1994**, p. 403-414 ; Cavicchioli M., Decortiat S., Bouyssi D., Gore J., Balme G., *Tetrahedron*, **1996**, p. 11463-11478 ; Cacchi S., *J. Organomet. Chem.*, **1999**, p. 42-64.
- [6] Abelman M.M., Overman L.E., *J. Am. Chem. Soc.*, **1988**, p. 2328-2329 ; Burns B., Grigg R., Ratanakul P., Sridharan V., Stevenson P., Sukirthalingam S., Worakun T., *Tetrahedron Lett.*, **1988**, p. 5565-5568.
- [7] Tsuda T., Ohashi Y., Nagahama N., Sumiya R., Saegusa T., *J. Org. Chem.*, **1988**, p.2650-2653 ; Arcadi A., Burini A., Cacchi S., Delmastro M., Marinelli F.,



- Pietroni B.R., *J. Org. Chem.*, **1992**, p. 976-982 ; Bouyssi D., Gore J., Balme G., *Tetrahedron Lett.*, **1992**, p. 2811-2814 ; Bouyssi D., Gore J., Balme G., Louis D., Wallach J., *Tetrahedron Lett.*, **1993**, p. 3129-3130.
- [8] Brown A.R., Hermkens P.H.H., Ottenheijm H.C.J., Rees D.C., *Synlett*, **1998**, p. 817-827 ; Booth S., Hermkens P.H.H., Ottenheijm H.C.J., Rees D.C., *Tetrahedron Report 476, Tetrahedron*, **1998**, p. 15385-15443.
- [9] Walters M.A., *Chemtracts-Organic Chemistry*, **1998**, p. 1013-1024 ; Hori M., Janda K.D., *J. Org. Chem.*, **1998**, p. 889-894 ; Booth R.J., Hodges J.C., *Acc. Chem. Res.*, **1999**, p. 18-26.
- [10] Kosugi M., Tamura H., Sano H., Migita T., *Tetrahedron*, **1989**, p.961-967 ; Koh J.S., Ellman J.A., *J. Org. Chem.*, **1996**, p. 4494-4495.
- [11] Larock R.C., Yum E.K., *J. Am. Chem. Soc.*, **1991**, p. 6689-6690 ; Smith A.L., Stevenson G.I., Swain C.J., Castro J.L., *Tetrahedron Lett.*, **1998**, p. 8317-8320 ; Zhang H-C., Brumfield K.K., Maryanoff B.E., *Tetrahedron Lett.*, **1997**, p. 2439-2442.
- [12] Burns B., Grigg R., Sridharan V., Stevenson P., Sukirthalingam S., Worakun T., *Tetrahedron Lett.*, **1989**, p. 1135-1138 ; Grigg R., Major J.P., Martin F.M., Whittaker M., *Tetrahedron Lett.*, **1999**, p. 7709-7711.
- [13] Larock R.C., Berrios-Pena N., Narayanan K., *J. Org. Chem.*, **1990**, p. 3447-3450 ; Wang Y., Huang T-N., *Tetrahedron Lett.*, **1998**, p. 9605-9608 ; Wang Y., Huang T-N., *Tetrahedron Lett.*, **1999**, p. 5837-5840.
- [14] Brase S., De Meijere A., *Angew. Chem. Int. Ed. Engl.*, **1995**, p. 2545-2547 ; De Meijere A., Nüske H., Es-Sayed M., Labahn T., Schroen M., Bräse S., *Angew. Chem. Int. Ed. Engl.*, **1999**, p. 3669-3672.
- [15] Arcadi A., Cacchi S., Marinelli F., *Tetrahedron Lett.*, **1992**, p. 3915-3918 ; Collini M.D., Ellingboe J.W., *Tetrahedron Lett.*, **1997**, p. 7963-7966.
- [16] Cavicchioli M., Sixdenier E., Derrey A., Bouyssi D., Balme G., *Tetrahedron Lett.*, **1997**, p. 1763-1766.

# La RMN haute résolution à l'angle magique en chimie combinatoire

**Guy Lippens**\*<sup>1</sup> chargé de recherche, **Jean-Michel Wieruszkeski**\* ingénieur de recherche, **Ralf Warrass**\* chercheur postdoctoral, **Pierre Rousselot-Pailley**\* thésard, **Gérard Montagne**\* technicien

**Summary :** *High resolution magic angle spinning NMR in combinatorial chemistry*

*The development of rapid and robust analytical techniques is essential for the introduction of solid phase organic chemistry in the field of combinatorial chemistry. We describe here the technique of high resolution magic angle spinning NMR, which allows to obtain high quality homo- and heteronuclear spectra of resin bound molecules. Solvent signal suppression is possible through the diffusion filter, allowing a reaction monitoring without any purification step. The same method gives access to the detection and identification of impurities coming from secondary reactions on the resin. Finally, we evaluate the sensitivity criteria necessary for identification of the molecules on one individual bead.*

**Mots clés :** *RMN HR MAS, chimie combinatoire, chimie organique en phase solide, suivi de réactions, impuretés.*

**Key-words :** *High resolution magic angle spinning (HR MAS) NMR, combinatorial chemistry, solid phase organic chemistry, reaction monitoring, impurity detection*

Les méthodes modernes de synthèse organique stéréosélective se trouvent au cœur de la découverte et du développement des médicaments. Afin de préparer les petites molécules qui seront des candidats médicaments pour demain, la chimie organique utilise des stratégies améliorées de synthèse, de meilleures méthodes de purification et des outils analytiques avancés. La synthèse organique en phase solide (SOPS), basée sur les idées originales de la synthèse peptidique en phase solide [1], contient des promesses pour augmenter d'une façon dramatique la diversité et le nombre de ces petites molécules [2].

SOPS est une méthode de choix pour la chimie combinatoire entre autre parce qu'elle mène facilement vers l'automatisation. Néanmoins, l'application de schémas de réaction connus en solution pour la chimie en phase solide n'est pas toujours évidente. A partir de cette expérience, il est clair que la détermination des conditions de réactions sera dans beaucoup de cas l'étape qui demandera le plus de temps, et ceci encore plus en chimie combinatoire où l'optimisation d'un schéma de réaction n'implique pas deux molécules mais deux classes de molécules.

Un désavantage important dans ce processus d'optimisation est le manque d'une méthode rapide et robuste pour analyser d'une façon quantitative les molécules sur le support solide, aussi bien pour les étapes intermédiaires que pour le produit final. L'analyse après coupure reste bien sûr possible, mais demande un investissement considérable en

temps. Pour des réactions plus complexes, la résonance magnétique nucléaire à haute résolution et à l'angle magique (RMN HR MAS) [3-8] se présente comme un nouvel outil analytique qui répond à cette question. En analysant les différentes étapes d'une banque combinatoire, Luo *et al.* ont démontré l'avantage décisif de la RMN HR MAS pour l'analyse des étapes plus complexes de la réaction [9] : avec les mêmes séquences d'impulsion que la RMN en solution, la méthode permet d'identifier tous les protons et carbones, comme la RMN en solution.

Après la publication initiale des spectres de bonne résolution d'une résine grâce à l'utilisation d'un gonflement correcte du gel et la rotation à l'angle magique (*cf. encadré*) [3-4], les groupes impliqués, dont le nôtre, ont cherché les critères physicochimiques du gonflement afin d'obtenir les meilleurs spectres. La conclusion basée sur l'expérience du chimiste est simple : le solvant qu'il utilise pour sa réaction est en général celui qui permet une bonne « injection de mobilité », et donc l'obtention de bons spectres. La mobilité rotationnelle des molécules sur le support solide n'est pas très différente de celle en solution, tandis que le polymère réticulé se comporte comme un vrai support solide, ce qui permet d'obtenir d'une façon sélective le spectre des premiers, tout en éliminant les signaux du polymère.

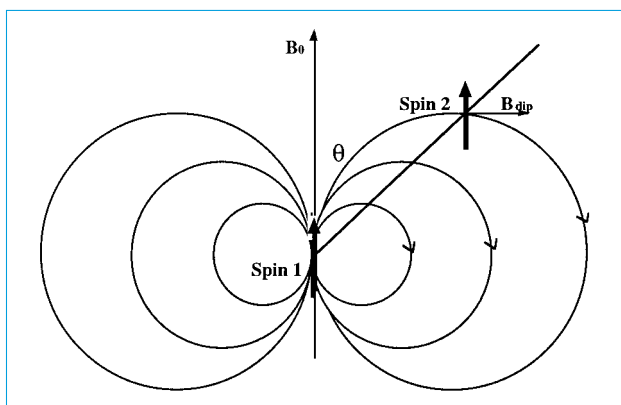
Nous avons montré que même avec une quantité de résine inférieure au milligramme (avec une charge normale de 0,4 mmole/g), on peut obtenir des spectres à une dimension satisfaisants dans la minute (*figure 1*). Dans une application classique de chimie combinatoire, on dispose facilement de plusieurs milligrammes de résine, et des spectres 2D (comme le TOCSY ou HSQC) peuvent être enregistrés en quelques heures, permettant une procédure d'identification

\* UMR 8525, Institut de Biologie de Lille, Université de Lille II, Institut Pasteur de Lille, 1, rue du Professeur Calmette, BP 447, 59021 Lille Cedex.

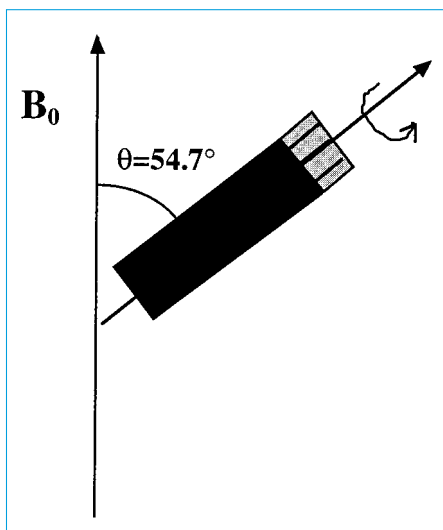
<sup>1</sup> Tél. : 03.20.87.12.29. Fax : 03.20.87.12.33.  
E-mail : guy.lippens@pasteur-lille.fr

### Encadré

La **qualité des spectres** obtenus sur les résines tient en ce que l'inhomogénéité présentée par l'échantillon de la résine, avec ses zones de polystyrène, de solvant, de molécule organique... peut être traitée formellement de la même façon que l'interaction dipolaire entre deux spins. Quand on représente le champs dipolaire graphiquement (voir ci-dessous), on s'aperçoit qu'il existe une orientation où il n'est pas nul (ce qu'il n'est qu'à l'infini), mais perpendiculaire à la direction imposée par le grand champs extérieur, réduisant ainsi à zéro l'interaction dipolaire, produit scalaire du champ et du moment magnétique associé au spin.



Lignes de champs dipolaire d'un spin 1 orienté selon la direction z de  $B_0$  et à l'origine. Ce champs n'a plus de composante selon z quand l'angle  $\theta$  est de  $54,7^\circ$ , solution de  $3 \cos^2(\theta) - 1 = 0$ .



Le rotor tourne autour d'un axe qui fait  $54,7^\circ$  avec la direction du champs magnétique  $B_0$ .

L'interaction dipolaire, qui conduit à des signaux très larges en RMN en phase solide (plusieurs dizaines de kHz), est moyennée à zéro en solution par suite des mouvements moléculaires rapides. De même, la **rotation à l'angle magique**

fait que chaque paire de spins se trouve en moyenne sous cette orientation magique. Si la vitesse de rotation est supérieure à l'interaction que l'on souhaite moyennner, des lignes fines résultent. L'élargissement des lignes dû aux inhomogénéités de l'échantillon, qui est de quelques centaines de Hz, impose une vitesse de rotation supérieure à 1 kHz, ce qui est facilement réalisable. Néanmoins, les largeurs de ligne obtenues par HR MAS NMR ne sont pas aussi fines que celles des mêmes molécules en solution, pour des raisons d'anisotropie de la susceptibilité magnétique.

### Ile-Leu-Asn-Gly-HMBA-PS

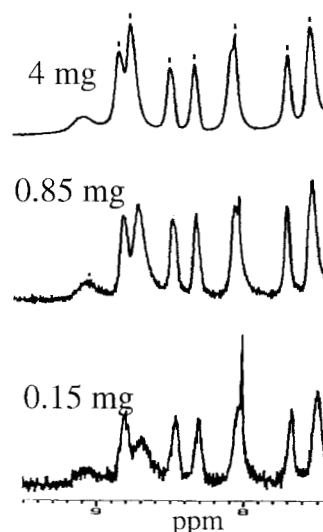


Figure 1 - Région amide du spectre proton d'un térapeptide Gly-Asn-Leu-Ile greffé par un bras moléculaire AHMB (acide hydroxyméthyle benzoïque) sur une résine de polystyrène/divinylbenzène 1%, gonflé dans le DMF-d7. L'acquisition a été faite en une minute sur un spectromètre Bruker DMX à 600 MHz.

identique à celle de la RMN en solution. L'intégration des spectres à une dimension, avec les précautions habituelles de temps de récupération suffisamment long, permet une quantification des molécules greffées.

L'identification et la quantification des produits de réaction étant possibles, on peut suivre l'avancement d'une réaction sans coupure des molécules de la résine. Nous avons démontré ceci par deux expériences différentes [10] : d'abord, une réaction d'Horner-Emmons directement dans le rotor, dans un solvant deutéré, et ensuite, la même expérience dans un récipient standard avec prélèvement de quelques milligrammes de résine à des moments précis de la réaction. La première expérience, malgré sa simplicité conceptuelle, ne représente pas de conditions réalistes de réaction - la petite taille du rotor (100  $\mu$ l, dans un rotor plein), l'impossibilité de contrôler l'atmosphère de réaction, l'effet de la rotation (3 000 tours par seconde !)... peuvent tous influencer la cinétique. Dans la deuxième expérience, la résine rincée et séchée était regonflée dans un solvant deutéré, et la comparaison de l'intégrale du signal correspondant au produit formé avec un standard soluble permet une quantification du produit de réaction.

Dans la robotisation de la synthèse parallèle, une fraction importante du temps est perdue dans les étapes de rinçage, lavage et séchage des résines. Ceci est également le cas pour l'analyse par RMN HR MAS, où les trois étapes prennent facilement quelques heures. L'étape de lavage et rinçage, même si elle enlève les molécules qui n'ont pas réagi, ne suffit pas, car les solvants organiques que l'on utilise couramment en SOPS possèdent facilement plusieurs signaux intenses, qui ne peuvent pas être supprimés facilement par les techniques standard de la RMN en solution. Afin de résoudre ce problème, nous avons exploité la nature insoluble du polymère : les molécules qui ont réagi ont, par définition, perdu leur degrés de liberté translationnelle, tout en

maintenant leur liberté rotationnelle (autrement, la RMN HR MAS ne serait pas possible), contrairement aux molécules restées en solution ou aux molécules du solvant. Utilisant un filtre de diffusion décrit pour la RMN en solution, nous pouvons supprimer sélectivement les signaux de ces dernières, sans sacrifier les signaux RMN des molécules ancrées (figure 2) [11]. Le même filtre de diffusion peut être introduit dans toute séquence 2D, permettant l'identification des molécules greffées avec un échantillon prélevé directement du ballon de réaction (figure 3). Ainsi, le suivi de la réaction devient facile : on transfère quelques milligrammes de suspension dans un rotor, on enregistre le spectre avec le filtre de diffusion, et après quelques minutes, on peut quantifier l'état d'avancement de la réaction en comparant l'intégrale du produit formé avec un signal commun au produit de départ et produit final. Cette facilité ne peut même pas être obtenue en chimie en phase homogène, car ici, la séparation des molécules attachées et réactifs solubles ne demande aucune étape de purification mais est faite par l'expérience RMN - les avantages chimiques de la SOPS se transforment en avantage analytique !

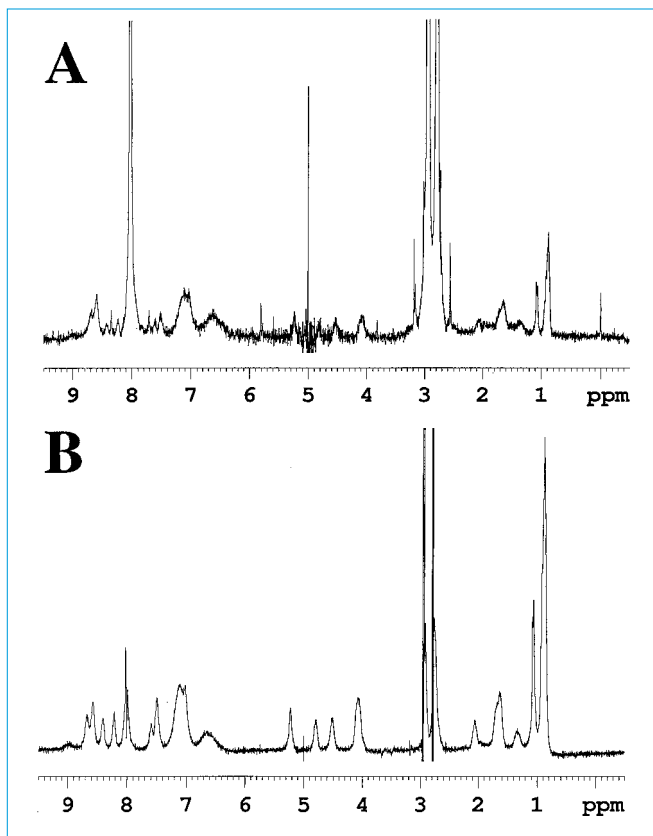


Figure 2 - Spectre de 0,15 mg de la même résine en DMF-h7, (a) avec une séquence à une impulsion et (b) avec le filtre de diffusion. Le nombre d'accumulations est identique pour les deux spectres.

Une autre question que l'on peut poser à une technique analytique est la détection, identification et quantification des impuretés. Avec un système artificiel de glutamate et pyroglutamate greffés simultanément sur un dipeptide préalablement synthétisé sur une résine, nous avons démontré qu'avec 10 mg de résine, 1 % d'impureté pouvait être détecté et identifié après 12 heures d'acquisition à la fréquence de 600 MHz (figure 4) [12]. Pouvoir discriminer très

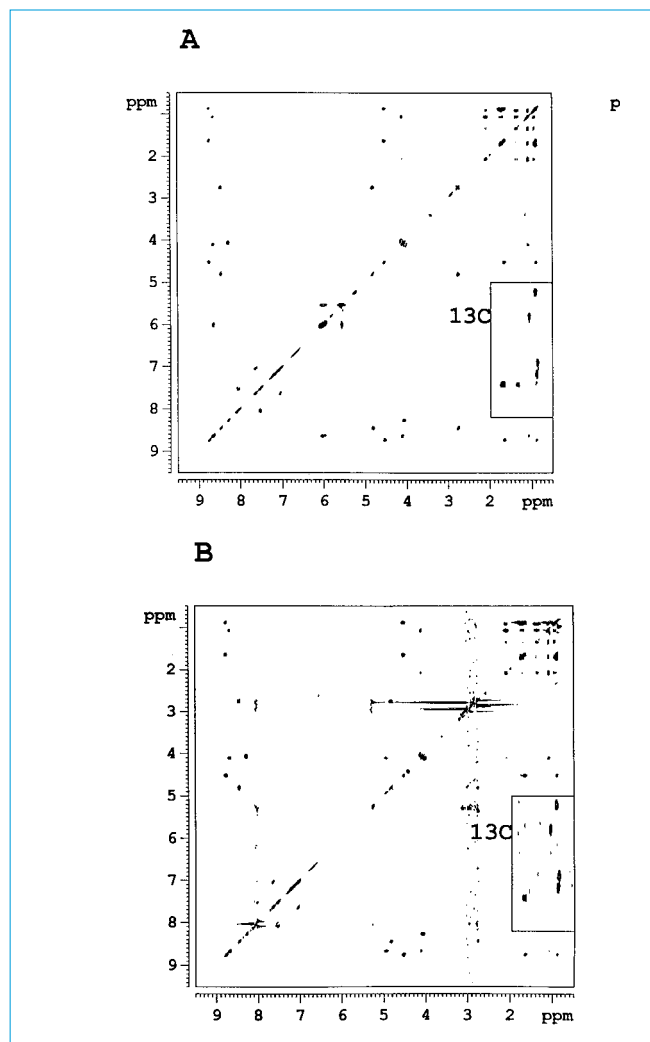


Figure 3 - Spectres à 2 dimensions (TOCSY, avec, en insert, la région des méthyles d'un spectre HSQC 1H-13C) de 2,5 mg de la même résine (a) en DMF-d7 et (b) en DMF-h7, où le solvant a été éliminé grâce au filtre de diffusion.

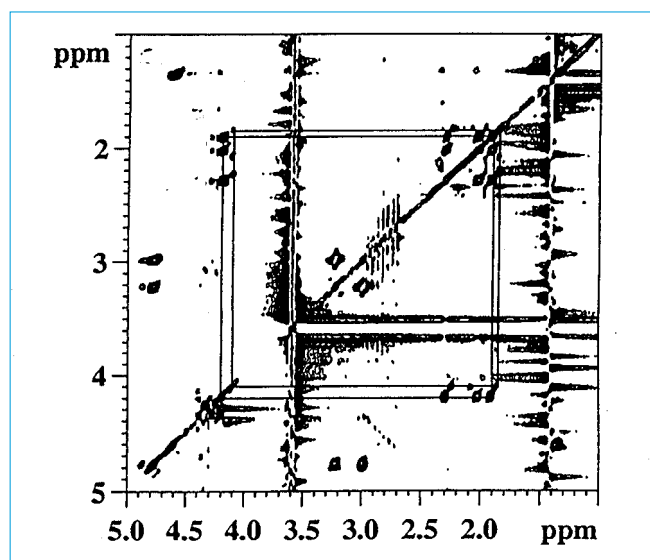


Figure 4 - Spectre TOCSY d'une résine contenant 99 % de tripeptide Ala-Phe-Glu, et 1 % de Ala-Phe-pyroglutamate. La détection et l'identification du système de spins de la pyroglutamate sont possibles après une nuit à 600 MHz, avec 10 mg de résine.

tôt dans le développement d'un produit des conditions de réactions qui mènent vers des réactions secondaires et donc des impuretés est d'une grande importance, car plusieurs cas sont connus où un résultat positif dans l'étape suivante de criblage n'était pas dû au produit attendu mais à un produit d'une réaction secondaire.

Les applications de la RMN HR MAS décrites ci-dessus concernent surtout l'optimisation des réactions de la SOPS - une autre application est l'identification des molécules sur une bille individuelle de résine. Dans le cadre d'une approche « split-and-mix » suivie d'un test biochimique sur les billes, comme développée par le groupe de Carlsberg [13], l'analyse de la structure moléculaire sur la bille pourrait éviter des stratégies de marquage de la bille. Malheureusement, la faible sensibilité de la RMN en général la rend a priori incompatible avec le nanomole de produit sur une bille individuelle. Travailler avec une molécule enrichie en  $^{13}\text{C}$  sur une résine permet de détecter le signal des molécules sur une bille [14], mais l'attribution totale d'une molécule sans connaissance préalable du résultat paraît difficile. Des approches avec des macrobilles ou avec des dendrimères augmentant la charge de la bille ont été décrites [15], mais il nous semble que, du fait de sa faible sensibilité, la RMN reste un mauvais compétiteur d'autres techniques analytiques comme la spectrométrie de masse ou la spectroscopie infrarouge à transformée de Fourier.

En **conclusion**, nous croyons qu'il existe de nombreuses applications potentielles pour la RMN HR MAS en synthèse organique en phase solide, et que l'outil accélèrera le pro-

cessus d'optimisation, étape cruciale dans son application à la chimie combinatoire. Ainsi, les méthodes de RMN HR MAS rentreront dans les laboratoires de SOPS comme un outil indispensable, équivalent de la RMN en solution pour la chimie organique en phase homogène.

## Références et note

- [1] Merrifield B., *J. Am. Chem. Soc.*, **1963**, 85, p. 2149-2154.
- [2] Une liste de références exhaustive de la SOPS en chimie combinatoire peut être trouvée à l'adresse <http://www.5z.com/divinfo/>.
- [3] Fitch W.L., Detre G., Holmes C.P., Shoorely J.N., Keifer P.A., *J. Org. Chem.*, **1994**, 59, p. 7955-7956.
- [4] Anderson R.C., Stokes J.P., Shapiro M.J., *Tetrahedron Lett.*, **1995**, 36, p. 5311-5314.
- [5] Pop I., Dhalluin C., Depréz B., Melnyk P., Lippens G., Tartar A., *Tetrahedron*, **1996**, 52, p. 12209 - 12222.
- [6] Keifer P.A., *J. Org. Chem.*, **1996**, 61, p. 1558-1559.
- [7] Riedl R., Tappe R., Berkessel A., *J. Am. Chem. Soc.*, **1998**, 120, p. 8994-9000.
- [8] Ruhland R., Andersin K., Pedersen H., *J. Org. Chem.*, **1998**, 63, p. 9204-9211.
- [9] Luo Y., Ouyang X., Armstrong R.W., Murphy M.M., *J. Org. Chem.*, **1998**, 63, p. 8719-8722.
- [10] Warrass R., Lippens G., *J. Org. Chem.*, **2000**, 65, p. 2946-2950.
- [11] Warrass R., Wieruszkeski J.-M., Lippens G., *J. Am. Chem. Soc.*, **1999**, 121, p. 3787-3788.
- [12] Rousselot-Pailley P., Maux D., Wieruszkeski J.-M., Martinez J., Aubagnac L., Lippens G., *Tetrahedron*, sous presse.
- [13] Rademann J., Grötli M., Meldal M., Bock K., *J. Am. Chem. Soc.*, **1999**, 121, p. 5459-5466.
- [14] Sarkar S.K., Garigipati R.S., Adams J.L., Keifer P.A., *J. Am. Chem. Soc.*, **1996**, 118, p. 2305-2306.
- [15] Porsch M., Schlotterbeck G., Tseng L.-H., Albert W., Rapp W., *Angew. Chemie*, **1996**, 108, p. 3034-3036.

# Contribution de la spectrométrie de masse au développement de la chimie combinatoire

**Pierre-Hervé Lambert\*** docteur, **Sophie Bertin\*** docteur, **Jean-Paul Volland\*** ingénieur chimiste

**Summary :** *Contribution of mass spectrometry to the development of combinatorial chemistry*

*In the field of the search for biological active compounds, the development of high throughput screening techniques, combinatorial chemistry and parallel synthesis have caused the need for efficient analytical methods able to handle complex mixtures or large numbers of compounds. In this article, various mass spectrometric technics allowing the analysis of libraries or collections of compounds will be discussed. New high throughput technologies like high-density sampling MALDI (matrix assisted laser desorption ionization) targets, coupling of parallel liquid chromatography and mass spectrometry, and new developments of mass spectrometry-based preparative chromatography, isotopic encoding and biological screening will be reported.*

**Mots-clés :** *Spectrométrie de masse, chimie combinatoire, bibliothèques, collections, haut débit.*

**Key-words :** *Mass spectrometry, combinatorial chemistry, libraries, collections, high-throughput.*

Durant la dernière décennie, les développements de la spectrométrie de masse (nouvelles méthodes d'ionisation douce, couplage avec la chromatographie en phase liquide, etc.) ont permis les progrès de chimies non conventionnelles comme la chimie combinatoire [1-3]. Selon les différentes stratégies retenues pour découvrir une molécule biologiquement active, le problème de l'identification moléculaire peut être posé de manière extrêmement variée. En effet, sous le vocable « chimie combinatoire », on distingue en réalité des problématiques très différentes : la nature même de l'échantillon peut être, soit un mélange, dénommé **bibliothèque**, d'un nombre plus ou moins grand de molécules obtenues par chimie combinatoire proprement dite (de quelques unités à plusieurs milliers d'exemplaires), soit, au contraire, un nombre tout aussi variable de molécules isolées obtenues par synthèse parallèle et l'on parlera alors de **collection** de produits. Ces principes ne seront pas revus ici, mais l'importance de la spectrométrie

de masse dans ce type de recherche y sera soulignée.

## Analyse de bibliothèques obtenues par synthèse combinatoire

Quelle que soit la stratégie choisie (déconvolution itérative, « positional scanning », etc.), une bibliothèque se compose d'un mélange plus ou moins complexe de molécules dont les structures possèdent une partie commune. Sur celles-ci sont greffés des substituants pris parmi différents ensembles, afin de synthétiser, par un jeu de synthèse combinatoire, le plus grand nombre de combinaisons possibles de molécules, le but étant de recouvrir une certaine diversité moléculaire. Les problèmes posés à l'analyste sont de différents ordres :

- contrôler la présence de chaque molécule attendue,
- estimer leur équimolarité,
- identifier les possibles produits secondaires.

Pour ce faire, le choix des techniques de spectrométrie de masse va dépendre essentiellement du nombre de composants de la bibliothèque ainsi que de sa dégradation, c'est-à-dire de la frac-

tion de molécules possédant la même masse moléculaire par rapport au nombre total de molécules.

## Analyse de bibliothèques par spectrométrie de masse basse résolution

Considérons une bibliothèque de tétra-peptides du type O1X2X3X4 synthétisée selon la méthode « mix and split » en vue d'une déconvolution itérative et pour laquelle X sera choisi parmi 20 acides aminés : cette bibliothèque est composée de  $20^3$  soit 8 000 peptides de nature différente. Étant donné qu'il est illusoire de tenter de séparer efficacement un tel nombre de composés, l'analyse la plus rapide est de comparer par spectrométrie de masse la distribution expérimentale des intensités des molécules protonées en fonction de leur rapport masse/charge avec la distribution théorique calculée. Le choix de la méthode d'ionisation sera donc limité à une méthode d'ionisation douce supposée conduire à un ion unique par molécule. Parmi celles-ci, l'électrospray (ESI : electrospray ionization), l'ionisation chimique à pression atmosphérique (APCI : atmospheric pressure chemical ionization) ou la désorption laser assistée par matrice (MALDI : matrix-assis-

\* Institut de Recherches Servier, division de Physico-chimie analytique, 11, rue des Moulineaux, 92150 Suresnes.  
Tél. : 01.55.72.22.68. Fax : 01.55.72.23.80.  
E-mail : libpca@francenet.fr

ted laser desorption ionization) sont généralement employées. Cependant, la précision d'une telle analyse reste limitée. En effet, ces méthodes d'ionisation ne conduisent pas nécessairement à un ion unique, par exemple de type  $[M+H]^+$  en mode positif, mais aussi, éventuellement, à des ions multichargés  $[M+nH]^{n+}$ , des ions adduits  $[M+H+\text{solvant}]^+$ , voire à des ions-fragments, selon des abondances qu'il est difficile de prévoir. L'équimolarité est à l'évidence délicate à estimer d'autant que les rendements d'ionisation de chacun des constituants peuvent varier d'un ordre de grandeur en fonction des propriétés physico-chimiques des molécules. Un certain bruit de fond chimique, dû à la matrice ou aux solvants utilisés, peut aussi ne pas être négligeable. La *figure 1* montre cependant qu'une librairie incomplètement synthétisée peut être mise en évidence par cette simple approche.

### Analyse de librairies par spectrométrie de masse haute résolution

Afin d'améliorer la spécificité de telles analyses, l'augmentation

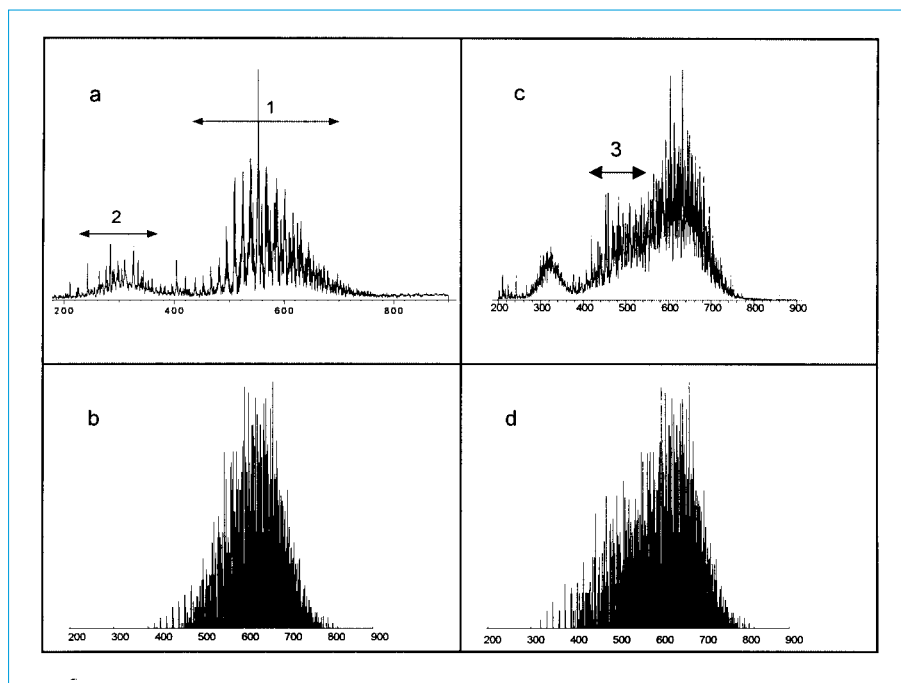
de la résolution, en séparant des ions de même masse nominale mais de compositions élémentaires différentes, peut être envisagée. Cependant, lorsque la dégénérescence de la librairie augmente, la résolution nécessaire pour séparer la plupart des ions devient vite hors de portée des spectromètres de masse les plus résolutifs. Il a été calculé [4] qu'une résolution procurant une précision de 5 ppm est nécessaire pour lever la dégénérescence d'une librairie non peptidique composée de seulement  $12 \times 96$  composés. Pour une librairie peptidique de  $12 \times 16 \times 9$  composés, pour laquelle on s'attend à une dégénérescence supérieure, une telle haute résolution ne permettrait l'identification que de 33 % de ses constituants. Lorsque l'on met en rapport le coût et éventuellement la difficulté de mise en œuvre de la haute résolution, le bénéfice de celle-ci ne paraît évident que pour des librairies de taille limitée à quelques dizaines de composés. Une façon de contourner cette difficulté est alors de construire des librairies non dégénérées. Hughes [5] propose un certain nombre de règles permettant d'en élaborer.

### Analyse de librairies par couplage spectrométrie de masse/spectrométrie de masse (MS/MS : mass spectrometry/mass spectrometry)

La spectrométrie de masse en tandem MS/MS ne peut être envisagée que pour rechercher plus ou moins spécifiquement dans des mélanges des familles de composés dont on suppose qu'une partie commune de la structure induirait une réaction prévisible et générale après ionisation en phase gazeuse. L'efficacité de certains modes de balayage MS/MS permet de détecter spécifiquement certains peptides incomplètement déprotégés. Ainsi, des peptides protégés par des groupements *tert*-butyloxy et Pmc (2,2,5,7,8-pentaméthylchroman-6-sulfonyl) sont détectés par des pertes de neutres de 56 et de 309 Da, tandis que des peptides encore tritylés peuvent l'être par des balayages sélectionnant les ions précurseurs du cation triphénylméthylum [6]. Les familles particulières de peptides de type O1O2-Arg-X4 peuvent aussi être détectées par MS/MS dans des librairies de type O1O2X3X4 de 24 x 24 tétrapeptides [7].

### Analyse de librairies par couplage chromatographie en phase liquide/spectrométrie de masse (LC/MS : liquid chromatography/mass spectrometry)

Bien entendu, les techniques de couplage chromatographie en phase liquide/spectrométrie de masse sont les plus adaptées à l'analyse de librairies à la condition qu'elles soient de taille modérée. L'acquisition simultanée du signal d'un détecteur de chromatographie en phase liquide et des spectres de masse apporte une information supplémentaire sur l'équimolarité de la librairie. Avec la spectrométrie de masse, il est possible, même lorsque de nombreux composés coéluent, de dénombrer le nombre de composés possédant une masse moléculaire donnée par le biais de courants ioniques reconstitués. Pour de très grandes librairies, les problèmes de coélution et de sensibilité rendent cette méthode inexploitable. Cependant, une image partielle de la librairie peut être obtenue par le mode d'acquisition d'ions



**Figure 1** - librairie Nip-X2X3X4-NH2 (Nip : 4-nitrophénylalanine).

a : spectre de masse ESI montrant les tétrapeptides protonés (zone 1) et doublement protonés (zone 2).

b : distribution théorique pour une librairie complète et équimolaire de tétrapeptides.

c : spectre de masse ESI d'une librairie incomplètement synthétisée. La zone 3 est révélatrice de la présence de tripeptides protonés.

d : distribution théorique des peptides protonés en supposant que l'étape de couplage de X2 ne s'est réalisée qu'avec un rendement de 20 %.

sélectionnés SIM (selected ion monitoring) où un (ou quelques) ion(s) de rapport  $m/z$  prédéfini(s) sont seulement détecté(s). Il a ainsi été montré [8] que la sensibilité de cette approche est suffisante pour détecter un peptide connu pour être le seul possédant une masse moléculaire donnée parmi  $24^3$  soit 13 824 autres peptides dans une librairie de type O1X2X3X4.

### Analyse à haut débit de collections obtenues par synthèse parallèle

Les problèmes posés par l'analyse de collections diffèrent notablement de ceux posés par l'analyse de bibliothèques puisque cette fois-ci, c'est le nombre, et non la complexité des échantillons, qui est en cause. Pour résoudre ce problème, l'ensemble des étapes de l'analyse, c'est-à-dire, la préparation des échantillons, leur introduction, leur analyse proprement dite ainsi que l'exploitation et la sortie des résultats, doit pouvoir être automatisé efficacement.

### Analyse de collections par injections en boucle (FIA : flow injection analysis)

Un des intérêts des méthodes d'ionisation à pression atmosphérique, ESI ou APCI, est qu'elles nécessitent l'introduction des échantillons en solution. De ce fait, la robotique existante pour des systèmes HPLC (high performance liquid chromatography), comme les préparateurs-injecteurs d'échantillons, est directement compatible avec des spectromètres équipés de ce type de source d'ionisation. Le nombre d'échantillons analysables par unité de temps n'est limité que par le cycle d'injection et par d'éventuels effets mémoire. Si l'information sur la masse moléculaire est seulement souhaitée, des cadences d'un échantillon par 0,5 à 2 minutes sont possibles. Par l'utilisation de vannes d'injection à entrées multiples, huit échantillons peuvent être introduits dans le flux de solvant en une seule injection. L'analyse totale d'une plaque de 96 puits est réalisée en seulement 12 minutes [9]. Les spectromètres de masse récents de type MS/MS sont aujourd'hui capables de reconnaître à chaque

balayage les ions les plus abondants et, convenablement programmés, d'entreprendre successivement l'enregistrement de leurs spectres d'ions-fragments produits par collision, apportant ainsi une information structurale complémentaire. D'autre part, l'avènement de spectromètres de masse à analyseur à temps de vol (TOF : time of flight), équipés de systèmes d'extraction d'ions retardée et de miroirs électrostatiques rend possible, par leur pouvoir résolutif, la détermination des compositions élémentaires de ces ions de façon routinière.

La principale limitation des systèmes d'analyse par FIA reste le fait que l'ion le plus abondant peut ne pas correspondre à l'espèce la plus abondante. En effet, lors de l'ionisation, des phénomènes de compétitivité, communément dénommés « effets de suppression », peuvent être responsables de la surreprésentation d'une espèce minoritaire par rapport à l'espèce principale. Ce problème est particulièrement aigu lorsqu'on travaille avec un échantillon provenant d'une bille unique (typiquement quelques dizaines de picomoles) et pouvant contenir, par exemple, un solvant comme le diméthyl sulfoxyde qui produit un effet de suppression puissant. Un mode d'introduction plus sophistiqué faisant intervenir après lavage à l'eau une élution sur une microcolonne capillaire résout ce type de difficulté [10].

### Analyse de collections par couplage chromatographie en phase liquide/ spectrométrie de masse

Malgré une mise en œuvre plus délicate et certaines limitations, les avantages de la LC/MS comparés à

l'injection en boucle paraissent évidents :

- raréfaction ou disparition d'effets de suppression évoqués plus haut,
- possibilité, par l'adjonction d'un deuxième type de détecteur, de disposer d'un critère de pureté qui peut être indispensable pour valider les résultats pharmacologiques.

En HPLC, par l'utilisation de nouvelles colonnes plus courtes à granulométrie plus faible, de gradients rapides, les analystes ont cherché un compromis entre la vitesse d'analyse, et une certaine préservation de la résolution chromatographique. Des cycles d'analyses inférieurs à 10 minutes sont couramment utilisés. Un gain de temps supplémentaire est réalisé par l'utilisation de systèmes utilisant deux colonnes en parallèle, l'une fonctionnant en mode analytique tandis que l'autre est rééquilibrée. L'emploi de chromatographie liquide supercritique couplée à la spectrométrie de masse SFC/MS permet de réduire ces cycles à 1 ou 2 minutes [11].

### Analyse de collections par couplage de la chromatographie en phase liquide en parallèle et de la spectrométrie de masse

Il est commercialisé depuis peu une interface LC/MS qui permet le couplage en parallèle de huit chromatographes en phase liquide à un seul spectromètre de masse équipé d'une source électrospray. Le principe de ce couplage est représenté *figure 2*. Un rotor sélectionne successivement par une fenêtre chacun des huit canaux et synchronise l'acquisition d'un spectre de masse et d'un signal UV dans chacun des huit fichiers d'acquisition. Le temps

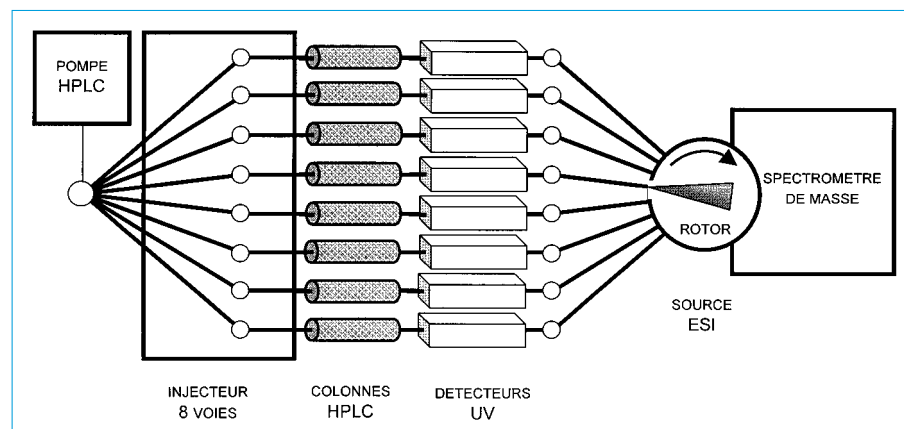


Figure 2 - Couplage LC/MS selon la technologie MUX™ développée par la société Micromass®.



de cycle de 1,6 seconde respecte l'intégrité des profils chromatographiques.

### Analyse de collections par désorption laser assistée par matrice (MALDI)

La désorption laser assistée par matrice est susceptible d'être automatisée pour répondre aux besoins de l'analyse de collections à haut débit. P. Little [12] décrit un système capable de dispenser à l'aide d'un robot équipé de micropipettes piézo-électriques des volumes de quelques nanolitres de solutions de matrice et d'échantillon dans chacun des 100 puits d'une cible de silicium de 1 cm<sup>2</sup>. L'ensemble est introduit dans le spectromètre de masse sur un support manœuvré par des moteurs pas à pas de telle sorte que chaque puits est placé successivement dans la zone de l'impact laser. L'acquisition des 100 spectres de masse est réalisée ainsi en moins de 6 minutes.

### Exploitation des données

Devant les grandes quantités de données analytiques générées par des systèmes d'analyse à haut débit, plusieurs laboratoires ont été amenés à développer leurs logiciels d'exploitation de données afin de réduire au maximum certaines tâches répétitives. Ces logiciels d'interprétation sont capables d'identifier différents types d'ions (ions multichargés, ions adduits, etc.), de proposer des structures pour certaines impuretés, de définir des critères de pureté [13]. Des logiciels adaptés au contrôle de qualité de collections analysées par plaques de 96 puits ont été commercialisés. Les puits sont visualisés par différentes couleurs selon qu'une molécule possédant la masse moléculaire (ou la formule brute) attendue a été détectée dans le puits correspondant tout en respectant certains critères de pureté.

## Autres développements

### Le codage isotopique

Une molécule greffée sur une bille de résine ayant réagi positivement à un test pharmacologique peut être, soit identifiée directement, soit indirectement par la lecture d'un code constitué

généralement de molécules cosynthétisées sur le support solide et dont l'analyse présente l'avantage d'une plus grande simplicité. Un exemple intéressant est donné par Geysen qui met en œuvre la cosynthèse de dipeptides Gly-Gly dont chaque glycine comporte aucun, un ou deux atomes de carbone 13. Le massif isotopique de ces dipeptides protonés constitue un véritable « code-barre » qui peut être décodé par spectrométrie de masse [14].

### La chromatographie préparative automatisée

La chromatographie liquide préparative assistée par spectrométrie de masse est aujourd'hui proposée commercialement [15-16]. Ces systèmes permettent de purifier automatiquement quelques milligrammes de grands nombres de produits, ce qui a pour effet d'accroître la fiabilité des tests pharmacologiques et de simplifier les mises au point de synthèses combinatoires.

### Le « screening »

Les technologies utilisant la spectrométrie de masse en tant que détecteur lors d'expériences de « screening » se développent rapidement. Des inhibiteurs de réactions enzymatiques entre plusieurs cibles et des bibliothèques peuvent être mis en évidence à raison de 4 000 échantillons disposés sur le porte-échantillons de 20 cm<sup>2</sup> d'un spectromètre de masse à temps de vol équipé d'une source MALDI [17]. Les ligands d'une bibliothèque les plus fortement liés à un récepteur peuvent être identifiés par spectrométrie de masse couplée à des techniques de sélection par affinité. Par exemple, les complexes récepteur-ligand, obtenus par incubation d'une bibliothèque avec un récepteur, après avoir été séparés des composés non liés par chromatographie d'exclusion, sont dissociés et les ligands résultants identifiés par spectrométrie de masse [18-19].

## Conclusion

La spectrométrie de masse intervient à toutes les étapes de la découverte d'une molécule active et prend toute sa part dans le développement de la chimie combinatoire. L'apparition de nouveaux

spectromètres de masse à temps de vol résolutifs équipés de source électrospray ou à désorption laser rend routinières les mesures précises de masse et la détermination des compositions élémentaires. De même, l'information structurale est aisément obtenue à partir d'expérience MS<sup>n</sup> à l'aide de trappes ioniques. Des systèmes à détecteurs multiples LC/UV/N/MS (N : détecteur d'azote à chimioluminescence) installés en ligne répondent aux questions de pureté, de quantification et d'identification des molécules constituant les bibliothèques ou les collections [20]. Des progrès en terme de miniaturisation, d'automatisation et de vitesse d'analyse sont encore imaginables en spectrométrie de masse et conditionneront le développement de la chimie combinatoire.

### Références

- [1] Süßmuth R.D., Jung G., *Journal of Chromatography B*, **1999**, 725, p. 49-65.
- [2] Loo J.A., *Eur. Mass Spectrom.*, **1997**, 3, p. 93-104.
- [3] Loo J.A., DeJohn D.E., Ogorzalek R.R., Andrews P.C., *Annual Reports in Medicinal Chemistry*, **1996**, p. 319-325.
- [4] Blom K.F., *Anal. Chem.*, **1997**, 69, p. 4354-4362.
- [5] Hughes I., *J. Med. Chem.*, **1998**, 41, p. 3804-3811.
- [6] Metzger J.W., Kempter C., Wiesmüller K.H., Jung G., *Anal. Biochem.*, **1994**, 219, p. 261-277.
- [7] Boutin J.A., Hennig P., Lambert P.-H., Bertin S., Petit L., Mahieu J.-P., Serkiz B., Volland J.-P., Fauchère J.-L., *Anal. Biochem.*, **1996**, 234, p. 126-141.
- [8] Lambert P.-H., Boutin J.A., Bertin S., Fauchère J.-L., Volland J.-P., *Rapid Commun. Mass Spectrom.*, **1997**, 11, p. 1971-1976.
- [9] Wang T., Zeng L., Strader T., Burtonet L., Kassel D.B., *Rapid Commun. Mass Spectrom.*, **1998**, 12, p. 1123-1129.
- [10] Marshall P.S., *Rapid Commun. Mass Spectrom.*, **1999**, 13, p. 778-781.
- [11] Ventura M.C., Farrell W.P., Aurigemma C.M., Greig M.J., *Anal. Chem.*, **1999**, 71, p. 4223-4231.
- [12] Little D.P., Cornish T.J., O'Donnell M.J., Braun A., Cotter R.J., Köster H., *Anal. Chem.*, **1997**, 69, p. 4540-4546.
- [13] Tong H., Bell D., Tabei K., Siegel M.S., *J. Am. Soc. Mass Spectrom.*, **1999**, 10, p. 1174-1187.
- [14] Geysen H.M., Wagner C.D., Bodnar W.M., Markworth C.J., Parke G.J., Schoenen F.J., Wagner D.S., Kinder D.S., *Chem. Biol.*, **1996**, 3, p. 679-688.
- [15] Zeng L., Burton L., Yung K., Shushan B., Kassel D.B., *Journal of Chromatography A*, **1998**, 794, p. 3-13.
- [16] Kiplinger J.P., Cole R.O., Robinson S., Roskamp E.J., Ware R.S., O'Connell H.J., Brailsford A., Batt J., *Rapid Commun. Mass Spectrom.*, **1998**, 12, p. 658-664.
- [17] Hsieh F., Keshishian H., Muir C., *J. Biomol. Screening*, **1998**, 3, p. 189-198.
- [18] Kaur S., McGuire L., Tang D., Dollinger G., Huebner V., *J. Protein Chem.*, **1997**, 16, p. 505-511.
- [19] Hsieh F., Gordon N., Regnier F., Afeyan N., Martin S.A., Vella G.J., *Mol. Diversity*, **1996**, 2, p. 189-196.
- [20] Taylor E.W., Qian M.G., Dollinger G.D., *Anal. Chem.*, **1998**, 70, p. 3339-3347.

# Les condensations à composants multiples et leur utilisation en synthèse combinatoire

Hugues Bienaymé\* docteur, Philippe Schmitt\*\* docteur

## Summary : Multi-component reactions in combinatorial chemistry

Since the emergence of high throughput screening techniques, corporate chemists were faced to one of their greatest challenge : respond to the overwhelming demand for generic molecule collections to « feed » these tests. Some reactions, characterized by an unusual synthetic efficiency, are particularly fitted to achieve this task. Multi-component reactions (M-CRs) such as the well-known Ugi reaction have been discovered and exploited to prepare large collections of generic molecules in a time- and cost effective manner. Over the last five years, corporate and academic researchers have made of this class of reactions one of the most efficient and cost-effective tools in the context of combinatorial and parallel synthesis.

**Mots clés :** Synthèse combinatoire, condensations à composants multiples, efficacité synthétique, réaction de Ugi, synthèse hétérocyclique.

**Key-words :** Combinatorial chemistry, multi-component reactions, synthetic efficiency, Ugi reaction, heterocyclic synthesis.

La pression économique imposée aux industries pharmaceutiques et agrochimiques a conduit les chercheurs industriels à adapter leurs méthodologies à des standards de productivité toujours croissants. Tandis que les biologistes ont introduit des techniques de criblage à haut débit (high throughput screening : HTS), les chimistes ont dû inventer une technique capable de produire de grandes collections de molécules, aussi diverses que possible en respectant des contraintes à la fois de temps et de prix : **la chimie combinatoire** [1].

L'aspect purement quantitatif de ce défi industriel est à présent quasi révolu. Avec la productivité des chimistes de synthèse augmentée d'un facteur 100, voire 1 000, la plupart des groupes pharmaceutiques et agrochimiques entretiennent actuellement des fonds moléculaires regroupant parfois plus d'un million de molécules.

Cependant, l'expérience industrielle récente démontre clairement que l'atteinte seule de cet objectif quantitatif est insuffisant.

D'une part, le problème de la représentativité des familles chimiques, ou chimiothèques, dans les fonds moléculaires se pose : sont-elles suffisamment diverses ? Sont-elles originales ? Ne vaut-il pas mieux posséder dix chimiothèques composées chacune de 5 000 molécules variées qu'une seule de 50 000 ?

Ces questions ne sont pas innocentes : l'industrie pharmaceutique s'attend à une augmentation importante du nombre de nouvelles cibles thérapeutiques dans les années à venir, avec la montée en puissance de la génomique et de la protéomique. Maximiser la diversité et l'originalité des fonds moléculaires est la garantie de trouver des touches, ou « hits », lors des campagnes de tests. Cet aspect qualitatif de la synthèse combinatoire (nouvelles familles chimiques), n'est que très imparfaitement pris en compte à l'heure actuelle.

D'autre part, un défi nouveau émerge : comment choisir parmi les centaines de touches ou « hits » générés par le couple « synthèse combinatoire/tests à haut débit » ceux qui ont un potentiel réel de développement ? [2]. La réponse à cette question est évidemment loin d'être triviale. La minimisation du risque lors de cette étape de

choix est cependant un des facteurs clé de succès des programmes de recherche pharmaceutique : rien ne sert de courir, si l'on ne suit pas la bonne direction ! Il faut donc décider très tôt de l'opportunité d'approfondir ou d'écarter ces pistes (« fast-kill strategy » [3]). L'enjeu est d'autant plus important que, dans l'histoire de l'industrie pharmaceutique, le temps et les ressources n'ont jamais été aussi précieux !

En pratique, afin de maximiser les chances de succès lors des phases avals d'optimisation, certaines contraintes portant sur les structures des molécules et leurs propriétés physico-chimiques seront prises en compte lors de la conception des chimiothèques génériques [4]. Le regain d'intérêt pour les familles hétérocycliques, actuellement constaté en synthèse combinatoire, relève de la même démarche [5].

On comprendra donc aisément que l'objectif de la synthèse combinatoire aujourd'hui n'est plus de préparer seulement de grandes collections de molécules (objectif quantitatif), mais bien de concevoir des chimiothèques qui répondent à la fois aux contraintes des tests à haut débit (nombre et diversité des molécules) et à celles de l'optimisation des « hits » (profil physico-chimique), tout en répondant à un souci d'origina-

\* CRIT Rhône-Poulenc Industrialisation, 24, avenue Jean-Jaurès, 69156 Décines. Tél. : 04.72.93.61.19. Fax : 04.72.93.53.37 E-mail :

hugues.bienayme@crit.rhone-poulenc.com

\*\* Tél. : 04.72.56.08.06.

E-mail : philippeschmitt@wanadoo.fr

lité des familles chimiques (brevetabilité). Il s'agit bien d'un objectif à la fois quantitatif et qualitatif.

L'objet de cet article est de montrer que certaines réactions chimiques, les condensations à composants multiples, sont des outils permettant de répondre de manière satisfaisante à ce nouveau défi de la synthèse combinatoire.

Les condensations à composants multiples (multi-component reactions, M-CR's), peuvent être brièvement décrites comme des réactions où au moins trois réactifs distincts contribuent à la structure du produit final (schéma 1) [6].

L'engouement récent en synthèse combinatoire, pour ces réactions, se comprend facilement :

- Des chimiothèques diverses et de grandes tailles sont accessibles à partir d'un petit nombre de blocs de construction.

- Ce sont des transformations « productives » (dans le sens du nombre d'échantillons qu'un chimiste peut préparer par unité de temps). Ceci est bien sûr le résultat du petit nombre de manipulations nécessaires pour préparer chaque molécule d'une chimiothèque (une seule étape chimique).

- Elles peuvent être conduites aussi

bien en solution que sur phase solide. Cette souplesse d'utilisation est tout à fait appréciable, le choix de la stratégie étant alors dicté par la nature des réactifs, la structure considérée, le niveau de pureté souhaité, etc.

- Le temps de mise au point de la méthode chimique sera plus court que pour une chimiothèque utilisant un schéma synthétique multi-étapes.

- Elles sont souvent facilement mises en œuvre (par exemple, compatibles avec des conditions d'inertage imparfaites).

D'un point de vue conceptuel, on peut constater que ces transformations possèdent, de fait, de nombreux attributs d'efficacité synthétique. En particulier, elles sont « hyper-convergentes », sélectives, générales, répondent au critère d'économie d'atome [7], et maximisent la complexité structurale entre réactifs et produits (schéma 2) [8]. Aucune autre classe de transformations chimiques ne rassemble, à notre connaissance, autant de qualités synthétiques.

Ces qualités se traduisent, en pratique, par la facilité relative avec laquelle il est possible de générer des collections importantes de molécules (chimiothèques génériques).

Mais leur intérêt ne s'arrête pas là :

dans la plupart des cas, la synthèse de chimiothèques ciblées autour d'un « hit » est également rapide. Il est ainsi possible d'évaluer son potentiel, et d'acquérir les premiers éléments de relation structure/activité, dans un temps relativement court.

Par ailleurs, un autre avantage, plus subtil, doit être cité : les condensations à composants multiples permettent l'accès à des familles chimiques dont l'invariant structural (le cœur de la chimiothèque) est minimisé.

L'expérience montre que pour les chimiothèques obtenues par fonctionnalisation séquentielle d'un « répartiteur moléculaire » [9], l'élément invariant à l'intérieur d'une famille, tend souvent à contribuer trop lourdement à la structure finale pour permettre d'optimiser facilement les « hits » obtenus. De la même façon que la masse élevée de ces répartiteurs imposera des contraintes fortes quant aux substituants disponibles pour conserver à l'ensemble une masse acceptable, leur caractère souvent lipophile tendra à déplacer la balance hydrophobe/hydrophile d'une famille donnée en dehors des normes de biodisponibilité orale admissibles [4]. L'intérêt de minimiser l'invariant structural au sein d'une famille chimique permet d'éviter ces dérives, et d'optimiser le profil physico-chimique des chimiothèques.

Ces avantages sont précieux lors des étapes d'optimisation des molécules actives.

Les condensations à composants multiples sont parmi les outils les plus performants de la synthèse organique, et les plus adaptés à la préparation de chimiothèques à fort « potentiel pharmaceutique ».

La réaction de Ugi constitue la plus connue d'entre elles (schéma 3) [10].

L'étape déterminante, pour cette réaction, est sans aucun doute liée à la propriété unique des isonitriles à subir une «  $\alpha$ -addition ». Au cours de la réaction, l'atome de carbone de l'isonitrile évolue d'une configuration divalente à une configuration tétravalente (imidate) en formant deux liaisons covalentes nouvelles. Dans le cas d'une réaction de Ugi, cette  $\alpha$ -addition de l'isonitrile intervient entre une espèce électrophile (un aldéhyde ou un cation immonium) et un nucléophile hétéroatomique (un

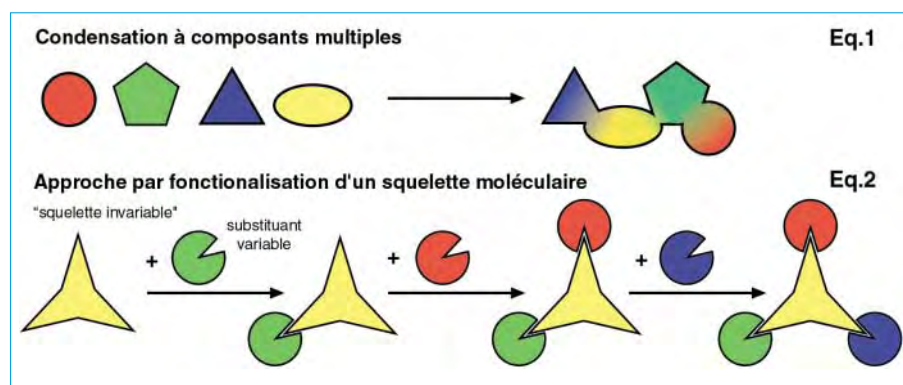


Schéma 1 - Les condensations à composants multiples.

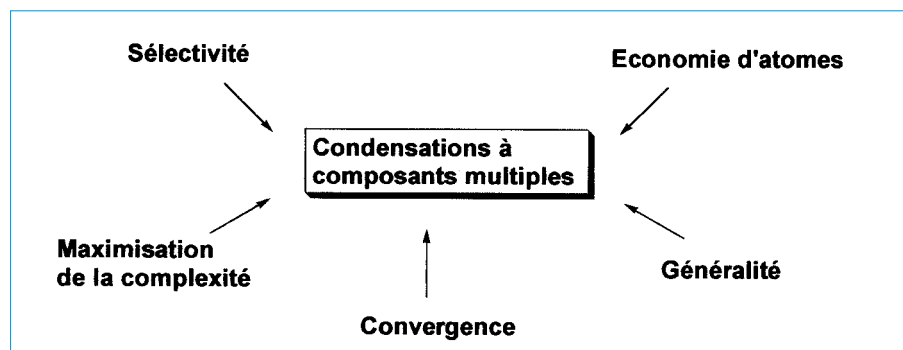


Schéma 2 - Propriétés des condensations à composants multiples.

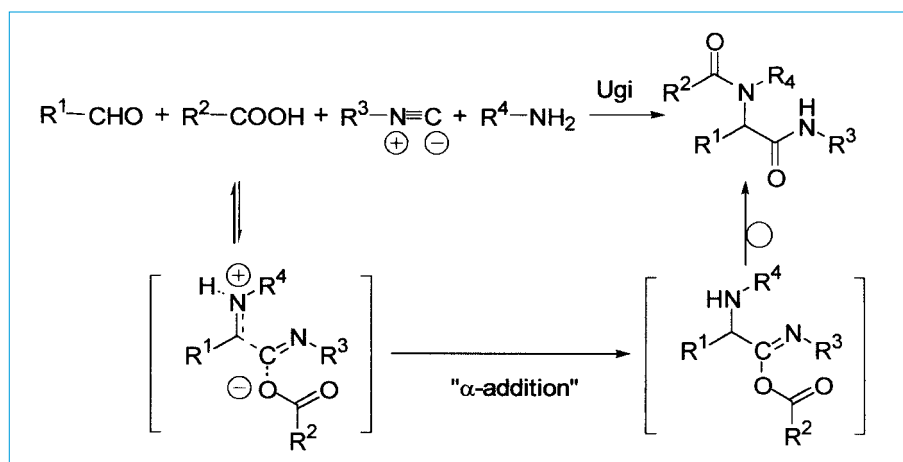


Schéma 3 - La réaction de Ugi.

anion carboxylate). De ce point de vue, la fonction isonitrile peut être apparentée à un analogue stable d'une espèce vinylidène-carbène. Cet événement, multi-composant par nature, est à la base du succès de la réaction de Ugi, ainsi que d'autres réactions basées sur l'utilisation d'espèces isonitriles.

Les adduits de réaction de Ugi sont pseudo-peptidiques par nature et peuvent, de ce point de vue, apparaître comme avoir des propriétés pharmacodynamiques relativement mal adaptées à des applications pharmacologiques. Il est en effet admis que des structures compactes, hétérocycliques notamment, possèdent un meilleur « potentiel pharmaceutique » lié à divers critères : meilleure biodisponibilité, métabolisation plus lente, etc.

Une manière simple d'obtenir des structures hétérocycliques, tout en gardant l'extraordinaire potentiel de cette condensation à quatre composants, est de faire en sorte qu'un des réactifs

contienne deux des quatre fonctionnalités nécessaires à la réaction. Par l'utilisation de céto-carboxylates, ont ainsi été notamment obtenus plusieurs familles de lactames de taille de cycle variable [11]. D'autres lactames ont également été obtenus selon une stratégie synthétique voisine par l'utilisation de  $\beta$ -amino-acides. L'utilisation d' $\alpha$ -amino-acides entraîne, quant à elle, la formation transitoire de *O*-acylimidates activés susceptible de réagir avec le solvant (ici le méthanol). Cette dernière réaction est souvent désignée par la réaction de Ugi à 5 centres et 4 composants (schéma 4) [12].

Si l'on néglige la diversité introduite par le solvant, le nombre de sites variables est réduit dans ce type d'approche de quatre à trois. Néanmoins, l'expérience montre que cette stratégie reste intéressante pour la production de bibliothèques de taille moyenne (5-10 000 composés) pour peu que la classe des réactifs bifonctionnels

contienne suffisamment de représentants aisément disponibles.

Il existe une stratégie alternative pour obtenir des composés cycliques qui consiste à préparer un adduit acyclique de type Ugi possédant la capacité d'être ultérieurement activé afin de subir une réaction de cyclisation. Dans cette approche, c'est la fonction amide secondaire créée au cours de la condensation multi-composant qui est la plus susceptible d'être activée. Pour ce faire, le recours au 1-cyclohexénylisonitrile permet l'obtention d'un énamide qui, après protonation et formation d'un cation transitoire *N*-acylimmonium, élimine une espèce cyclohexyle-immonium et conduit à la formation d'une münchnone intermédiaire. Cette münchnone peut être piégée par un nucléophile interne pour conduire à des espèces aussi intéressantes et variées que des 1,4-benzodiazépine-2,5-diones [13], des céto-pipérazines, dicéto-pipérazines et des lactames [14]. En absence de nucléophile interne, l'addition d'un dipolarophile conduit à la formation de pyrroles polysubstitués possédant un excellent potentiel d'application pharmaceutique ou agrochimique (schéma 5) [15].

De nombreuses variantes sont possibles. Ainsi, en utilisant un isonitrile contenant un fragment  $\beta$ -diméthylaminoacrylate en combinaison avec une imine et de l'acide hydrazoïque, il a été possible d'obtenir, après à une étape d'addition-élimination, une classe de tétrazoles bicycliques d'un grand intérêt pharmacologique [16]. Cette procédure, que l'on peut décrire comme relevant d'une méthode « two-step one-pot », démontre que des classes de molécules élaborées, hétérocycliques et rigides peuvent aisément être obtenues à partir de précurseurs relativement simples et disponibles (schéma 6).

Hormis la réaction de Ugi et celle de Passérini qui lui est apparentée, les remarquables propriétés des isonitriles ont été récemment exploitées dans d'autres transformations. Ainsi, une grande variété d'amino-3-imidazoles fusionnés ont été obtenus par une condensation à trois composants, mettant en jeu une cyclo-addition [4+1] remarquablement efficace (schéma 7) [17]. Cette réaction est en plusieurs points remarquable, et illustre parfaite-

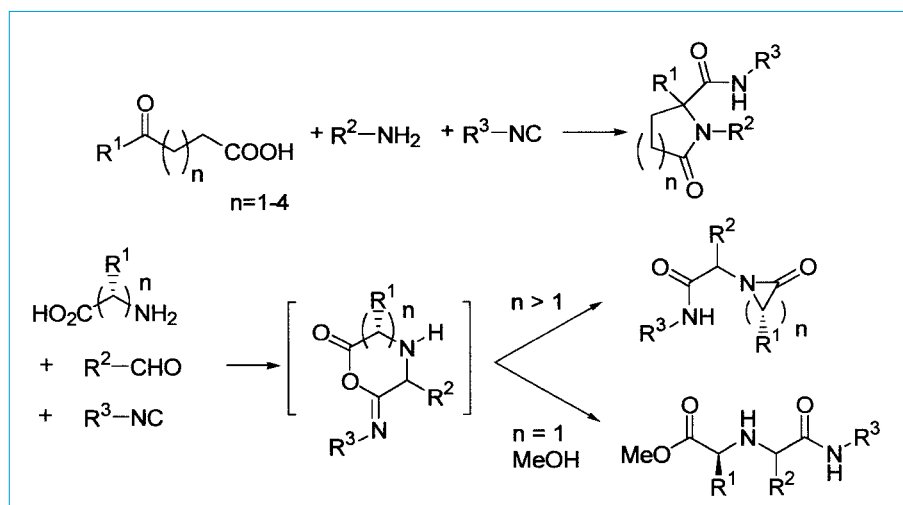


Schéma 4 - Synthèse directe d'hétérocycles par réaction de Ugi.

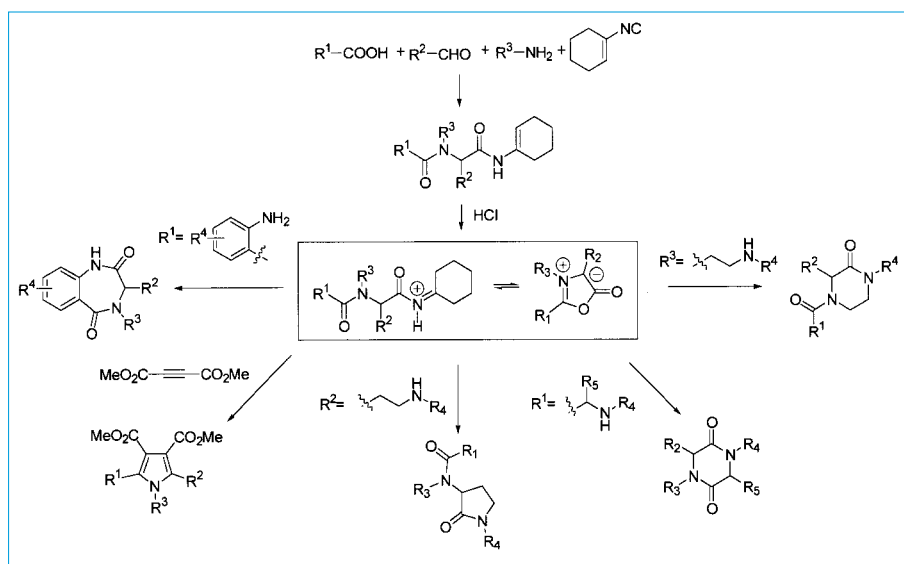


Schéma 5 - Hétérocycles obtenus par des cascades Ugi/cyclisations.

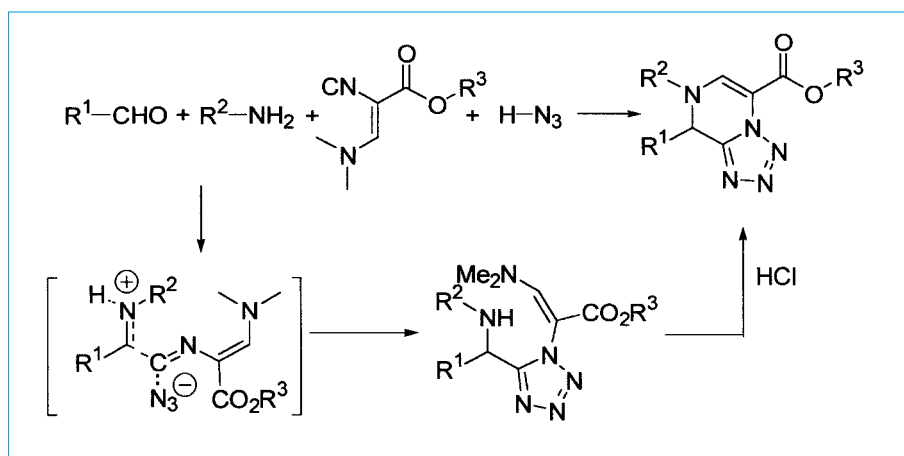


Schéma 6 - Une synthèse de tétrazoles bicycliques.

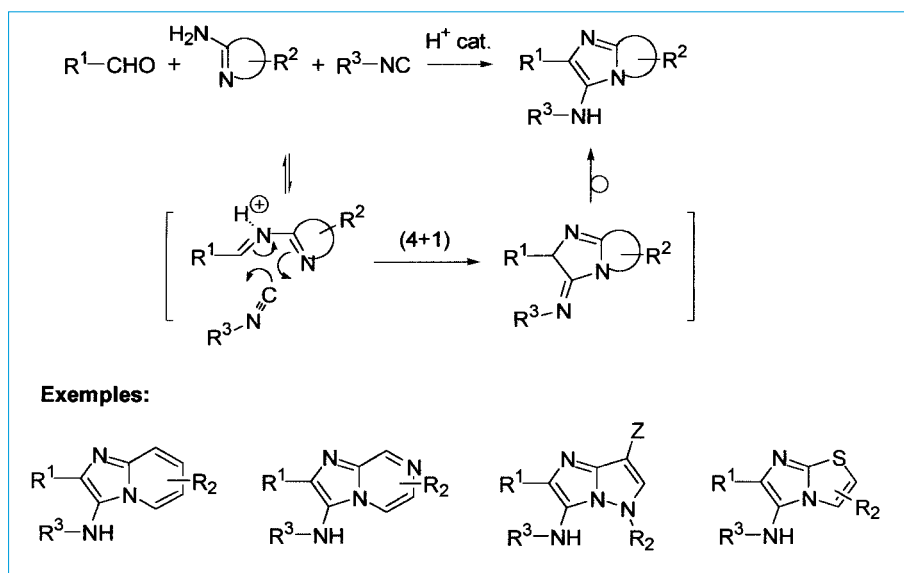


Schéma 7 - Une synthèse d'amino-3-imidazoles à trois composants.

ment le potentiel des condensations à composants multiples.

Premièrement, par sa généralité et sa simplicité de mise en œuvre, plus de 30 000 hétérocycles élaborés ont ainsi été obtenus dans notre laboratoire en un temps réduit. D'autre part, l'invariant structural au sein de la chimiothèque est composé de seulement sept atomes : deux carbones, trois azotes et un hydrogène, constituant le noyau amino-3-imidazole (masse molaire = 67 g mol<sup>-1</sup>). Les représentants de cette chimiothèque sont non seulement structurellement très divers, comme illustré sur le schéma 7, mais leurs propriétés physico-chimiques sont parfaitement optimisables par le choix judicieux des substituants.

Relevant de réactions mécanistique-ment variées, il apparaît, des exemples jusqu'ici présentés, qu'au centre de leur efficacité synthétique se trouve la fonction isonitrile et son profil réactionnel unique. Encore sous-exploitée, cette fonction sera sans doute au cœur de nouveaux développements méthodologiques dans les années à venir.

Au-delà des réactions basées sur la fonction isonitrile, d'autres condensations à composants multiples existent.

Ces transformations mettent souvent en jeu des intermédiaires réactifs obtenus par condensations entre deux précurseurs simples. Par exemple, une imine obtenue par condensation d'une aniline avec un aldéhyde, ou une cétone α,β-insaturée obtenue par réaction de Knoevenagel entre un composé à méthylène activé et un aldéhyde, peuvent être engagées dans des réactions de cycloadditions [4+2], pour fournir des hétérocycles d'intérêt pharmaceutique (schéma 8, eq. 2 et 3) [18].

De même, la cycloaddition dipolaire entre une oléfine activée et un ylure d'azométhyne, est à la base d'une méthode élégante d'accès à des pyrrolidines substituées. La « déconnexion » de l'espèce dipolaire en deux précurseurs stables, un α-aminoacide et un aldéhyde (ou une cétone), permet d'envisager la déclinaison de cette réaction classique en une variante à trois composants. Par cette astuce, 25 600 composés ont été préparés (sur un total potentiellement réalisable de plus de 400 000) à partir d'isatines, d'α-aminoacides et de chalcones commerciaux (schéma 8, eq. 1) [19].

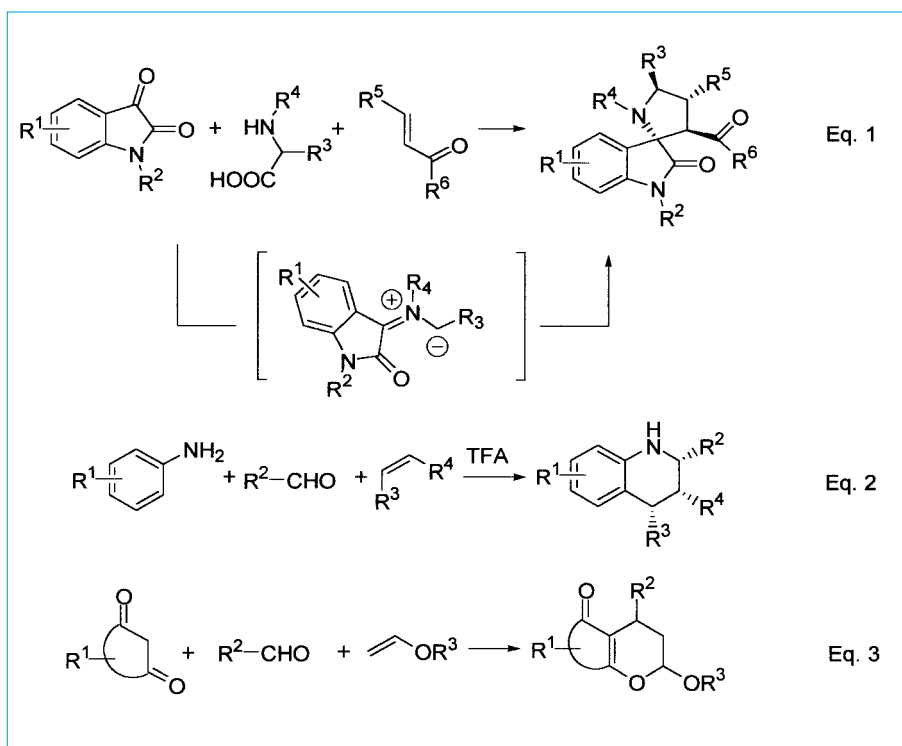


Schéma 8 - Synthèses multi-composants hétérocycliques par cycloadditions.

Encore rares mais certainement promises à un développement important dans l'avenir, plusieurs réactions impliquent la participation d'un catalyseur organométallique. Ainsi, la combinaison en cascade d'une addition de Michael suivie d'une fonctionnalisation de type Heck d'une oléfine a conduit à la préparation de tétrahydrofurannes diversement substitués (schéma 9) [20].

Outre les avantages déjà décrits, les condensations à composants multiples bénéficient de qualités de productivité incomparables. Hyper-convergentes,

répondant aux critères d'économie d'atome et de sélectivité habituellement recherchés dans le cadre de développement industriel, ces réactions sont rapides et se prêtent aisément au changement d'échelle (scale-up). Elles sont souvent robustes, compatibles avec des conditions d'inertage imparfaites et éventuellement transposables en synthèse sur phase solide. Autant de paramètres indispensables à l'usage d'automates de distribution XY lors de la production de chimiothèques (Tomtech™, Rapid Plate™, Hydra96™, Gilson... figure 1).

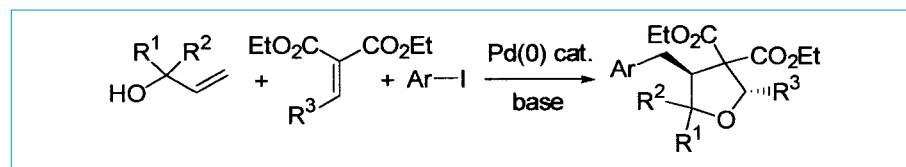


Schéma 9 - Une cascade à trois composants métallo-catalysée.

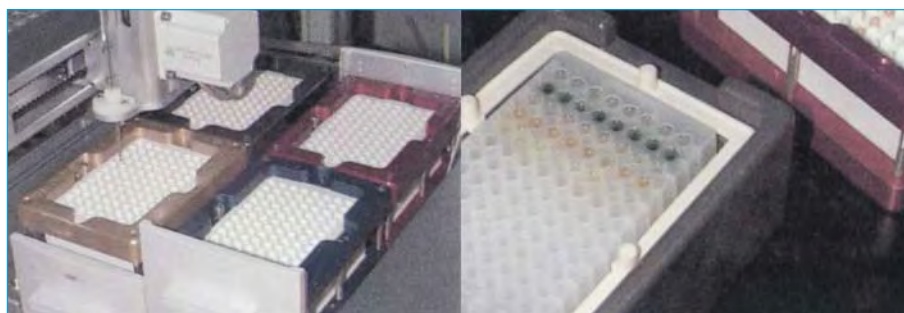


Figure 1 - Appareillage typique utilisé en synthèse parallèle.

De ces considérations particulièrement importantes lorsque des chimiothèques doivent être produites, chacun comprendra que les réactions multi-composants constituent un outil synthétique puissant dont il serait préjudiciable de se passer.

## Conclusion

Quarante ans se sont écoulés depuis la découverte de la réaction de Ugi. Des structures originales pseudo-peptidiques, jusqu'aux hétérocycles à fort « potentiel pharmaceutique » qui forment le noyau dur des molécules d'intérêt thérapeutique et agrochimique, de nombreux progrès ont été accomplis.

Il est frappant de constater qu'une part importante des condensations à composants multiples est basée sur la chimie remarquable des isonitriles. L'exploitation d'autres fonctions chimiques porteuses d'un potentiel similaire, encore à identifier, initierait sans aucun doute un enthousiasme renouvelé pour ce domaine de recherche. L'engouement récent constaté pour ces réactions se comprend aisément au vu de leurs immenses qualités, et il ne fait pas de doute que la recherche de nouvelles condensations à composants multiples constituera un axe important de développement en synthèse combinatoire dans les années à venir [21].

## Références et notes

- [1] a) Wilson S.R., Czarnik A.W., *Combinatorial Chemistry. Synthesis and Applications*, Ed. J. Wiley, New York, **1997** ; b) Balkenhohl F., von den Bussche-Hünnefeld C., Lansky A., Zechel C., *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, p. 2288-2337 ; c) Thompson L.A., Ellman J.A., *Chem. Rev.*, **1996**, *96*, p. 555-600 ; d) Gordon E.M., Barrett R.W., Dower W.J., Fodor S.P.A., Gallop M.A., *J. Med. Chem.*, **1994**, *37*, p. 1233-1251 et 1385-1401.
- [2] Ou potentiel pharmaceutique (« drug likeness »). Il n'existe pas de définition scientifiquement satisfaisante de ce concept. Néanmoins, le lecteur comprendra que sous ce vocable, on exprime le potentiel d'une molécule à être optimisée, non seulement vis-à-vis de son affinité/sélectivité pour une cible biologique donnée, mais aussi vis-à-vis de ses propriétés pharmaco-dynamiques et de toxicité. Voir par exemple : Walters W.P., Ajay Murcko M.A., *Curr. Op. Chem. Biol.*, **1999**, *3*, p. 384-387.

- [3] La tendance actuelle est d'évaluer le plus tôt possible les propriétés de biodisponibilité, métabolisation et de toxicité (ADMET) des molécules actives.
- [4] Il est maintenant reconnu que certaines corrélations statistiques existent entre les propriétés physico-chimiques d'une molécule et ses propriétés pharmacodynamiques. En particulier, il a été montré que des paramètres tels que la masse molaire, la lipophilie (mesurée par le cLogP), le nombre de sites donneurs ou accepteurs de liaisons H, les degrés de liberté conformationnels d'une molécule ont un impact sur sa biodisponibilité orale. Ces corrélations sont connues sous le nom de règle mnémotechnique de Lipinski. Lipinski C.A., Lombardo F., Dominy F.W., Feeney P.J., *Adv. Drug Deliv. Rev.*, **1997**, 23, p. 3-25.
- [5] Outre le fait que les hétérocycles ont « fait leurs preuves » en pharmacie humaine, il est souvent plus facile de comprendre les relations structure/activité, et de modéliser les interactions protéine/ligand avec des molécules à faible degré de liberté conformationnel.
- [6] Quelques revues : a) Dax S.A., McNally J.J., Youngman M.A., *Curr. Med. Chem.*, **1999**, 6, p. 255-270 ; b) Tietze L.F., Lieb M.E., *Curr. Opin. Chem. Biol.*, **1998**, 2, p. 363-381 ; c) Armstrong A.P., Combs A.P., Tempest P.A., Brown S.D., Keating T.A., *Acc. Chem. Res.*, **1996**, 29, p. 123-131 ; d) Posner G.H., *Chem. Rev.*, **1986**, 86, p. 831-844.
- [7] Voir par exemple : a) Trost B.M., *Angew. Chem. Int. Ed. Engl.*, **1995**, 34, p. 259-281 ; b) Trost B.M., *Science*, **1991**, 254, p. 1471-1477.
- [8] Le bénéfice tiré d'une réaction fortement inductrice de complexité structurale est souvent compréhensible intuitivement mais plus difficile à rationaliser. D'autant plus difficile à rationaliser que la notion même de complexité structurale est loin d'être établie de façon non équivoque (voir par exemple : a) Whitlock H.W., *J. Org. Chem.*, **1998**, 63, p. 7982-7989 ; b) Bertz S.H., *J. Am. Chem. Soc.*, **1982**, 104, p. 5801-5803. Dans le contexte de recherche d'activité pharmacologique, il suffit d'accepter qu'à un niveau d'activité et de sélectivité donné pour une molécule active, correspond un degré d'élaboration lui permettant d'optimiser son interaction avec le récepteur visé tout en minimisant les interactions parasites avec d'autres cibles potentielles. La sophistication d'une tâche pharmacologique est donc corrélée à la sophistication de la structure moléculaire du vecteur de cette activité.
- [9] Un « répartiteur moléculaire » possède trois ou quatre fonctionnalités séquentiellement et orthogonalement adressables par des classes de réactifs suffisamment divers (généralement par formation d'une liaison amide). Voir : Bunin S.A., *The Combinatorial Index*, Academic Press, London, **1998**.
- [10] a) Ugi I., Dömling A., Hörl W., *Endeavour*, **1994**, 18, p. 115-122 ; b) Ugi I., *Isonitrile Chemistry*, Academic Press, New York, **1971**.
- [11] a) Hanush-Kompa C., Ugi I., *Tetrahedron Lett.*, **1998**, 39, p. 2725-2728 ; b) Harriman G.C.B., *Tetrahedron Lett.*, **1997**, 38, p. 5591-5594 ; c) Short K.M., Mjalli A.M.M., *Tetrahedron Lett.*, **1997**, 38, p. 359-362.
- [12] a) Ebert B.M., Ugi I., Grosche M., Herdtweck E., Herrmann W.A., *Tetrahedron*, **1998**, 54, p. 11887-11898 ; b) Demharter A., Hörl W., Herdtweck E., Ugi I., *Angew. Chem. Int. Ed. Engl.*, **1996**, 35, p. 173-175.
- [13] a) Hulme C., Peng J., Tang S.-Y., Burns C.J., Morize I., Labaudinière R., *J. Org. Chem.*, **1998**, 63, p. 8021-8023 ; b) Keating T.A., Armstrong R.W., *J. Org. Chem.*, **1996**, 61, p. 8935-8939.
- [14] a) Hulme C., Peng J., Louridas B., Menard P., Krolkowski P., Vasant Kumar N., *Tetrahedron Lett.*, **1998**, 39, p. 8047-8050 ; b) Hulme C., Morissette M., Volz F., Burns C.J., *Tetrahedron Lett.*, **1998**, 39, p. 1113-1116.
- [15] a) Keating T.A., Armstrong R.W., *J. Am. Chem. Soc.*, **1996**, 118, p. 2574-2583 ; b) Strocker A.M., Keating T.A., Tempest P.A., Armstrong R.W., *Tetrahedron Lett.*, **1996**, 37, p. 1149-1152 ; c) Mjalli A.M.M., Sarshar S., Baiga T.J., *Tetrahedron Lett.*, **1996**, 37, p. 2943-2946.
- [16] Bienaymé H., Bouzid K., *Tetrahedron Lett.*, **1998**, 39, p. 2735-2738.
- [17] Bienaymé H., Bouzid K., *Angew. Chem. Int. Ed. Engl.*, **1998**, 37, p. 2234-2237.
- [18] Tietze L.F., *Chem. Rev.*, **1996**, 96, p. 115-136.
- [19] Fokas D., Ryan W.J., Casebier D.S., Coffen D.L., *Tetrahedron Lett.*, **1998**, 39, p. 2235-2238.
- [20] Cavicchioli M., Sixdenier E., Derrey A., Bouyssi D., Balme G., *Tetrahedron Lett.*, **1997**, 38, p. 1763-1766.
- [21] A tel point que le premier congrès international dédié exclusivement aux M-CR se tiendra en octobre 2000 à Munich.

# L'approche combinatoire en catalyse hétérogène : quelles stratégies et perspectives pour la recherche publique ?

**Claude Mirodatos\*** directeur de recherche CNRS

**Summary :** *Combinatorial approaches in heterogeneous catalysis : strategy and perspectives for academic research*

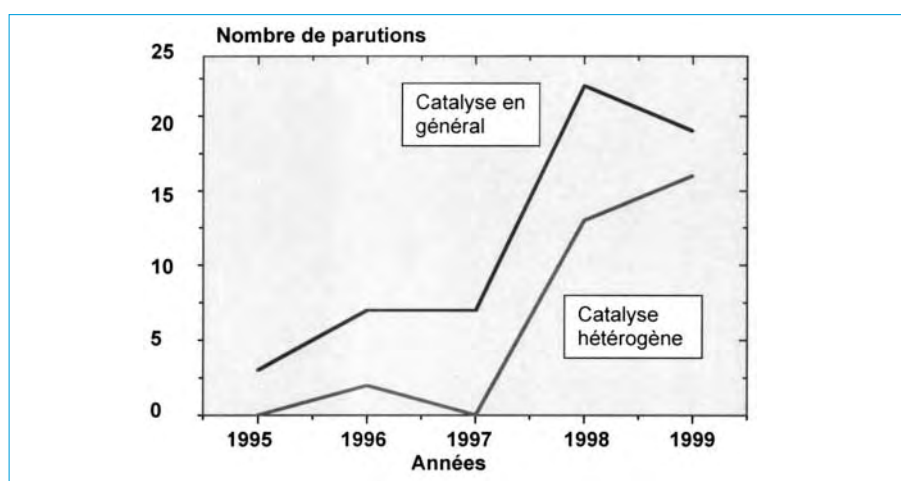
*Applying combinatorial chemistry to heterogeneous catalysis is analysed in term of present strategies and perspectives on the industrial and academic level. Potential methodologies for academic research laboratories are proposed with emphasis on both theoretical and practical specificity for that rapidly developing process.*

**Mots clés :** *Catalyse hétérogène, chimie combinatoire.*

**Key-words :** *Heterogeneous catalysis, combinatorial chemistry.*

La progression rapide du nombre d'articles et revues traitant de la chimie combinatoire appliquée à la catalyse hétérogène (figure 1) atteste d'un phénomène certain : l'intérêt de la communauté scientifique pour cette méthodologie nouvelle visant à la découverte de matériaux susceptibles de surpasser par leurs propriétés les meilleures formulations d'usage industriel. De la fascination, voire crédulité pour les uns, à un fort scepticisme, voire rejet catégorique pour les autres, la très large palette d'avis qui s'exprime aujourd'hui au sein des universités, instituts, centres de recherche publique ou industrielle relève pour large part de l'image très brouillée des stratégies en cours.

Il existe tout d'abord une grande confusion dans les champs traités ou envisagés par l'approche combinatoire. La synthèse de structures organiques nouvelles (CHONX) peut explorer un champ en apparence quasi illimité mais, en fait, borné par nos connaissances en chimie organique de base. En revanche, la synthèse de matériaux inorganiques nouveaux par combinaison des éléments utiles du tableau périodique



**Figure 1** - Évolution du nombre d'articles consacrés à la chimie combinatoire appliquée à la catalyse en général et à la catalyse hétérogène en particulier.

(environ 75) ouvre à la créativité combinatoire un domaine plus vaste encore. Dans ce domaine, il convient de différencier la préparation de matériaux à propriétés physiques déterminées, comme par exemple la supraconductivité [1], de matériaux plus spécifiquement adaptés à la catalyse hétérogène, dont la finalité reste une réactivité chimique ciblée vis-à-vis de mélanges gazeux ou liquides donnés.

Un deuxième élément majeur de confusion découle de l'information opaque que produisent certaines sociétés nouvelles (Symyx, Aventium, HTE...) spécialisées dans la chimie combinatoire. Ces dernières annoncent

ainsi, sans données aujourd'hui vérifiables, des progrès considérables accomplis en quelques années (généralement amalgamés avec les avancées certaines obtenues dans le domaine de la chimie moléculaire pharmaceutique), tout en menant des politiques de protection industrielle agressives (cf. les débats en cours sur la validité des brevets américains Symyx dans l'espace européen).

Sur le front de la recherche publique, des prises de positions individuelles sous forme d'articles généralistes [2-3], l'organisation tout azimut de conférences et d'écoles spécialisées [4] témoignent d'une activité évidente mais

\* Institut de Recherches sur la Catalyse, 2, avenue Albert Einstein, 69626 Villeurbanne Cedex. Tél. : 04.72.44.53.66. Fax : 04.72.44.53.99. E-mail : mirodato@catalyse.univ-lyon1.fr



fortement polymorphe dont il est mal aisé de cerner la réalité des efforts consentis et des objectifs retenus. La divulgation *a minima* d'informations sur des regroupements en réseaux/consortiums nationaux, transnationaux et internationaux (États-Unis, Europe, Japon) n'ajoute en rien à la clarté de l'ensemble.

Enfin, le dernier élément propre à engendrer la confusion relève de la quasi-absence de publications académiques détaillées sur la réalité de la recherche menée en laboratoire (développement et mise en œuvre d'outils de synthèse et d'évaluation rapide des performances, logiciels de gestion des données et de pilotage dans la préparation des générations de catalyseurs, application concrète à des réactions cibles en catalyse hétérogène). Ce dernier point permet simplement de constater que cette recherche est encore balbutiante, ne serait-ce que par l'ampleur des ouvertures possibles pour appliquer cette démarche.

### Analyse de la démarche combinatoire en catalyse hétérogène

Sans prétendre à une quelconque analyse exhaustive du phénomène (aujourd'hui impossible du fait du médiocre niveau de communication évoqué précédemment), il est quelques éléments d'analyse qui peuvent permettre d'en cerner de plus près les contours et peut-être d'éviter tout excès de critique ou d'espoir irraisonné. Pour ce faire, nous limiterons cet article à l'analyse des questions que peut légitimement se poser un laboratoire de recherche publique en catalyse hétérogène « pour y aller ou pas », sur la base de notre propre expérience en la matière [5].

La démarche combinatoire peut, en première approximation, être envisagée selon deux méthodologies complémentaires :

- **l'investigation primaire** (primary screening) qui revendique la découverte de nouveaux matériaux *ab initio*, par l'exploration systématique de combinaisons aléatoires d'éléments constitutifs de matériaux multi-composants. Cette progression « empirique et pragmatique »

relève généralement de la recherche industrielle. Ainsi, la société Symyx [6] annonce-t-elle la redécouverte en quelques mois de la meilleure formulation pour la réaction de déshydrogénation oxydante de l'éthane en éthylène mise au point après plusieurs années de recherche [7] : un catalyseur à base de Mo, V, Nb, O. Cette investigation primaire aurait ensuite conduit à des nouveaux matériaux quadruplant le rendement en éthylène, c'est-à-dire susceptibles de directement concurrencer les procédés industriels de vapocraquage des naphas. Une telle démarche nécessite objectivement des moyens techniques considérables, avec des bancs robotisés de préparation des catalyseurs, couplés à des systèmes de contrôle des performances catalytiques susceptibles de tester des quantités astronomiques de formulations en des temps records (plus de 10 000 par mois). Notons que la préparation des solides dans ce cas s'apparente plus à la science des surfaces (dépôt de multicouches par vapodéposition physique ou chimique) qu'aux techniques chimiques conventionnelles (imprégnation, coprécipitation ou synthèse sol-gel).

- **l'investigation secondaire** (secondary screening) qui cible l'optimisation de formulations déjà existantes, sans toutefois exclure la découverte, en passant, de matériaux imprévus. La principale caractéristique de cette seconde démarche est d'associer *un potentiel de préparation/évaluation de matériaux à débit élevé* (au moins plusieurs dizaines de formulations par jour) et *dans des conditions aussi réalistes que possible* (type microréacteurs en parallèle) à *la démarche rationnelle et cognitive* qui oriente l'investigation sur des voies prédéterminées. La référence dans ce domaine de l'approche intelligente reste la science des surfaces qui explore les propriétés intrinsèques fondamentales de solides modèles (monocristaux, alliages homogènes et structurés) par des spectroscopies aptes à visualiser le processus catalytique à l'échelle moléculaire. De même, la microcinétique, par modélisation et couplage avec des techniques spectroscopiques *in situ*, analyse les étapes élémentaires de la réaction et notamment celles qui *déterminent* le déroulement global de l'acte catalytique.

Cette classification en investigation primaire et secondaire présente évidemment un caractère arbitraire, non nécessairement adapté à chaque type de catalyse. C'est cependant la seconde voie qui associe approche fondamentale et exploration programmée de domaines insuffisamment ou jamais étudiés qui nous semble correspondre au mieux aux possibilités techniques d'un laboratoire de recherche publique à vocation fondamentale (CNRS, université).

Quelles en sont les principales caractéristiques ?

### Stratégies et technologies

La clé de l'approche combinatoire en catalyse hétérogène consiste à disposer d'un ensemble cohérent qui associe :

- une définition des objectifs chimiques à atteindre (par exemple un rendement en produit ciblé ou une sélectivité ou une résistance à l'empoisonnement par le coke, etc.) et des éléments chimiques de base constitutifs des catalyseurs envisagés (nature et nombre de composants élémentaires ou ensembles de base, mode de préparation et d'activation...). Cette définition initiale souligne d'emblée la nécessité d'une prédétermination implicite des domaines à explorer (fondée sur l'expérience et la littérature, ce qui n'exclut pas d'y intégrer des éléments marginaux jamais ou rarement envisagés) :

- un système de préparation des solides catalytiques de composition variable en vue de constituer des « bibliothèques » de catalyseurs ;

- un système d'évaluation rapide des performances ou propriétés ciblées des solides préparés en vue d'une sélection selon un ou des critères (ou fonctions d'objectifs) qui renvoient directement ou indirectement aux objectifs chimiques prédéterminés ;

- une gestion/programmation des données qui permette à partir de la sélection des catalyseurs d'une génération donnée de préparer la génération suivante ou d'évoluer vers une deuxième boucle de confirmation ou rejet des formulations sélectionnées après synthèse en plus grande quantité et test selon des conditions proches des conditions industrielles.

La figure 2 expose le principe de ces boucles d'optimisation adaptées aux spécificités de la catalyse hétérogène. Explicitons-en les étapes clé.

## Préparation de « bibliothèques » de catalyseurs

### Au plan technique

La préparation d'un grand nombre d'échantillons de composition différente impose à terme une automatisation complète. Il existe aujourd'hui une offre de robots de synthèse de catalyseurs susceptibles de préparer de manière précise et reproductible des bibliothèques d'échantillons (quelques mg par échantillon). Ces robots doivent idéalement posséder un système de prélèvement et d'injection automatisé x,y,z de liquides et de solides pour réaliser des collections d'échantillons (jusqu'à une centaine de compositions par série), avec des fonctions programmables d'agitation, filtration, chauffage, et manipulation des phases solides, éventuellement sous atmosphère contrôlée. Ces fonctions correspondent en fait aux différents actes effectués manuellement lors de la préparation de solides par méthodes conventionnelles (coprécipitation, coimpregnation, sol-gel). Un premier écueil sera de constater que les offres commerciales du moment ne présentent pas toutes ces fonctions simultanément et qu'une adaptation au niveau du laboratoire est requise. Un second écueil sera leur coût élevé lié à leur caractère prototypique. L'ouverture du marché et la mise en concurrence entre de nouvelles PME spécialisées, voire une commercialisation par les laboratoires publics eux-mêmes, devraient dans les années à venir rendre plus accessibles ces outils. Notons cependant que la plupart des laboratoires précurseurs en la matière ont débuté avec des micro-pipettes manuelles et des batteries des tubes à essai bon marché.

### Au plan conceptuel

Il ne s'agit pas de préparer « n'importe quoi » dans cette démarche d'investigation secondaire, mais d'avoir pour cible des matériaux qui présentent une concentration maximisée de sites actifs et sélectifs. Illustrons

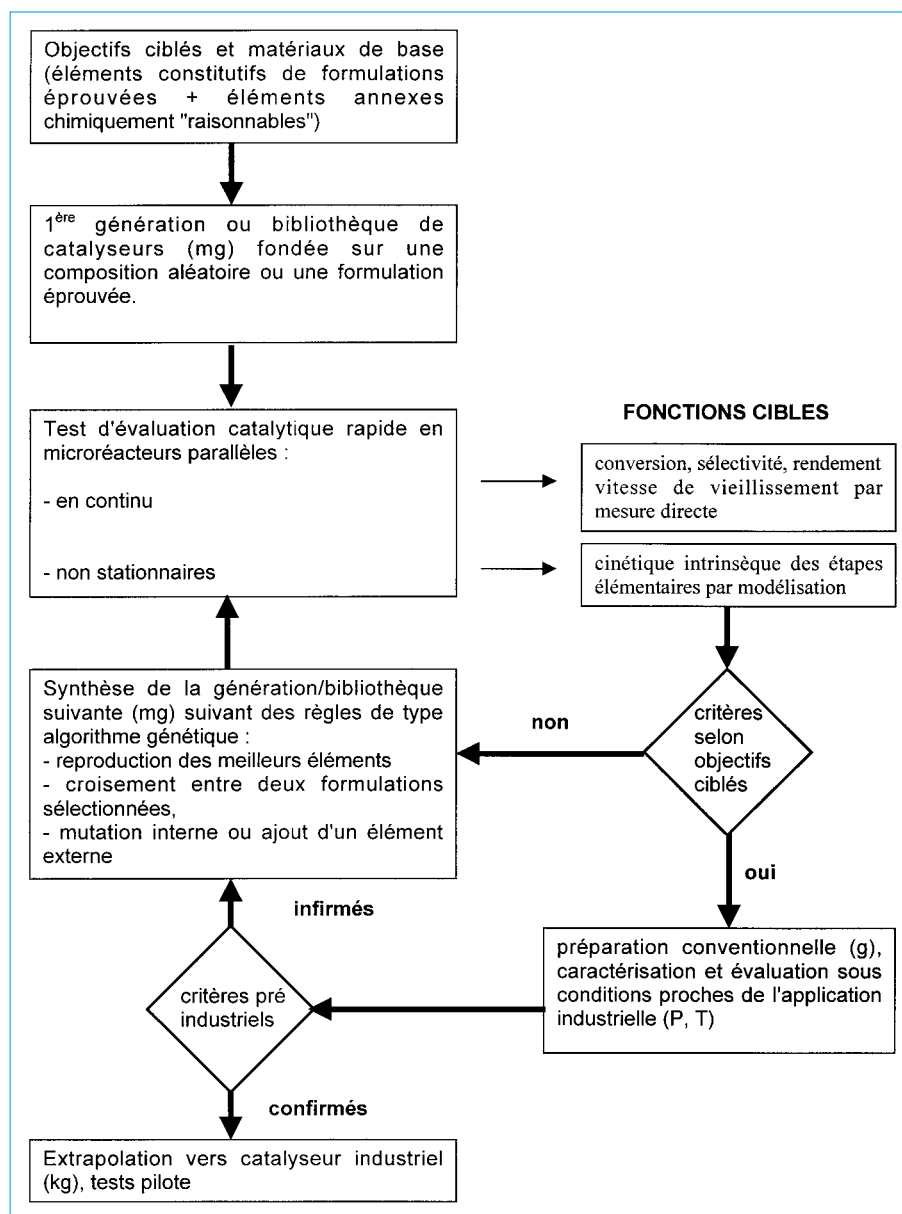


Figure 2 - Boucles d'optimisation de catalyseurs hétérogènes à l'échelle du laboratoire.

d'un exemple la complexité de cet objectif.

Dans le domaine très sensible des réactions de déshydrogénation oxydante (ODH) d'alcane légers (production d'alcènes dont la demande mondiale ne cesse d'augmenter), une idée de base largement développée dans la littérature associe la sélectivité catalytique à un type de structure bien définie. Ainsi la phase orthovanadate d'un catalyseur VMgO serait plus sélective que des phases pyro- ou ortho-, bien que des propositions contraires aient également été formulées [8]. La réalité d'une telle loi chimique limiterait sans conteste toute démarche combinatoire rationnelle, vu le nombre limité de structures cristallines existantes pour un

ensemble ternaire V-Mg-O. Cependant, de nombreuses évidences tendent à démontrer que c'est plutôt la structure superficielle des phases supportées (qui peut être totalement différente des structures de cœur) qui « règle » les performances ODH. Ainsi, en optimisant la teneur en vanadium, nous avons démontré que les meilleurs rendements en alcène C<sub>3</sub> ou C<sub>4</sub> correspondaient à de très faibles teneurs pour lesquelles il n'existe plus de structures définies mais seulement une forte concentration de tétraèdres distordus VO<sub>4</sub><sup>3-</sup> (avec l'évolution rapide d'une structure amorphe dispersée vers une structure cristallisée de type spinelle selon les conditions réactionnelles [9-10]). De même, une étude fondamentale d'oxydes de type

Mo-V-Nb pour l'ODH de l'éthane a montré que les meilleures formulations (déduites d'une approche précombinatoire sur une dizaine de solides) sont multiphasiques avec un rôle important des phases amorphes aisément réductibles [11]. Un tel constat élargit considérablement les perspectives combinatoires en matière d'ODH puisqu'il ne s'agit plus de découvrir de nouvelles structures cristallines massives, mais de combiner des éléments en vue de créer des structures superficielles pseudo cristallines, denses en lacunes et oxygènes de réseau mobiles et sélectifs.

### Comment constituer une bibliothèque de catalyseurs par combinaison d'éléments ?

Cette étape cruciale de la démarche combinatoire se réfère généralement aux règles de sélection et d'évolution que l'on observe dans la nature sous le terme générique d'algorithmes génétiques [12]. Une excellente adaptation de ces règles à la catalyse hétérogène, qui combine statistique aléatoire et organisation rationnelle, est proposée par Wolf *et al* [13].

Une première génération est tout d'abord réalisée par combinaison aléatoire d'éléments de base choisis pour leur fonctionnalité propre, jugée nécessaire pour un catalyseur performant. Ainsi, dans le cas de l'ODH, les éléments de base seront des oxydes primaires présentant des fonctions diverses d'acidité, de basicité, de fort potentiel redox ou d'aptitude à facilement dissocier l'oxygène, comme des oxydes de type,  $V_2O_5$ ,  $MoO_3$ ,  $Fe_2O_3$ ,  $GaO$ ,  $MgO$ ,  $La_2O_3$ . Chaque combinaison/formulation est codée, en incluant la concentration de chaque élément et des données complémentaires sur le mode de préparation ou d'activation (figure 3).

Puis, après évaluation des performances de ces différentes combinaisons initiales (nous reviendrons sur le mode d'évaluation), les trois principaux opérateurs mathématiques qui permettent la création évolutive d'une nouvelle génération de solides sont appliqués :

- la sélection (tri des « bonnes » et « mauvaises » formulations par rapport à un critère ou une fonction d'objectif),

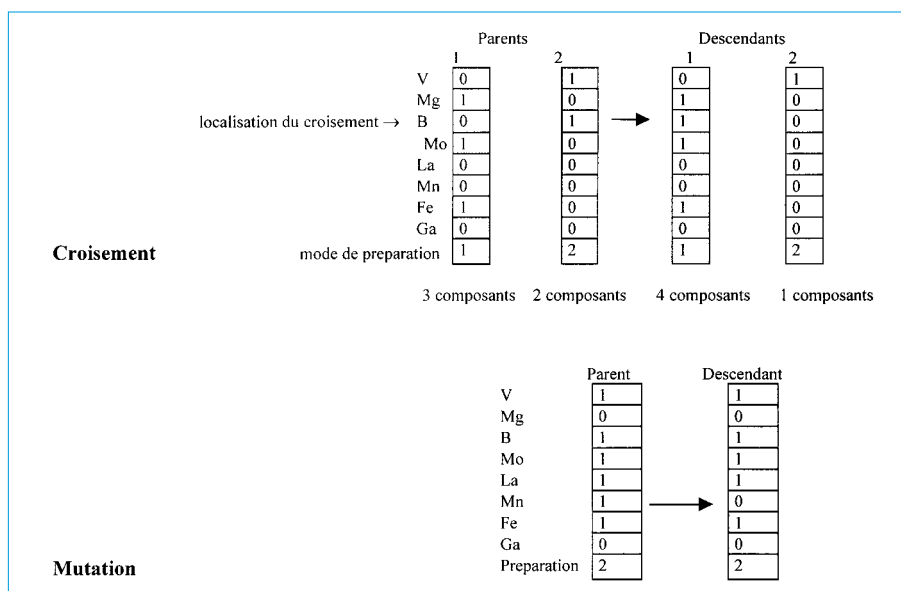


Figure 3 - Représentation de formulations en catalyse de déshydrogénation oxydante par codage binaire et schéma des deux opérateurs de croisement et de mutation qui permettent le passage d'une génération de catalyseurs à une autre (selon [13]).

- le croisement (échange d'éléments constitutifs entre deux formulations sélectionnées),

- la mutation (changement de concentration d'éléments constitutifs ou dans l'ordre d'addition pour précipitation séquentielle ; ajout d'un nouvel élément, par exemple en ODH, un nouvel oxyde sans fonctionnalité *a priori* importante pour la chimie considérée).

Ainsi, de génération en génération, une évolution rapide vers des formulations optimisées est théoriquement envisageable. Soulignons cependant **l'importance primordiale des choix initiaux du chimiste** dans la constitution des éléments de base qui servent à la combinaison des formulations complètes. Ces choix sont aussi dictés par la possibilité ou non de combiner physiquement et chimiquement ces éléments (le cas favorable de la combinaison d'oxydes en ODH ne l'est pas forcément pour la combinaison de différents métaux, préparés par réduction de sels précurseurs, CVD, plasma, etc.). On voit là toute la complexité de la démarche et surtout l'évidence que des démarches combinatoires pourront rapidement aboutir dans certains cas mais également rapidement faillir dans d'autres.

### Évaluation et sélection des catalyseurs

De nombreux systèmes d'évaluation rapide des catalyseurs ont déjà été pro-

posés dans la littérature, reflet de l'avancement des techniques et du type d'information recherché. Ainsi, la thermographie infrarouge sur plaques d'échantillons [14-15] conduit à une évaluation de la seule activité thermique d'un solide. Le développement de réacteurs parallèles sous flux continu, couplés à des systèmes d'analyse rapide (comme la microchromatographie en phase gazeuse ou la spectrométrie de masse) donne accès aux données cinétiques globales d'une réactivité (conversion, sélectivité, vieillissement). Enfin, la cinétique non stationnaire [16-17] permet, grâce à la modélisation des signaux transitoires, un accès rapide aux paramètres cinétiques intrinsèques relatifs aux étapes élémentaires de la réaction (figure 2). D'autres modes d'évaluation, mettant en œuvre par exemple des mesures physico-chimiques *in situ* (comme la conductivité électrique), sont techniquement envisageables à l'échelle du laboratoire. Nous privilégierons *les modes d'évaluation les plus informatifs* car ils permettent des critères de sélection directement fondés sur des propriétés essentielles pour l'acte catalytique. Là encore, la connaissance fondamentale de la réaction envisagée et notamment de sa cinétique intrinsèque reste déterminante pour l'efficacité de la méthodologie [18].

**Concrètement**, la technique d'évaluation doit évidemment présenter un degré

d'automatisation compatible avec le débit de rotation souhaité dans la première boucle d'optimisation représentée dans la *figure 2*. La sélection se fera par analyse de convergence (type moindres carrés) entre les données recueillies par le test d'évaluation et les critères préétablis (fonction d'objectifs = rendement, vitesse de désactivation, chaleur d'adsorption d'un réactif ou de désorption d'un produit cible, coefficient de collage, vitesse de diffusion de l'oxygène de réseau, constante redox, etc.).

Notons que la commercialisation de ces techniques d'évaluation (microréacteurs en parallèle, DRX en balayage, thermographie infrarouge, etc.) reste encore extrêmement confidentielle par comparaison avec les techniques de préparation, mais devrait rapidement envahir les sites « .com » du marché mondial.

### Extrapolation vers des matériaux adaptés aux exigences industrielles

Une critique majeure de l'approche combinatoire est que celle-ci ne conduirait qu'à des formulations pas ou peu reproductibles et non extrapolables à de réelles formulations industrielles (à noter que l'on retrouve là une antienne sur la préparation des seuls catalyseurs modèles dans les laboratoires universitaires). La mise en place d'une deuxième boucle d'optimisation (à rotation plus lente et sans processus automatisé), basée sur la préparation (quantité de l'ordre du g ou kg), la caractérisation et l'évaluation conventionnelle des meilleures sélections de la première boucle nous semble pouvoir répondre à cette interrogation (*cf. figure 2*).

## Conclusions

Au-delà d'une certaine confusion en terme de concepts et stratégies, l'approche combinatoire nous apparaît comme une réelle opportunité à saisir dans cette phase initiale d'ouverture au vaste champ de la catalyse hétérogène, et ce notamment en matière de recherche publique. Les outils de base de cette recherche, tant techniques que théoriques, existent et des avancées sans doute très inégales selon les chimistes abordées peuvent être raisonnablement envisagées à court terme. Quant à la « rentabilité » attendue d'investissements humains ou matériels non négligeables, il convient de souligner que cette approche qui conjugue *découverte et compréhension* s'inscrit au cœur même des objectifs de notre métier. De ce fait, elle ne peut que renforcer la créativité de nos laboratoires.

## Références et note

- [1] Xiang X.D., Sun X., Briceno G., Lou Y., Wang K.A., Chang H., Wallace-Freedman W.G., Chen S.W., Schultz P.G., *Science*, **1995**, 268, p. 1738-1740.
- [2] Schlögl R., *Angew. Chem. Int. Ed.*, **1998**, 37, p. 2333.
- [3] Maier W.F., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 1216.
- [4] Nato Advanced Study Institute : Combinatorial Catalysis and High Throughput Catalyst Design and Testing, Vilamora, Portugal, July 11-24, **1999**.
- [5] COMBI 2000 : Combinatorial Approaches for New Materials Discovery, San Diego, California, USA, January 23-25, **2000**
- [6] COMBI Europe, Combinatorial Approaches and High Throughput Screening for New Materials Discovery, Frankfurt am Main, Germany, June 28-July 1, 2000.
- [5] C. Mirodatos est coordinateur scientifique d'un consortium européen « COMBI-CAT », 2000-2003, comprenant les partenaires suivants : IRC-CNRS (France), Institute for Applied Chemistry, Berlin-Adlershof (ACA) (Germany), Instituto de Tecnologia Química (Spain), Institut Français du Pétrole (France), Energy Research Foundation and Environment (The Netherlands), DSM Research B.V. (The Netherlands), Engelhard De Meern B.V. (The Netherlands), ENITECNOLOGIE (Italy), Cambridge Discovery Chemistry (UK), AMTEC GmbH (Germany).
- [6] Cong P., Dehestani A., Doolen R., Giaquinta D.M., Guan S., Markov V., Poojary D., Self K., Turner H., Weinberg W., *Appl. Phys. Sci.*, **1999**, 96, p. 11077.
- [7] Thorsteinson E.M., Wilson T.P., Young F.G., Kasai P.H., *J. Catal.*, **1978**, 52, p. 116.
- [8] Cavani F., Trifiro F., *Catal. Today*, **1995**, 24, p. 307 ; Mamedov E.A., Cortès Corberan V., *Appl. Catal. A : General*, **1995**, 127, p. 1.
- [9] Pantazidis A., Mirodatos C., *Heterogeneous Hydrocarbon Oxidation*, Warren BK and Oyama ST Ed, ACS Symposium Series, Washington, **1996**, 638, p. 207-222.
- [10] Pantazidis A., Burrows A., Kiely C.J., Mirodatos C., *J. Catal.*, **1998**, 177, p. 325-334.
- [11] Ruth K., Kieffer R., Burch R., *J. Catal.*, **1998**, 175, p. 16 et 27.
- [12] Goldberg D.E., *Genetic Algorithms*, Addison-Wesley, **1989**.
- [13] Wolf D., Buyevskaya O.V., Baerns M., *Appl. Catal.*, à paraître.
- [14] Holzwarth A., Schmidt H.W., Maier W.F., *Angew. Chem. Int. Ed.*, **1998**, 37(19), p. 2644.
- [15] Reetz M.T., Becker M.H., Kühling K.M., Holzwarth A., *Angew. Chem. Int. Ed.*, **1998**, 37(19), p. 2647.
- [16] Mirodatos C., *Catal. Today*, **1991**, 9, p. 83-95.
- [17] Schuurman Y., Décamp T., Jalibert J.C., Mirodatos C., « Reaction Kinetics and the Development of Catalytic Processes », G.F. Froment and K.C. Waugh, Eds, *Studies in Surface Science and Catalysis*, Elsevier, Amsterdam, **1999**, 122, p. 133-140.
- [18] Boudard M., *Catal. Lett.*, **2000**, 65, p. 3.

# Catalyse combinatoire

## Nouveau concept et nouvelles techniques

**Olivier Lavastre\*** chargé de recherche

**Summary :** *New concept and new methods in combinatorial catalysis*

*Recent progress in medicinal chemistry for the rapid identification of active substances in drug design has increased the demand for new molecules. Several methods of combinatorial synthesis connected to new analytical techniques allow synthesis and high throughput screening of huge number of molecules. Nowadays, this concept is almost applied exclusively for drug discovery. Pioneering researchers are showing that combinatorial chemistry can be applied for other research areas such as materials or catalysis.*

*The **combinatorial catalysis** is a new and powerful method to generate and evaluate a large number of catalysts in very short delay.*

**Mot clés :** *Diversité de catalyseurs, chimie combinatoire, évaluation à haut débit.*

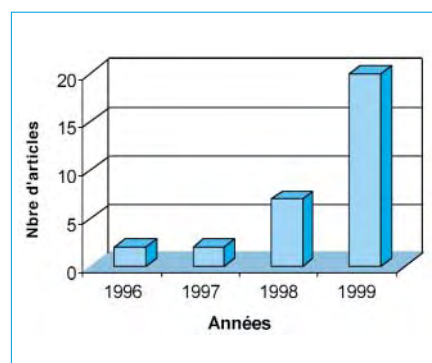
**Key-words :** *Catalysts diversity, combinatorial chemistry, high throughput screening.*

La chimie combinatoire est maintenant largement utilisée en recherche pharmaceutique [1]. En effet, de récentes études [2] ont démontré que l'approche combinatoire permet un gain **facteur temps** de 1 000 et un gain **facteur coût de production** de 600 par rapport à l'approche traditionnelle visant à isoler et purifier chaque substance avant d'en tester, dans une première évaluation, le potentiel thérapeutique. Dans le sillage de ce développement révolutionnaire en chimie pharmaceutique, la chimie combinatoire est maintenant appliquée à d'autres domaines, comme la mise au point de nouveaux matériaux [3] ou catalyseurs [3-4].

La catalyse combinatoire représente une approche totalement innovante pour la détection et la mise au point de nouveaux systèmes catalytiques dans des délais ultra-rapides.

L'objectif de cet article n'est pas tant de faire une liste exhaustive des différents résultats en catalyse combinatoire, mais plutôt d'attirer l'attention sur le formidable impact de cette technique pour la recherche en catalyse.

Bien qu'il s'agisse d'une approche en pleine émergence, il est important de parler d'impact, plutôt que de potentiel car les résultats sont déjà là. Plusieurs banques de catalyseurs ont déjà été élaborées [5] ; de nombreuses réactions catalytiques ont été testées en allant de l'homogène [6] à l'hétérogène [7], incluant des réactions de polymérisation [8] ou des réactions énantiosélectives [9-11]. Des techniques d'analyse à haut débit, adaptées à la catalyse, ont été développées. L'analyse du *graphe 1* démontre clairement que nous sommes au point de départ d'un phénomène en croissance exponentielle. De 4 articles entre 1996 et 1997, la catalyse combinatoire a fait l'objet au minimum de 20 articles en 1999, sans inclure les brevets.



**Grappe 1** - Croissance du nombre d'articles de catalyse combinatoire.

### Présentation du concept et des techniques de catalyse combinatoire

Les réactions organiques catalysées par des complexes de métaux peuvent être décrites comme des problèmes à multi-composants : type de réactions, nature du métal et des ligands associés, solvant, température et durée de réaction. Par l'approche traditionnelle lorsqu'un catalyseur a été imaginé, il est synthétisé, purifié, caractérisé puis testé. Puis, selon le résultat, le catalyseur est légèrement ou complètement modifié. Enfin, pour le meilleur catalyseur, les conditions expérimentales sont alors optimisées par de fines variations du solvant, de la température et du temps de réaction. Si **restrictivement** seulement trois précurseurs organométalliques, cinq ligands, trois solvants, deux températures et deux durées de réaction sont sélectionnés pour une réaction catalytique donnée,  $3 \times 5 \times 3 \times 2 \times 2 = 180$  différentes combinaisons devraient être testées. L'approche « une réaction catalytique après une réaction catalytique » n'est pas adaptée pour évaluer toutes ces possibilités. Aussi, très souvent, le chimiste doit restreindre le nombre de combinaisons en diminuant le nombre de paramètres en

\* UMR 6509 CNRS, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex.  
Tél. : 02.99.28.16.35. Fax : 02.99.28.69.39.  
E-mail : lavastre@univ-rennes1.fr

étudiant par exemple l'association d'un seul métal à deux ou trois ligands pour un solvant donné.

La chimie combinatoire peut permettre d'éviter cette restriction :

- 1) par la diversité de catalyseurs qui peuvent être créés,
- 2) par la détection rapide des catalyseurs les plus actifs,
- 3) pour l'optimisation d'une réaction catalytique en variant les conditions expérimentales.

La catalyse combinatoire repose à la fois sur une nouvelle conception de la recherche en catalyse et sur de nouveaux « outils » de synthèse et de caractérisation.

### Nouveau concept

La stratégie catalyse combinatoire consiste à créer et évaluer simultanément un grand nombre de catalyseurs. Le fait d'utiliser des banques de catalyseurs est en effet fondamentalement différent de l'approche traditionnelle en catalyse qui vise la synthèse, la caractérisation et l'évaluation d'un catalyseur après l'autre.

### Outils

Pour créer et évaluer cette diversité catalytique, des techniques de synthèse et d'analyse à haut débit sont indispensables.

### Techniques pour créer de la diversité

Les méthodes de synthèse, appliquées en chimie combinatoire pour la recherche pharmaceutique, sont directement applicables pour générer très rapidement de nombreux catalyseurs.

- Une première technique est basée sur l'utilisation de billes de polymères fonctionnalisées (technique Split and Pool). Ces billes, de taille homogène et de concentration contrôlée en fonctions réactives, sont commercialisées par de nombreuses compagnies (ex : billes de polystyrène, diamètre 300 µm, teneur en fonction amino -NH<sub>2</sub> : 1 mmol/g billes). A la surface de ces billes, via leurs fonctions réactives, un ligand peut être construit en plusieurs étapes par une succession de réactions (chimie organique supportée COS) (schéma 1).

La diversité provient du nombre d'étapes et du nombre de variables pour



Schéma 1 - Synthèse multi-étapes de molécules triblocs ABC.

chacune de ces étapes. Par exemple, à partir de 10 synthons A<sub>1-10</sub>, 10 synthons B<sub>1-10</sub> et 5 synthons C<sub>1-5</sub>, une banque de 500 ligands constitués de trois blocs A-B-C peut être créée en seulement trois étapes (fixations successives de A, B puis C à la surface des billes). La technique Split and Pool offre de plus l'avantage de ne générer qu'un seul type de molécules à la surface d'une bille donnée ; par exemple, une bille supportera le ligand A<sub>3</sub>-B<sub>5</sub>-C<sub>2</sub> et aucune autre combinaison. A chaque étape de synthèse, il y a une opération de marquage (fixation par liaison covalente d'une molécule précise à hauteur de 1 % par rapport aux fonctions réactives ; un exemple de marqueur est reporté dans le schéma 2).

Lorsqu'une bille est repérée comme très active en catalyse, il suffit de libérer et de lire les marqueurs correspondants. Ceci permet de déterminer la nature du ligand qui, associé au précurseur organométallique étudié, présente l'activité catalytique maximale. Une banque de ligands se présente donc sous la forme d'une multitude de billes, chacune correspondant à un seul ligand (photo 1). Un gramme de billes peut contenir 10 000 billes. Cette technique



Photo 1 - Banques de catalyseurs sous forme de billes.

a démontré toute sa puissance aussi bien pour la détection de nouveaux catalyseurs que pour l'optimisation d'un ligand donné [5b, 12].

- A cette première technique, il faut ajouter la **synthèse en parallèle** pour laquelle la nature de la combinaison étudiée va dépendre de sa position dans un repère orthonormé direct XY. Cette approche est tout à fait applicable à la catalyse homogène. Une plaque présentant 96 puits (photo 2) permettra par exemple de combiner en solution 12 précurseurs métalliques (1 par colonne) à 8 ligands (1 par rangée) pour générer potentiellement 96 combinaisons métal/ligand différentes. L'utilisation d'automates de synthèse ouvre des perspectives vers des préparations à grande échelle : nombre d'exemplaires d'une même banque ou nombre de banques différentes.



Photo 2 - Répartition de réactifs dans une plaque 96 puits.

### Évaluation d'activité catalytique à haut débit

De nombreuses techniques ont été élaborées pour comparer l'activité catalytique d'une banque de catalyseurs dans des délais ultra-rapides.

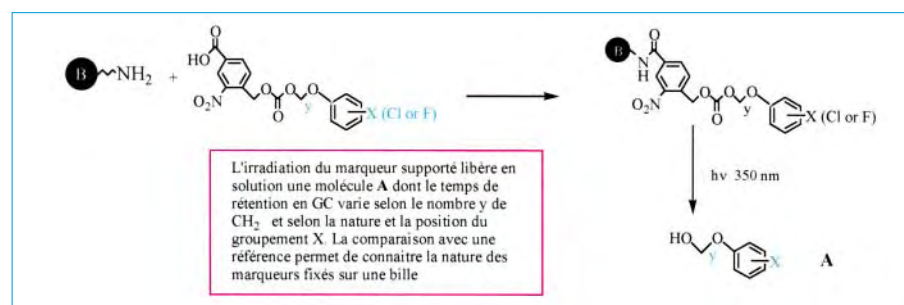


Schéma 2 - Exemple de marqueur avec détection par chromatographie en phase gazeuse.

Les techniques classiques de chromatographie (GC [3], HPLC [6]) ou de spectrométrie de masse [3, 13-14] ont été modifiées par l'adjonction d'analyseurs automatiques d'échantillons (microprélèvement, injection et analyse automatiques, gestion informatisée des résultats). Des méthodes permettant de visualiser instantanément l'activité de l'ensemble d'une banque de catalyseurs ont été utilisées comme techniques ultra-rapides d'évaluation. Par exemple, la thermographie infrarouge, capable de mesurer des variations de température de l'ordre de 0,02 °C, s'est avérée particulièrement performante pour détecter tout dégagement de chaleur associée à des réactions catalytiques hétérogènes [15], homogènes [5b] ou énanti-sélectives [16].

Cependant, la catalyse combinatoire n'est pas obligatoirement dépendante d'une instrumentation complexe et onéreuse. La « boîte à outils » du chimiste peut s'avérer particulièrement performante en catalyse combinatoire. Par exemple, l'association d'un changement de couleur à la présence d'une activité catalytique représente un test visuel très simple, d'un coût sans commune mesure avec les différents équipements cités ci-dessus, pour détecter tout catalyseur actif [4a]. Par exemple, sur la *photo 3*, les catalyseurs actifs pour la réaction d'alkylation allylique sont facilement repérés sous la forme de spots de couleur rouge.

## Résultats et perspectives

De nombreuses réactions catalytiques ont déjà été évaluées par l'approche catalyse combinatoire. En catalyse hétérogène, l'hydrogénation de l'hexyne [15], l'oxydation du monoxyde de carbone [7a] et la réduction du monoxyde d'azote [13], la déhydrogénation du cyclohexane [7b] ainsi que l'oxydation du propène [7c] ont déjà été étudiées. La liste de l'*encadré 1*, non exhaustive, donne un aperçu du nombre et de la diversité des réactions de catalyse homogène déjà étudiées.

### Encadré 1 - Diversité de réactions catalytiques étudiées par l'approche chimie combinatoire

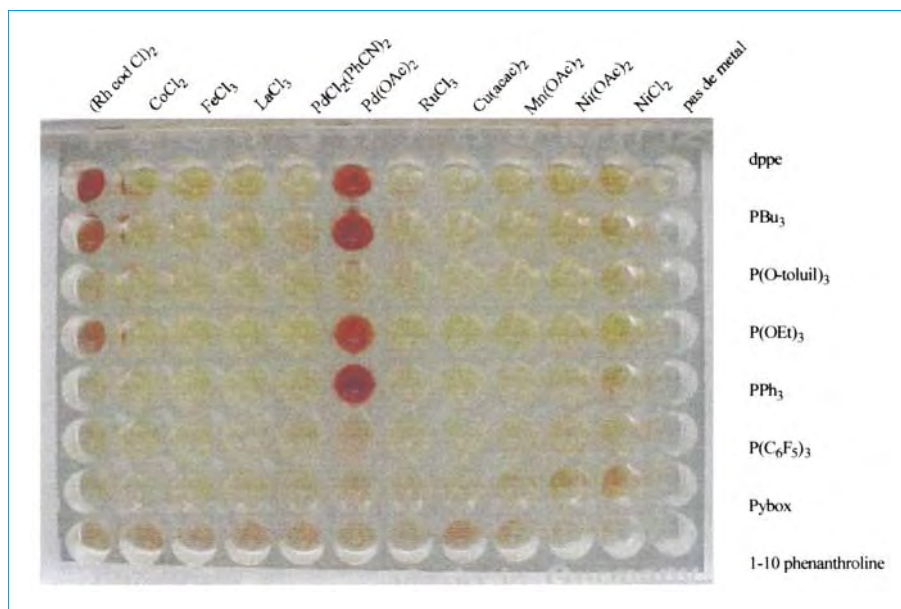
#### Catalyse combinatoire (références)

Alkylation allylique [4a]  
Hydrosilylation d'alcènes et d'imines [4b]  
Couplage de Heck [4c]  
Réaction d'acylation [5b]  
Hydrocyanation d'imines [6a]  
Ouverture d'époxydes en diols [16]  
Époxydation d'oléfines [12]  
Hydrogénation d'énamides [17]  
Alkylation d'aldéhydes [9]  
Réaction de métathèse [18]  
Insertion dans des liaisons C-H [6c]  
Polymérisations d'oléfines [9, 19]

Ces premiers résultats démontrent clairement l'étendue des domaines accessibles par la catalyse combinatoire. Toutes ces recherches ont été menées dans des délais beaucoup plus courts que l'approche traditionnelle en catalyse. A titre d'exemple, sur le principe du test visuel présenté dans la *photo 3*, un tout nouveau catalyseur pour la réaction d'alkylation allylique a été découvert en moins de deux mois entre l'idée de départ et la mise en évidence du résultat [4a].

La « vitesse » d'exécution de la catalyse combinatoire représente un atout indéniable pour les chimistes et ouvre de nombreuses perspectives. Elle peut permettre, par de fines et nombreuses variations, d'accéder rapidement au meilleur catalyseur pour une famille donnée. Mais elle offre surtout aux chimistes l'opportunité de tester un grand nombre de systèmes catalytiques différents et d'oser des combinaisons métal/ligands inhabituelles, **ce qui représente une porte ouverte vers la découverte de catalyseurs totalement innovants**. Si de nombreux résultats attestent de l'efficacité de cette approche, il faut néanmoins noter qu'il s'agit d'un domaine en pleine émergence. Aussi, afin de généraliser cette technique, un effort important de prospection et d'innovation doit être fourni pour repousser les limites actuelles de cette méthode. Par exemple, la diversité des catalyseurs accessibles actuellement par la chimie combinatoire n'est pas aussi étendue que celle proposée par l'approche traditionnelle et doit donc être élargie. D'autre part, les techniques classiques utilisées pour évaluer une activité catalytique ne sont pas toutes adaptées pour évaluer à haut débit ; en particulier, les techniques chromatographiques sont difficilement envisageables pour des banques importantes de composés. Pour repousser ces limites, de nouvelles stratégies de synthèse et de caractérisation doivent être imaginées et proposées, ce qui représente un domaine d'investigation totalement nouveau.

Le présent article est axé sur la catalyse combinatoire, mais il est important de signaler que la chimie combinatoire est aussi en train de révolutionner la façon de découvrir et de mettre au point de nouveaux matériaux [3], autre



*Photo 3* - 96 combinaisons différentes (12 complexes de métaux et 8 ligands). Détection des catalyseurs actifs par colorimétrie.

secteur économique d'importance. Il semble donc que, dans un proche futur, **l'expérimentation à haut débit** puisse devenir une technique incontournable pour bien des domaines scientifiques et, en tout cas, pour la plupart des domaines de la chimie. Ces prévisions sont renforcées par le fort engouement et les investissements du secteur industriel pour cette nouvelle approche ainsi que par la création et le développement de plusieurs sociétés spécialisées dans cette thématique.

### Remerciements

Je remercie chaleureusement le Dr J. Morken pour m'avoir accueilli dans son équipe spécialisée en chimie combinatoire organique, ainsi que le CNRS pour une mise à disposition d'une année comme chercheur associé à l'Université de Chapel Hill (NC, États-Unis). Je suis aussi très reconnaissant à la société Elf pour son soutien et sa confiance dans ce projet de catalyse combinatoire.

### Références

- [1] Wilson S.R., Czarnik A.W., *Combinatorial Chemistry : Synthesis and applications*, Ed. Wiley, **1997**.
- [2] Persidis A., A business to bank on, *Chemistry & Industry*, **1998**, p. 782-784.
- [3] Jandeleit B., Schaefer D.J., Powers T.S., Turner H.W., Weinberg W.H., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 2494-2532.
- [4] a) Lavastre O., Morken J.P., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 3163-3165 ; b) Cooper A.C., McAlexander L.H., Lee D.-H., Torres M.T., Crabtree R.H., *J. Am. Chem. Soc.*, **1998**, 120, p. 9971-9972 ; c) Shaughnessy K.H., Kim P., Hartwig J.F., *J. Am. Chem. Soc.*, **1999**, 121, p. 2123-2132.
- [5] a) Francis M.B., Jamison T.F., Jacobsen E.N., *Current Opinion in Chemical Biology*, **1998**, 2, p. 422-428 ; b) Taylor S.J., Morken J.P., *Science*, vol. 280, **1998**, p. 267-270.
- [6] a) Sigman M.S., Jacobsen E.N., *J. Am. Chem. Soc.*, **1998**, 120, p. 4901-4902 ; b) Shimizu K.D., Cole B.M., Krueger C.A., Kuntz K.W., Snapper M.L., Hoveyda A.H., *Angew. Chem. Int. Ed. Engl.*, **1997**, 36, p. 1704 ; c) Burgess K., Lim H.-J., Porte A.M., Sulikowski G.A., *Angew. Chem. Int. Ed. Engl.*, **1996**, 35, p. 220-222.
- [7] a) Hoffmann C., Wolf A., Schüth F., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 2800-2803 ; b) Senkan S.M., Ozturk S., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 791-795 ; c) Orschel M., Klein J., Schmidt H.-W., Maier W.F., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 2791-2799.
- [8] Hinderling C., Chen P., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 2253-2256.
- [9] Ding K., Ishii A., Mikami K., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 497-500.
- [10] Bein T., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 323-326.
- [11] Crabtree R.H., *Chemtech*, April **1999**, p. 21-26.
- [12] Francis M.B., Jacobsen E.N., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 937-941.
- [13] Cong P., Doolen R.D., Fan Q., Giaquinta D.M., Guan S., McFarland E.W., Poojary D.M., Self K., Turner H.W., Weinberg W.H., *Angew. Chem. Int. Ed.* **1999**, 38, p. 484-485.
- [14] Senkan S., Krantz K., Ozturk S., Zengin V., Onal I., *Angew. Chem. Int. Ed.*, **1999**, 38, p. 2794-2799.
- [15] Holzwarth A., Schmidt H.-W., Maier W.F., *Angew. Chem. Int. Ed.*, **1998**, 37, p. 2644-2647.
- [16] Reetz M.T., Becker M.H., Kühling K.M., Holzwarth A., *Angew. Chem. Int. Ed.*, **1998**, 37, p. 2647-2650.
- [17] Gilbertson S.R., Wang X., *Tetrahedron Letters*, **1996**, vol. 37, n° 36, p. 6475-6478.
- [18] Reetz M.T., Becker M.H., Liebl M., Furstner A., *Angew. Chem. Int. Ed.*, **2000**, 39, p. 1236-1239.
- [19] Boussie T.R., Coutard C., Turner H., Murphy V., Powers T.S., *Angew. Chem. Int. Ed.*, **1998**, 37, p. 3272-3275.



# Tails and steeds versus heads and beads

## The case for fluororous phase methodology in combinatorial and parallel syntheses

Christian Rocaboy\* *doctorand*, John A. Gladysz\* *professor*

**Summary :** The development of fluororous phase synthetic chemistry is reviewed from its recent origins in strategies for recoverable catalysts through emerging applications in high-throughput library synthesis. The roles of « pony tails »,  $(\text{CH}_2)_m(\text{CF}_2)_{n-1}\text{CF}_3$ , as solubility control or « phase labelling » elements, are described. The synthesis and properties of fluororous aliphatic amines are detailed, and lead references to fluororous protecting groups and chromatographic methods are given.

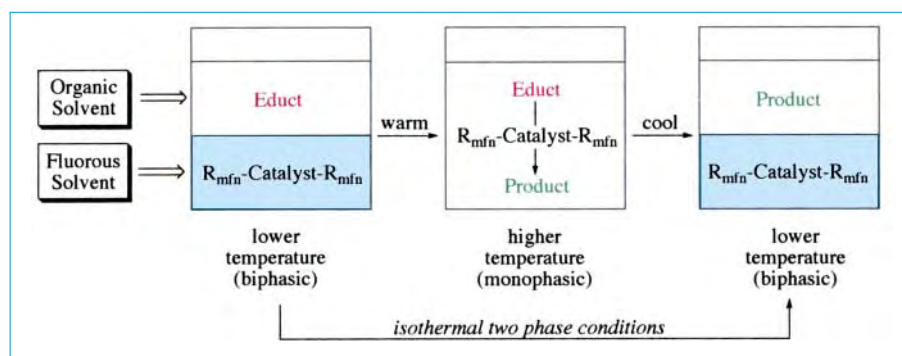
**Mots clés :** Fluoré, pony tail, coefficients de partition, amines, basicité.

**Key-words :** Fluororous, pony tail, partition coefficients, amines, basicity.

Le récent développement de la chimie en phase fluorée est revu depuis ses premières utilisations en catalyse jusqu'à ses applications émergentes en chimie combinatoire. Le rôle des « pony tails »,  $(\text{CH}_2)_m(\text{CF}_2)_{n-1}\text{CF}_3$ , comme éléments de contrôle de la solubilité, est décrit. La synthèse et les propriétés d'amines aliphatiques fluorées sont détaillées, ainsi que les premières références concernant les groupes protecteurs et méthodes chromatographiques fluorés.

The term « fluororous » was introduced by Horváth in 1994 as an analog to « aqueous » for highly fluorinated alkane, ether, and tertiary amine solvents [1]. Such solvents are very non-polar, and many are commercially available [2]. They commonly give bilayers with organic solvents at room temperature, as illustrated by the first container in *scheme 1*. At the same time, many such solvent combinations become miscible at elevated temperatures, as illustrated by the second container in *scheme 1*. Many synthetic chemists had been unaware of these characteristics.

Organic compounds - even hydrocarbons such as dodecane - normally have low affinities for fluororous solvents relative to organic solvents. However, compounds that consist mainly of per-fluoro-alkyl segments show high affinities. This reflects a « like dissolves



**Scheme 1** - Some possibilities for catalysis with fluororous solvents ( $R_{mfn} = (\text{CH}_2)_m(\text{CF}_2)_{n-1}\text{CF}_3$ ).

like » effect, and similar strategies are used to design dyes that can adhere to Teflon. Accordingly, Horváth noted that high fluororous affinities could be imparted to common catalysts and reagents by appending « pony tails »  $(\text{CH}_2)_m-(\text{CF}_2)_{n-1}\text{CF}_3$  (abbreviated  $(\text{CH}_2)_m\text{R}_{fn}$ ) in sufficient numbers or lengths. Compilations of fluororous/organic phase partition coefficients are available [2], and continually updated on the authors' web page (<http://www.organik.uni-erlangen.de/gladysz/research/partition.html>).

The  $(\text{CH}_2)_m$  segments of the pony tails can be used to fine-tune electronic properties. When sufficiently long, they insulate the active site from the electron withdrawing fluorines. When short, catalysts and reagents have enhanced Lewis acidities.

However, the most important function of the pony tail is as a solubility control device, which furthermore provides the basis for an innovative new approach to recoverable catalysts and reagents, combining all advantages of one-phase chemistry with biphasic pro-

\* Institut für Organische Chemie, Friedrich-Alexander Universität Erlangen-Nürnberg, Henkestrasse 42, 91054 Erlangen, Germany. Phone : +49 (0)9131-8522540. Fax : +49 (0)9131-8526865. E-mail : gladysz@organik.uni-erlangen.de

duct separation. Reactions can be effected under homogeneous conditions in the high temperature monophasic limit, and the organic products disassociated from the fluoruous catalyst or transformed reagent in the low temperature biphasic limit (*scheme 1*, middle and right containers). Many applications of this protocol have been developed over the last few years [3-4].

A variant involving (trifluoromethyl)toluene,  $\text{CF}_3\text{C}_6\text{H}_5$ , deserves note. This solvent is able to dissolve both fluoruous and non-fluoruous substances. It can be removed at the end of a reaction, and a non-fluoruous or fluoruous solvent added to extract the desired material.

In more general terms, fluoruous chemistry represents a new use of « orthogonal phases » in separation strategy. As nicely articulated by Curran [5], organic workups commonly involve manipulations of target molecules and by-products through various states and media : (1) solid phases, (2) gas phases, (3) aqueous liquid phases, (4) organic liquid phases, and (5) fluoruous liquid phases. These can be idealized as mutually exclusive, but various approximations are obvious. For example, some water is soluble in most organic liquids, and gases dissolve to some extent in all liquids (particularly fluoruous liquids, which readily accommodate small molecules irrespective of polarity) [2].

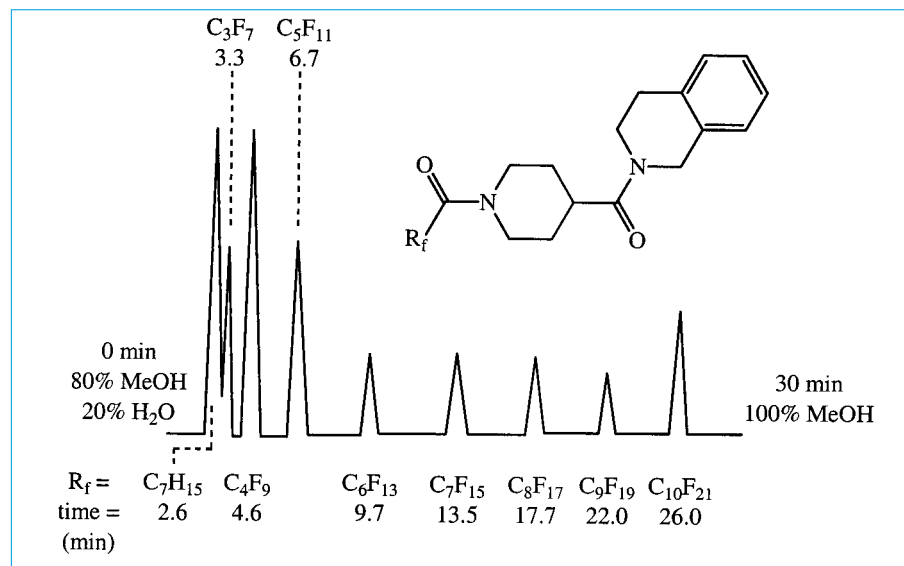
High-throughput combinatorial and parallel syntheses entail an even more demanding series of manipulations involving different orthogonal phases. Hence, it is not surprising that fluoruous chemistry was rapidly extended in this direction [6]. The applicability to solution-phase methodologies is particularly obvious. Advantages and disadvantages of solution methodologies

versus the presently dominant solid-phase technologies have been extensively discussed and debated [7]. However, as analyzed below, fluoruous chemistry has impressive potential for all types of high-throughput synthesis.

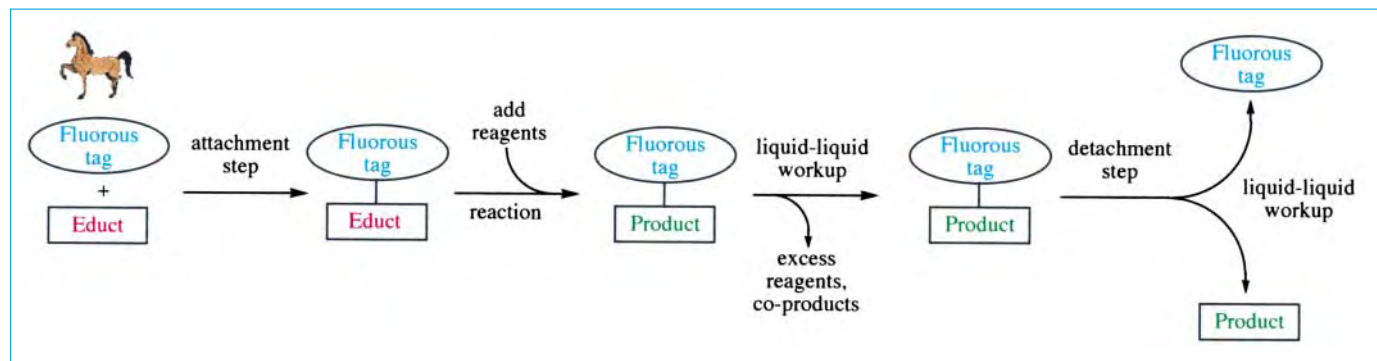
A current goal of polymer bead-based synthesis is to be able to reproduce - as much as possible within certain intrinsic limitations - the entire repertoire of organic reactions. Fluoruous chemistry has a similar goal : to develop a complete « parallel universe » of fluoruous functional organic molecules, reagents, and catalysts. Although much work remains to be done, this « fluoruous library » will be achieved in the foreseeable future. Fluoruous analogs of, for example, Swern oxidations or Mitsunobo reactions would give fluoruous sulfur, phosphorus, and nitrogen-containing by-products that could be removed under the conditions of *scheme 1*. These would have immediate application in solution phase parallel synthesis.

Curran and Wipf have developed a « tagging » or « phase labeling » protocol (*scheme 2*) for fluoruous variants of the Ugi and Biginelli multicomponent condensations [8]. The basic idea is to derivatize the educt with some type of removable assembly of pony tails. A series of liquid phase reactions are then conducted, and the labeled product recovered by means of its enhanced fluoruous phase affinity. The partition coefficients do not need to be as high as those required in *scheme 1*. At the end of the sequence, the tag or phase label is removed.

Fluoruous silica gel is easy to prepare, and analytical columns are commercially available [9]. As illustrated in *scheme 3*, efficient reverse-phase separations are possible, with homologous compounds eluting in inverse order of fluoruous content. This provides an alternative to extraction, especially for compounds with low fluoruous phase affinities. Curran prepared a 4 x 4 amide library where each member was puri-



*Scheme 3* - Chromatogram of homologous fluoruous amides on a fluofix 120E column.



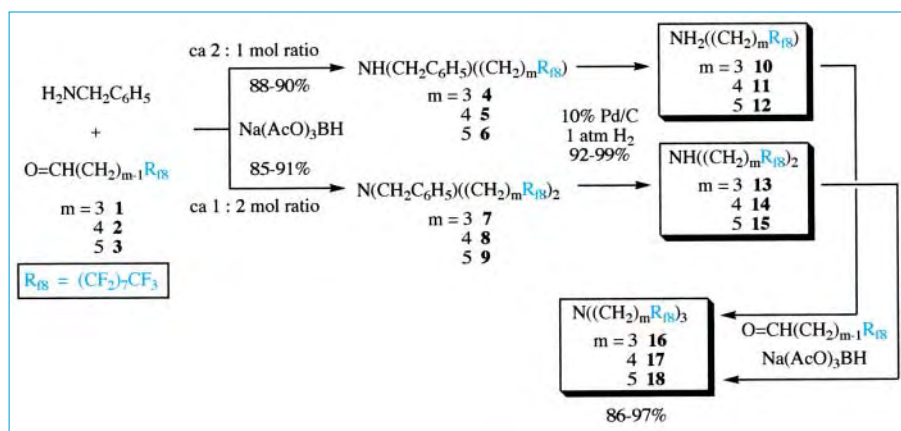
*Scheme 2* - Reversible « phase labelling » of educts and products.

fied by a simple fluoros silica gel filtration [9].

Protecting groups that have been phase labeled are also of obvious utility in high-throughput synthesis. To date, fluoros analogs of benzyl, THP, and alkoxy ethyl ether protecting groups for alcohols have been developed (*scheme 4A*) [10-11]. Fluoros strategies for the rapid purification of non-fluorous [12] and fluoros [4d] organometallic compounds have been described (*scheme 4B*). Inhibitors are sometimes employed in high-throughput syntheses. Noteworthy in this regard is a fluoros diaryl diselenide that, in conjunction with stannanes, retards certain radical rearrangements (*scheme 4C*) [13].

Scavengers play important roles in both solid and solution phase parallel synthesis. For example, the quest for optimal covalent scavengers for primary and secondary amines has been detailed in a recent review [14]. The use of fluoros amines for scavenging excess isocyanates has been reported [15]. These and other factors prompted us to develop systematic syntheses of fluoros aliphatic amines with fine-tuned fluoros phase affinities and basicities. These data are summarized in the final portion of this article [4i].

The fluoros aldehydes **1-3** shown in *scheme 5* are easily prepared, and reductive aminations can be conducted



*Scheme 5* - Versatile syntheses of fluoros primary, secondary and tertiary amines.

with  $\text{Na}(\text{AcO})_3\text{BH}$  in THF at different aldehyde/amine stoichiometries. When two-fold excesses of benzyl amine are used, workups give the secondary amines **4-6** in 88-90 % yields. When two-fold excesses of aldehydes are used, the **4-6** generated can condense further. Workups give the tertiary amines **7-9** in 85-91 % yields. The benzyl protecting groups in **4-9** can be removed under mild hydrogenolysis conditions, giving the primary amines **10-12** and the secondary amines **13-15** in > 99-92 % yields.

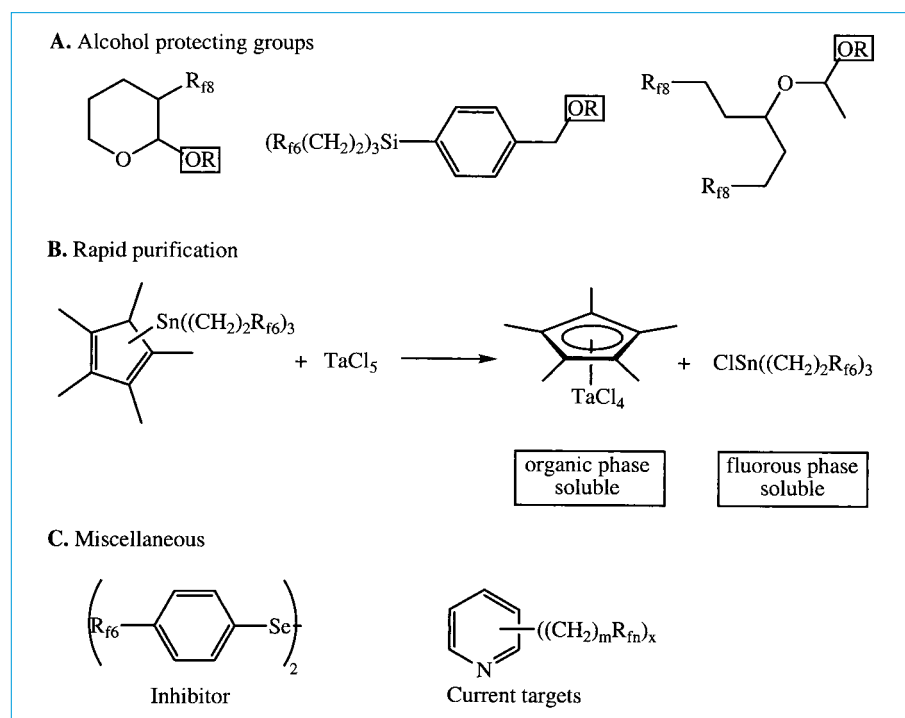
As shown in *scheme 5*, the secondary amines **13-15** are then condensed with 1.0-1.2 equivalents of aldehydes **1-3** under identical reductive amination conditions. Workups give the corres-

ponding tertiary amines **16-18** in 97-86 % yields. The fluoros amines **10-18** dissolve in quite a broad range of solvents. All are highly soluble in  $\text{CHCl}_3$ ,  $\text{CF}_3\text{C}_6\text{H}_5$ , and fluoros solvents such as  $\text{CF}_3\text{C}_6\text{F}_{11}$ . The primary and secondary amines **10-15** are also very soluble in methanol. The highly fluoros tertiary amines **16-18** remain sparingly soluble in methanol. None of the amines are soluble in water or DMSO.

Quantitative data on relative fluoros phase affinities are given by the  $\text{CF}_3\text{C}_6\text{F}_{11}$ /toluene partition coefficients in *table I*. As would be expected from a simple « like dissolves like » model, affinities decrease as the methylene chain is lengthened (**10** vs **11** vs **12** ; **13** vs **14** vs **15** ; **16** vs **17** vs **18**). They also decrease when pony tails are replaced by hydrogen atoms (**16** vs **13** vs **10** ; **17** vs **14** vs **11** ; **18** vs **15** vs **12**). The two amines with the highest partition coefficients (**16**, **17**) show no detectable residual quantity in the organic phase (< 0.3 %). This represents, with respect

*Table I* - Partition coefficients (24 °C)

analyte	$\text{CF}_3\text{C}_6\text{F}_{11}$ /toluene
<b>10</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})$	70.0:30.0
<b>11</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})$	63.2:36.8
<b>12</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})$	56.9:43.1
<b>13</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_2$	96.5:3.5
<b>14</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_2$	95.1:4.9
<b>15</b> $\text{NH}_2(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_2$	93.0:7.0
<b>16</b> $\text{N}(\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_3$	>99.7:<0.3
<b>17</b> $\text{N}(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_3$	>99.7:<0.3
<b>18</b> $\text{N}(\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{R}_{f8})_3$	99.5:0.5



*Scheme 4* - Additional fluoros compounds and applications.

to the protocol in *scheme 1*, a very high degree of immobilization.

Brønsted basicity measurements show that the inductive effect of the perfluoroalkyl groups is strongly felt in tertiary amine **16**, which has three methylene groups in each pony tail. Amine **18**, which features longer five-methylene segments, remains a weaker base than tri(dodecyl)amine (ca. 25-30:75-70 protonation ratio in  $\text{CDCl}_3$ ). The corresponding fluorous trialkylphosphines exhibit similar behavior [4c, h]. Hence, more than five methylene groups are required to fully insulate heteroatom lone pairs from electro-negative perfluoroalkyl groups.

In conclusion, a two-part summary/outlook is offered. First, the work detailed from our laboratory has provided three series of nearly isosteric fluorous amines - primary, secondary, and tertiary - with finely modulated fluorous phase affinities and basicities. Similar families of pyridines, sketched in *scheme 4C*, will be reported in the near future. These constitute valuable « toolkits » for the systematic development of new catalytic and stoichiometric reactions, including high-throughput synthesis, based upon fluorous nitrogen bases and nucleophiles.

Second, the many developments outlined above clearly indicate a growing role for fluorous techniques in all of combinatorial and parallel synthesis. Additional innovative applications can be expected at a rapid pace. Some cognoscente caution, however, that commercial uses are already covered by an extremely broad patent [16], such that licensing inquiries would be prudent.

### Acknowledgment

We thank the Deutsche Forschungsgemeinschaft (DFG ; GL 301/3-1) for support of this research.

### References

- [1] Horváth I.T., *Acc. Chem. Res.*, **1998**, *31*, p. 641-650.
- [2] Barthel-Rosa L.P., Gladysz J.A., *Coord. Chem. Rev.*, **1999**, *190-192*, p. 587-605.
- [3] Review literature since 1999 : (a) de Wolf E., van Koten G., Deelman B.-J., *Chem. Soc. Rev.*, **1999**, *28*, p. 37-41 ; (b) Fish R.H., *Chem. Eur. J.*, **1999**, *5*, p. 1677-1680 ; (c) Cavazzini M., Montanari F., Pozzi G., Quici S., *J. Fluorine Chem.*, **1999**, *94*, p. 183-193 ; (d) Diederichsen U., *Nachr. Chem. Tech. Lab.*, **1999**, *47*, p. 805-809 ; (e) Hope E.G., Stuart A.M., *J. Fluorine Chem.*, **1999**, *100*, p. 75-83.
- [4] Papers from our laboratory : (a) Guillevic M.-A., Rocaboy C., Arif A.M., Horváth I.T., Gladysz J.A., *Organometallics*, **1998**, *17*, p. 707-717 ; (b) Rutherford D., Juliette J.J.J., Rocaboy C., Horváth I., Gladysz J.A., *Catalysis Today*, **1998**, *42*, p. 381-389 ; (c) Alvey L.J., Rutherford D., Juliette J.J.J., Gladysz J.A., *J. Org. Chem.*, **1998**, *63*, p. 6302-6308 ; (d) Juliette J.J.J., Rutherford D., Horváth I.T., Gladysz J.A., *J. Am. Chem. Soc.*, **1999**, *121*, p. 2696-2704 ; (e) Klose A., Gladysz J.A., *Tetrahedron-Asymmetry*, **1999**, *10*, p. 2665-2674 ; (f) Dinh L., Gladysz J.A., *Tetrahedron Lett.*, **1999**, *40*, p. 8995-8998 ; (g) Rocaboy C., Rutherford D., Bennett B.L., Gladysz J.A., *J. Phys. Org. Chem.*, **2000**, in press ; (h) Alvey L.J., Meier R., Soós T., Bernatis P., Gladysz J.A., *Eur. J. Inorg. Chem.*, **2000**, in press ; (i) Rocaboy C., Bauer W., Gladysz J.A., *Eur. J. Org. Chem.*, **2000**, p. 2621-2628.
- [5] Curran D.P., *Angew. Chem., Int. Ed. Engl.*, **1998**, *37*, p. 1174-1196 ; *Angew. Chem.*, **1998**, *110*, p. 1230-1255.
- [6] Studer A., Curran D.P., *Tetrahedron*, **1997**, *53*, p. 6681-6696.
- [7] Baldino C.M., *J. Combinatorial Chemistry*, **2000**, *2*, p. 89-103.
- [8] Studer A., Jeger P., Wipf P., Curran D.P., *J. Org. Chem.*, **1997**, *62*, p. 2917-2924.
- [9] Curran D.P., Luo Z., *J. Am. Chem. Soc.*, **1999**, *121*, p. 9069-9072.
- [10] Wipf P., Reeves J.T., *Tetrahedron Lett.*, **1999**, *40*, p. 4649-4652 and 5139-5142.
- [11] Curran D.P., Ferritto R., Hua Y., *Tetrahedron Lett.*, **1998**, *39*, p. 4937-4940.
- [12] Spetseris N., Hadida S., Curran D.P., Meyer T.Y., *Organometallics* **1998**, *17*, p. 1458-1459.
- [13] Crich D., Hao X., Lucas M.A., *Organic Letters*, **1999**, *1*, p. 269-271.
- [14] Hodges J.C., *Synlett*, **1999**, p. 152-158.
- [15] Linclau B., Sing A.K., Curran D.P., *J. Org. Chem.*, **1999**, *64*, p. 2835-2842.
- [16] US Patent 5,463,082, **1995**.

# Chimie combinatoire et chimie supramoléculaire

Ivan Huc\* chargé de recherche, Jean-Marie Lehn\*\* professeur

## Summary : Combinatorial chemistry and supramolecular chemistry

The concept of dynamic combinatorial chemistry has developed from the consideration that self-assembly in a multicomponent system is a combinatorial process. Combinatorial libraries based on the covalent or non-covalent reversible assembly of initial building blocks allow a continuous variation of the proportions between library products in response to a particular assay. Such dynamic libraries have recently emerged as an efficient tool to detect and to evaluate molecular recognition between some library products and a target molecule.

**Mots clés :** Réversibilité, dynamique, auto-assemblage, reconnaissance moléculaire, combinatoire.

**Key-words :** Reversibility, dynamic, self-assembly, molecular recognition, combinatorial.

La chimie combinatoire repose sur la mise au point de méthodes de synthèse et de criblage de vastes mélanges construits par combinaison systématique d'unités de base. Cependant, si l'on définit une combinaison comme un « arrangement dans un certain ordre de divers éléments », force est de constater que la chimie est par essence une science combinatoire. Les molécules sont des combinaisons d'atomes, les (bio)polymères sont des combinaisons de monomères, les supermolécules sont des combinaisons de molécules, etc. L'outil du chimiste de synthèse est un ensemble hiérarchisé d'éléments à combiner : groupes fonctionnels, synthons, molécules complexes, macromolécules et supermolécules, et l'objet de la chimie structurale est précisément de définir dans quel ordre ces éléments sont arrangés au sein d'une combinaison. De la synthèse organique aux matériaux, en passant par la catalyse et les biomolécules, il est donc naturel que les développements récents du combinatoire trouvent échos et ramifications dans toutes les branches de la chimie. En chimie supramoléculaire, une mise en perspective de l'auto-assemblage

comme un processus combinatoire a abouti à l'idée de bibliothèques dynamiques, dans lesquelles des liaisons réversibles s'établissent entre les unités de bases [1]. Le principal avantage de tels systèmes est de permettre une variation continue des proportions entre les produits d'une bibliothèque en réponse à un changement des conditions, comme par exemple l'introduction d'un composé cible pouvant se lier à certains produits de la bibliothèque. Le criblage du mélange se fait simplement en analysant les modifications de la composition du mélange, notamment le changement des proportions entre les produits.

## L'auto-assemblage comme processus combinatoire

La formation dans des conditions thermodynamiques d'assemblages bien définis de molécules par association et agencement spontané est un outil puissant pour la préparation de complexes édifices moléculaires et supramoléculaires. Deux exemples sont représentés figure 1, illustrant d'une part l'auto-assemblage d'une grille supramoléculaire par liaisons non covalentes entre quatre ions  $\text{Ag}^+$  et quatre ligands pyridiniques [2] et, d'autre part, l'auto-assemblage d'un macrocycle par liai-

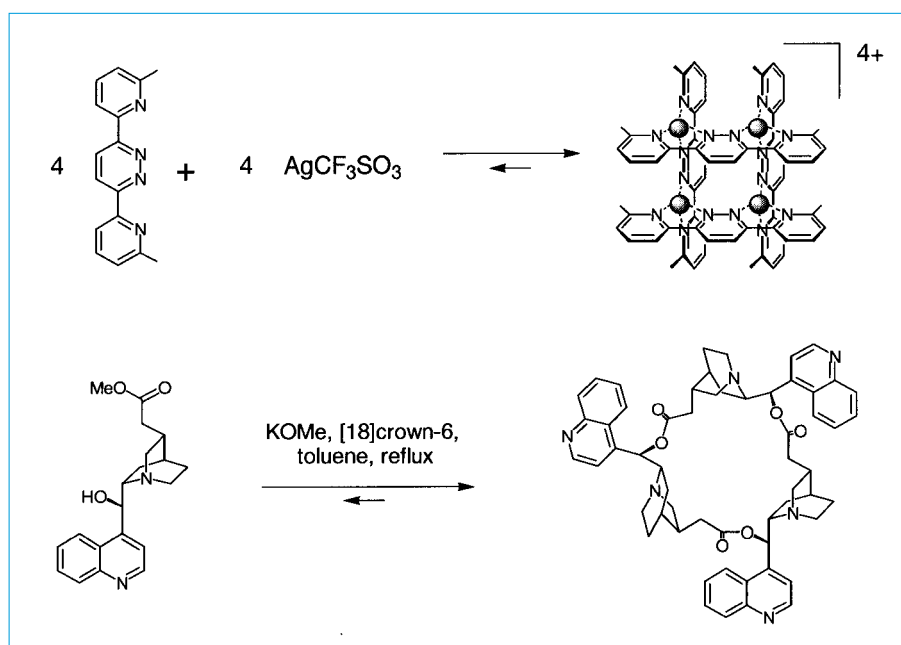


Figure 1 - Exemples d'auto-assemblages quantitatifs non covalents et covalents.

\* Institut Européen de Chimie et Biologie, ENSCPB, av. Pey Berland, 33402 Talence Cedex.  
Tél. : 05.57.96.22.29. Fax : 05.57.96.22.26.  
E-mail : ivan.huc@iecb-polytechnique.u-bordeaux.fr

\*\* Institut de Science et d'Ingénierie Supramoléculaires, 4, rue Blaise Pascal, 67000 Strasbourg.  
Tél. : 03.88.41.60.62. Fax : 03.88.41.10.20.  
E-mail : lehn@chimie.u-strasbg.fr

sons covalentes entre trois unités d'un hydroxy-ester dérivé de la cinchonidine [3]. Ces assemblages se produisent à l'équilibre thermodynamique de façon quantitative. Toutefois, les structures obtenues ne sont pas les seuls produits envisageables. L'hydroxy-ester aurait pu former d'autres oligomères cycliques ou linéaires, de même pour le système ligand pyridinique et ion  $Ag^+$ . Le caractère combinatoire de ces assemblages covalents ou non covalents apparaît donc lorsque l'on considère non pas seulement les produits formés, mais tous les produits potentiellement accessibles au système. Chacun d'entre eux représente une combinaison, c'est-à-dire une manière d'agencer les unités de base du système en nombre, en nature et en orientations relatives. L'ensemble de ces combinaisons définit la bibliothèque des produits potentiellement accessibles au système.

Les deux exemples de la *figure 1* constituent d'un certain point de vue des cas extrêmes dans lesquels une combinaison unique parmi toutes celles possibles est réellement exprimée. Mais il est facile d'imaginer des systèmes moins « efficaces », dans lesquels plusieurs structures, voire un mélange complexe de produits seraient formés en quantité significative. D'une façon générale, l'expression plus ou moins complète de la variété des structures accessibles est dirigée par les paramètres thermodynamiques et cinétiques associés aux composants et à leurs connexions : les réactivités intrinsèques des fonctions connectrices, les facteurs structuraux aboutissant éventuellement à une prédisposition ou une préorganisation (énergies des conformations et configurations), les termes de solvation des différentes structures, et les interactions non covalentes spécifiques attractives ou répulsives intra- et inter-moléculaires.

Finalement l'auto-assemblage peut être envisagé de deux manières :

- dans une approche réciproque, l'expression préférentielle et inattendue d'assemblages à partir d'un mélange combinatoire de composants devrait permettre d'identifier des associations originales et les interactions non covalentes spécifiques mises en jeu dans leur formation (reconnaissance moléculaire).

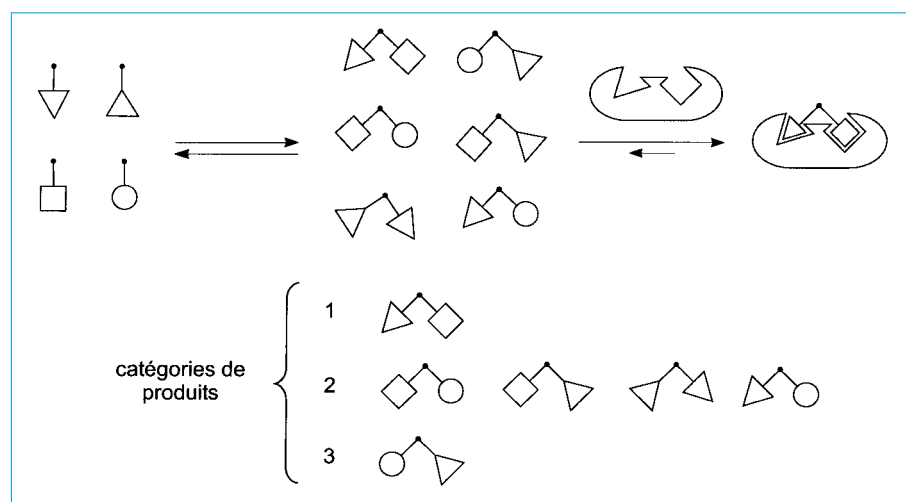
Alors que la première démarche est depuis longtemps établie, cette deuxième voie ne s'est vraiment développée que dans les cinq dernières années, amenant au concept de chimie combinatoire dynamique qui met en œuvre des bibliothèques dynamiques virtuelles [1, 4].

### La chimie combinatoire dynamique

Le principe fondamental d'une bibliothèque combinatoire dynamique est la réversibilité de l'assemblage des unités de base (*figure 2*). On peut envisager des liaisons réversibles covalentes comme par exemple des imines, des borates ou des disulfures, menant à des bibliothèques de molécules. On peut également utiliser des liaisons réversibles non covalentes comme des liaisons hydrogène, des liaisons ioniques, ou des liaisons de coordination à des métaux, menant à des bibliothèques de supermolécules. Les proportions entre produits de la bibliothèque reflètent ini-

tialement les paramètres cinétiques et thermodynamiques associés aux unités de bases et à leur connexions. Mais ces proportions peuvent varier sous l'effet d'un agent extérieur comme l'ajout d'un composé cible, une irradiation, ou l'application d'une contrainte mécanique, électrique ou magnétique. La variation des proportions renseigne sur l'effet de l'agent et ses interactions avec la bibliothèque. La formation préférentielle d'un produit (d'une combinaison) au détriment des autres s'apparente au criblage spontané du mélange. Une bibliothèque dynamique se distingue donc essentiellement d'une bibliothèque classique dans le fait que la composition du mélange n'est pas préétablie, et notamment que la proportion des produits recherchés peut être amplifiée au cours du criblage, effectué simultanément à la synthèse. En fait, il se peut fort bien que parmi tous les composants possibles d'une telle bibliothèque, seuls s'expriment ceux qui se lient à la cible, d'où la notion de bibliothèque combinatoire virtuelle [1, 4].

Dans les travaux décrits jusqu'à présent dans la littérature, l'agent externe est généralement un composé cible ajouté, dont la présence peut conduire à un augmentation de la fraction molaire (et à l'identification) des produits qui lui sont le mieux complémentaires (*figure 2*). Cette stabilisation spécifique de produits par le biais des forces de la reconnaissance moléculaire consiste en



**Figure 2** - La chimie combinatoire dynamique. Principe du coulage d'un substrat par un récepteur. Trois catégories de produits sont distinguées : 1- ceux qui lient fortement la cible (deux sites d'ancrage), 2- ceux qui lient faiblement la cible (un site d'ancrage), et 3- ceux qui ne lient pas la cible. Le coulage aboutit à une augmentation de la fraction molaire de produits de la première catégorie aux dépens des produits de la deuxième catégorie. La fraction des produits de la troisième catégorie peut également augmenter substantiellement. L'amplitude de la stabilisation par la cible de produits de la première catégorie est limitée par l'association des produits de la deuxième catégorie qui, bien que plus faible, devient significative si ces produits sont abondants.

un coulage par effet matrice d'un substrat par un récepteur (comme pour un bronze coulé dans un moule), ou réciproquement, en un moulage d'un récepteur sur un substrat. Au cours des cinq dernières années, de nombreux systèmes combinatoires dynamiques organiques, inorganiques, et biochimiques ont été décrits. Des articles de revues récents font l'inventaire de ces travaux [1], nous nous contenterons de citer quelques exemples représentatifs (figure 3).

Un certain nombre de bibliothèques dynamiques est basé sur des équilibres thermiques ou photochimiques entre isomères de configuration [5-6] ou de conformations [7]. Dans ces systèmes, la stabilisation d'un isomère par interaction avec le substrat s'apparente à l'adaptation conformationnelle du site actif des enzymes. De nombreux

exemples d'auto-assemblage de récepteurs supramoléculaires induits par la reconnaissance du substrat ont été traités [4]. De multiples unités de bases s'associant par liaisons hydrogène ou par coordination d'un métal peuvent être introduites dans ces systèmes avec le dessein d'augmenter le nombre de combinaisons [8-9]. Des bibliothèques de macrocycles ont été construites par connexion covalente réversible d'un nombre variable de différents monomères [10]. D'autres systèmes mettent en jeu la reconnaissance moléculaire de peptides [11], d'acides nucléiques [12], ou d'anticorps [13]. Plus proches d'applications d'intérêt pharmaceutique, des exemples de moulage de sites actifs d'enzymes par des petites bibliothèques dynamiques ont été décrits pour l'anhydrase carbonique [14] des lectines [15], ou la trypsine [16].

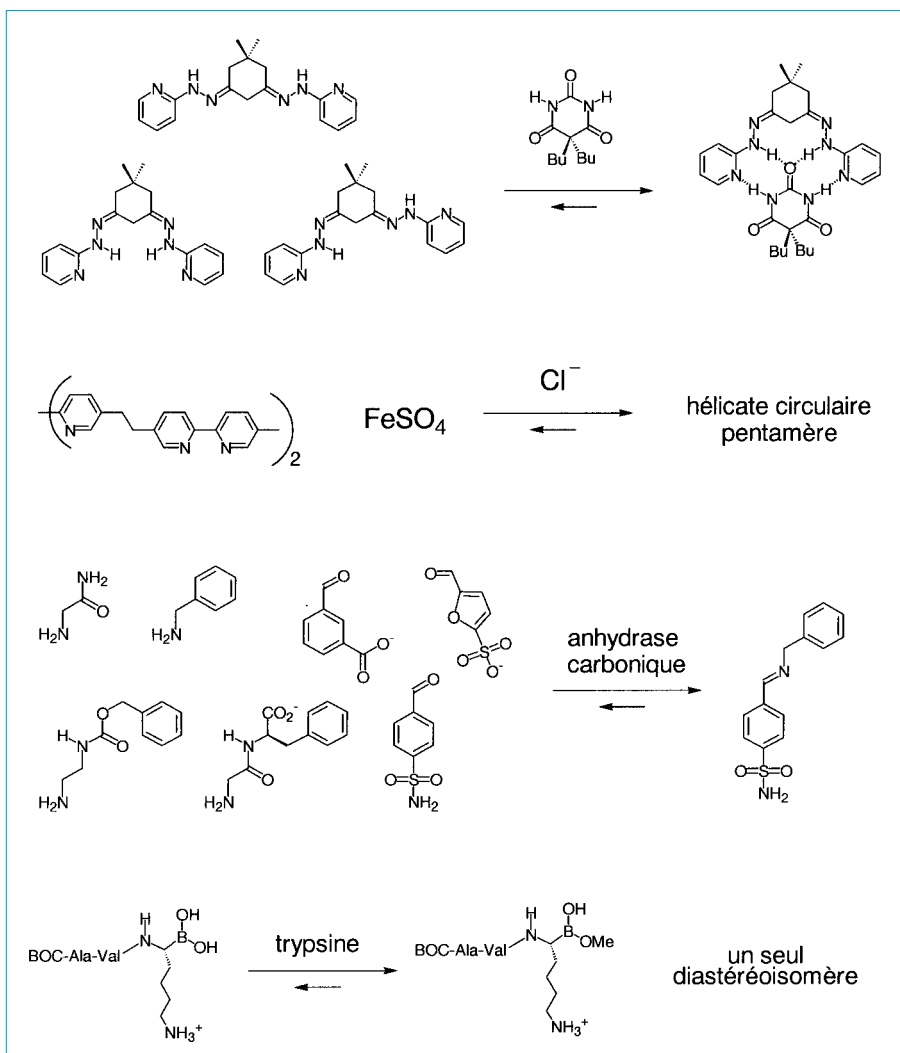


Figure 3 - Exemples d'effet matrice dans des systèmes combinatoires dynamiques tirés respectivement des références [6], [4], [14], et [16]. Les équilibres représentent schématiquement les unités de base, la molécule cible, et l'espèce la mieux complémentaire stabilisée

## Aspects pratiques

Un certain nombre d'aspects pratiques pour la mise en œuvre d'un système combinatoire dynamique méritent d'être discutés.

### Réactions

De multiples réactions permettent de former réversiblement des liaisons covalentes entre unités de bases : Diels-Alder, condensation aldoliques, etc. Ces réactions, souvent délaissées par les chimistes comme étant une source d'instabilité des produits, trouvent maintenant un champ d'application nouveau. Outre les réactions effectivement réversibles connues de la chimie organique et de la chimie de coordination, la catalyse enzymatique offre tout un arsenal de processus réversibles. Les liaisons peptidiques, par exemple, peuvent être formées et hydrolysées réversiblement en présence de protéases [13].

Une forte contrainte s'applique sur le choix des réactions réversibles dans le fait que la reconnaissance moléculaire du composé cible (le criblage) a généralement lieu dans le même milieu que la formation du mélange combinatoire. Même si des phénomènes de reconnaissance moléculaire ont été caractérisés dans l'éthylène glycol à reflux (196 °C) [4], les interactions non covalentes s'expriment en général dans des conditions douces. Pour des applications pharmacologiques, le nombre de réactions réversibles en milieu physiologique n'est pas très grand. Une solution consiste à effectuer la reconnaissance de la cible et l'équilibre dynamique dans deux chambres différentes entre lesquelles circulent les produits de la bibliothèque [5, 13].

### Réactivité

Dans un mélange combinatoire dynamique, toutes les unités de bases sont continuellement en compétition pour la formation des produits. Des réactivités trop différentes les unes des autres conduisent à une distribution initiale de produits très inégale. Ces déséquilibres sont indésirables dans une synthèse combinatoire classique et généralement minimisés par des méthodes

de séparation et recombinaison des produits. Dans un mélange dynamique, l'effet matrice du composé cible peut compenser de faibles réactivités et promouvoir une combinaison qui autrement serait quantitativement absente.

### Réversibilité

La réversibilité des processus est constitutive d'une bibliothèque dynamique, mais elle peut devenir un obstacle à l'identification ou l'utilisation des produits recherchés : que se passe-t-il si ceux-ci ne sont pas stables en absence de la cible ? Pour effectuer une analyse chromatographique d'un mélange sans que les produits ne se désassemblent, il peut être utile de bloquer les équilibres en refroidissant, en changeant le pH, ou par un traitement chimique, comme la réduction irréversible d'imines en amines [14]. De même, un produit sélectionné à partir d'un mélange dynamique ne sera pas considéré comme un candidat direct à une utilisation pharmaceutique, mais plutôt comme un point de départ (tête de série) pour concevoir des composés stables.

### Facteurs entropiques

Une grande part des bibliothèques combinatoires dynamiques décrites dans la littérature met en jeu un nombre relativement limité de combinaisons ( $< 10^2$ ). Dans quelques systèmes, la diversité accessible, sans pouvoir être réellement exprimée, est virtuellement infinie. De plus grands mélanges devraient en principe permettre d'augmenter la probabilité de trouver des composés satisfaisants aux critères de sélection. Cependant, comme dans toutes approches combinatoires, une augmentation inconsidérée de la taille des bibliothèques aboutit à des difficultés. Dans le cas de bibliothèques combinatoires dynamiques, le moulage de produits complémentaires par un composé cible est limité par des facteurs entropiques quand la taille du mélange augmente, du fait qu'un nombre croissant de composés pouvant lier faiblement la cible entrent en compétition pour l'association avec celle-ci (voir explication *figure 2*). D'un point de vue pratique, une grande bibliothèque pourra être envisagée quand une sélection forte et sélective est possible. Pour évaluer des différences d'énergies plus faibles, une bibliothèque petite et focalisée sera préférée. Comme tout mélange combinatoire, un mélange combinatoire dynamique devra être élaboré avec attention afin, en particulier, d'éviter toute redondance dans l'information structurale portée par les unités de base.

Le concept de chimie combinatoire dynamique a déjà été appliqué dans de multiples contextes organiques, inorganiques ou biochimiques, principalement pour identifier des phénomènes de reconnaissance moléculaire. Un développement attendu est le criblage de bibliothèques virtuelles pour d'autres fonctions que la reconnaissance,

comme les propriétés mécaniques, électriques, et magnétiques des produits, ou encore la catalyse. Il n'y a pas si longtemps, « un mélange d'espèces instables » était pour un chimiste un cul-de-sac dont on ne savait quoi faire. C'est maintenant un ensemble dynamique de produits utile à la découverte de structures ou de propriétés nouvelles.

### Références

- [1] Lehn J.-M., *Chem. Eur. J.*, **1999**, *5*, p. 2455 ; Eliseev A.V., Lehn J.-M., *Current Topics in Microbiology and Immunology*, **1999**, *243*, p. 159 ; Huc I., Nguyen R., *Combinatorial Chemistry & High Throughput screening*, sous presse.
- [2] Youinou M.-T., Rahmouni J., Fisher J., Osborn J.A., *Angew. Chem. Int. Ed. Engl.*, **1992**, *31*, p. 733.
- [3] Rowan S.J., Brady P.A., Sanders J.K.M., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, p. 2143.
- [4] Hasenknopf B., Lehn J.-M., Kneisel B.O., Baum G., Fenske D., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, p. 1838 ; Hasenknopf B., Lehn J.-M., Boumediene N., Dupont-Gervais A., Van Dorsselaer A., Kneisel B., Fenske D., *J. Am. Chem. Soc.*, **1997**, *119*, p. 10956.
- [5] Eliseev A.V., Nelen M.I., *Chem. Eur. J.*, **1998**, *4*, p. 825.
- [6] Berl V., Huc I., Lehn J.-M., DeCian A., Fischer J., *Eur. J. Org. Chem.*, **1999**, p. 3089.
- [7] Berl V., Krische M.J., Huc I., Lehn J.-M., Schmutz M., *Chem. Eur. J.*, **2000**, p. 1938.
- [8] Huc I., Krische M.J., Funeriu D.P., Lehn, J.-M., *Eur. J. Inorg. Chem.*, **1999**, p. 1415.
- [9] Jolliffe K.A., Crego Calama M., Fokkens R., Nibbering N.M.M., Timmerman P., Reinhoudt D.T., *Ang. Chem. Int. Ed.*, **2000**, *39*, p. 755.
- [10] Cousins G.R.L., Poulsen S.-A., Sanders J.K.M., *Chem. Commun.*, **1999**, p. 2279.
- [11] Hioki H., Still W.C., *J. Org. Chem.*, **1998**, *63*, p. 904 ; Shibakami M., Inagaki M., Regen S.L., *J. Am. Chem. Soc.*, **1998**, *120*, p. 3758 ; Cha X., Ariga K., Onda M., Kunitake T., *J. Am. Chem. Soc.*, **1995**, *117*, p. 11833.
- [12] Klekota B., Miller B.L., *Tetrahedron*, **1999**, *55*, p. 11687.
- [13] Swann P.G., Casanova R.A., Desai A., Frauenhoff M.M., Urbancic M., Slomczynska U., Hopfinger A.J., Le Breton G.C., Venton D.L., *Biopolymers*, **1996**, *40*, p. 617.
- [14] Huc I., Lehn J.-M., *Proc. Natl. Acad. Sci. États-Unis*, **1997**, *94*, p. 2106.
- [15] Ramström O., Lehn J.-M., *ChemBioChem*, **2000**, *1*, sous presse.
- [16] Katz B.A., Finer-Moore J.S., Mortezaei R., Rich D.H., Stroud R.M., *Biochemistry*, **1995**, *34*, p. 8264.



# Synthèse et application de polymères fonctionnels à structure poreuse pour la chimie combinatoire

**Anthony Mercier<sup>\*\*\*</sup>** doctorant, ingénieur chimiste, **Hervé Deleuze<sup>\*</sup>** maître de conférences, **Olivier Mondain-Monval<sup>\*\*</sup>** maître de conférences

**Summary :** *Syntheses and applications of emulsion-derived ultraporous functional polymers*

*We describe the preparation and post-functionalization of an emulsion-derived polymeric foam, called polyHIPE<sup>®</sup>, bearing pendant vinyl functionalities. Particularly of great interest is the low resistance to flow, due to large pores and large channels, making possible the functionalization of a monolithic column of polyHIPE<sup>®</sup> by simple circulation of reactants with low applied pressure. In the presence of a free-radical initiator, compounds such as HBr and thiols underwent a regioselective anti-Markovnikov addition to the residual vinyl groups. These functional porous polymers find applications in the field of polymer-supported catalysis, solid-phase organic synthesis, quenching and combinatorial chemistry.*

**Mots clés :** (Vinyl)polystyrène PolyHIPE<sup>®</sup>, groupement vinylbenzyle, écoulement, thiols.

**Key-words :** (vinyl)polystyrene PolyHIPE<sup>®</sup>, vinylbenzene groups, cross-flow, thiols.

La polymérisation d'une émulsion inverse hautement concentrée, composée uniquement dans sa phase continue de divinylbenzène commercial, conduit à une résine ultraporeuse, interconnectée, constituée d'un grand nombre de groupements vinylbenzyles. Ces groupes vinyliques résiduels font l'objet d'addition radicalaire conduisant à des fonctions supportées séparées du squelette polymère par des espaceurs diméthylènes. D'autre part, la structure poreuse connectée permet une circulation en continu d'espèces chimiques à travers le polymère. L'accessibilité des fonctions greffées ouvre le champ à de nombreuses applications en catalyse supportée, synthèse organique sur support solide et chimie combinatoire.

Les matériaux poreux ont déjà fait l'objet de nombreuses études et sont produits selon des techniques différentes, de la solubilisation de particules inorganiques à l'expansion de gaz, en passant par la séparation de phase. La structure des matériaux obtenus par ces méthodes est souvent irrégulière et difficile à contrôler. Une méthode récente, basée sur la polymérisation d'une émulsion inverse hautement concentrée (HIPE = High Internal Phase Emulsion), a été développée par les chercheurs d'Unilever, conduisant à un matériau

poreux de structure plus régulière. Le terme générique polyHIPE<sup>®</sup> a été attribué à cette mousse polymère [1].

Un travail considérable a été réalisé par Williams *et al.* afin de déterminer les facteurs qui influent sur la structure cellulaire [2] et la taille des cellules [3] du polyHIPE, obtenue avec le système styrène/divinylbenzène comme mélange de monomère constituant la phase continue de l'émulsion inverse hautement concentrée. Sherrington *et al.* ont décrit la préparation de matériaux polyHIPE à haute surface spécifique interne, utilisant une méthodologie similaire à celle utilisée pour la préparation de billes de polymères poreuses en présence d'un porogène [4].

Les polyHIPE sont caractérisés par une densité très faible inférieure à 0,1 g.cm<sup>-3</sup>, due à un volume poreux

élevé et une complète interconnexion entre les cellules adjacentes [5].

La formation, les propriétés, la stabilité d'une émulsion inverse hautement concentrée ainsi que la méthode générale de préparation d'un polyHIPE ont fait l'objet d'un précédent article [6].

A travers les années, de nombreuses tentatives ont été entreprises pour fonctionnaliser des polymères réticulés, insolubles par modification de précurseurs polymères plutôt que par copolymérisation de monomères fonctionnels [7-8].

**Précurseur  
de fonctionnalisation :  
le (vinyl)polystyrène polyHIPE**

En ce qui concerne la composition de la matrice polymère, le système le

\* Laboratoire de chimie organique et organométallique, UMR 5802, 351, cours de la Libération, 33405 Talence Cedex.  
Tél. : 05.56.84.24.25. Fax : 05.56.84.69.94.  
E-mail : a.mercier@lcoo.u-bordeaux.fr

\*\* Centre de Recherche Paul Pascal CNRS, avenue A.Schweitzer, 33600 Pessac.  
Tél. : 05.56.84.56.69. Fax : 05.56.84.56.00.  
E-mail : mondain@crpp.u-bordeaux.fr

plus étudié est celui du styrène/divinylbenzène avec très peu d'agent réticulant, le divinylbenzène.

Cependant, des études récentes ont montré que la polymérisation de cet agent réticulant n'est que partielle et que l'on peut tirer profit des groupements vinyliques résiduels, n'ayant pas été réticulés, pour une fonctionnalisation a posteriori [9]. C'est pourquoi nous avons eu l'idée de fabriquer un polyHIPE composé uniquement de divinylbenzène commercial (80 % divinylbenzène/20 % éthylvinylbenzène) (figure 1).

Nous avons appelé cette matrice polymère (vinyl)polystyrène polyHIPE, noté Ps-CH=CH<sub>2</sub>. Elle comporte de nombreux groupements vinylbenzyles révélés et quantifiés par spectroscopie infrarouge à transformée de Fourier : le taux de groupes vinyliques pendants a été estimé entre 3,0 et 4,0 mmol/g.

## Propriétés physiques et chimiques

Les propriétés chimiques de ces matériaux sont pour l'essentiel identiques à celles de leurs analogues non poreux, préparés par les voies classiques de la polymérisation.

Par contre, les propriétés physiques sont considérablement changées. Les propriétés les plus remarquables sont les suivantes :

- Densité macroscopique très faible : de 0,02 g/cm<sup>3</sup> à 0,25 g/cm<sup>3</sup> selon la valeur de  $\phi$  imposée.
- Volume de vide correspondant au volume des pores supérieur à 74 % (par exemple 90 % de volume de pore pour 10 % de polymère pour  $\phi = 0,90$ ).
- Macroporosité avec des tailles de pores allant de 5 à 25  $\mu\text{m}$ .
- Structure complètement interconnectée et uniforme avec toutes les cellules connectées à leurs voisines. Cette structure est clairement mise en évidence par microscopie à balayage électronique. On constate le nombre important de petites fenêtres de taille 1 à 5  $\mu\text{m}$ , entre les pores adjacents.

• Possibilité de percolation de liquide sans qu'il soit nécessaire d'appliquer de fortes pressions. L'écoulement à travers le polymère obéit à l'équation de Hagen-Poiseuille sui-

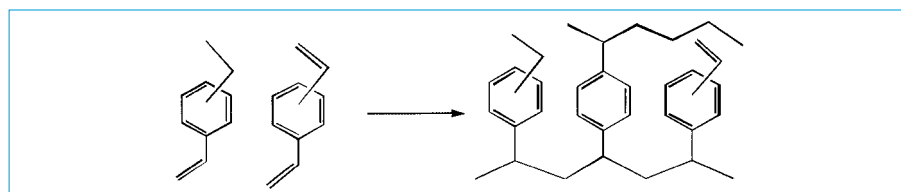


Figure 1 - Polymérisation d'une émulsion inverse hautement concentrée de divinylbenzène commercial (méta:para 2:1).

vante, reliant la pression exercée à la vitesse d'écoulement :

$$\Delta P = \frac{8\eta L}{\Phi r^2} v$$

( $\Delta P$  : pression exercée sur le liquide ;  $\eta$  : viscosité de la phase mobile ;  $L$  : longueur du monolithe ;  $\Phi$  : fraction volumique de pore ;  $r$  : rayon moyen d'une connexion entre deux pores adjacents).

• Grande accessibilité des sites internes rendant possible une modification chimique à l'intérieur du polymère.

## Fonctionnalisation du précurseur

### Écoulement à travers la structure à cellule ouverte

La taille relativement élevée des pores associée au caractère interconnecté de la structure des polyHIPE permet la circulation de réactifs en solution à travers le monolithe, sous faible pression. Nous avons ainsi pu réaliser au laboratoire les fonctionnalisations du précurseur (vinyl)polystyrène polyHIPE placé dans une colonne, par écoulement continu de réactifs dans des solvants adéquates.

### Addition radicalaire régiosélective

Nous avons réalisé des additions de molécules fonctionnelles sur les groupements vinyliques du précurseur (vinyl)polystyrène polyHIPE® par ce procédé d'écoulement. Ces additions peuvent se produire avec des réactifs électrophiles, avec des boranes, avec des thiols ou encore d'autres radicaux libres [10-11]. Les additions de type Markovnikov, en  $\alpha$  du noyau

benzénique, conduisent à des fonctions supportées sur polymère du type Ps-CHX-CH<sub>3</sub>, et sont indésirables du fait de la proximité des groupements fonctionnels greffés vis-à-vis du squelette polymère : les liaisons C-X sont relativement fragiles du fait de la stabilisation par résonance des charges potentielles créées lors d'éventuelles ruptures [12]. Par contre, les additions anti-Markovnikov, en  $\beta$  du noyau benzénique, qui représentent la plupart des additions radicalaires, créent un espaceur diméthylène entre la fonction X et le noyau benzénique, augmentant ainsi la stabilité et l'activité des fonctions supportées du fait d'une plus grande mobilité et accessibilité [13].

En particulier, l'addition radicalaire de thiols HS-R présente l'intérêt de conduire exclusivement à l'isomère anti-Markovnikov, dû à la formation en premier lieu du radical intermédiaire le plus stable (figure 2).

L'addition radicalaire sur les groupements vinylbenzyles nous a ainsi permis d'obtenir un large panel de polyHIPE fonctionnels, comme le montre la figure 3, avec des taux de fonctionnalisations relativement élevés [14].

## Applications en chimie combinatoire : chimie organique sur support solide, scavenging

La diversité des polymères fonctionnels obtenus ouvre le champ à de nombreuses applications en chimie organique sur support solide et scavenging.

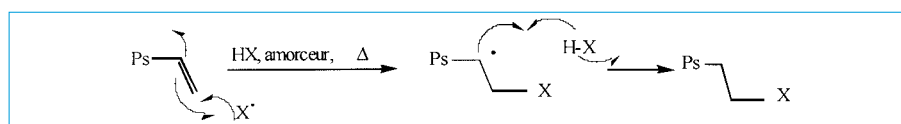


Figure 2 - Addition radicalaire anti-Markovnikov sur le (vinyl)polystyrène polyHIPE®.

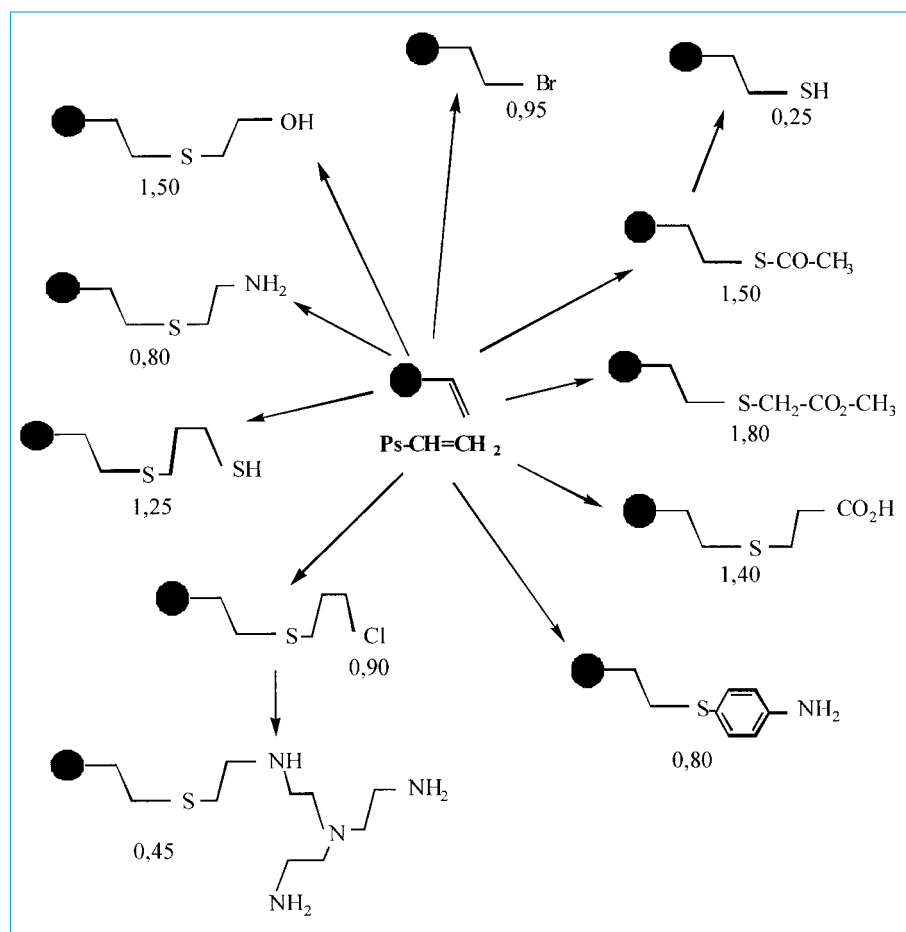


Figure 3 - (Vinyl)polystyrène polyHIPE® fonctionnels synthétisés par la méthode par écoulement continu (taux de fonctionnalisation en mmol/g).

### Synthèse organique sur phase solide (SOPS)

Il est possible de modifier les polyHIPE fonctionnels obtenus par addition radicalaire de petites molécules, en utilisant les voies classiques de la chimie. Nous avons ainsi pu modifier certaines fonctions supportées (le plus souvent substitutions nucléophiles, réductions, oxydations, formation de liaison C-C) et avoir accès à des groupes fonctionnels supplémentaires, ainsi que des entités plus complexes

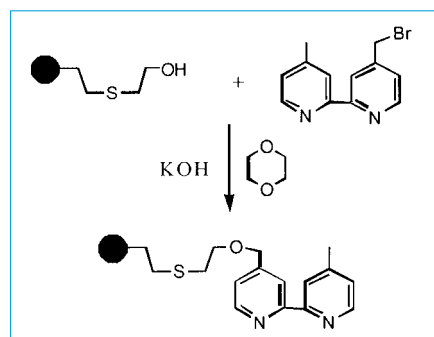


Figure 4 - Synthèse sur phase solide d'un ligand bipyridinique pour la complexation de métaux de transition.

comme des ligands pour la complexation de métaux de transition (figure 4) [15] ou encore des liaisons peptides supportées (figure 5) [16].

### Extraction sélective

L'intérêt majeur de la SOPS est la facilité avec laquelle la purification peut être réalisée : un simple lavage permet la séparation des espèces en solution du polymère, rendant alors possible l'emploi d'excès de réactifs pour atteindre une réaction complète. Dans cette approche, des produits de réaction peuvent être formés à partir

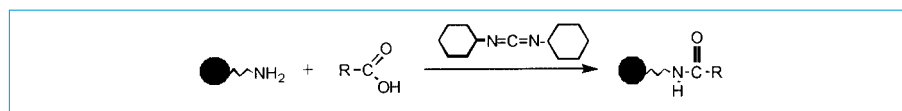


Figure 5 - Formation d'une liaison peptidique sur phase solide.

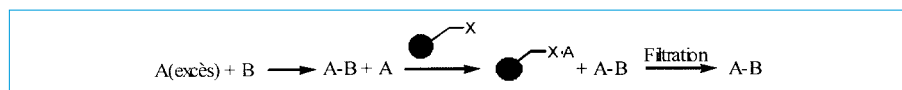


Figure 6 - Piégeage d'un réactif A par formation de liaison covalente avec un polyHIPE® fonctionnel.

d'une réaction en solution, et l'excès de réactif nécessaire à un bon taux de conversion peut être facilement éliminé grâce à une étape ultérieure dite de « scavenging » impliquant la formation d'une liaison covalente avec une espèce électrophile ou nucléophile supportée sur le polymère (figure 6).

Au laboratoire, nous avons choisi comme réaction test de scavenging, l'élimination de l'isocyanate 2,6-diméthylphénylisocyanate (DMPI) par formation d'un liaison covalente sur un scavenger nucléophile, l'aminoéthylpolystyrène polyHIPE [17]. Au bout de 48 heures, le contrôle par chromatographie gazeuse montre une capacité de piégeage de l'isocyanate de 0,85 mmol par gramme de polymère, soit 75 % des sites aminés du support.

### Conclusion

La polymérisation d'émulsion inverse hautement concentrée permet de fabriquer des polymères ultraporeux à cellules ouvertes. Facile à préparer et peu coûteux, les polyHIPE ont des pores de largeur élevée offrant une accessibilité plus grande des sites, par comparaison avec les billes mésoporeuses préparées par suspension. Nous avons montré que la largeur des cellules et la complète connexion permettent une circulation aisée de liquides à travers la structure poreuse.

Les polyHIPE peuvent ainsi faire l'objet d'automatisation utile en chimie combinatoire, en les utilisant sous forme de monolithes à travers lesquels des réactifs ont la possibilité de circuler et de réagir avec les sites actifs. Un simple lavage suivi d'une filtration permet de séparer l'excès de réactif du polymère, évitant ainsi les procédés souvent longs et coûteux de la purification en solution.

Nous avons montré dans un premier temps que les polyHIPE peuvent être facilement fonctionnalisés par voie radicalaire, ou éventuellement par copolymérisation de monomères fonctionnels. Nous avons mis en évidence dans un second temps la possibilité d'utiliser ce nouveau type de support pour la synthèse organique sur phase solide. Les espèces réactives immobilisées sur le réseau polymère peuvent être utilisées dans des réactions de purification (« scavenger »), où les produits de réaction peuvent être facilement séparés, isolés et purifiés, souvent par simple filtration et lavage du support.

Nos efforts se concentrent aujourd'hui sur la mise au point d'un système d'automatisation de circulation de réactifs en solution en utilisant une pompe péristaltique qui réinjecte le mélange

réactionnel de bas en haut de la colonne de polyHIPE. Ce procédé en boucle devrait nous permettre de réduire les temps de réaction et d'augmenter l'efficacité de nos supports fonctionnels.

### Références

- [1] Barby D., Haq Z., *European Patent 0,060,138*, **1982** (to Unilever).
- [2] Williams J.M., Wroblewski D.A., *Langmuir*, **1988**, *4*, p. 656.
- [3] Williams J.M., Gray A.J., Wilkerson M.H., *Langmuir*, **1990**, *6*, p. 437.
- [4] Hainey P., Huxham I.M., Rowatt B.R., Sherrington D.C., Tetley L., *Macromolecules*, **1991**, *24*, p. 117.
- [5] Cameron N.R., Sherrington D.C., *Macromolecules*, **1997**, *30*, p. 5860.
- [6] Mercier A., Deleuze H., Mondain-Monval O., *L'Actualité Chimique*, mai **2000**, p. 10.
- [7] Sherrington D.C., *Chem. Commun.*, **1998**, p. 2275.
- [8] Akelah A., Moet A., *Functionalized Polymers and Their Applications*, Chapman and Hall, New York, **1990**.
- [9] Hubbard K.L., Finch J.A., Darling G.D., *Reactive and Functional Polymers*, **1998**, *36*, p. 1-16.
- [10] Romani F., Passaglia E., Aglietto M., Ruggeri G., *Macromol. Chem. Phys.*, **1999**, *200*, p. 524.
- [11] Stranix B.R., Gao J.P., Barghi R., Salha J., Darling G.D., *J. Org. Chem.*, **1997**, *62*, p. 8987.
- [12] Darling G.D., Fréchet J.M.J., *Chemical Reactions on Polymers*, Benham, Kinstle Eds, ACS, Washington, vol. 364, **1988**.
- [13] Sylvain C., Wagner A., Miokowski C., *Tetrahedron Letters*, **1998**, *30*, p. 9679.
- [14] Mercier A., Deleuze H., Mondain-Monval O., submitted to *Reactive & Functional Polymers*.
- [15] Gosh P.K., Spiro T.G., *J. Am. Chem. Soc.*, **1980**, *102*, p. 5543.
- [16] Bodanszky M & A., *The Practice of Peptide Synthesis*, Springer-Verlag, Berlin/Heidelberg/New York/Tokyo, vol. 21, **1984**.
- [17] Booth R.J., Hodges J.C., *J. Am. Chem. Soc.*, **1997**, *119*, p. 4882-4886.

# Le screening virtuel à haut débit (v-HTS)

Jacques R. Chrétien\* professeur, Marco Pintore\* Ph. D, Frédéric Ros\* docteur

## Summary : Virtual high throughput screening (v-HTS)

Combinatorial chemistry and high-throughput screening (HTS) research programs are widely used in medicinal chemistry and agro-chemistry, in order to select new leads. But, due to the high cost associated to synthesizing and screening a very large number of compounds, there is an increasing need of efficient tools allowing to design and to classify large chemical libraries in order to get enhanced information content. This objective, defined as « data base mining » (DBM), can be achieved by analyzing molecular diversity in large databases with help of the most up-to-date methods based on Kohonen Self Organizing Maps, Fuzzy Logic techniques and Genetic Algorithm. In fact, the proposed methods allow to get a friendly representation of the compound distribution in the hyperspace derived from their molecular descriptors. These derived models of virtual high-throughput screening (v-HTS), linking the structures of the compounds with their biological properties, are suitable to predict activity values for new untested molecules.

**Mots clés :** Screening virtuel, data base mining, DBM, diversité moléculaire, chimie combinatoire

La chimie combinatoire (CC) et le screening à haut débit ou « high throughput screening » (HTS) font l'objet de programmes de recherche ambitieux dans l'industrie pharmaceutique et l'agrochimie en vue de repérer de nouvelles têtes de séries ou « leads » [1-3]. Le but de ces travaux est d'accroître la diversité moléculaire des composés envisageables par la constitution de grandes collections de composés mettant simultanément à profit toutes les combinaisons de synthèses possibles. Ce sont ces nouvelles têtes de série, ainsi détectées, qui vont faire l'objet, dans un second temps, de pharmaco-modulations appropriées pour optimiser au mieux l'activité biologique recherchée. Ces optimisations résultent de modélisation moléculaire par docking, si toutefois la protéine, du récepteur ou de l'enzyme impliquée, a pu faire l'objet d'études 3D cristallographique préalable. Les modélisations moléculaires sont complétées par des études de « relations quantitatives structure activité » (QSAR), grâce aux analyses de champs de potentiels, avec par exemple la méthode « comparative molecular field analysis » (CoMFA) [4-6] qui reste la plus utilisée.

Mais, pour réduire le coût de la synthèse et des tests *in vitro* de grandes collections de composés, un besoin croissant d'une plus grande rationalité dans la conception et la gestion des plans d'expérience de chimie combinatoire se fait sentir. De plus, il y a, souvent encore, un grand chemin à parcourir entre une molécule active, dans un test *in vivo*, et une molécule candidate, en tant que principe actif d'un médicament, surtout lorsqu'elle résulte de synthèse peptidique, en particulier pour des raisons de stabilité ou de biodisponibilité. C'est là que les chimiothèques des industries

concernées, ou les grandes bases de données de molécules accessibles sur le plan commercial ou technique, prennent tout leur sens, comme source de molécules pouvant rentrer dans le champ d'investigation ouvert par le nouveau « lead ». Le rapprochement CC/chimiothèques/grandes bases se situe, alors, en terme d'activité biologique visée, et non en terme de recherche sur une série structurale reposant sur une facilité de synthèse ou un savoir faire particulier en chimie organique classique. C'est ainsi que le concept de diversité moléculaire et sa maîtrise est devenu l'élément central et fédérateur des recherches actuelles de nouvelles molécules bio-actives.

La diversité moléculaire est impliquée à des titres divers dans les trois pôles de la créativité pour la recherche systématique de molécules bio-actives. Ces trois pôles, indiqués à la figure 1, sont les suivants :

- 1. La chimie combinatoire/et le HTS ;
- 2. Le data base mining/ou virtual-HTS (v-HTS) ;
- 3. La modélisation moléculaire/ou docking et QSAR.

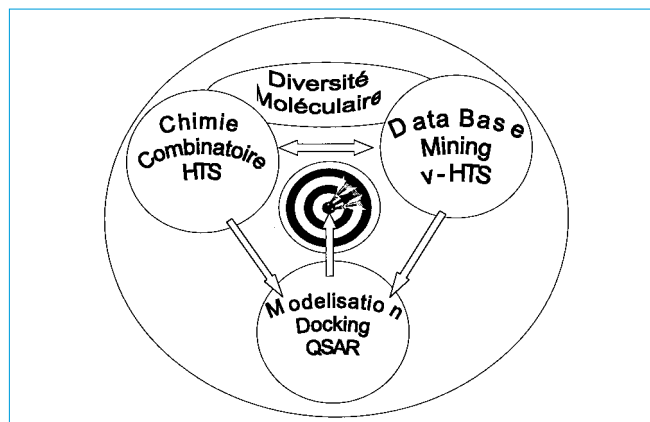


Figure 1 - Les trois pôles de la créativité pour la recherche de nouvelles molécules bio-actives.

\* Laboratoire chimométrie et bioinformatique, Faculté des sciences, Université d'Orléans, BP 6759, 45067 Orléans Cedex 2. Tél. : 02.38.41.70.76. Fax : 02.38.41.72.21. E-Mail : Jacques.Chrétien@univ-orleans.fr

Certes, très souvent, ces trois pôles sont dissociés et mis en œuvre indépendamment, pour des raisons de spécialités, technologiques ou conceptuelles. Et, pourtant, ils présentent de nombreux points de recoupements. En effet, une activité biologique donnée implique des interactions récepteur/ligand qui se ramènent elles-mêmes à des calculs d'interactions physico-chimiques. La maîtrise de ces interactions constitue le fondement de la modélisation moléculaire mais, aussi, des stratégies d'analyse de la diversité moléculaire qui sont l'essence du data base mining (DBM) de grandes bases qui est en plein développement (figure 2).

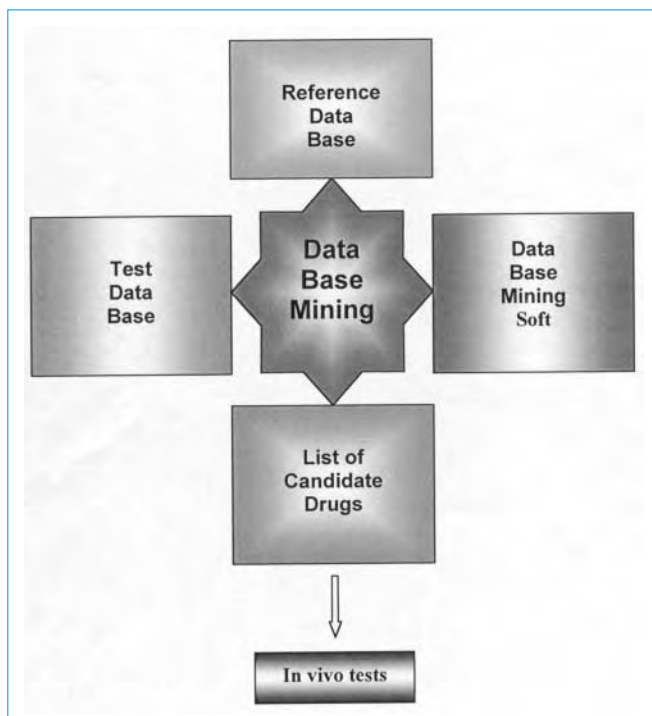


Figure 2 - Principes du data base mining (DBM).

L'explosion de la chimie combinatoire au cours des cinq dernières années a conduit à renforcer l'intérêt général pour la diversité moléculaire. C'est l'exploitation de la diversité moléculaire qui sert de trait d'union entre la chimie combinatoire et le DBM.

Si la chimie combinatoire s'est imposée dans les grands groupes industriels ou dans les sociétés spécialisées au milieu des années 90, la modélisation moléculaire a commencé à s'imposer il y a une quinzaine d'années. Par contre, les bases de données en chimie, documentaires ou factuelles, ont commencé à se développer il y a deux ou trois décennies. Elles correspondaient surtout à des besoins de stockage et de gestion de l'information. Cependant, certaines d'entre elles ont pu servir de bases de connaissances à des systèmes experts, par exemple dans le cas des psychotropes [7] et servir de précurseurs au data mining, en vue de pouvoir estimer, automatiquement, si une molécule nouvelle est susceptible de présenter l'activité neuroleptique ou non. Après l'engouement porté à la modélisation moléculaire, puis à la chimie combinatoire, c'est donc un juste retour des choses, qu'en terme d'épistémologie scientifique, le data base mining permette de restaurer l'intérêt porté aux « grandes

bases maison », ou de mieux appréhender l'intérêt de grandes bases extérieures par la mise en œuvre de véritables stratégies d'exploitation innovantes. C'est là tout l'enjeu du DBM, avec la mise en œuvre de nouveaux concepts et de nouvelles procédures de traitement de l'information chimique, en vue d'aboutir au v-HTS.

## Le data base mining (DBM)

### Le screening virtuel : son objet.

La CC vise à générer, de façon la plus systématique possible, l'ensemble des combinaisons réalisables, cependant avec des limitations d'ordre cinétiques et thermodynamiques. La CC, dans sa phase de mise en œuvre, ne sait pas a priori si les molécules générées seront spécialement intéressantes pour atteindre la cible biologique visée. C'est le HTS qui permettra de conclure. En DBM, l'analyse de la diversité moléculaire vise à générer systématiquement de larges ensembles de descripteurs capables de traduire la complexité des interactions biologique possibles. Les moyens lourds en robotique de la CC sont remplacés, en DBM, par une nécessité de concepts nouveaux et d'algorithmes performants soigneusement validés. L'intérêt fondamental du screening virtuel est d'aboutir à une liste très réduite et la plus sélective possible de molécules présentant l'activité souhaitée, seules ces molécules seront soumises à des tests *in vitro* ultérieurs pour valider la prédiction. En terme d'effort, de coût et de gain de temps, l'enjeu est donc fondamental !

### Générer la diversité moléculaire

#### Descripteurs moléculaires

La diversité moléculaire vise à une approche fragmentaire d'une information chimique globale. Cette information fragmentaire peut être établie sur des bases structurales comme dans le système DARC [8] ou les fingerprints. Ces approches fragmentaires structurales sont excellentes dans un but documentaire. Mais un nombre relativement élevés de fragments est généré, ce qui peut nuire aux besoins de généralité impliqué dans la maîtrise de la diversité moléculaire en vue de la prédiction d'activités biologiques, lorsque les composés sont en nombre très divers et présentent une diversité structurale très grande.

L'information fragmentaire peut être établie sur des bases comportementales, c'est l'objectif de l'utilisation des descripteurs moléculaires. Ils appartiennent en fait à trois classes : descripteurs topologiques, stériques et électroniques. Mais un descripteur moléculaire est par définition réducteur. Le fait de résumer le potentiel d'information biochimique qui se cache dans une structure 3D en un nombre est obligatoirement réducteur et ce, quelle que soit la nature de l'algorithme de calcul mis en jeu. Il engendre obligatoirement une perte d'information. On y remédie en testant de façon systématique un très grand nombre de descripteurs.

Une dégénérescence plus faible de l'information peut être obtenue en prenant mieux en compte la structure moléculaire.

laire 3D, mais il ne faut pas oublier que plus les descripteurs deviennent spécifiques, plus ils perdent de leur pouvoir généraliste dans un but de screening virtuel de grandes bases. Un bon compromis doit donc être trouvé dans le choix entre spécificité et généralité, c'est pourquoi les descripteurs de type 2D sont souvent retenus par rapport aux descripteurs 3D.

### Métrie et organisation

Dans les bases de données de chimie, les composés sont représentés sous forme de graphes qui ne sont pas directement exploitables pour des comparaisons quantitatives. Pour pouvoir estimer a priori l'activité biologique de ces molécules, il faut être capable de les comparer les unes aux autres et de mettre en œuvre des raisonnements analogiques. Il faut passer de représentations sous forme de graphes, à une représentation numérisée, que l'on puisse comparer entre elles sous forme vectorielle, par exemple. Pour cela, on va passer d'une collection de graphes à une distribution dans un espace de descripteurs moléculaires, comme le montre la figure 3.

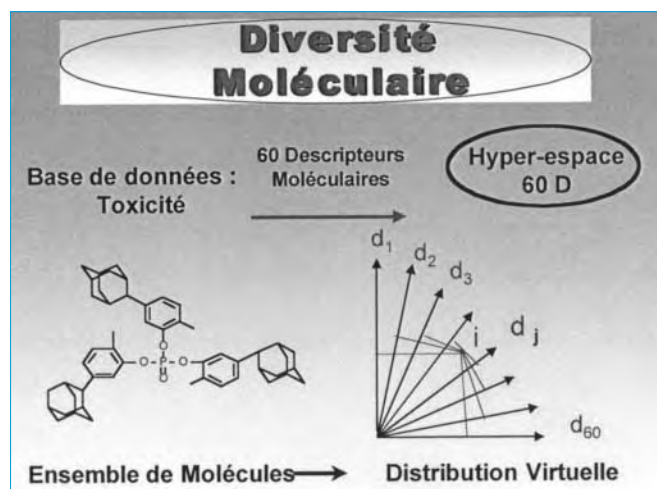


Figure 3 - L'objectif de la diversité moléculaire est de distribuer les composés dans l'hyperespace de leurs descripteurs.

Ainsi, chaque molécule correspond à une position dans l'hyperespace créé. Cet espace virtuel est muni d'une certaine organisation qui n'est accessible que par voie mathématique, en traitant les distances relatives entre les points représentatifs des milliers de molécules constitutives de cette base.

L'objectif de créer une métrique est d'induire dans la base de structure un ordre, une organisation qui reflète au mieux la diversité de la population moléculaire envisagée. Certes, cette organisation reste tributaire du pouvoir discriminant des descripteurs envisagés.

### Visualiser ou modéliser la structure de grandes bases

Un hyperespace est un concept mathématique pour lequel on a du mal à imaginer une représentation intuitive pouvant conduire à une exploitation rationnelle. Une analyse en composante principale (ACP) nécessite facilement une dizaine de plans factoriels, même dans des bases à l'organisation

peu complexe, pour prendre en compte 95 % de l'information. C'est ce qui fait l'intérêt de méthode de projection non linéaire par réseaux de neurones et notamment les « self organizing map » (SOM) de Kohonen [9-12]. La technique SOM conduit à une représentation de la distribution des molécules dans un hyperespace à une représentation dans un espace 2D. L'exemple de la figure 4 illustre la possibilité d'offrir une méthodologie simple pour comparer visuellement l'originalité relative de deux bases de données [13]. La figure 5 montre la possibilité de pouvoir cribler des bases de données en discriminant simultanément quatre mécanismes d'action de molécules anticancéreuses.

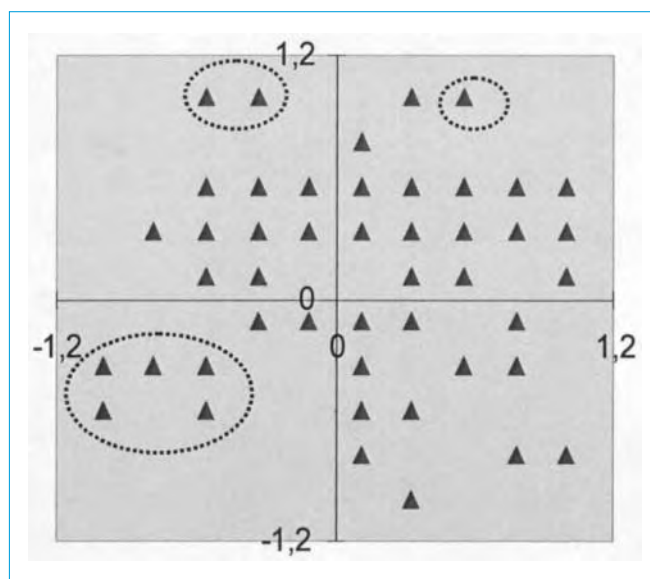


Figure 4 - Comparaison de deux bases de données sur les pesticides organophosphorés, l'une privée, l'autre correspondant aux pesticides commerciaux à l'aide d'une carte SOM de Kohonen. Le degré d'originalité plus grand de la base privée correspond aux composés qui se projettent dans les zones en pointillé de cette carte.



Figure 5 - Screening virtuel et prédiction d'activité pour des molécules anti-cancéreuses :

■ inhibiteurs de l'aromatase, ▲ inhibiteurs de la farnésyl transférase, ● inhibiteurs de la thymidylate synthase, ◆ inhibiteurs de la topoisomérase.

La technique SOM en tant que méthode de projection non linéaire par réseaux de neurones conduit à une certaine distorsion entre les distances relatives intermoléculaires et les distances propres dans l'hyperespace d'origine. C'est pour-

quoi des études directes dans l'hyperespace d'origine ont été entreprises par des méthodes déduites de la logique floue.

Les concepts de logique floue sont également très utiles pour traiter simultanément plusieurs propriétés. Ce cas a été abordé initialement pour les parfums [14-15] où une même molécule peut présenter simultanément plusieurs notes olfactives.

## Conclusion

Les méthodes de travail en chimie organique dans la majorité des laboratoires de synthèse sont en train de subir une révolution sous l'influence de la chimie combinatoire. Elles se manifestent sous la forme de la miniaturisation des procédés de synthèse, des méthodes de séparation et d'analyse par le développement de synthèse sur support solide...

La chimie combinatoire comporte deux niveaux : (i) la constitution de « bibliothèques de composés » et (ii) le screening in vitro (HTS). La chimie informatique y intervient surtout pour dépouiller les séquences de synthèse proprement dites et remonter à la combinaison efficace qui a donné naissance à une molécule active, donc à une possible tête de série.

Il existe toute une méthodologie applicable au DBM qui n'en est qu'à ses débuts. Mais, si nul chimiste organicien ne penserait se soustraire, aujourd'hui, de l'apport des méthodes séparatives (GC, HPLC) ou spectroscopiques (RMN  $^1\text{H}$ ,  $^{13}\text{C}$ , SM...), nous pensons que le screening virtuel, indispensable à l'exploitation de grandes bases de données structurales, deviendra un outil complémentaire de la chimie combinatoire.

Au cours du temps, en chimie pharmaceutique et en agrochimie, l'intérêt s'est déplacé des bases de données vers la modélisation puis, plus récemment, vers la chimie combinatoire. Par un effet de balancier, le data base mining, avec la mise en œuvre de concepts nouveaux et puissants [16], redonne une nouvelle modernité à une compétence qui se banalisait : celle de la gestion des grandes bases de données. Gestion et exploitation poussée des grandes bases de sociétés, c'est l'objectif du data base mining qui est en train de conquérir une place de choix au cœur du processus d'innovation et de drug design. Le fond documentaire qui s'offre au DBM est immense : d'un côté, en terme rétrospectif, les *Chemical abstracts* ont déjà inventorié de l'ordre de 19 millions de molécules diverses, et une action concertée a permis de générer une base de molécules organiques plausibles du milliard de composés. Le DBM a donc encore de beaux jours en perspective au service du drug design !

Si le DBM est en plein développement et atteint sa phase de maturité avec des taux de molécules correctement pré-

dités supérieurs à 80 %, il reste à mettre en place des méthodologies interactives totalement intégrées avec les robots de HTS.

La CC et le DBM procèdent d'une même démarche intellectuelle : générer une hypercomplexité dans la phase préliminaire, éventuellement avec un nombre presque infini de solutions, puis imaginer, dans la phase intermédiaire, des procédures de simplification, reposant soit sur des bases technologiques et robotiques coûteuses, ou des bases mathématiques et informatiques, elles mêmes très complexes, délicates à concevoir mais de coût infiniment moindre dans la mise en œuvre quotidienne. De toute façon, dans les deux cas, les objectifs se recoupent pour aboutir à des réponses, en nombre très limité, en l'occurrence à quelques molécules ou quelques dizaines de molécules.

Une intégration très poussée et des échanges bidirectionnels entre les trois pôles CC/DBM/modélisation sont une nécessité. Pour profiter au mieux de la synergie entre ces trois pôles qui ont leurs contraintes propres, une stratégie unitaire s'impose.

## Références

- [1] Dagani R., *C&EN*, **1999**, March 8, p. 51-60.
- [2] Borman S., *C&EN*, **1999**, March 8, p. 33-48.
- [3] Oldenburg K., *Annual Rep. in Med. Chem.*, **1998**, p. 301-311.
- [4] Kubinyi H., Folkers G., Martin Y.C. *3D QSAR in Drug Design*, vol. 3, Part I « 3D QSAR Methodology. CoMFA and Related Approaches », p. 3-113.
- [5] Cramer III R.D., Patterson J.D., Bunce J.D., *J. Amer. Chem. Soc.*, **1988**, *110*, p. 5959-5967.
- [6] Golbraikh A., Bernard P., Chrétien J.R., *Eur. J. Med. Chem.*, **2000**, *35*, p. 1-14.  
CoMFA / Tripos
- [7] Chrétien J.R., Szymoniak J., Dubois J.E., *Eur. J. Med. Chem.*, **1985**, *20*, p. 315-325.
- [8] Dubois J.E., in *Computer Representation and Manipulation of Chemical Information*, Wipke W.T., Heller S., Feldmann R., Hyde E., Editeurs, Wiley, New York, **1974**, p. 239.
- [9] Kohonen T., *Self Organization and Associative Memory*, Springer-Verlag, Berlin, **1988**.
- [10] van Osdol W.W., Myers T.G., Paull K.D., Kohn K.W., Weinstein J.N., *J. Natl. Cancer Inst.* **1994**, *86*,
- [11] Kireev D.B., Ros F., Bernard P., Chrétien J.R., Rozhkova N., in *Computer-Assisted Lead Finding and Optimisation*, van de Waterbeemd H., Testa B., Folkers G., Wiley-VCH, **1997**, p. 255-264.
- [12] Kireev D.B., Chrétien J.R., Bernard P., Ros F., *SAR and QSAR in Envir. Res.* **1998**, *8*, 93-107.
- [13] Bernard P., Golbraikh A., Kireev D., Chrétien J.R., Rozhkova N., *Analisis*, **1998**, p. 333-341.
- [14] Ros F., Audouze K., Pintore M., Chrétien J.R., *SAR and QSAR in Env. Res.* **2000**, *11* (sous presse).
- [15] K. Audouze, F. Ros, M. Pintore, J.R. Chrétien, *Eur. J. Anal. Chem.*, **2000** (sous presse).
- [16] Data Base Mining Soft. Logiciel développé au Laboratoire de Chimométrie et de BioInformatique de l'Université d'Orléans.



# Similarité moléculaire et criblage virtuel

## Méthodes de recherche *in silico* d'analogues actifs pour la découverte de composés thérapeutiques

**Dragos Horvath\*** Ph.D, head of Molecular Modeling Department, **Catherine Jeandenans\*\*** Ph.D, molecular modeler

**Summary :** *Molecular similarity and virtual screening. In silico methods to retrieve active analogs in the context of discovering therapeutic compounds*

The « similarity paradigm », currently used by the medicinal chemist to design analogs of a known active compound, inferring a property similarity on the basis of structural relatedness, is actually a statistical law : the density of actives in a set of molecules that are similar to an active lead will exceed the density of actives in a set of randomly chosen molecules - if the current definition of « molecular similarity » is a valid one. In the context of combinatorial chemistry and high throughput screening, *in silico* similarity-based queries rely on molecular descriptors and similarity metrics to select analogs of active compounds out of huge electronic molecular databases.

This paper is a brief review of the high throughput similarity search methodology, including a concrete comparative example of the ability of different descriptors and similarity metrics to retrieve active analogs from a library.

**Mots clés :** *Similarité, descripteurs, métriques, échantillonnage conformationnel, librairies combinatoires.*

**Key-words :** *Similarity, 2D and 3D descriptors, conformational sampling, combinatorial libraries.*

Définir sans ambiguïté le concept de « similarité moléculaire » s'avère une tâche très délicate du fait de la coexistence de plusieurs niveaux d'interprétation d'une structure moléculaire, chacun permettant d'aborder le problème de la similarité sous un angle différent. En effet, on peut parler de la similarité « 2D (topologique) » faisant référence à la manière dont les atomes sont interconnectés dans la molécule, d'une similarité « 3D (stérique) » se rapportant à la ressemblance de forme, d'une similarité « pharmacophorique » se basant sur la comparaison des distributions dans l'espace des groupes fonctionnels pouvant interagir avec un récepteur macromoléculaire, enfin d'une similarité « quantique » des densités des nuages électroniques. Traditionnellement, le chimiste médicinal fait appel au concept heuristique de similarité moléculaire pour concevoir des analogues apparentés à un composé actif. En se

fixant comme but l'optimisation de l'activité de ces analogues, il apprend, au fur et à mesure qu'il accumule de l'expérience, la « bonne » pondération de ces divers aspects définissant la similarité moléculaire. Néanmoins, son expertise sera biaisée envers des aspects intuitivement appréhendés par un humain comme la connectivité moléculaire.

La puissance de l'approche informatique de la similarité moléculaire réside en deux aspects :

- La puissance de calcul, permettant le criblage de vastes collections de structures chimiques.

- La possibilité de traiter des aspects clés de la structure moléculaire difficilement intelligibles par un chimiste, comme le motif pharmacophorique ou les « champs » moléculaires de différents types.

Pratiquement, l'exploitation des bases de données moléculaires par « criblage virtuel » exige une méthodologie mathématique et informatique formelle qui doit permettre :

- L'énumération des structures des composés contenus dans les bases de données ;

- La représentation de chaque structure par un descripteur moléculaire, soit un vecteur localisant chaque composé dans un point d'un « espace structural » ainsi défini.

- La définition d'une métrique de (dis)similarité entre deux composés, représentant la « distance » entre les vecteurs descripteurs de ces molécules dans l'espace structural.

Le contexte combinatoire, soit la gestion de millions de composés potentiels, entraîne certaines contraintes dans le choix des outils, en favorisant des approches empiriques et rapides, tout en conservant la faculté de mettre en évidence les aspects importants de la structure moléculaire.

### Méthodologie du « criblage virtuel »

#### Modélisation des produits combinatoires

La gestion informatique de « l'explosion combinatoire » - la génération d'un grand nombre de composés égal au **produit** des nombres de syn-

\*e1\*\* CEREP, 128, rue Danton, BP 50601, 92506 Rueil-Malmaison Cedex.

\* Tél. : 01.55.94.84.00. Fax : 01.55.94.84.10  
E-mail : d.horvath@cerp.fr

\*\* Tél. : 01.55.94.84.27. Fax : 01.55.94.84.10.  
E-mail : c.jeandenans@cerp.fr

thons impliqués dans la bibliothèque combinatoire - nécessite un algorithme de génération des structures chimiques à partir des motifs structuraux présents dans les réactifs de départ, qui simule *in silico* les ruptures et formations des nouvelles liaisons chimiques observées pendant la réaction de couplage.

Les approches de type « Analog Builder » [1] procèdent en déterminant une sous-structure commune à toute une librairie qui sera « décorée » des différents groupes fonctionnels apportés par chaque synthon. L'étape initiale consiste à extraire ces groupes des réactifs initiaux [2]. L'étape de construction des produits est une opération purement topologique qui consiste à mettre à jour une table de connectivité entre les atomes pour y inclure les nouvelles liaisons créées lors de l'accrochage de ces fragments à la sous-structure centrale. Les groupes interconnectés sont souvent des fragments bidimensionnels des dessins moléculaires et, de toute manière, leurs positions relatives après connexion ne sont pas optimisées afin d'éviter les mauvais contacts stériques. Si la connaissance des aspects stériques s'impose, il est impératif de soumettre les produits issus de la concaténation des synthons, à des algorithmes de conversion « 2D-3D » pour construire une géométrie correcte à partir de la table de connectivité et des valeurs standard des longueurs de liaison, angles de valence, etc. (module CatConf de Catalyst [2]).

Les méthodes orientées vers la définition chimique d'une réaction permettent de décrire n'importe quelle transformation chimique, même complexe, y compris les cyclisations [3]. L'environnement chimique commun à tous les réactifs doit alors être défini pour permettre au logiciel de détecter tous les groupes fonctionnels dans les réactifs initiaux. De plus, des règles de correspondance doivent être établies entre les groupes fonctionnels inclus dans les réactifs et dans les produits finals. L'énumération des produits est à nouveau une opération purement topologique.

Dans le cas particulier où la structure finale du composé combinatoire peut être formellement représentée comme résultant du couplage de deux sous-structures par la formation d'une **seule**

nouvelle liaison ( $A_1 + B_j \rightarrow A_1 - B_j$ ), une méthodologie développée récemment [4] permet l'accès direct à de multiples conformations raisonnables des produits A—B, construites quasi instantanément à partir des géométries prédéfinies des sous-structures A et B. Ces dernières sont issues d'une méthode spécifique d'échantillonnage conformationnel générant des géométries dans lesquelles les points d'ancrage des deux sous-structures A- et -B sont complètement dépourvus de gênes stériques afin que le couplage générant le produit A—B soit stériquement possible. Avec cet algorithme, l'utilisation des modèles multiconformationnels pour décrire des chimiothèques de dizaines de millions de composés combinatoires n'est plus prohibitive en terme d'effort de calcul. L'approche bénéficie donc de « l'avantage combinatoire » qui consiste à effectuer, autant que possible, toutes les étapes coûteuses en temps de calcul au niveau des  $N_A + N_B$  synthons, afin de pouvoir rapidement extrapoler les propriétés désirées aux  $N_A \times N_B$  produits de synthèse sur la base des résultats ainsi acquis.

### Descripteurs moléculaires : 2D ou 3D ?

La discussion reste ouverte sur les performances comparées des descripteurs 2D relativement aux descripteurs 3D [5-6]. L'utilisation de descripteurs 3D peut en effet être source d'artefacts dus à la variance de ces descripteurs par rapport à l'ensemble des conformations d'un composé. Si on génère, par exemple, la géométrie d'un composé à partir de sa table de connectivité avec deux logiciels différents, on pourra obtenir deux conformations soit deux valeurs assez différentes pour un même descripteur 3D - jusqu'au point où ces deux conformères seront perçus comme des espèces complètement dissimilaires par une métrique de similarité 3D. La reproductibilité des descripteurs 3D augmente s'ils sont pris comme valeur moyenne par rapport à un ensemble de conformations représentatives du composé. Ceci exige de procéder à un échantillonnage conformationnel de chaque composé, d'où l'intérêt pour les approches du style de celle présentée dans la référence [4]. Actuellement, les logiciels dédiés au traitement d'un

grand nombre de molécules se limitent plutôt au calcul de descripteurs physico-chimiques (logP, pKa) ou topologiques à partir de la structure 2D des composés [1, 7].

Dans le but d'accélérer encore plus le calcul des descripteurs moléculaires des chimiothèques combinatoires, l'exploitation de « l'avantage combinatoire » a été envisagée par plusieurs auteurs, notamment pour générer rapidement des descripteurs 2D pour les librairies combinatoires dans le contexte des représentations de Markush [8], ou pour définir des descripteurs pseudo-3D des produits comme une concaténation des vecteurs décrivant chaque synthon (sans néanmoins générer une vraie structure 3D du produit fini [9]). L'accès quasi instantané aux conformations des produits [4] ouvre la voie à l'exploitation des descripteurs 3D dans le contexte de chimiothèques de très grande taille [10]. La caractérisation complète de 80 millions de composés combinatoires en termes d'un descripteur pharmacophorique « FBPA - Fuzzy Bipolar Pharmacophore Autocorrelograms » - a pu se faire en un mois de temps CPU sur une station de travail Silicon Graphics.

Le descripteur FBPA [10] représente les densités de distribution des paires d'éléments pharmacophoriques (tels que « hydrophobe-cation », « hydrophobe-aromatique », « donneur-accepteur d'hydrogène »...) en terme des distances géométriques séparant ces éléments, moyennées sur l'ensemble des conformations. Ceci est un cas particulier de descripteur pharmacophorique, se basant sur la logique floue afin de minimiser l'impact des artefacts d'échantillonnage conformationnel. D'autres descripteurs pharmacophoriques « binaires » décrivent l'existence ou absence des motifs structuraux à 3, voire 4 éléments de pharmacophore [11]. Chaque motif, défini par les éléments qui le composent et les intervalles des distances entre ces éléments, est codifié par un bit dans une séquence binaire : un bit « allumé » (1) signalant la présence tandis qu'un bit « éteint » (0) signale l'absence du motif. Ce caractère binaire rend ces descripteurs assez sensibles aux artefacts géométriques, car des bits différents peuvent

être allumés par des conformères différents (un bit sera allumé dans le descripteur moléculaire final si le motif codé par celui-ci est au moins présent dans une conformation).

## Métriques de similarité

En fonction de la nature des descripteurs moléculaires, il existe une grande variété de métriques qui sont des fonctions traduisant une « distance » entre les points occupés par deux molécules dans l'espace structural défini par ces descripteurs. Généralement [12], ces métriques sont soit de type euclidien (mieux adaptés pour la comparaison des vecteurs de descripteurs 2D et 3D classiques), soit des coefficients de corrélation entre deux vecteurs (mieux adaptés pour les descripteurs binaires). Enfin, toujours dans le but de réduire l'impact des artefacts géométriques, une approche introduisant la logique floue est utilisée pour comparer les FBPA. Certaines approches peuvent inclure des paramètres ajustables dans le calcul de la métrique de similarité comme c'est le cas pour la métrique associée aux FBPA. Les métriques peuvent alors être calibrées afin d'optimiser des problèmes spécifiques dans une approche purement mathématique ou, au contraire, de manière à corrélérer l'intuition du chimiste qui examine les composés trouvés similaires par la métrique.

## Validation

Une expérience test a été conçue pour déterminer :

- si l'utilisation de ces outils permettait de diminuer l'effort d'expérimentation à haut débit des chimiothèques dans le but d'y découvrir des composés actifs,
- la performance relative des différentes métriques de similarité dans la détection de composés analogues actifs.

Nous avons considéré deux familles d'inhibiteurs A et B de la protéine farnesyl transférase FT [13-14] (figure 1) pour traiter le problème suivant : supposons qu'un seul inhibiteur de la famille A soit connu et qu'une chimiothèque contienne plusieurs représentants de la classe B, les outils utilisés sont-ils capables, à partir de la

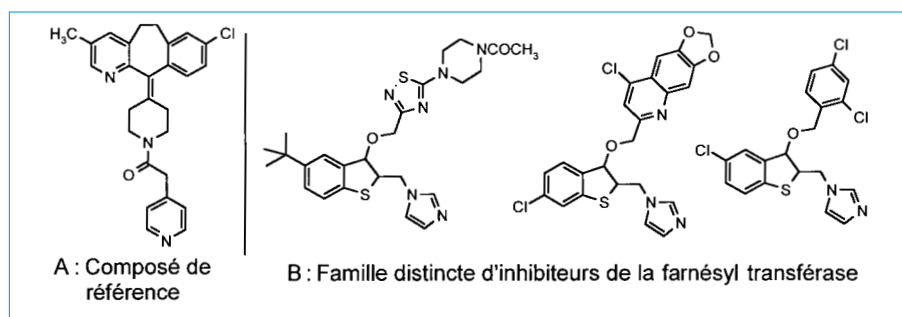


Figure 1 - Composé de référence A inhibiteur de la farnésyl transférase à droite utilisé pour retrouver les composés de type B (exemples des structures à droite) « cachés » dans une librairie combinatoire.

structure de A de classer les composés de types B comme plus proches voisins de A - prouvant ainsi que la stratégie du test sélectif de ces plus proches voisins est supérieure au test à l'aveugle de toute la chimiothèque. Cette chimiothèque est un sous-ensemble de 629 médicaments de la pharmacopée des États-Unis, ne contenant pas d'autres composés dont le rôle thérapeutique majeur soit inhibiteur de FT (bien qu'on ne puisse exclure que certaines de ces molécules aient une activité résiduelle sur cette cible). Parmi ces 629 composés, nous avons réparti 37 inhibiteurs de type B. Les familles d'inhibiteurs de type A et B, bien qu'agissant sur le même site de la protéine, sont structurellement très différentes, leur similarité supposée n'étant pas intuitivement reconnaissable par un chimiste médical.

Le degré de similarité de chaque molécule dans la chimiothèque par rapport à la référence A a été estimé par 4 méthodes différentes - dans les 3 der-

niers cas, les descripteurs ont été calculés à partir des ensembles d'au plus 20 conformations pour chaque molécule :

1. descripteurs 2D, 3D (de forme) et topologiques « standard » calculés par le programme Cerius<sup>2</sup> [1], incluant des informations relatives aux hétéroatomes (états électrotopologiques [7]) ; métrique euclidienne,
2. descripteurs pharmacophoriques binaires à 3 éléments ; métrique de Dice [15],
3. descripteurs pharmacophoriques binaires à 4 éléments ; métrique de Dice,
4. descripteurs et métrique de Dice « floue » FBPA [10].

Les composés des chimiothèques ont été classés par ordre de similarité décroissante avec A et la fraction des composés B « cachés » (y %) trouvés parmi les composés classés comme les plus similaires vis-à-vis de A (x %) est représentée en fonction de x % sur la figure 2. Plus la pente de cette courbe est importante, moins on devra tester de

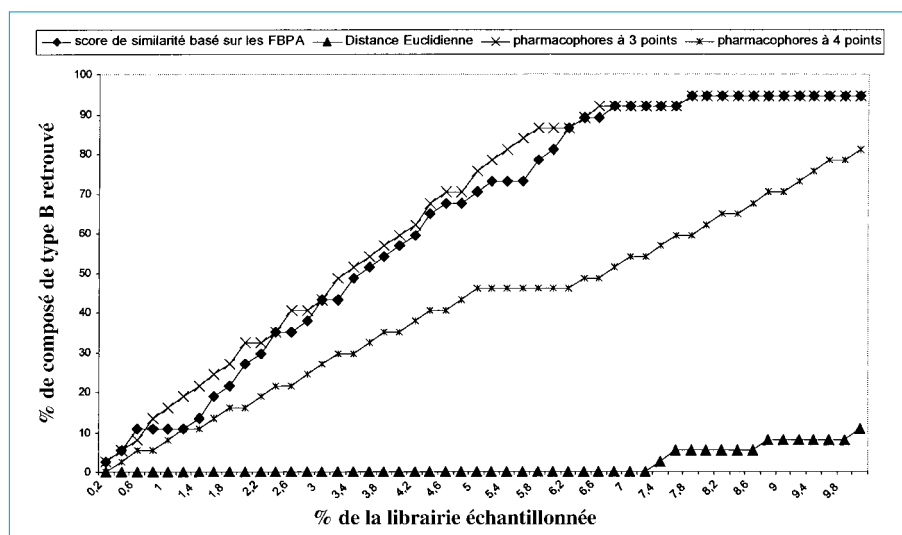


Figure 2 - Capacité de différentes métriques de similarité à retrouver les composés inhibiteurs de la farnésyl transférase de type B, disséminés parmi 629 composés de la pharmacopée US par sélection des plus proches voisins du composé de référence A (figure 1).

composés si on suit l'ordre indiqué par l'indice de similarité pour découvrir une fraction donnée des actifs « cachés » dans la chimiothèque.

En utilisant la métrique de Dice « floue » associée aux FBPA, nous aurions à tester seulement 10 % de la chimiothèque, dans l'ordre de similarité décroissante par rapport à A, pour trouver 90 % des composés inhibiteurs de type B. La stratégie utilisant le criblage virtuel basé sur l'approche FBPA est donc plus efficace que le test direct de l'ensemble de la chimiothèque.

En revanche, la métrique euclidienne associée aux descripteurs 2D classiques ne découvre pas la similarité existante entre la référence A et les composés de type B. Pire, les composés de type B apparaissent encore plus éloignés de A que les 629 composés de la librairie, en principe dépourvus d'activité sur FT. Bien que certains descripteurs (comme les clés électrotopologiques) contiennent des informations sur le type et le nombre des hétéroatomes, ils sont impuissants à traduire l'arrangement relatif des groupes fonctionnels des molécules. Tandis que l'analyse des motifs pharmacophoriques avec la technique FBPA arrive à mettre en évidence une similarité peu évidente entre A et B, l'approche basée uniquement sur topologie et forme moléculaire échoue.

Les empreintes pharmacophoriques basées sur la définition de pharmacophores à 3 points classent correctement les inhibiteurs de type B dans les composés similaires à A, avec une performance globale comparable à la méthode FBPA. Curieusement, les empreintes pharmacophoriques basées sur la définition de pharmacophores à 4 points sont **moins** performantes que les descripteurs 3D codant la position de paires ou triplets d'atomes. Si on examine les scores de similarité des composés identifiés comme plus proches voisins de A par les descripteurs à 4 points, on constate que le score de similarité est de 0,98 traduisant une similarité presque nulle (score de similarité maximum = 0, dissimilarité maximum = 1). On en déduit qu'aucun des composés de la librairie (y compris les inhibiteurs de type B) ne possède pratiquement de motifs pharmacophoriques à 4 points en commun avec la référence. Ils possèdent néan-

moins des motifs **apparentés**, mais qui sont représentés par d'autres bits allumés dans les empreintes. On touche ici à la limitation de ces descripteurs 3D qui doivent contenir une certaine notion de « flou » pour compenser les problèmes dus à la comparaison de valeurs discrètes dans une métrique de similarité. Même si les descripteurs à 4 points contiennent potentiellement plus d'information sur l'arrangement 3D des éléments pharmacophoriques, et donc devraient permettre de classer plus précisément les composés en matière de similarité, les problèmes dus aux artefacts géométriques annulent complètement l'avantage de la définition précise des pharmacophores. L'augmentation de la taille de la famille des conformations utilisées pour la construction de cette empreinte peut partiellement atténuer ce problème.

## Conclusion

L'apport des algorithmes de reconnaissance de la similarité moléculaire à la recherche d'analogues actifs peut, au-delà de la rapidité de l'outil informatique, se concrétiser par la découverte de composés qui présentent l'activité souhaitée, même si leur ressemblance avec le composé de référence n'est pas évidente. En effet, en utilisant des approches basées sur l'analyse des motifs pharmacophoriques 3D, le criblage virtuel met en évidence des similitudes significatives pour l'activité biologique, mais difficilement perceptibles par le chimiste. Bien que coûteuse en temps de calcul, cette analyse, nécessitant la construction de modèles conformationnels 3D, peut désormais être appliquée au traitement de grandes chimiothèques combinatoires. Les artefacts engendrés par l'utilisation des descripteurs 3D peuvent être minimisés par une définition appropriée de la métrique de similarité utilisant la logique floue par exemple.

## Références

[1] *Cerius<sup>2</sup> v 4.0*, Molecular Simulations Inc., San Diego, California.  
 [2] *Catalyst v 3.5*, Molecular Simulations Inc., San Diego, California.  
 [3] *The Daylight Theory Manual*, <http://www.daylight.com>.

[4] Horvath D., Deprez B., Tartar A.T., High throughput molecular modeling using « Fast 3D » descriptors, *Act. Chim. Ther.*, **1997**, 23, p. 55-69.  
 [5] Matter H., Selecting optimally diverse compounds from structure databases : a validation study of two-dimensional and three-dimensional molecular descriptors, *J. Med. Chem.*, **1997**, 40, p. 1219-1229.  
 [6] Brown R.D., Martin Y.C., The information content of 2D and 3D structural descriptors relevant to ligand-receptor binding, *J. Chem. Inf. Comp. Sci.*, **1996**, 36, p. 572-584.  
 [7] Hall L.H., Kier L.B., Electrotopological state indices for atom types : a novel combination of electronic, topologic and valence state information, *J. Chem. Inf. Comp. Sci.*, **1995**, 35, p. 1039-1045.  
 [8] Brown R.D., Downs G.M., Barnard J.M., Use of Markush structure-analysis techniques for rapid processing of large combinatorial libraries, Conference at the 218<sup>th</sup> National ACS Meeting, New Orleans, Louisiana, Aug. 22-26, **1999**.  
 [9] Sybyl 6.5, module Legion<sup>TM</sup>, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, États-Unis.  
 [10] Horvath D., « High throughput conformational sampling and fuzzy similarity metrics : a novel approach to similarity searching and focused combinatorial library design and its role in the drug discovery laboratory », chapitre à paraître dans le livre de Ghose & Viswanadan édité par Marcel Dekker, N-Y.  
 [11] Mason J.S., Morize I., Menard P.R., Cheney D.L., Hulme C., Labaudiniere R.F., New 4-point pharmacophore method for molecular similarity and diversity applications : overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures, *J. Med. Chem.*, **1998**, 38, p. 144-150.  
 [12] Menard P.R., Mason J.S., Morize I., Bauerschmidt S., Chemistry space metrics in diversity analysis, library design and compound selection, *J. Chem. Inf. Comp. Sci.*, **1998**, 38, p. 1204-1213.  
 [13] Njoroge F.G., Vibulbhan B., Rane D.F., Bishop W.R., Petrin J., Patton R., Bryant M.S., Chen K.-J., Nomeir A.A., Lin C.-C., King I., Liu M., Chen J., Lee S., Yaremko B., Dell J., Lipari P., Malkowski M., Li Z., Catino J., Doll R.J., Girijavallabhan V., Ganguly A.K., Structure-activity relationship of 3-substituted N-(pyridinylacetyl)-4-(8-chloro-5,6-dihydro-11H-benzo[5,6]cyclohepta [1,2-b]pyridin-11-ylidene)-piperidine inhibitors of farnesyl-protein transferase : design and synthesis of in vivo active antitumor compounds, *J. Med. Chem.*, **1997**, 40, p. 4290-4301.  
 [14] Kaminski J.J., Rane D.F., Snow M.E., Weber L., Rothofsky M.L., Anderson S.D., Lin S.L., Identification of novel farnesyl-protein transferase inhibitors using three-dimensional database searching methods, *J. Med. Chem.*, **1997**, 40, p. 4103-4112.  
 [15] Willett P., Barnard J.M., Downs G.M., Chemical similarity searching, *J. Chem. Inf. Comp. Sci.*, **1998**, 38, p. 983-996.

## ...un outil complémentaire pour l'instant

Les différents articles de ce numéro spécial amènent certains éléments de réponse à la question soulevée dans l'éditorial « La chimie combinatoire : nouvelle discipline ou stratégie de développement ? ».

L'utilisation d'équipements capables de travailler à haut débit (robot de synthèse ou d'évaluation), qui permet en fait simplement un gain de temps, ne constitue pas en soit une nouvelle approche de la recherche. Par contre, le temps gagné peut être valorisé pour tester de très nombreuses combinaisons qui n'auraient pu être étudiées par l'approche traditionnelle faute de temps. Aussi, les outils de la chimie combinatoire sont de plus en plus employés aussi bien dans le milieu académique qu'industriel. Dans ce sens, il est possible de parler d'une **stratégie de développement avec comme priorité la compétitivité**.

Néanmoins, il existe une première différence importante entre l'approche traditionnelle et l'approche combinatoire. L'approche traditionnelle se place dans une logique itérative où les conclusions de chaque expérience servent à planifier la suivante jusqu'à l'obtention du résultat visé. Le raisonnement de l'approche combinatoire est inverse : sur la base d'une définition exacte de l'objectif, un test de sélection est mis en place ; ensuite, une banque est préparée et évaluée. Par cette dernière approche, seules les molécules qui passent au travers de ce test, et qui présentent donc le profil adapté à l'objectif visé, seront resynthétisées et caractérisées. A cette première différence, il faut ajouter une approche expérimentale différente. En effet, si la base scientifique et la compétence des chercheurs dans leur domaine est indispensable pour ces deux types de recherche, il faut, dans le cas de la chimie combinatoire, ajouter un raisonnement supplémentaire qui prend en compte le nombre très important de molécules qui sont préparées et testées simultanément. S'assurer de la qualité des bibliothèques et de la pertinence du test s'effectue en effet différemment de la caractérisation et de l'évaluation d'une seule molécule. Néanmoins, ces différences ne suffisent pas pour parler réellement d'une nouvelle discipline. L'approche combinatoire pourrait donc, **pour l'instant**, plutôt être décrite comme **un outil complémentaire pour la recherche traditionnelle**. Un outil nouveau, particulièrement efficace mais **qui doit encore progresser pour atteindre une large gamme d'applications**.

**Pourquoi pour l'instant ?** En fait, l'énorme quantité de données générées par la chimie combinatoire représente une mine d'informations qui n'est, pour le moment, utilisée que d'un point de vue qualitatif (système intéressant/système non intéressant). De récentes études visent à appliquer la chimiométrie d'une part pour extraire, par des traitements informatiques, les informations cachées parmi une multitude de résultats et, d'autre part, pour planifier les protocoles de préparation des bibliothèques. Ceci pourrait entraîner très prochainement les chimistes à prévoir leurs expériences non plus uniquement sur la base de leur compétences mais en s'appuyant en plus sur l'analyse statistique des résultats précédemment obtenus par chimie combinatoire. Dans ce sens, l'association chimie combinatoire/chimiométrie pourrait être décrite comme une nouvelle discipline.

*Au nom de tous les auteurs ayant participé à ce numéro spécial, je tiens à remercier chaleureusement Anna Proust et Bernard Sillion pour nous avoir donné l'opportunité de présenter cette approche qui concerne l'ensemble de la chimie.*

Olivier Lavastre