

bu 2010 341

50376
1999
391

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

Université de Lille I



**Thèse pour obtenir le grade de Docteur en Chimie Organique et
Macromoléculaire**

Présentée et soutenue publiquement le 15 Octobre 1999

par

XAVIER WILLIARD

**MISE AU POINT ET UTILISATION DE TECHNIQUES
DE SYNTHÈSE COMBINATOIRE ET DE SYNTHÈSE
PARALLELE A HAUT DEBIT POUR LA
DECOUVERTE D'ACTIVITES BIOLOGIQUES**

Directeur de thèse : Pr. André TARTAR

JURY :

Président :

Rapporteurs :

Examineurs :

Pr. Olivier Lafont

Pr. Henri Moskowitz

Pr. Jean-Claude GESQUIERE

Dr. Frédéric REVAH

Dr. Lydie Pelinski

A Monique,

Mes parents, Vincent et Ingrid

Nicolas, Nathalie et Stéphane, Nathalie et

Arnaud

REMERCIEMENTS

La première partie de ce travail a été effectuée au sein du service Chimie des Biomolécules URA CNRS 1309 (Institut Pasteur de Lille) dirigé par André Tartar puis, pour la deuxième partie, au sein de la société CEREP.

Ce travail a fait l'objet d'une collaboration avec les laboratoires GLAXO-WELLCOME, Centre de Recherche des Ulis, dirigé par François Hyafil.

Il a été réalisé grâce à une bourse mixte Conseil régional du Nord pas de Calais / laboratoires GLAXO-WELLCOME

Je tiens à exprimer ma profonde reconnaissance à Monsieur le Professeur André Tartar pour m'avoir accueilli dans son laboratoire et pour avoir fait le pari de confier à une jeune équipe le thème de recherche naissant qu'était la chimie combinatoire, dès 1992. Il a toujours réussi à orienter nos recherches dans des directions fructueuses tant sur le plan scientifique que sur le plan humain.

Je tiens à exprimer mon profond respect à Benoît Déprez pour avoir brillamment réussi à encadrer des doctorants de sa génération, en leur faisant profiter à la fois de sa culture scientifique, de sa culture générale mais aussi de son goût pour la réussite. Il a toujours constitué un des éléments clé du succès de nos projets.

Monsieur le Professeur Jean-Claude Gesquière, je tiens à vous remercier pour m'avoir donné goût à la chimie organique lors de vos enseignements à la faculté de Pharmacie de Lille, pour avoir accepté de faire partie de mon jury de thèse mais aussi, en tant que gardien de but et de par vos arrêts spectaculaires, pour avoir évité à l'équipe de football de l'Institut Pasteur de Lille de perdre par de trop grandes différences de buts.

Monsieur le Professeur Olivier Lafont, je suis profondément honoré que vous ayez accepté d'évaluer et de commenter ce travail en tant que rapporteur.

Madame le Professeur Annie Marcincal, je tiens à vous remercier pour m'avoir donné goût à la chimie organique lors de vos enseignements à la faculté de Pharmacie de Lille mais aussi pour m'avoir aidé à constituer ce jury de qualité.

Monsieur le Professeur Henri Moskowitz, je suis profondément honoré que vous ayez accepté d'évaluer et de commenter ce travail en tant que rapporteur.

Madame Lydie Pelinski, je vous remercie vivement pour avoir accepté d'évaluer et de commenter ce travail en tant qu'examinatrice.

Monsieur Frédéric Revah, je tiens à vous remercier pour avoir accepté de faire partie de mon jury de thèse mais aussi pour la culture industrielle que vous apportez à la société CEREP.

Je tiens aussi à remercier Monsieur Thierry Jean, PDG de la société CEREP, pour m'avoir permis de rédiger une partie de ce manuscrit pendant mes heures de travail.

Une liste exhaustive des personnes m'ayant contribué de près ou de loin à la rédaction et la soutenance de cette thèse serait impossible à dresser ici. La première personne que je voudrais remercier est Monique, merci pour ta présence au quotidien et le bien être qu'elle me procure. Les autres personnes ayant participé activement à ce travail sont Anick Lemaire, Marie-Isabelle Steeland et Valérie Dureux.

Je tiens aussi à remercier les premières personnes avec lesquelles j'ai pu enrichir mes connaissances dans le domaine de la synthèse peptidique : Hervé Drobecq, Marc Bossus, Eric Diésis, Corinne Rommens et tous les autres membres du SCBM. D'autres personnes avec lesquelles je partage ou j'ai partagé, d'agréables moments professionnels doivent évidemment être citées : Romuald, Corinne, Patricia, Iuliana, David, Line, Guillaume, Luc, Maxence, Gilles, Sébastien, Valérie (D) et (M), Marie-Isabelle, Christophe, Catherine, Véronique, Manuel, Fred, Kathy, Béatrice, Pierre... pardon pour celles ou ceux oubliés malencontreusement.

TRAVAUX

PUBLICATIONS

B. Déprez, X. Williard, L. Bourel, H. Coste, F. Hyafil and A. Tartar, « Orthogonal Combinatorial Chemical Libraries » *Journal of the American Chemical Society*, **1995**, *117*, 5405-5406

L. Bourel, X. Williard, I. Pop, R. Baudelle, D. Horvath, B. Déprez, P. Melnyk et A. Tartar, « Synthèse combinatoire : les autoroutes de la diversité » *Actualité Chimique*, déc. **1995**, 33

X. Williard, I. Pop, D. Horvath, R. Baudelle, P. Melnyk, B. Déprez and A. Tartar, « Combinatorial Chemistry : a rational approach to chemical diversity » *Eur. J. Med. Chem.* **1996**, *31*, 87-98

X. Williard and A. Tartar, « A Practical Guide to Combinatorial Chemistry », **1997**, published by the American Chemical Society, 249-277

X. Williard I. Pop et B. Déprez « Synthèse de chimiothèques combinatoires basées sur les squelettes oxadiazolinones et oxadiazolinthiones », demande de brevet déposée le 30 Avril 1999 sous le numéro 99 05553

X. Williard B. Déprez and A. Tartar, « A mild and efficient method for the synthesis of Δ^2 -1,3,4-oxadiazolin-5-ones », à soumettre dans *tetrahedron letters*

COMMUNICATIONS

B. Déprez, X. Williard, L. Bourel, H. Coste, F. Hyafil and A. Tartar, « Orthogonal Combinatorial Libraries », XIIIth Symposium on Medicinal Chemistry, Paris 19-23 Septembre 1994 (poster)

L. Bourel, X. Williard, B. Déprez, H. Coste, F. Hyafil, A. Tartar « Synthèse et activité biologique d'analogues d'un antagoniste de la vasopressine découvert au moyen d'une banque combinatoire », Groupe d'Etude Structure Activité, GESA XXV, Saint Valéry en Caux, 9-12 Mai 1995 (poster)

SOMMAIRE

Remerciements	3
Travaux	5
Sommaire	6
Avant-propos	11

Introduction -Les différentes stratégies de découvertes de têtes de série-	13
---	-----------

A- Les découvertes fortuites	13
B- Les découvertes par une approche rationnelle	17
C- Les découvertes par criblage de molécules sur des tests biologiques	19

Première partie -La synthèse combinatoire en mélange : la première étape d'une course à la diversité-	23
--	-----------

Introduction	23
--------------	----

I Définition et principe de la synthèse combinatoire	26
---	-----------

II Les méthodes chimiques utilisées en synthèse combinatoire	28
---	-----------

II-A Phase solide ou solution ?	28
---------------------------------	----

II-B Les stratégies d'assemblage de la diversité	29
--	----

II-B-1 Assemblage de type oligomérique	29
--	----

II-B-2 Assemblage par décoration d'un squelette de base	31
---	----

II-C Exemple	32
--------------	----

III Organisation des chimiothèques et élucidations structurales des produits conférant l'activité aux mélanges	34
---	-----------

III-A Elaboration d'un modèle général de description des méthodes de déconvolution	36
---	----

III-A-1 Le modèle Matriciel	36
-----------------------------	----

III-A-2 Evaluation biologique d'une chimiothèque	37
--	----

III-B Application du modèle matriciel aux méthodes de déconvolution	38
---	----

III-B-1 La déconvolution itérative	38
------------------------------------	----

III-B-2 Le « balayage positionnel » ou « positional scanning »	45
--	----

III-B-3 Les chimiothèques orthogonales	49
--	----

III-C Comment assurer l'équimolarité des composés synthétisés en mélange ? _____	55
IV <u>Les méthodes d'élucidations structurales des produits conférant l'activité aux mélanges, ne faisant pas appel à la déconvolution</u> _____	61
V <u>Découverte d'un ligand du récepteur V₂ à la vasopressine au moyen de chimiothèques orthogonales</u> _____	63
V-A Conception de la chimiothèque généraliste _____	63
V-A-1 L'objectif _____	63
V-A-2 Les moyens _____	63
V-A-2-1 Choix des molécules à synthétiser _____	63
V-A-2-2 Utilisation d'une nouvelle méthode de déconvolution développée au laboratoire _____	68
V-A-2-3 Définition de l'architecture de la chimiothèque _____	68
V-B Préparation des chimiothèques orthogonales _____	70
V-B-1 Protections choisies pour les différents acides aminés _____	70
V-B-2 Ancrage des premiers acides aminés au support solide : accès aux aminoacyl-résines protégées _____	74
V-B-3 Couplage des acides aminés de la position 2 : accès aux dipeptidyl-résines _____	76
V-B-4 Couplage des acides aminés de la position 3 : accès aux tripeptidyl-résines _____	78
V-B-5 Déprotection et clivage des tripeptides du support solide _____	79
V-C Résultats biologiques _____	80
V-C-1 Conditions générales des tests biologiques _____	80
V-C-2 Résultats obtenus sur le récepteur V ₂ à la vasopressine _____	81
V-C-2-1 Intérêt thérapeutique de la cible biologique _____	81
V-C-2-2 Identification d'un produit actif au cours des tests de déplacement de la vasopressine tritiée du récepteur V ₂ porcin _____	82
V-C-2-3 Discussion _____	85
VI <u>Conclusion de la première partie</u> _____	87

Seconde partie -La synthèse parallèle à haut débit : l'outil de choix pour la découverte de nouvelles molécules biologiquement actives-	89
--	-----------

Chapitre A : L'automatisation en chimie organique	89
Chapitre B : La synthèse robotisée d'urées asymétriques	98
I Etude Bibliographique	98
A Les différentes voies d'accès aux urées	98
A-1 Réactions des amines avec l'urée	98
A-2 Réactions des amines avec l'acide cyanique	99
A-3 Réactions des amines avec les isocyanates	100
A-4 Réactions des amines avec les carbamates	100
A-5 Réactions des amines avec les dérivés carbonylés activés	101
A-5-1 Utilisation du Phosgène et de ses succédanés	101
A-5-1-1 Utilisation du Phosgène	101
A-5-1-2 Utilisation du Triphosgène	103
A-5-1-3 Utilisation du Carbonyldiimidazole	104
A-5-1-4 Utilisation des Chloroformates	105
A-5-2 Utilisation du dioxyde de carbone et de ses dérivés	106
B L'utilisation du carbonyldiimidazole pour la synthèse d'urées	106
B-1 Premières mises en évidence de la réactivité des imidazolides	107
B-2 Réactivité des azolides	107
B-3 Synthèse d'urées au moyen du carbonyldiimidazole	109
II Mise au point de la synthèse robotisée d'urées	113
A Etude comparative du triphosgène et du carbonyldiimidazole	114
A-1 Utilisation du triphosgène	115
A-1-1 Mise au point du protocole expérimental	115
A-1-2 Evaluation de l'importance des conditions anhydres	117
A-1-3 Evaluation de l'influence de la vitesse d'introduction de la première amine	119
A-1-4 Utilisation du DMF comme solvant de réaction	122
A-2 Utilisation du carbonyldiimidazole	123
A-2-1 Evaluation de l'importance des conditions anhydres	123
A-2-2 Optimisation des temps de réaction et des vitesses d'introduction des réactifs	125
A-2-2-1 Mise au point de l'étape d'addition de l'amine A sur le carbonyldiimidazole	125
A-2-2-2 Détermination du nombre d'équivalents de CDI à utiliser	131
A-2-2-3 Etude de l'influence de la vitesse d'addition de l'amine A au CDI, sur la formation des urées symétriques ou asymétriques	134
A-2-2-4 Détermination du temps de réaction de l'amine B	136

A-2-3 Evaluation des possibilités d'extension du protocole de synthèse à une amine A aliphatique _____	137
A-2-3-1 Préparation des urées symétriques _____	138
A-2-3-2 Préparation des urées asymétriques _____	141
A-3 Conclusion _____	144
B <u>Mise au point du protocole final de la synthèse robotisée d'urées au carbonyldiimidazole</u> _____	145
B-1 Objectifs et contraintes de la production robotisée _____	145
B-2 Etude de la mise en solution des amines _____	145
B-3 Etude de la stabilité du CDI en solution dans différents solvants _____	152
B-4 Mise au point des tests d'évaluation de la réactivité du CDI _____	154
B-5 Evaluation de la réactivité de différentes familles de N-nucléophiles _____	155
B-5-1 Préparation des urées symétriques ou des imidazolides issus des représentants des familles de N-nucléophiles _____	158
B-5-2 Préparation d'urées asymétriques aux conditions de dilutions robotiques, à partir des représentants des familles de N-nucléophiles _____	161
B-5-2-1 Adaptation du protocole pour la synthèse d'urées à partir d'anilines _____	161
B-5-2-2 Détermination des N-nucléophiles réactifs _____	165
B-6 Validation du protocole final par la synthèse robotisée de 11 urées de référence _____	168
C <u>Intérêts potentiels d'une réaction secondaire conduisant aux Oxadiazolinones</u> _____	170
C-1 Intérêt biologique des oxadiazolinones _____	170
C-2 Intérêt combinatoire des oxadiazolinones _____	171
C-3 Protocole de synthèse robotisée d'oxadiazolinones _____	172
C-4 Conclusion _____	172
D <u>Application du protocole de synthèse robotisée d'urées à la synthèse de deux chimiothèques</u> _____	173
D-1 Synthèse d'une chimiothèque de 154720 urées pour un partenaire de l'industrie pharmaceutique _____	173
D-1-1 Choix des monomères et architecture de la Chimiothèque _____	173
D-1-2 Planification de la production _____	174
D-1-3 Contrôle qualité de la production _____	175
D-1-4 Améliorations à apporter _____	177

D-2 Synthèse d'une chimiothèque de 4000 urées pour des clients de l'industrie pharmaceutique ou biotechnologique	179
D-2-1 Choix des monomères et architecture de la Chimiothèque	179
D-2-2 Planification de la production	184
D-2-3 Contrôle qualité de la production	184
D-2-4 Résultats de tests biologiques	187
III Conclusion de la seconde partie	189
Conclusion générale	192
Partie expérimentale	195
Glossaire	234
Bibliographie	238

AVANT-PROPOS

L'homme pour survivre, doit à la fois être capable, de se protéger, de se nourrir et de se guérir. L'apport en lipides, protides, glucides, minéraux et oligo-éléments ainsi qu'en substances à vertus médicamenteuses reste essentiellement garanti par la nature. Celle-ci constituant un réservoir impressionnant de structures chimiques aux propriétés physiques et biologiques très diverses fût la première source de produits pharmacologiquement actifs. La présence de tels composés dans les plantes, les champignons ou les organismes marins est souvent corrélée à une fonction de défense contre les effets de diverses toxines, d'agents carcinogènes, mutagènes ou contre l'attaque d'un prédateur. Les premiers produits utilisables en thérapeutique furent découverts de manière fortuite, par les individus ayant su associer une feuille, un fruit, un champignon, la partie d'un animal... à un effet sur le corps humain, par ingestion ou par contact. Il est d'ailleurs à noter que la transmission de ce savoir à d'autres membres d'une espèce, n'est pas l'apanage de la race humaine. Chez les chimpanzés par exemple un comportement d'automédication a pu être observé par ingestion de plantes sélectionnées lorsqu'ils se sentent malades (Lozoya 1997). Puis la chimie thérapeutique a fait son apparition au XIX^e siècle lorsque les physiologistes, chimistes et pharmaciens ont essayé d'isoler les différents produits contenus dans les extraits végétaux ou tissus animaux et de relier les fractions obtenues aux propriétés observées. De nombreux efforts ont alors été consacrés à la conception de modèles animaux pour mimer les pathologies cliniques et pour tester de nouvelles substances. Ces tests s'avérant difficiles à mettre œuvre et souvent coûteux, une tendance actuelle consiste à l'élaboration de tests in vitro facilement automatisables, les progrès scientifiques et techniques permettant d'identifier et de produire des cibles pharmacologiques, pour lesquelles un ligand est connu ou non. La méthode du criblage (screening) de molécules a ainsi pu voir le jour, celle-ci consistant à tester, au hasard, un grand nombre de molécules sur une batterie de tests biologiques. Certaines sociétés, spécialisées dans le criblage à haut débit (highthroughput screening) sont actuellement capables de tester, chaque jour plusieurs milliers de molécules. Toute la difficulté revient maintenant aux chimistes qui tentent «d'alimenter» ces tests, en accélérant les cadences de production. Il semble cependant impossible, en utilisant les techniques classiques de la synthèse organique, d'atteindre la production de plusieurs milliers de molécules par jour dans un même laboratoire. La synthèse combinatoire et la synthèse automatisée laissent entrevoir des solutions à ce défi. Afin de mieux comprendre l'impact que peuvent avoir ces techniques

sur la recherche pharmaceutique, un rappel des différentes étapes de la découverte de nouveaux médicaments introduira cette thèse. Celle-ci traitera ensuite des résultats obtenus grâce à des méthodes de chimie combinatoire et finira par une description des modes opératoires et des techniques.

INTRODUCTION

Les différentes stratégies de découvertes de têtes de série

(Wermuth 1996, Sneader 1996)

Le processus de découverte de médicaments peut se décomposer en deux phases : la recherche et le développement. Dans un cas idéal, le point de départ de la recherche serait l'identification d'une maladie et de la cible thérapeutique associée, se poursuivrait par la caractérisation d'une nouvelle molécule active (tête de série ou «lead») puis par l'optimisation de cette molécule en dérivés plus actifs et plus sélectifs. Le développement débiterait par des études précliniques et par la formulation du composé, puis les études effectuées chez l'animal et enfin chez l'homme lors des phases cliniques aboutiraient à la mise sur le marché du médicament. Afin de minimiser les sommes d'argent engagées dans ces recherches longues et coûteuses, les industries pharmaceutiques tentent de réduire systématiquement la durée de chacune de ces étapes. L'étape à laquelle nous avons choisi de nous intéresser est la découverte des têtes de série.

Jusqu'aux années 70, les découvertes de structures actives originales reposaient sur la connaissance moléculaire des processus biologiques ou sur des événements imprévus, leur observation et la pertinence des interprétations qui s'en suivaient. Les progrès de la biologie, de la biochimie, des techniques analytiques et de l'informatique ont permis ensuite l'apparition d'une démarche rationnelle dans le processus de découverte. Cependant les années 90 resteront sans doute marquées par un engouement tout particulier pour les techniques d'automatisation et de miniaturisation accordant à un «hasard maîtrisé», une place prépondérante dans la découverte des médicaments de demain.

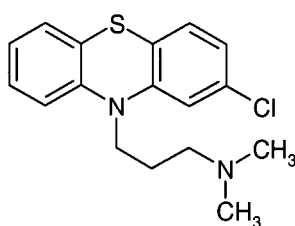
A Les découvertes fortuites

- suite à des observations réalisées sur l'homme

Une source précieuse de substances actives réside dans les **pharmacopées** indigènes. Celles-ci furent constituées au fil des siècles, et permettent encore la découverte de composés très originaux. Les curares, les glycosides cardiotoniques, les opiacés ainsi que les alcaloïdes issus

du quinquina en sont quelques exemples. Afin d'expliquer les effets observés, la théorie des signatures fut proposée. Elle se basait sur le principe que la nature, de part l'apparence ou l'habitat d'une plante symbolisait l'effet pharmacologique de celle-ci : le lierre s'enroulant autour des arbres, les empêchant de prendre de l'envergure devait traiter l'obésité, le saule vivant les pieds dans l'eau devait traiter les maladies causées par les pieds humides, les racines en forme de phallus comme l'asperge européenne et le ginseng chinois devaient traiter les troubles sexuels. Il reste cependant difficile de distinguer les découvertes qui furent réalisées par hasard et auxquelles une signature fut attribuée a posteriori, de celles résultant de l'extrapolation d'une activité, en fonction d'une signature établie a priori.

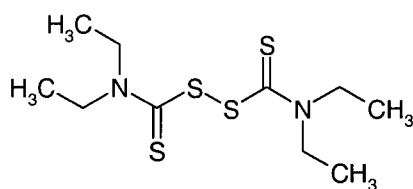
L'observation d'**effets secondaires de médicaments** mis sur le marché, constitue une seconde source de découverte d'activités originales. L'effet sédatif marqué de la prométhazine (utilisé comme antihistaminique) observé par le clinicien H. Laborit en avril 1949, a conduit la société Rhône-Poulenc à mettre en œuvre la synthèse d'analogues de ce composé afin de renforcer cet effet sédatif. Dès le printemps 1951, la chlorpromazine était utilisée par Laborit pour limiter les chocs opératoires, en association avec la prométhazine et la pethidine (association dénommée cocktail lytique).



Chlorpromazine
(neuroleptique)

Laborit a alors constaté que ses patients, non seulement se portaient mieux pendant et après l'opération mais qu'avant l'opération, ceux-ci semblaient particulièrement détendus et détachés des événements extérieurs. Une étude réalisée ensuite par Delay et Deniker a montré que la chlorpromazine, administrée seule, avait un rôle sédatif et soulageait les problèmes de confusion mentale. Chez les schizophrènes, elle entraînait une diminution de leur agressivité, de leur activité et de leurs hallucinations. Cette molécule a révolutionné les traitements utilisés alors en psychiatrie, qui ne consistaient qu'en l'administration d'électrochocs et le port de camisoles de force.

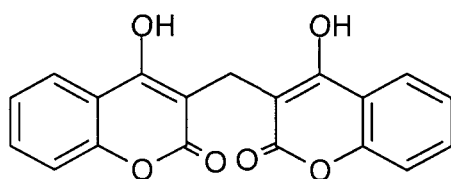
Les effets secondaires observés chez des **personnes exposées** à des produits issus de **l'industrie chimique** peuvent aussi être mis à profit pour la découverte de molécules biologiquement actives. Le disulfiram par exemple, utilisé comme antioxydant dans l'industrie du caoutchouc provoquait chez les personnes manipulant ce produit un dégoût marqué pour l'alcool. Cet effet a été expliqué par l'inhibition d'une enzyme clé du catabolisme de l'alcool éthylique, l'aldéhyde déshydrogénase par le disulfiram. Cette enzyme permettant normalement l'oxydation de l'acétaldéhyde en acide acétique, la prise de disulfiram provoque une accumulation d'acétaldéhyde qui conduit à une sensation de chaleur au niveau de la face, des palpitations, des dyspnées et des nausées. Ce produit a donc été proposé pour les cures de sevrage alcoolique.



disulfiram
(antabuse)

- suite à des observations réalisées sur l'animal

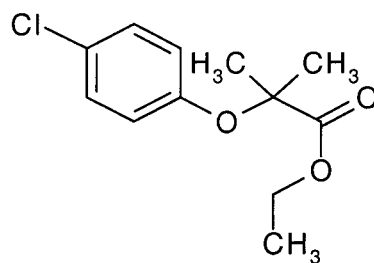
En 1922, aux USA, du bétail fût retrouvé mort après avoir ingéré du foin moisi, préparé à partir de Mélilot. Roderick a alors montré qu'une substance anticoagulante inconnue, avait causé la mort du bétail. En 1939 Campbell a pu isoler des cristaux de ce puissant anticoagulant et Huebner, un an plus tard, a confirmé en le synthétisant, qu'il s'agissait de la 3,3'-methylenebis(4-hydroxycoumarine) ou dicoumarol. Ce composé s'était formé dans le foin fermenté, par oxydation de la coumarine en 4-hydroxycoumarine suivi d'une dimérisation par le formaldéhyde. Les dérivés coumariniques sont actuellement utilisés comme anticoagulant dans le traitement des maladies thromboemboliques.



dicoumarol
(anticoagulant)

- suite à des travaux réalisés sur des végétaux

Au cours d'études menées sur des analogues de l'hormone de croissance végétale : l'acide indolylacétique, des chercheurs ont montré que des dérivés chlorés de l'acide phénoxyacétique, possédaient des propriétés phytohormonales apparentées à celles de l'acide indolylacétique. En 1954, des chercheurs de la société I.C.I ont découvert que certains de ces dérivés agissaient sur le métabolisme humain des lipides et se manifestaient notamment par une diminution plasmatique du cholestérol, des triglycérides et des VLDL. La molécule la plus utilisée actuellement est le clofibrate (prodrug de l'acide chlorophénoxyisobutyrique)

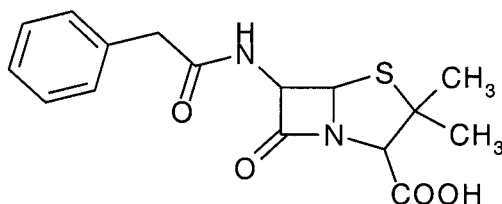


clofibrate
(hypocholestérolémiant)

- suite à des observations réalisées sur des micro-organismes

En 1878, Pasteur proposa le concept de «l'antagonisme», entre deux souches différentes de micro-organismes, le développement de l'une des souches empêchant l'autre de proliférer (des bactéries aérobies non pathogènes limitent le développement de Staphylocoques dorés). Une vingtaine d'années plus tard, Roux et un de ses étudiants, Ernest Duchesne jeune médecin militaire, étudièrent ce concept et montrèrent qu'il s'étendait à l'antagonisme entre moisissures (ex. : *penicillium glaucum*) et bactéries (ex. : *bacillus coli*, *bacillus typhosus* ...). Les expériences permettaient déjà de supposer la présence de substances antibiotiques, cependant Ernest Duchesne, pour des problèmes de santé et pour sa carrière militaire et Roux ayant axé ses recherches sur d'autres thématiques ne permirent pas la découverte de la pénicilline. En septembre 1928, qu'Alexander Fleming remarqua que dans de vieilles boîtes de culture, des colonies microbiennes avaient complètement disparu et que ce phénomène était corrélé à la présence de moisissures dans ces boîtes. Ayant ensuite reproduit, in vitro, la mort de ces bactéries en utilisant le jus de fermentation de ces moisissures il montra la présence d'une substance bactéricide qu'il appela Pénicilline, en référence au *penicillium*, l'une des moisissures les plus répandues. Bien que publiés en 1929, ses travaux n'eurent que peu d'impact sur la communauté médicale, les tests in vivo s'étant avérés peu concluants. Il

fallut attendre les travaux de Ernst Chain en 1935, (s'étant d'abord intéressé aux travaux de Flemming sur les lysozymes puis, par hasard, à ceux concernant la pénicilline) pour connaître la nature chimique de la pénicilline et vérifier son activité in vivo.



benzylpénicilline (pénicilline G)
(antibiotique)

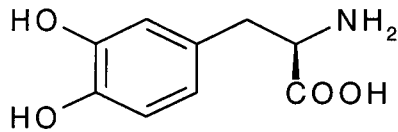
B Les découvertes par une approche rationnelle

Les importants progrès réalisés à la fois en biochimie et en pharmacologie ont permis une connaissance approfondie de l'**étiologie** de différentes maladies, base indispensable à toute conception rationnelle de médicaments.

- le traitement de la maladie de Parkinson

Dans les années 60, Hornykiewicz mit en évidence un déficit important en dopamine dans les cerveaux provenant de personnes ayant développé la maladie de Parkinson. Plutôt que d'administrer à ses patients de la dopamine, trop polaire pour franchir la barrière hématoencéphalique, il leur administra, de la Levodopa (L-dihydroxyphénylalanine), précurseur de la dopamine. Les résultats obtenus bien que satisfaisants, furent ensuite optimisés en ajoutant au traitement, un inhibiteur de dopa-décarboxylase (enzyme décarboxylant la Levodopa en dopamine), ne passant pas la barrière hématoencéphalique afin de réduire les doses de Levodopa absorbée. (Lorsque la Levodopa est administrée, sans inhibiteur de dopa-décarboxylase, la proportion de Levodopa décarboxylée avant son passage au niveau du cerveau est estimée à 95 %).

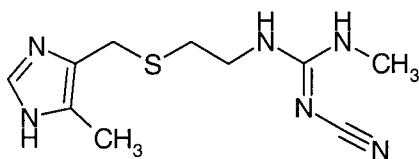
Cette démarche constitue un exemple de conception rationnelle de médicament, dans laquelle l'activité de la Levodopa fut anticipée sur la base de connaissances biologiques et biochimiques.



Levodopa
(antiparkinsonien)

- la découverte des antagonistes des récepteurs histaminergiques H₂

L'observation pharmacologique initiale, ayant amorcé le processus de découverte des antagonistes H₂, est l'absence d'inhibition des sécrétions gastriques (provoquées par l'histamine) par les antagonistes des récepteurs histaminergiques H₁ (antagonistes H₁). Black a alors envisagé l'existence d'un autre type de récepteur histaminergique (ultérieurement appelé H₂) et lancé, à partir de 1964 chez Smith Kline un programme de recherche d'antagonistes de ce nouveau récepteur. Pendant quatre ans, Ganellin assisté de Durant et Emmett ont synthétisé et testé plus de 200 analogues de l'histamine mais aucun composé ne se liait aux récepteurs stomacaux, sans déclencher une libération d'acide. Le projet allait être abandonné quand un composé synthétisé au début du programme, la guanylhistamine, montra un effet antagoniste sur les sécrétions d'acides gastriques. Ce composé n'avait pas été retenu au début du programme, car il mimait aussi l'effet de l'histamine au niveau des récepteurs stomacaux, il s'agissait en fait d'un agoniste partiel de l'histamine. Quelques jours plus tard, Emmett montra que le remplacement isostérique de la fonction guanidine par la fonction thiourée augmentait l'effet antagoniste mais son activité était insuffisante pour pouvoir être administré par voie orale. Un autre composé le metiamide, dix fois plus puissant que le précédent a été testé cliniquement mais fut responsable d'agranulocytoses. La fonction thiourée a donc été remplacée par d'autres fonctions polaires qui ont abouti à la découverte de la cimétidine. Cette molécule s'est rapidement inscrite comme la meilleure vente de produit pharmaceutique en atteignant un chiffre d'affaire de 1.000.000.000 de dollars entre 1976 et 1983.



Cimetidine
(antiulcéreux)

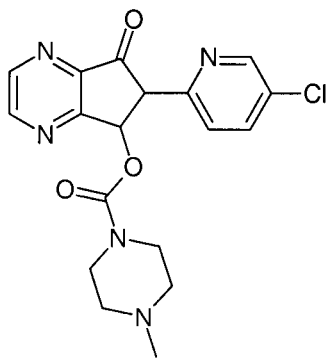
C Les découvertes par criblage de molécules sur des tests biologiques

Cette méthode consiste en un criblage systématique de molécules, quelques soient leurs structures, sur des tests biologiques tels que des mesures de binding, d'inhibitions d'enzymes, d'activités sur organes isolés ou sur cultures cellulaires. Ces techniques de tests in vitro, permettent de simplifier et d'accélérer la mise en évidence d'activités biologiques, grâce notamment à la facilité de leur mise en œuvre et de leur automatisation. Un grand nombre de composés (appelé chimiothèque) peuvent être testés au hasard, les activités découvertes servant de point de départ à la recherche et au développement d'un futur médicament.

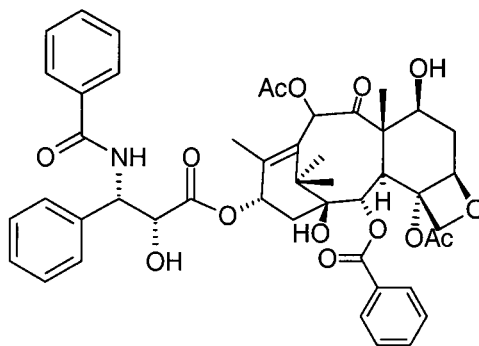
- le criblage systématique (“Random screening”)

Cette technique consiste à tester, systématiquement, des chimiothèques de plusieurs milliers de composés sur une cible prédéfinie. Elle a d'abord été utilisée pour la recherche d'antibiotiques et d'antifongiques, puis s'est étendue à la recherche d'antipaludiques, d'anticancéreux puis d'antiépileptiques. Les progrès de la biologie moléculaire permettent de disposer actuellement de nombreux récepteurs humains clonés, permettant la mise en place de tests de screening.

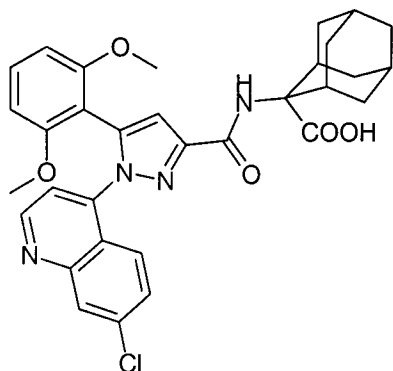
Exemples de molécules trouvées par criblage systématique :



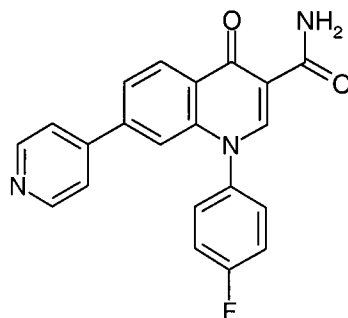
Zopiclone
(hypnotique)



Taxol
(anticancéreux)



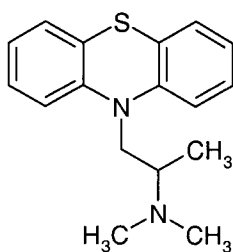
SR 48692
(antagoniste de la neurotensine)



1-(4-fluorophenyl)-1,4-dihydro-4-oxo-7-(4-pyridinyl)-3-quinolinecarboxamide
(antiherpétique)

- le criblage extensif

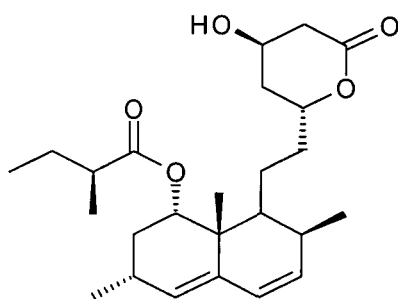
Contrairement à la technique précédente, celle-ci consiste à cribler une petite série de molécule sur une batterie de tests biologiques. Le criblage extensif s'applique surtout à des structures totalement nouvelles, issues de recherches chimiques approfondies ou de sources naturelles, qu'il est intéressant de valoriser. Les phénothiazines par exemple, d'abord synthétisées dans le but d'obtenir des composés antipaludiques, trypanocides ou anthelmintiques, ont révélé des propriétés antihistaminiques au cours de criblages extensifs.



promethazine
(antihistaminique)

- le criblage à haut débit (ou HTS : High Throughput Screening)

L'avènement des techniques robotiques et de miniaturisation des tests in vitro, ont permis, dans les années 80, de tester plusieurs milliers de composés sur un grand nombre de cibles biologiques. Cette technique correspond donc à une combinaison des deux approches précédentes. La lovastatin (inhibiteur de l'hydroxyméthylglutaryl-CoA reductase) a notamment été découverte par HTS.



lovastatin
(hypocholestérolémiant)

La limite de cette stratégie s'est d'abord située au niveau de l'approvisionnement en molécules à tester. Les sociétés pharmaceutiques ont, dans un premier temps, criblé leurs collections historiques (molécules accumulées au cours de différents programmes de recherche) puis se sont orientées vers des sources externes telles que l'approvisionnement en composés naturels ou l'achat de molécules à différents instituts de recherche publique ou privée. Cependant ces différentes sources ont vite montré leur incapacité à « alimenter » les tests de criblage. Les efforts se sont donc tournés vers la synthèse rapide de chimiothèques (ou banques ou bibliothèques) de plusieurs milliers de composés. Les premières chimiothèques se sont construites sur la base de la chimie peptidique en phase solide, du fait de l'excellente documentation relative à celle-ci. La chimie sur support solide a ensuite été appliquée à des réactions de synthèse organique classique pour aboutir à une large variété de réactions réalisées sur support solide ou en solution. Les cadences de synthèse sont maintenant telles que la limite de cette stratégie se situe au niveau des capacités de criblage. Le défi actuel est donc de choisir parmi les composés chimiquement accessibles, les composés les plus prometteurs.

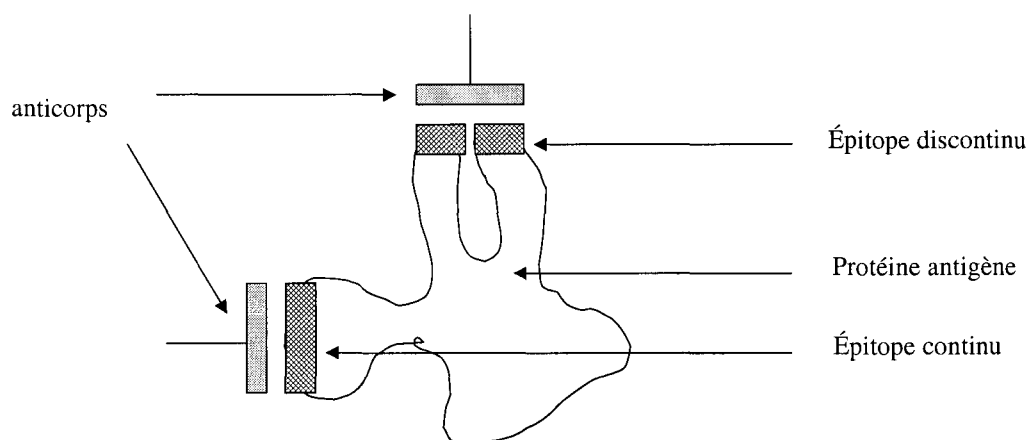
Première partie

La synthèse combinatoire en mélange

la première étape d'une course à la diversité

Introduction :

La publication à l'origine de l'intérêt porté à la synthèse de produits en mélange, ayant marqué le début des années 1990, est attribuée à Mario Geysen (Geysen 1986). Ses travaux concernaient alors l'interaction antigène / anticorps, et plus particulièrement l'identification d'épitopes reconnus par des anticorps monoclonaux. Avant cette publication, les méthodes de détermination d'épitopes, s'appuyaient sur la connaissance de la séquence d'une protéine antigène reconnue par un anticorps monoclonal. Elles consistaient à substituer différents acides aminés de cette protéine et à mesurer l'influence de ces substitutions sur la liaison protéine / anticorps monoclonal. Cette méthode permettait de déterminer les acides aminés jouant un rôle important pour établir cette liaison et d'en déduire la localisation de l'épitope. Cependant, les anticorps monoclonaux spécifiques de la protéine native, ne se fixaient que dans de rares cas à des fragments de cette protéine, ce phénomène confirmant la notion d'épitope discontinu, dans lequel intervient la structure tertiaire de la protéine. Dans ce cas, l'épitope reconnu peut être composé d'acides aminés éloignés au niveau de la structure primaire mais très proches dans l'espace après que la protéine se soit repliée.



L'étude par rayons X de complexes antigène / anticorps permet de résoudre ce problème, cependant ces études restent très complexes. Geysen a donc cherché à mettre au point une technique alternative pour déterminer l'épitope correspondant à un anticorps monoclonal et l'a appliquée à la détermination de l'épitope d'une protéine du virus de la fièvre aphteuse. Un épitope étant classiquement constitué d'un enchaînement de 8 acides aminés et les acides aminés naturels étant au nombre de 20, il fallait retrouver parmi ces 20^8 (25,6 milliards) combinaisons possibles les épitopes les plus spécifiques. La synthèse individuelle de toutes ces combinaisons d'octapeptides n'étant pas envisageable, Geysen a proposé la synthèse de

mélanges d'octapeptides. Ayant montré dans une étude antérieure, que pour obtenir une liaison efficace entre un épitope et un anticorps, au moins 3 acides aminés de l'épitope ne pouvaient être ni substitués ni déplacés et qu'au moins 2 de ces acides aminés étaient adjacents, Geysen construisit sa chimiothèque de la manière suivante :

Etape 1 : synthèse en phase solides des 400 mélanges (de 20^6 octapeptides) permettant de définir les positions 3 et 4 :

Mélange 1 : Acetyl-*-*AA₁-AA'₁-*-*-*-*Polymère

Mélange 2 : Acetyl-*-*AA₂-AA'₁-*-*-*-*Polymère

Mélange 20 : Acetyl-*-*AA₂₀-AA'₁-*-*-*-*Polymère

Mélange 21 : Acetyl-*-*AA₁-AA'₂-*-*-*-*Polymère

Mélange 22 : Acetyl-*-*AA₂-AA'₂-*-*-*-*Polymère

Mélange 40 : Acetyl-*-*AA₂₀-AA'₂-*-*-*-*Polymère

Mélange 400 : Acetyl-*-*AA₂₀-AA'₂₀-*-*-*-*Polymère

* = combinaison de 20 acides aminés naturels introduits en mélange lors de la synthèse

AA₁ à AA₂₀ = un des 20 acides aminés naturel introduit individuellement lors de la synthèse

Le mélange le plus « actif » parmi ces 400 est alors identifié et les deux acides aminés définis seront retenus pour les synthèses ultérieures.

Etape 2 : synthèse en phase solides des 20 mélanges (de 20^5 octapeptides) permettant de définir la position 2 :

Mélange 1 : Acetyl-* AA₁'-AA_D-AA'_D-*-*-*-*Polymère

Mélange 2 : Acetyl-*AA₂'-AA_D-AA'_D -*-*-*-*Polymère

Mélange 20 : Acetyl-*AA₂₀'-AA_D-AA'_D -*-*-*-*Polymère

AA_D : acide aminé défini lors d'une étape précédente de sélection

Le mélange le plus « actif » parmi ces 20 est alors identifié et l'acide aminé défini en position 2 sera retenu pour les synthèses ultérieures.

Les 5 autres positions seront ensuite définies une à une en utilisant cette même technique et le meilleur épitope sera ainsi déterminé.

Cette technique a par la suite pris le nom de chimie combinatoire, car elle consistait en l'incorporation de *combinaisons* d'acides aminés (plutôt que d'un seul acide aminé) au cours de différentes étapes de la synthèse de ces peptides.

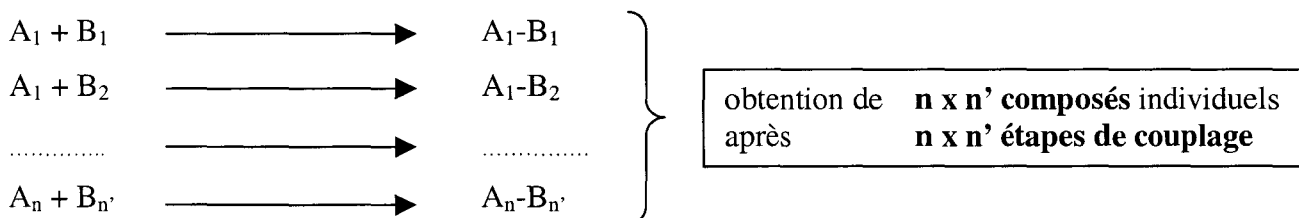
I Définition et principe de la synthèse combinatoire

Définition : La synthèse combinatoire est une technique consistant à incorporer, à au moins une des étapes d'une synthèse, un mélange de composés de réactivité apparentée, pour conduire à plusieurs combinaisons en mélange. Par extension, ce terme est maintenant employé pour l'ensemble des techniques permettant de générer rapidement un grand nombre de molécules.

Principe :

Pour obtenir toutes les combinaisons possibles entre une série de réactifs A (A_1 à A_n) et une série de réactifs B (B_1 à $B_{n'}$) :

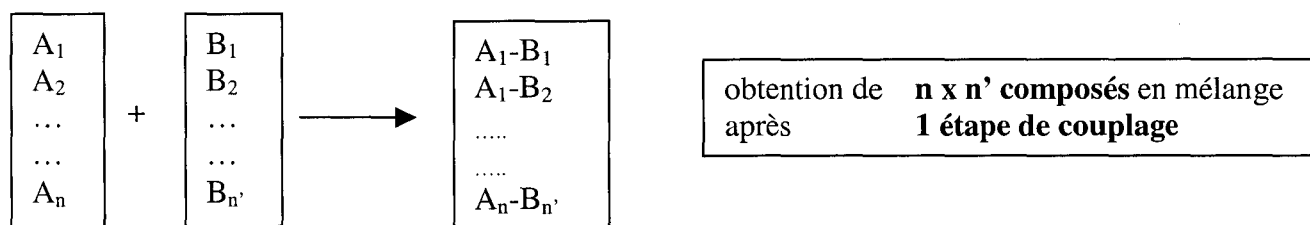
en synthèse classique les composés seront préparés individuellement :



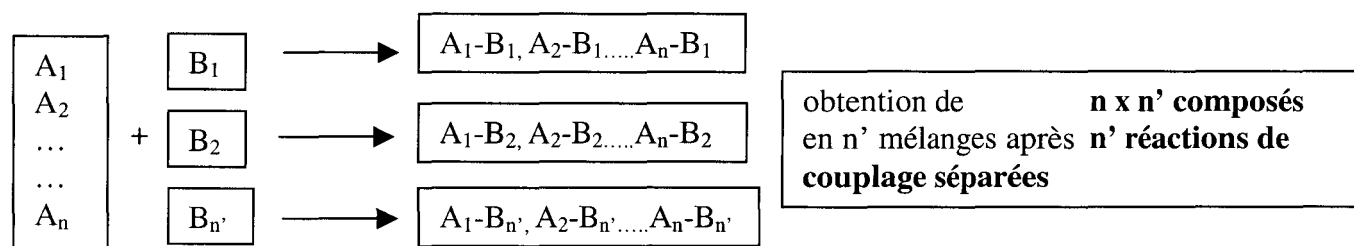
en synthèse combinatoire les composés seront préparés en mélange :

Deux méthodes peuvent alors être utilisées

Méthode 1 : à un mélange de réactifs A est ajouté un mélange de réactifs B.



Méthode 2 : à un mélange de réactifs A divisé en n' parties sont ajoutés séparément les n' réactifs B. Méthode intitulée « division mélange » ou « Split and Mix ».



Cette technique, surtout utilisée pour la synthèse en phase solide permet d'éviter les incorporations préférentielles de certains monomères B, ceux-ci réagissant séparément. D'autres part, lors de synthèses en phase solide cette méthode permet d'utiliser des excès de réactifs (réactifs B sur le schéma précédent) afin de forcer leurs couplages et ainsi d'augmenter les taux de conversion.

Si une troisième position combinatoire est désirée, les différents mélanges obtenus sont réunis, divisés à nouveau, puis les nouveaux monomères incorporés.

Une fois les mélanges de composés préparés, quelque soit la méthode combinatoire, leur activité biologique potentielle est ensuite recherchée grâce à différents tests biologiques. **L'étape ultérieure sera l'identification du (ou des) produits conférant l'activité au(x) mélange(s) « positif(s) ».**

Remarques :

1) La taille des mélanges obtenus croît très vite avec le nombre de réactifs introduits à chaque position combinatoire, mais aussi avec le nombre de positions combinatoires.

Nombre de positions combinatoires	Nombre de produits formés	
	si 10 réactifs sont introduits à chaque position	si 100 réactifs sont introduits à chaque position
2	100	10.000
3	1.000	1.000.000
4	10.000	100.000.000

Tableau 1

2) Deux types de chimiothèques peuvent être distinguées :

Les chimiothèques **généralistes** : celles-ci consistent en la synthèse d'un grand nombre de molécules très diverses devant permettre la découverte de composés porteurs d'activités biologiques originales.

Les chimiothèques **focalisées** : celles-ci consistent en la synthèse d'un grand nombre de molécules apparentées devant permettre d'optimiser l'activité biologique de composés connus.

II Les méthodes chimiques utilisées en synthèse combinatoire

II-A Phase solide ou solution ?

Le but de la synthèse combinatoire étant d'engendrer des chimiothèques de taille importante, les réactions utilisées doivent être adaptées à un nombre important de matières premières (appelées aussi monomères ou building-blocks) et doivent engendrer le minimum de produits secondaires. L'expression « réactions robustes » est couramment employée pour décrire ce type de réaction.

Les premiers efforts se sont concentrés sur le développement de réactions en phase solide car :

- la synthèse peptidique en phase solide est à l'origine de la synthèse combinatoire (Geysen 1986, Furka 1988, Lam 1991, Houghten 1991)
- les rendements réactionnels et la pureté finale des produits sont souvent très satisfaisants, des excès de réactifs pouvant être utilisés. Actuellement de nombreuses recherches portant sur la conception de nouveaux polymères et de nouveaux types d'ancrage, permettent d'élargir le champ des réactions applicables (Bunin 1998). L'utilisation de la chimie en solution est ensuite apparue (Carell 1994, Pirrung 1995) mais n'a bénéficié que d'un succès limité pour la synthèse de produits en mélange. En effet, l'élimination de solvants et de réactifs étant dans ce cas très fastidieuse, les synthèses en solution ne permettent généralement d'accéder qu'à des composés obtenus en une voire deux étapes de synthèse.

	Phase solide	Solution
Avantages		
Synthèses multiétapes	****	*
Nombre de réactions chimiques utilisables	**	***
Complexité des produits obtenus	***	**
Facilité d'automatisation	**	***
Analyses en cours de synthèse	**	****
Inconvénients		
Consommation en solvants et réactifs	****	*
Matériel de synthèse spécial	***	
Nécessité de purifications intermédiaires		***
Etape de fixation puis clivage du support solide	**	

Tableau 2 : comparaison des avantages et inconvénients de la synthèse en phase solide et en solution

Un regain d'intérêt pour la synthèse en solution a actuellement lieu mais pour la synthèse parallèle à haut débit de composés individuels.

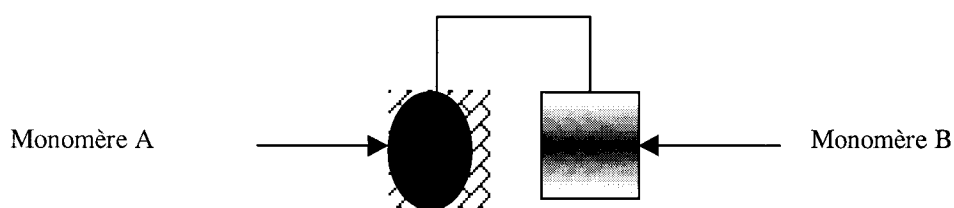
II-B Les stratégies d'assemblage de la diversité (Ecker 1995)

II-B-1 Assemblage de type oligomérique (Pinilla 1994, Pinilla 1995, Ostresh 1994 a, Zuckermann 1994)

Cette stratégie consiste à enchaîner de manière séquentielle, plusieurs monomères sans leur imposer d'orientation particulière.

- Les **dimères** constituent les premiers représentants de ce type d'assemblage.

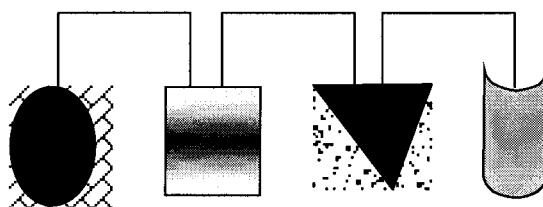
Leur synthèse nécessite uniquement la présence **d'une position réactive sur chacun des 2 monomères** de départ.



Les espèces potentiellement présentes dans le milieu réactionnel en fin de réaction étant peu nombreuses (les produits attendus, les produits de départ, éventuellement un agent de couplage), facilement identifiables et accessibles, la synthèse de dimères ne nécessite généralement **pas de purification préalable** aux tests biologiques, et s'adapte parfaitement à la **chimie en solution**.

L'obtention de chimiothèques de taille conséquente, par dimérisation, n'est possible que lorsque la chimie permet d'utiliser un grand nombre de monomères de départ. Les familles les plus souvent utilisées sont : les acides, les amines, les alcools, les aldéhydes, les cétones, les sulfochlorures, les isocyanates et les chlorures d'acides.

- Les **oligomères** nécessitent l'utilisation de **monomères plus sophistiqués** comportant généralement deux fonctions réactives (dont l'une doit parfois être protégée puis libérée).



Monomères :

A

B

C

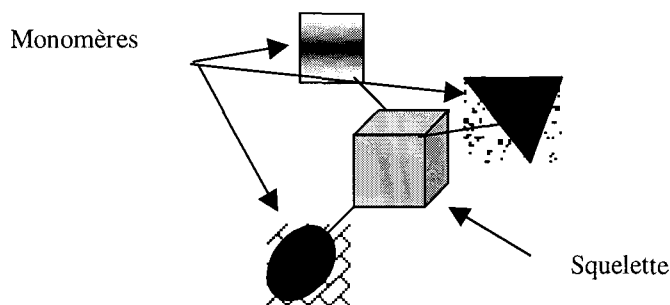
D

L'accroissement du nombre de position combinatoire exige l'**utilisation de la phase solide** afin d'éliminer les réactifs n'ayant pas réagi et ainsi de limiter le nombre de produits secondaires potentiels. Mais l'utilisation de la phase solide requiert des dérivés bifonctionnels ainsi qu'une liaison au support. Les familles de monomères bifonctionnels étant relativement peu peuplées, les chimiothèques de grande taille sont souvent obtenues en **multipliant les positions combinatoires**, ou en incorporant, généralement lors de la dernière étape, une famille de monomères particulièrement peuplée. Les oligomères décrits dans la littérature, sont essentiellement issus de travaux visant à substituer les enchaînements polypeptidiques par des enchaînements conférant aux produits une meilleure biodisponibilité.

De nombreux exemples faisant intervenir la synthèses de peptides (Pinilla 1994, Pinilla 1995) de peptides perméthylés (Ostresh 1994 a) ou d'oligo Glycinamides-N-substitués (peptoïdes) (Zuckermann 1994) ont été décrits.

II-B-2 Assemblage par décoration d'un squelette de base (Pirrung 1995)

Dans cette stratégie les différents monomères seront ajoutés à un squelette de base polyfonctionnel.



La nature du squelette est généralement choisie pour son **activité biologique intrinsèque**, ou pour la **géométrie** de l'espace qu'il permet d'explorer. Lorsqu'une importance prépondérante est accordée au squelette de base, cette méthode s'apparente à la chimie médicinale classique et servira essentiellement à optimiser une structure dont le squelette de base constitue le pharmacophore. En revanche, lorsque la diversité apportée par les différents monomères revient à « masquer » la structure du squelette, celui-ci ne sert qu'à orienter dans l'espace les éléments de diversité et conduira à une chimiothèque généraliste. De nombreux médicaments actuellement commercialisés présentant une nature hétérocyclique, ce type de structure reste très souvent recherché pour la synthèse de squelettes décorés.

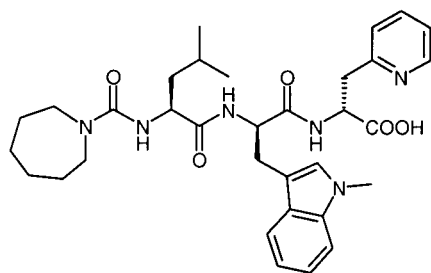
Remarques :

- 1) les dimères décrits au paragraphe précédent, peuvent aussi être considérés comme une forme particulière de squelette décoré, dans lequel le squelette de base serait constitué par le lien chimique et la diversité serait introduite de part et d'autre de ce lien
- 2) le squelette de base peut exister avant l'incorporation des monomères ou bien être formé après une étape de cyclisation

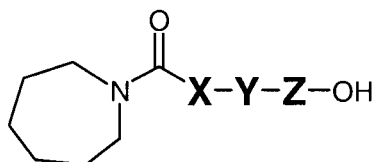
II-C Exemple

Synthèse sur support solide d'une bibliothèque oligomérique de 30.752 composés (Terrett 1995)

But : optimiser l'activité d'un agoniste du récepteur ET_A à l'endothéline : le FR-139,317.



FR-139,317



X = monomères 1 à 32

Y = monomères 1 à 32

Z = monomères 1 à 31

Figure 1 : Structure de base de la chimiothèque

Monomères	Structure
1	
2	
3	
4	
5	
6	D-Lys
7	L-Lys
8	D-Ser
9	L-Ser
10	D-Asp
11	L-Asp

Monomères	Structure
12	D-Asn
13	L-Asn
14	
15	
16	
17	

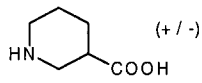
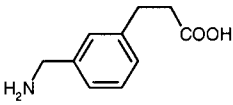
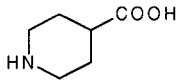
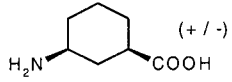
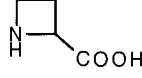
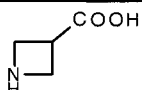
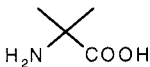
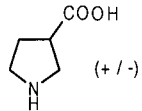
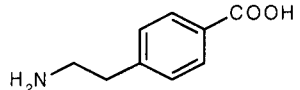
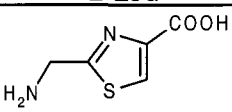
Monomères	Structure	Monomères	Structure
18	 (+ / -)	25	
19		26	D-Phe
20	 (+ / -)	27	
21	D-N-Me-Trp	28	
22		29	 (+ / -)
23		30	L-Leu
24	L-Val	31	
		32	D-2-Pyr-Ala

Figure 2 : Structure des monomères

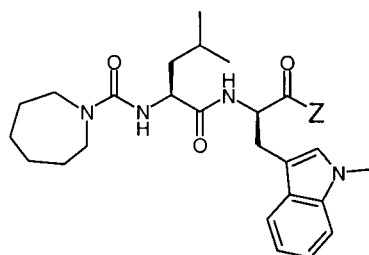
Les monomères retenus permettaient d'explorer d'une part l'influence de la distance entre l'amine et la fonction acide carboxylique et d'autre part l'importance de diverses fonctions présentes sur les chaînes latérales d'acides aminés naturels ou non.

Caractéristiques de la synthèse :

La synthèse a été effectuée en phase solide après couplage de chacun des premiers acides aminés sur une résine de Wang. La protection temporaire de la fonction amine était un groupement Fmoc. Les réactifs d'activation des fonctions acide carboxylique étaient l'HOBt, TBTU.

Résultats :

Parmi les 30.752 composés synthétisés, 15 présentaient une IC_{50} inférieure à $3,5 \mu M$, le produit de référence figurant parmi ces composés (IC_{50} $5,3 nM$). Trois nouveaux composés ayant une IC_{50} inférieure à $10 nM$ ont été découverts, leurs structures étaient cependant très proches de la structure de départ cf figure 3. Les résultats obtenus ont aussi permis d'établir des relations structure / activité, compatibles avec celles déjà décrites dans des travaux précédents.



Z = D-Phe, IC₅₀ = 2,5 nM

Z = D-Ser, IC₅₀ = 8 nM

Z = D-N-Me-Trp, IC₅₀ = 8 nM

Figure 3 : structures des trois composés les plus actifs (différents du produits de référence) présents dans la chimiothèques

III Organisation des chimiothèques et élucidations structurales des produits conférant l'activité aux mélanges

La recherche de composés actifs, présents au sein de mélanges, a d'abord concerné les travaux consacrés à l'étude des substances naturelles. La composition des mélanges étudiés étant le plus souvent inconnue, la première opération consiste à diviser le mélange « père » complexe, en plusieurs mélanges « fils » de complexité moindre, les techniques les plus utilisées, étant l'extraction liquide / liquide ou la chromatographie (Hostettmann 1998). Les mélanges « fils » sont ensuite testés et celui ayant conservé l'activité est à son tour fractionné en mélanges moins complexes. Ces cycles successifs aboutissent généralement à l'identification du (des) produit(s) actif(s). Cette méthode itérative est souvent fastidieuse et nécessite la mise au point de conditions expérimentales pour chaque nouvel extrait.

Un des avantages offerts par la chimie combinatoire est la possibilité de définir, a priori, la composition de chacun des mélanges testés et donc d'organiser leurs synthèses de telle sorte que la méthode de déconvolution (détermination du (des) composé(s) actif(s) au sein d'un mélange lui-même actif, par fractionnement de ce mélange) soit la plus rapide possible.

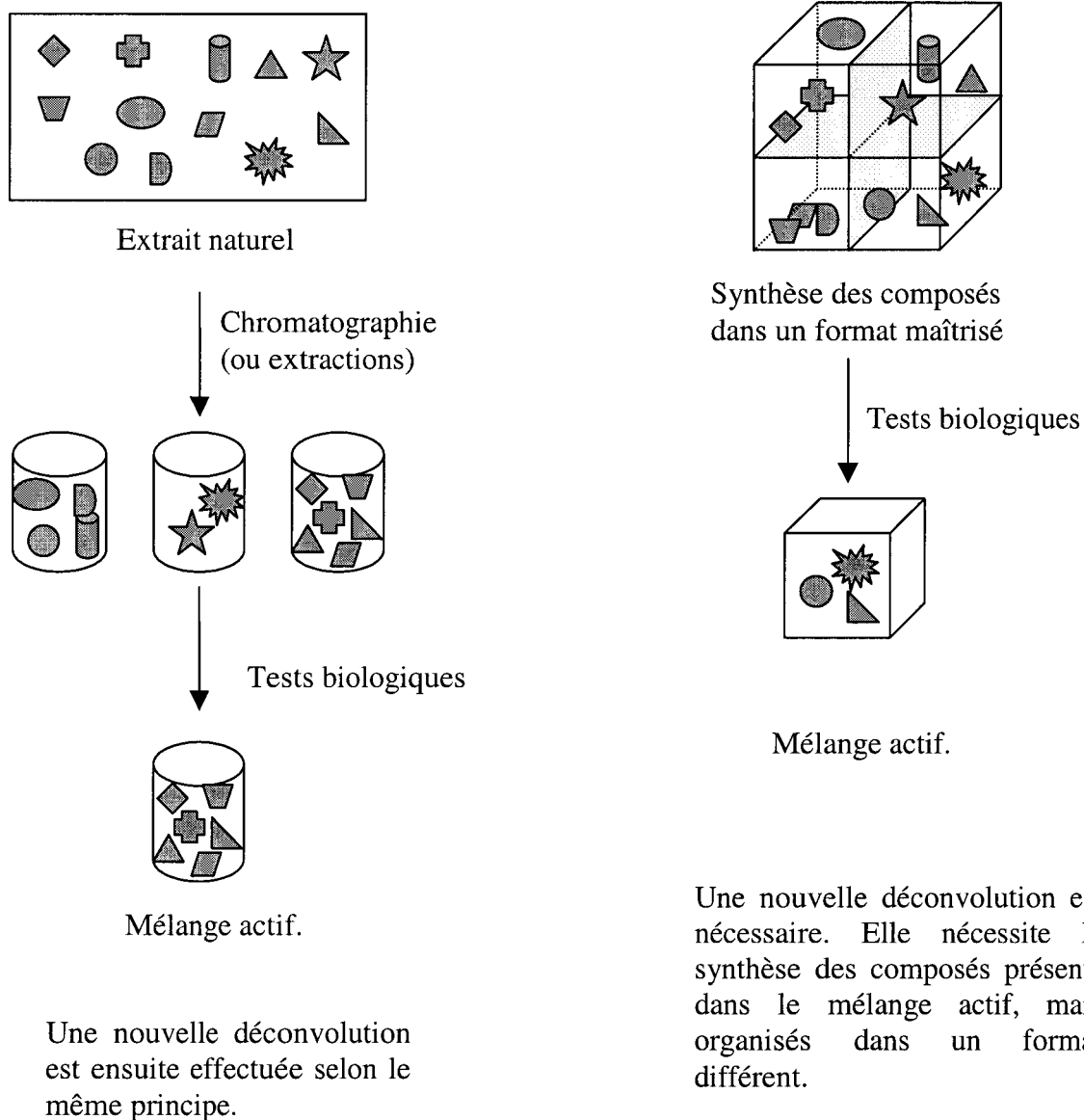


Figure 4 : comparaison de la déconvolution d'un mélange actif provenant d'un extrait naturel ou d'une chimiothèque combinatoire

Déconvolution de mélanges

Produits naturels :

Faible connaissance :

- de la taille et la composition des mélanges
- des caractéristiques physico-chimiques des constituants (ex : stabilité, absorbance UV, masse moléculaire...)

Quantités disponibles et réapprovisionnement pouvant être problématiques

Les techniques de purification sont utilisées pour fractionner des mélanges actifs.

Produits combinatoires :

Maîtrise :

- du format des chimiothèques (taille et composition des mélanges)
- des quantités produites et du réapprovisionnement

Anticipation possible des caractéristiques physico-chimiques des constituants (ex : stabilité, absorbance UV, masse moléculaire...)

Le format de synthèse des chimiothèques est utilisé pour la déconvolution des mélanges actifs.

III-A Elaboration d'un modèle général de description des méthodes de déconvolution

X. Williard and A. Tartar, « A Practical Guide to Combinatorial Chemistry », 1997, published by the American Chemical Society, 249-277.

III-A-1 Le modèle Matriciel

Afin de disposer d'un modèle général pour la description d'une banque combinatoire et des différents moyens d'en assurer sa déconvolution, nous avons proposé de décrire toute chimiothèque combinatoire, sous la forme d'une matrice à N-dimensions, où N représente le nombre de positions combinatoires.

Une chimiothèque comportant deux positions combinatoires A et B, impliquant respectivement n et n' monomères représentée sous la forme d'une matrice à deux dimensions peut être obtenue (figure 5).

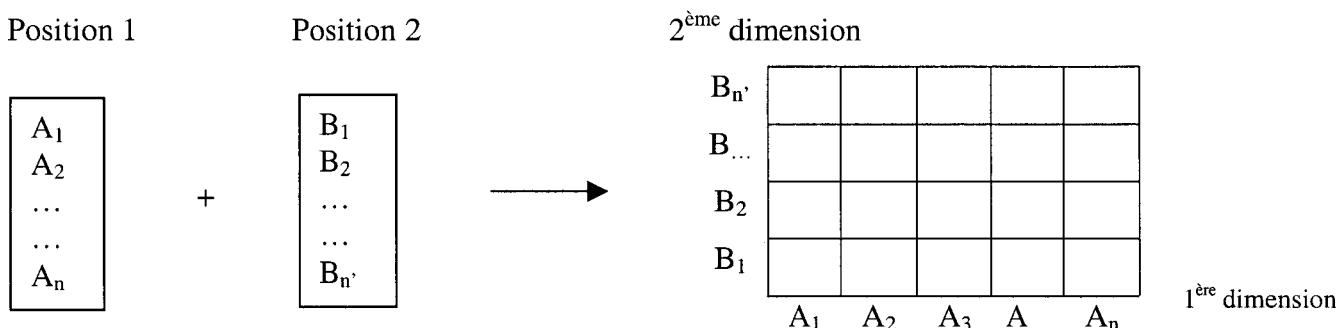


Figure 5 : matrice représentant une chimiothèque à deux positions combinatoires

Chaque direction de la matrice possède autant d'éléments que de monomères impliqués dans la position combinatoire qu'il représente. Cette représentation n'est pas limitée aux matrices à deux dimensions. La représentation matricielle d'une chimiothèque combinatoire incluant 3 positions combinatoires A, B et C ($N = 3$) impliquant chacune 4 monomères ($n_a = n_b = n_c = 4$). Nombre total de composé : $4 \times 4 \times 4 = 64$ est illustrée figure 6. Chaque composé peut être localisé dans une des cellules de la matrice et défini par la combinaison : A_x, B_y, C_z .

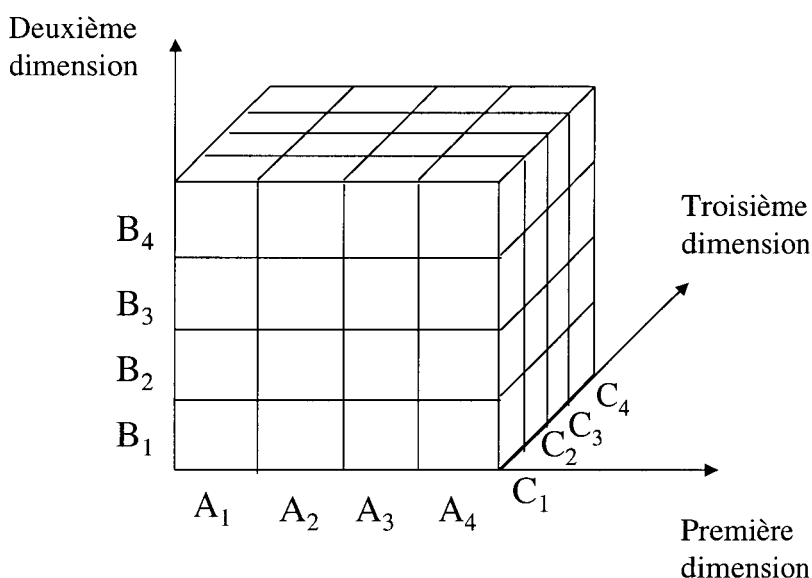


Figure 6 : matrice représentant une chimiothèque à trois positions combinatoires

III-A-2 Evaluation biologique d'une chimiothèque

Dans le cas où l'information relative à chacune de ces cellules (ex : l'activité biologique portée par chaque composé) est nécessaire, il faudra pouvoir isoler chaque cellule et les traiter individuellement. (cf. synthèses en parallèle, seconde partie). Cependant, dans la plupart des cas, le criblage primaire de chimiothèques sert à répondre à la question : existe-t-il au moins un composé actif dans cette chimiothèque ? Pour répondre à cette première question, il suffit de tester le mélange (sous réserve que la dilution du produit actif parmi les autres constituants du mélange ne l'amène pas à une concentration inférieure à la limite de détection du test utilisé). La seconde question qui fait immédiatement suite à une réponse positive à la première est : quelle est la structure du composé actif ? Là encore, l'information requise est moindre que celle contenue dans l'ensemble de la matrice et pourra s'obtenir au moyen d'un nombre limité d'expériences bien choisies. Différentes stratégies élaborées à cette fin ont été regroupées sous le vocable de « méthodes de déconvolution ». Nous nous attacherons dans la

suite de cet exposé à démontrer que toutes ces méthodes peuvent être comparées en utilisant une description matricielle unique.

III-B Application du modèle matriciel aux méthodes de déconvolution

Trois méthodes de déconvolution ont été décrites : la déconvolution itérative, le « balayage positionnel » (positional scanning) et les chimiothèques orthogonales. Celles-ci sont toutes basées sur la partition d'une chimiothèque en sous-chimiothèques qui seront synthétisées puis testées individuellement. L'évaluation comparative de l'efficacité de ces méthodes sera faite sur la base de quatre critères :

- le nombre et la durée des étapes permettant l'identification du produit actif
- le nombre de composés synthétisés au cours des étapes de déconvolution
- la fiabilité de la méthode considérée
- la génération d'informations complémentaires

III-B-1 La déconvolution itérative

Définition : méthode de déconvolution consistant à réduire progressivement, par synthèses successives, le nombre de composés présents dans un mélange actif jusqu'à l'obtention du composé actif.

Principe : (illustré figure 7)

Soit une chimiothèque de $4 \times 4 \times 4 = 64$ trimères issue de l'incorporation de 4 monomères ($n_a = n_b = n_c = 4$) différents en 3 positions combinatoires distinctes ($N = 3$) (monomères A_1 à A_4 pour la position 1, monomères B_1 à B_4 pour la position 2, monomères C_1 à C_4 pour la position 3)

Selon les conventions proposées par Houghten, (Houghten 1991) la lettre « O » est utilisée pour représenter une position combinatoire à laquelle est incorporé un monomère unique, la lettre « X » est utilisée quand il s'agit d'un mélange de monomères. Pour la chimiothèque de 64 trimères, la matrice complète est donc symbolisée par X-X-X (soient $4 \times 4 \times 4 = 64$ combinaisons). Une sous-chimiothèque pour laquelle la position 2 serait fixée à un seul monomère, les combinaisons seraient symbolisées par X-O-X (soient $4 \times 1 \times 4 = 16$ combinaisons).

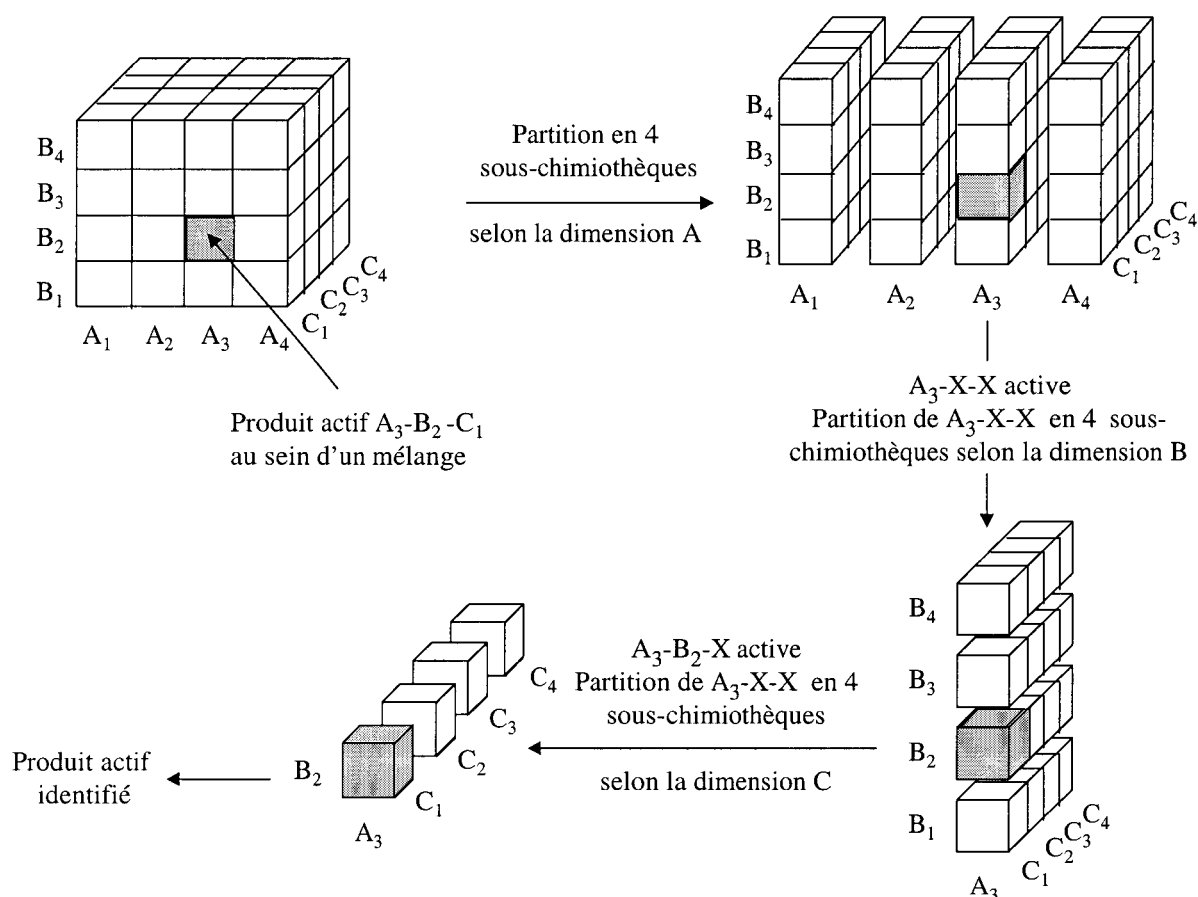


Figure 7 : application du modèle matriciel à la description d'une déconvolution itérative

Selon ce modèle, la déconvolution s'opèrera en trois étapes successives :

Étape 1

Si une activité biologique est détectée au sein du mélange X-X-X, la première étape sera de définir pour la position 1, le monomère présent dans le produit actif. Pour cela, les quatre sous-chimiothèques suivantes, symbolisées par O-X-X, seront préparées :

- A₁-X-X (1 x 4 x 4 = 16 combinaisons)
- A₂-X-X (1 x 4 x 4 = 16 combinaisons)
- A₃-X-X (1 x 4 x 4 = 16 combinaisons)
- A₄-X-X (1 x 4 x 4 = 16 combinaisons)

Les 64 combinaisons initiales seront donc réparties en quatre sous-chimiothèques (de 16 composés) n'ayant aucun élément commun.

Un test biologique permettra d'identifier la sous-chimiothèque contenant le produit actif (par exemple A₃-X-X figure 7)

Etape 2

La sous chimiothèque trouvée active à l'étape 1 sera ensuite « resynthétisée » dans un format permettant d'identifier le monomère occupant la position 2 dans le produit actif.

Dans la figure 7, les 16 combinaisons de la sous-chimiothèque A3-X-X seront donc réparties dans quatre sous-chimiothèques A3-O-X (de 4 composés) dont la position 2 est fixée, n'ayant aucun élément commun.

Un test biologique permettra d'identifier la sous-chimiothèque contenant le produit actif (par exemple A₃-B₂-X, figure 7).

Etape 3

La sous chimiothèque trouvée active à l'étape 2 sera ensuite « resynthétisée » dans un format permettant d'identifier le monomère occupant la position 3 dans le produit actif.

Dans la figure 7, les 4 combinaisons de la sous-chimiothèque A₃-B₂-X seront synthétisées individuellement (A₃-B₂-O) et testées séparément. Ce test biologique permettra d'identifier le produit actif (par exemple A₃-B₂-C₁, figure 7).

Discussion :

Sur la base des quatre critères définis précédemment nous avons analysé l'efficacité de cette méthode :

Le nombre et la durée des étapes permettant l'identification du produit actif

Le nombre d'étapes est directement fonction du nombre de positions combinatoire (N), chaque itération permettant d'identifier, successivement, les monomères présents dans le produit actif à chacune des positions. Ainsi, quel que soit le nombre de monomères (n) introduits à chaque position combinatoire le nombre d'étapes de déconvolution reste constant. En revanche, le nombre de sous-chimiothèques synthétisées à chaque étape est égal au nombre de monomères (n_i) occupant la position combinatoire -i- à définir.

Le nombre de composés présent dans une chimiothèque croît de manière multiplicative avec les monomères (n₁ à n_i) introduits à chaque étape combinatoire :

Nombre total de composés présents dans la chimiothèque = n₁ x n₂ x n₃ x ... n_i

alors que le nombre de sous-chimiothèques nécessaires à la déconvolution itérative croît de manière additive :

Nombre de sous-chimiothèques nécessaires à la déconvolution = n₁ + n₂ + n₃ + ... n_i.

Le tableau 3 permet de comparer le nombre de sous-chimiothèques nécessaires à la déconvolution itérative de chimiothèques de tripeptides et d'hexapeptides impliquant 15 ou 30 acides aminés à chaque position combinatoire. Il apparaît clairement que **le nombre de sous-chimiothèques à synthétiser pour des chimiothèques de petite taille ou de grande taille ne varie que faiblement**, ce qui rend cette méthode particulièrement attractive pour la déconvolution de grandes chimiothèques.

		Tripeptides	Hexapeptides
A partir de 15 acides aminés	Nombre total de peptides	3.375 (15 ³)	11.390.625 (15 ⁶)
	Nombre de sous-chimiothèques*	45 (3 x 15)	90 (6 x 15)
A partir de 30 acides aminés	Nombre total de peptides	27.000 (15 ³)	729.000.000 (30 ⁶)
	Nombre de sous-chimiothèques*	90 (3 x 30)	180 (6 x 30)

* nécessaires à une déconvolution itérative

Tableau 3 : comparaison du nombre de sous-chimiothèques nécessaires à la déconvolution itérative de chimiothèques de tripeptides et d'hexapeptides impliquant 15 ou 30 acides aminés à chaque position variable

La technique de déconvolution itérative souffre principalement de la durée de chacune des étapes. En effet, ce type de déconvolution nécessite des aller-retours entre les phases de synthèse et les phases de tests biologiques, pouvant nécessiter plusieurs jours, voire plusieurs semaines. Ce **processus discontinu** (et non réversible, caractéristique particulièrement dommageable si un choix erroné a été fait à une étape) fournit la structure du composé actif de façon progressive.

Le nombre de composés synthétisés à l'issue des étapes de déconvolution

Il nous faut ici définir l'expression taille d'une chimiothèque ou d'une sous-chimiothèque. Celle-ci fait référence au nombre de composés que cette chimiothèque ou sous-chimiothèque contiennent :

- pour une chimiothèque l'expression taille limitée s'emploie pour un nombre de composés inférieur à 10.000, l'expression taille moyenne pour un nombre compris entre 10.000 et 100.000 et l'expression taille importante pour un nombre supérieur à 100.000
- pour une sous-chimiothèque l'expression taille limitée s'emploie pour un nombre de composés inférieur à 100, l'expression taille moyenne pour un nombre compris entre 100 et 1000 et l'expression taille importante pour un nombre supérieur à 1000.

Le nombre de composés synthétisés à l'issue des étapes de déconvolution est étroitement lié au nombre de positions combinatoires ainsi qu'au nombre de monomères occupant chaque

position : la déconvolution d'une chimiothèque de taille importante nécessitera la préparation de sous-chimiothèques de taille importante. D'autre part, la taille des sous-chimiothèques se réduit considérablement au cours du processus de déconvolution itérative. Le tableau 4 permet d'évaluer le nombre de composés nécessaires à la déconvolution itérative des 11.390.625 hexapeptides possibles synthétisés à partir de 15 acides aminés : si aux 11.390.625 composés de la chimiothèque de départ sont ajoutés ceux issus des étapes de déconvolution ((15 x 50.625) + (15 x 3.375) + (15 x 225) + (15 x 15) + 15) le nombre total de composés synthétisés pour effectuer la déconvolution itérative est de 12.204.240, soit un **faible « excès de synthèses »** de 813.615 composés ou de **7,14 %** lorsqu'il est exprimé en valeur relative, par rapport au nombre de composés de la chimiothèque de départ. Ce tableau permet aussi de mettre en évidence la décroissance de la taille des sous-chimiothèques, au fur et à mesure du processus de déconvolution. Il est important de remarquer qu'en parallèle à cette diminution, l'analogie des composés au sein d'une même sous-chimiothèque ainsi que l'analogie des chimiothèques entre elles, croissent considérablement.

	Etape 1	Etape 2	Etape 3	Etape 4	Etape 5	Etape 6	Total
Nombre de sous-chimiothèques	15	15	15	15	15	15	90
Taille des sous-chimiothèques	759.375 composés	50.625 composés	3.375 composés	225 composés	15 composés	1 composé	12.204.240 composés

Tableau 4 : nombre et taille des sous-chimiothèques synthétisées au cours de la déconvolution itérative d'une chimiothèque d'héxapeptides impliquant 15 acides aminés à chaque position combinatoire

La taille importante des sous-chimiothèques nécessaires aux premières étapes du processus de déconvolution itérative, pose le problème de la **dilution de l'activité** d'un composé actif au sein d'un mélange complexe de composés inactifs. Cependant, dans la plupart des cas, l'activité d'un mélange combinatoire ne repose pas uniquement sur un composé (ligand optimal), mais sur un ensemble de composés ayant des structures apparentées (ligands suboptimaux) et pouvant engendrer une activité par **effet cumulatif**. Le problème devient : comment retrouver le composé le plus actif ?

En utilisant des résultats de la littérature, Freier et al (Freier 1995) ont analysé l'activité biologique de sous-chimiothèques préparées au cours d'étapes successives de déconvolution. Ils ont ensuite défini le SBF : facteur de liaison suboptimale (Suboptimal Binding Factor) qui permet d'évaluer un éventuel effet cumulatif de l'activité de plusieurs sous-chimiothèques. Pour cela, après avoir déterminé l'activité d'une chimiothèque puis décomposé celle-ci en

sous-chimiothèques, pour ne retenir que celle apparaissant comme la plus active, le ratio de l'activité théorique de la chimiothèque de départ (calculée en supposant que seule la sous-chimiothèque retenue est active) sur son activité observée est calculé, la valeur obtenue correspond au SBF.

$$\text{SBF} = \frac{\text{Activité théorique de la chimiothèque} * (\text{IC}_{50} \text{ exprimée en molarité})}{\text{Activité observée de la chimiothèque} (\text{IC}_{50} \text{ exprimée en molarité})}$$

* calculée en considérant que la sous-chimiothèque retenue comme la plus active au sein de la chimiothèque de départ soit la seule à être active

Il est important de remarquer que plus une chimiothèque sera active, plus son IC_{50} sera faible.

un $\text{SBF} \gg 1$ indique la présence d'autres sous-chimiothèques actives que celle retenue

un SBF proche de 1 indique la faible incidence d'autres sous-chimiothèques actives que celle retenue. Plus le SBF sera proche de 1, plus l'enrichissement en produit actif lors de la déconvolution sera efficace. Idéalement, si la chimiothèque est parfaitement diverse, seul un composé est actif et le SBF sera égal à 1 tout au long de la déconvolution.

Les SBF obtenus d'après les données de la littérature s'échelonnent de 3,5 à 1561. Trois facteurs principaux semblent influencer sur cette valeur :

- la taille des sous-chimiothèques testées : pour des sous-chimiothèques de taille réduite, le SBF diminue, car le risque de trouver plusieurs structures actives sur une même cible est plus faible que pour des sous-chimiothèques de grande taille
- la spécificité de la cible : il semble beaucoup plus difficile par exemple, d'obtenir des peptides interagissant avec l'HIV-protéase (SBF proche de 1) qu'avec un récepteur aux opiacés (augmentation du SBF)
- la nature des composés présents dans la sous-chimiothèque : plus les composés présents dans la sous-chimiothèque seront divers (nécessité de choisir des monomères de départ

très divers), plus le risque d'engendrer des analogues d'un produit actif est limité (diminution du SBF).

La fiabilité de la méthode

La présence de ligands suboptimaux peut conduire, lors d'une déconvolution itérative, à sélectionner pour les étapes ultérieures, une sous-chimiothèque autre que celle contenant le produit le plus actif. En effet les sous-chimiothèques seront comparées en fonction de leurs activités respectives et non en fonction de l'activité du composé le plus actif. Afin d'éviter la **seule** sélection d'une sous-chimiothèque (la plus active) résultant de l'effet cumulatif de plusieurs ligands suboptimaux, une analyse de **plusieurs** sous-chimiothèques sera généralement effectuée. Elle pourra permettre de continuer la déconvolution sur plusieurs sous-chimiothèques, notamment quand celles-ci sembleront emprunter deux voies bien distinctes pouvant conduire à des composés très différents.

Cette méthode de déconvolution, amplement utilisée, a permis par exemple l'identification d'un inhibiteur nanomolaire du récepteur $\alpha 1$ -adrénergique, dans une chimiothèque de peptoïdes (Zuckermann 1994).

La génération d'informations complémentaires

Au cours des différentes étapes de déconvolution, de nombreuses données biologiques relatives aux différentes sous-chimiothèques testées sont engendrées et peuvent être utilisées dans des études plus tardives de relations structure-activité et d'optimisation de composés actifs.

Conclusion :

La méthode de déconvolution itérative est une méthode basée sur un principe simple et qui nécessite un effort de synthèse modéré lorsqu'elle n'est effectuée que pour le criblage sur une cible. Cette méthode reste très flexible et peut être conduite de différentes façons (ex : déconvolution de plusieurs sous-chimiothèques plutôt qu'une seule) en fonction des résultats obtenus, de la cible biologique étudiée et des informations désirées sur les relations structure-activité. De plus, si cette méthode est menée correctement, l'évolution du SBF demeure un bon critère d'évaluation de la diversité des composés et de l'efficacité du processus de déconvolution. Cependant les points faibles de cette méthode résident dans la nécessité de préparer de nouvelles sous-chimiothèques pour chaque nouvelle déconvolution, ainsi que des allers-retours entre ces synthèses et les tests biologiques, pouvant la rendre particulièrement

coûteuse en temps. Elle sera donc plutôt retenue pour la déconvolution de chimiothèques testées sur un nombre limité de cibles biologiques.

Remarque : une méthode consistant à préparer les différentes sous-chimiothèques, non plus en incorporant, un par un, les monomères occupant la position à définir, mais incorporant tous les monomères sauf un, et ceci pour chacun des monomères a été décrite (Furka 1994). Il s'agit alors de détecter toute perte d'activité liée à l'absence d'un monomère particulier.

III-B-2 Le « balayage positionnel » ou « positional scanning » (Pinilla 1994, Dooley 1993)

Définition : méthode de déconvolution consistant à préparer une chimiothèque en plusieurs sous-chimiothèques balayant, en parallèle, chaque position combinatoire et permettant l'identification directe du composé actif.

Principe : (illustré figure 8)

Soit une chimiothèque de $4 \times 4 \times 4 = 64$ trimères issue de l'incorporation de 4 monomères différents en 3 positions combinatoires distinctes (monomères A_1 à A_4 pour la position 1, monomères B_1 à B_4 pour la position 2, monomères C_1 à C_4 pour la position 3). $N=3$, $n_a = n_b = n_c = 4$.

Pour représenter les positions combinatoires des différentes sous-chimiothèques, les conventions proposées par Houghten (cf. paragraphe III-B-1) seront utilisées.

Le « positional scanning » ou balayage positionnel conduira dans ce cas à la synthèse des sous-chimiothèques suivantes : O-X-X, X-O-X et X-X-O. Les 4 sous-chimiothèques représentées par O-X-X vont permettre de définir la position 1, les 4 sous-chimiothèques représentées par X-O-X, la position 2 et les 4 sous-chimiothèques représentées par X-X-O, la position 3. La synthèse de tous les composés de la chimiothèques a donc été effectuée trois fois mais dans des formats différents. Ces 12 sous-chimiothèques sont ensuite testées sur une cible biologique et aboutiront à l'identification du monomère présent dans le produit actif pour chacune des positions combinatoires, ce qui permettra une identification directe du composé actif.

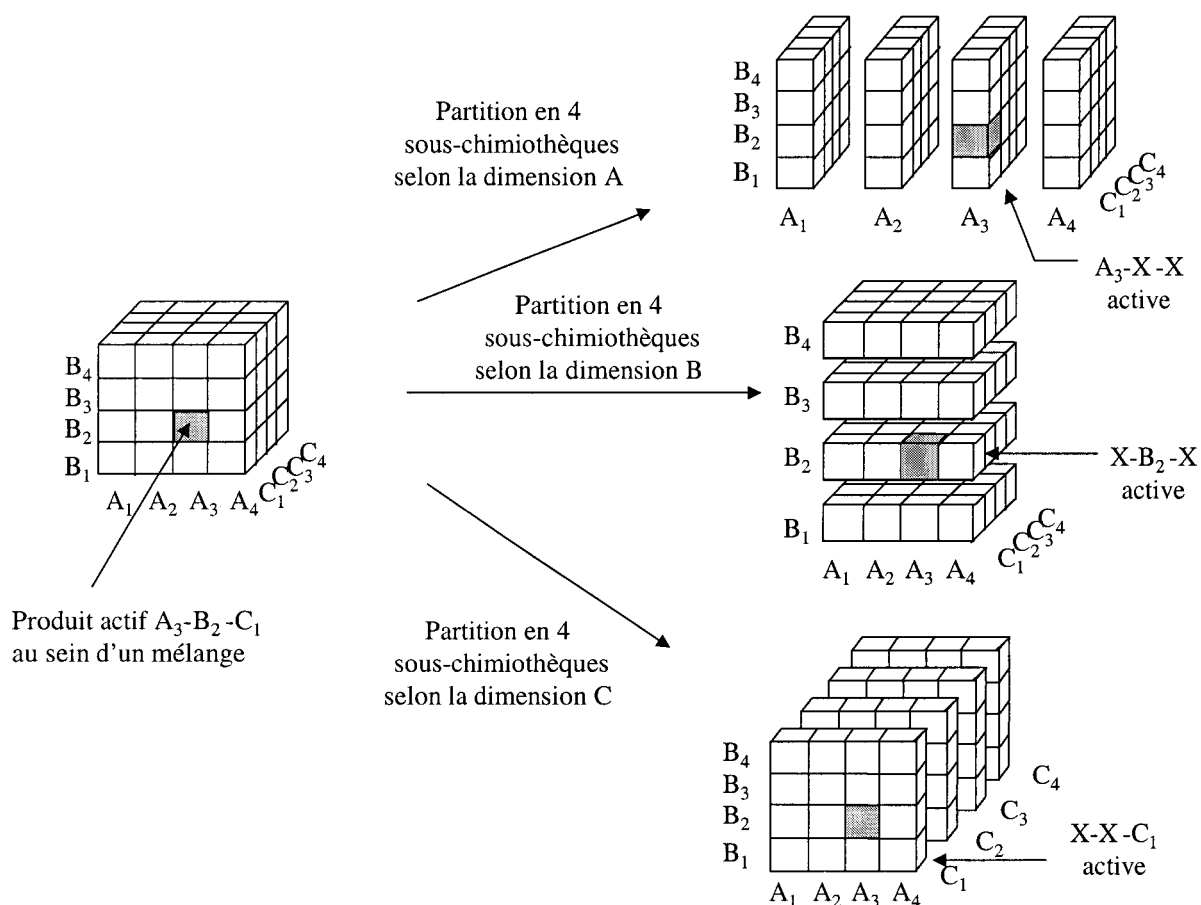


Figure 8

Discussion :

Le nombre et la durée des étapes permettant l'identification du produit actif

L'intérêt majeur du balayage positionnel est la possibilité en une seule « étape », d'identifier les monomères conférant l'activité à chacune des positions. En effet les sous-chimiothèques sont synthétisées une seule fois et les mêmes sous-chimiothèques seront utilisées pour des phases de déconvolution sur plusieurs cibles. Aucune synthèse additionnelle n'est nécessaire cette méthode, ce qui rend cette méthode très rapide.

Comme dans le cas de la déconvolution itérative, le nombre de composés présents dans une chimiothèque croît de manière multiplicative avec les monomères (n_1 à n_i) introduits à chaque étape combinatoire :

Nombre de produits présents dans la chimiothèque = $n_1 \times n_2 \times n_3 \times \dots \times n_i$

alors que le nombre de sous-chimiothèques nécessaires à la déconvolution itérative croît de manière additive :

Nombre de sous-chimiothèques nécessaires à la déconvolution = $n_1 + n_2 + n_3 + \dots + n_i$.

La comparaison du nombre de sous-chimiothèques nécessaires à la déconvolution par balayage positionnel de chimiothèques de tripeptides et d'hexapeptides impliquant 15 ou 30 est identique à celle décrite au tableau 3. **Le nombre de sous-chimiothèques à synthétiser pour des chimiothèques de petite taille ou de grande taille ne varie que faiblement**, ce qui rend cette méthode particulièrement attractive pour la déconvolution de grandes chimiothèques.

Le nombre de composés synthétisés à l'issu des étapes de déconvolution.

Le nombre de composés présents dans les sous-chimiothèques ne décroît pas comme dans le cas d'une déconvolution itérative.

Le tableau 5 permet d'évaluer le nombre de composés nécessaires à la déconvolution par balayage positionnel des 11.390.625 hexapeptides possibles, synthétisés à partir de 15 acides aminés. Celui-ci est de $6 \times 11.390.625$, soit un « **excès de synthèses** » de **500 %**. Le balayage de chaque position nécessite la synthèse de la totalité des composés de la chimiothèque. Ainsi, lorsque N positions sont à définir, la totalité des composés de la chimiothèque sera synthétisée N fois, mais en répartissant à chaque fois les composés dans des sous-chimiothèques distinctes. Aucun effet de concentration du produit actif dans les différentes sous-chimiothèques ne peut être obtenu car toutes les sous-chimiothèques comportent le même nombre de composés.

	Position 1	Position 2	Position 3	Position 4	Position 5	Position 6	Total
Nombre de sous-chimiothèques	15	15	15	15	15	15	90
Taille des sous-chimiothèques	759.375 composés	759.375 composés	759.375 composés	759.375 composés	759.375 composés	759.375 composé	68.343.750 composés

Tableau 5 : nombre et taille des sous-chimiothèques synthétisées au cours de la déconvolution par balayage positionnel d'une chimiothèque d'hexapeptides impliquant 15 acides aminés à chaque position combinatoire

La taille importante des sous-chimiothèques pose le problème de la **dilution de l'activité** d'un composé actif au sein d'un mélange complexe de composés inactifs. Cependant, comme nous l'avons déjà évoqué dans le paragraphe précédent, dans la plupart des cas, l'activité d'un mélange combinatoire ne repose pas uniquement sur un composé (ligand optimal), mais sur un ensemble de composés ayant des structures apparentées (ligands suboptimaux) et pouvant engendrer une activité par **effet cumulatif**. Les sous-chimiothèques produites dans le cadre du

balayage positionnel étant de taille importante, le risque d'obtenir un SBF $\gg 1$ reste à prendre en compte.

La fiabilité de la méthode

Le plus souvent, plusieurs monomères sont trouvés actifs pour chaque position. Deux stratégies sont alors possibles :

- la première consiste à ne retenir que le monomère qui confère l'activité la plus élevée pour chaque position et de synthétiser la combinaison a priori la plus active. Le risque d'une telle stratégie est de ne retenir, pour chaque position, que le monomère le plus « tolérant » vis à vis des positions adjacentes, et non pas le monomère présent dans le(s) produit(s) le(s) plus actif(s).

- la seconde, plus fiable, consiste à prendre en compte le fait que plusieurs produits peuvent être actifs et qu'en chaque position, plusieurs monomères doivent être retenus. Dans le cas des hexapeptides, lorsque 3 monomères sont retenus pour chaque position, 729 (3^6) peptides possibles seront synthétisés. Il a aussi été proposé par Dooley et al (Dooley.1993) d'effectuer une nouvelle déconvolution sur cette chimiothèque de 729 combinaisons possibles, soit de manière itérative, soit par balayage positionnel.

Contrairement à la déconvolution itérative, aucun effet d'enrichissement n'est obtenu au cours du processus de déconvolution par balayage positionnel.

Cette méthode de déconvolution, a permis l'identification de peptides antibactériens Houghten et al 1991 (Houghten 1991, Owens 1991), d'inhibiteurs d'HIV protéase (Owens 1991), d'inhibiteurs d'adhésion des plaquettes à la fibronectine (Houghten 1992) ...

La génération d'informations complémentaires

Les activités biologiques obtenues lors des tests des différentes sous-chimiothèques restent difficiles à exploiter du fait qu'aucune position (autre que celle sur laquelle est effectué le balayage) n'est définie. Aucune information biologique ne peut être corrélée à une connectivité définie.

Conclusion :

La méthode de déconvolution par balayage est une méthode nécessitant un effort de synthèse relativement important, tous les composés de la chimiothèques étant synthétisés autant de fois qu'il existe de positions combinatoires dans la chimiothèque désirée. Les résultats biologiques obtenus sont parfois difficilement exploitables et nécessitent généralement la synthèse de

plusieurs produits individuels (pouvant s'avérer inactifs) afin d'aboutir finalement à la structure du (des) produit(s) actif(s). Cependant l'avantage indéniable de cette procédure reste la rapidité, les différentes positions étant définies en parallèle et les sous-chimiothèques synthétisées une seule fois, pouvant être utilisées à la déconvolution d'une même chimiothèque sur plusieurs tests biologiques.

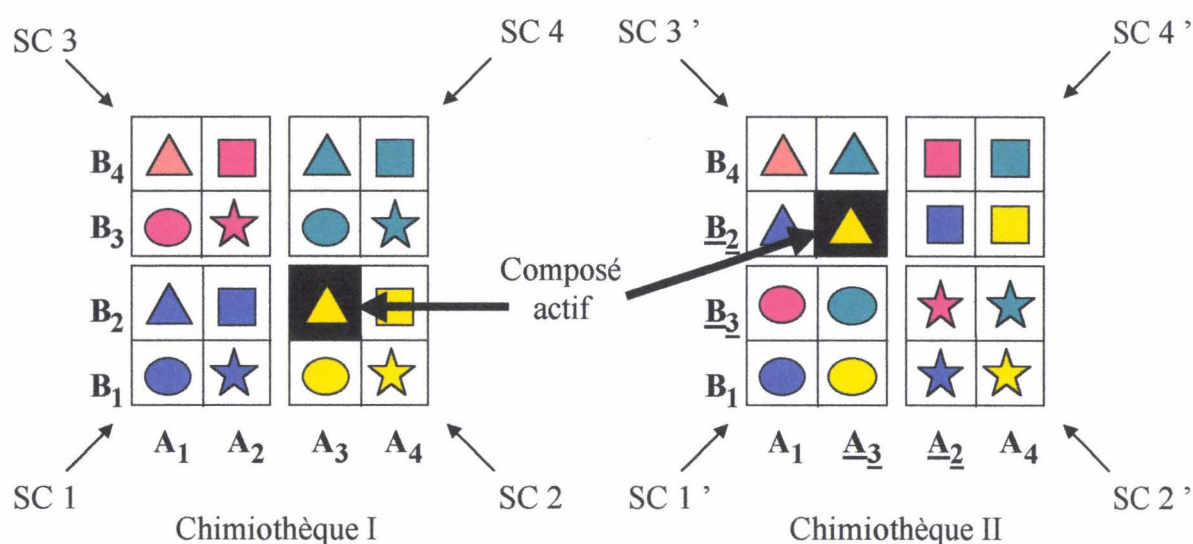
III-B-3 Les chimiothèques orthogonales

Benoît Déprez, Xavier Williard, Line Bourel, Hervé Coste, François Hyafil and André Tartar, J. Am. Chem. Soc. 1995, 117, 5405-5406

Définition : développée au sein de notre équipe, cette méthode de déconvolution consiste à préparer une même chimiothèque sous deux formats différents I et II, de telle sorte que chaque sous-chimiothèque composant I ne possède qu'un seul composé commun avec chaque sous-chimiothèque composant II.

Principe : (illustré figures 9 et 12)

Afin de schématiser plus facilement ce principe, nous prendrons l'exemple d'une chimiothèque de $4 \times 4 = 16$ dimères, issue de l'incorporation de 4 monomères différents en 2 positions combinatoires distinctes (monomères A_1 à A_4 pour la position 1, monomères B_1 à B_4 pour la position 2) $N=2$, $n_a = n_b = 4$ (figure 9). La chimiothèque I contient 16 composés répartis dans 4 sous-chimiothèques de 4 composés. La chimiothèque II contient les mêmes composés répartis également dans 4 sous-chimiothèques, mais de façon orthogonale aux sous-chimiothèques de I, c'est à dire que toute sous-chimiothèque de II possédera un seul composé commun et un seul avec toute sous-chimiothèque de I. Les sous-chimiothèques de I seront d'abord testées. Si une sous-chimiothèque s'avère active, l'identité du produit actif sera obtenue en testant les sous-chimiothèques de II. En effet, après avoir identifié la sous-chimiothèque de II porteuse de l'activité la structure du produit actif, correspondant au seul composé commun, sera déduite. La figure 9 montre que le produit actif est présent à la fois dans une sous-chimiothèque de I (SC2) et dans une sous-chimiothèque de II (SC3') et qu'il constitue le seul composé partagé.



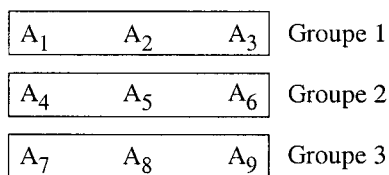
SC = sous-chimiothèque

Figure 9 : Composition de deux chimiothèques orthogonales de formule générale A_n-B_n (n = 4)

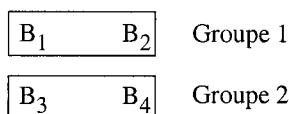
Le principe de la partition orthogonale est facilement applicable aux chimiothèques issues de la chimie combinatoire. Dans ce cas, la synthèse de la première chimiothèque sera réalisée en utilisant, lors des étapes de couplage, des groupes prédéfinis de monomères, de composition constante pour une même position combinatoire (cf. figure 10). Pour synthétiser les sous-chimiothèques composant une deuxième chimiothèque, partition orthogonale de la première, il suffit de constituer de nouveaux groupes de monomères, de composition constante, **ne partageant au maximum, qu'un seul monomère** avec tout groupe défini pour la synthèse de la première chimiothèque (cf. figure 10). La conséquence, sur la composition des groupes, sera que le nombre de monomères par groupe soit inférieur ou égal au nombre de groupes (définis pour chaque position combinatoire) pour au moins une des 2 partitions (cf. figure 11). Cette règle peut ne plus être valable si les groupes d'une même partition ne sont pas de taille identique.

Chimiothèque I

Position A : 9 monomères

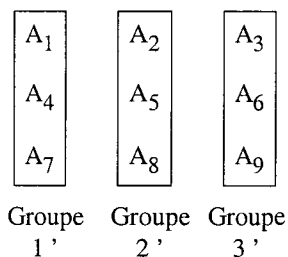


Position B : 4 monomères

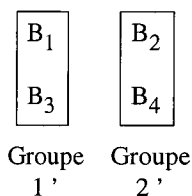


Chimiothèque II (orthogonale)

Position A



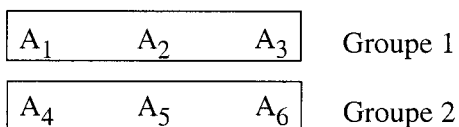
Position B



Rq : les monomères d'un même groupe seront toujours associés

Figure 10 : constitution de groupes orthogonaux de monomères pour une chimiothèque de formule générale : A_n - B_{n'} (n = 9, n' = 4)

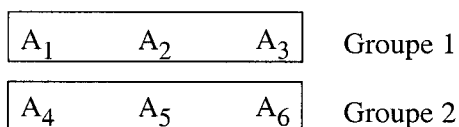
Partion P de 6 monomères en 2 groupes de 3



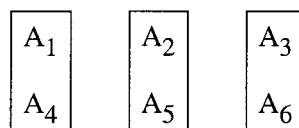
*Partion P' de 6 monomères
en 2 groupes de 3 ne
partageant au maximum
qu'un seul monomère avec
les groupes 1 et 2*

impossible

Partion P de 6 monomères en 2 groupes de 3



*Partion P' de 6 monomères
en 3 groupes de 2 ne
partageant au maximum
qu'un seul monomère avec
les groupes 1 et 2*



possible si le le nombre de monomères par groupe (= 2) est inférieur ou égal au nombre de groupes (= 3)

Figure 11 : contrainte à considérer lors de la constitution de groupes, pour permettre la synthèse de chimiothèques orthogonales

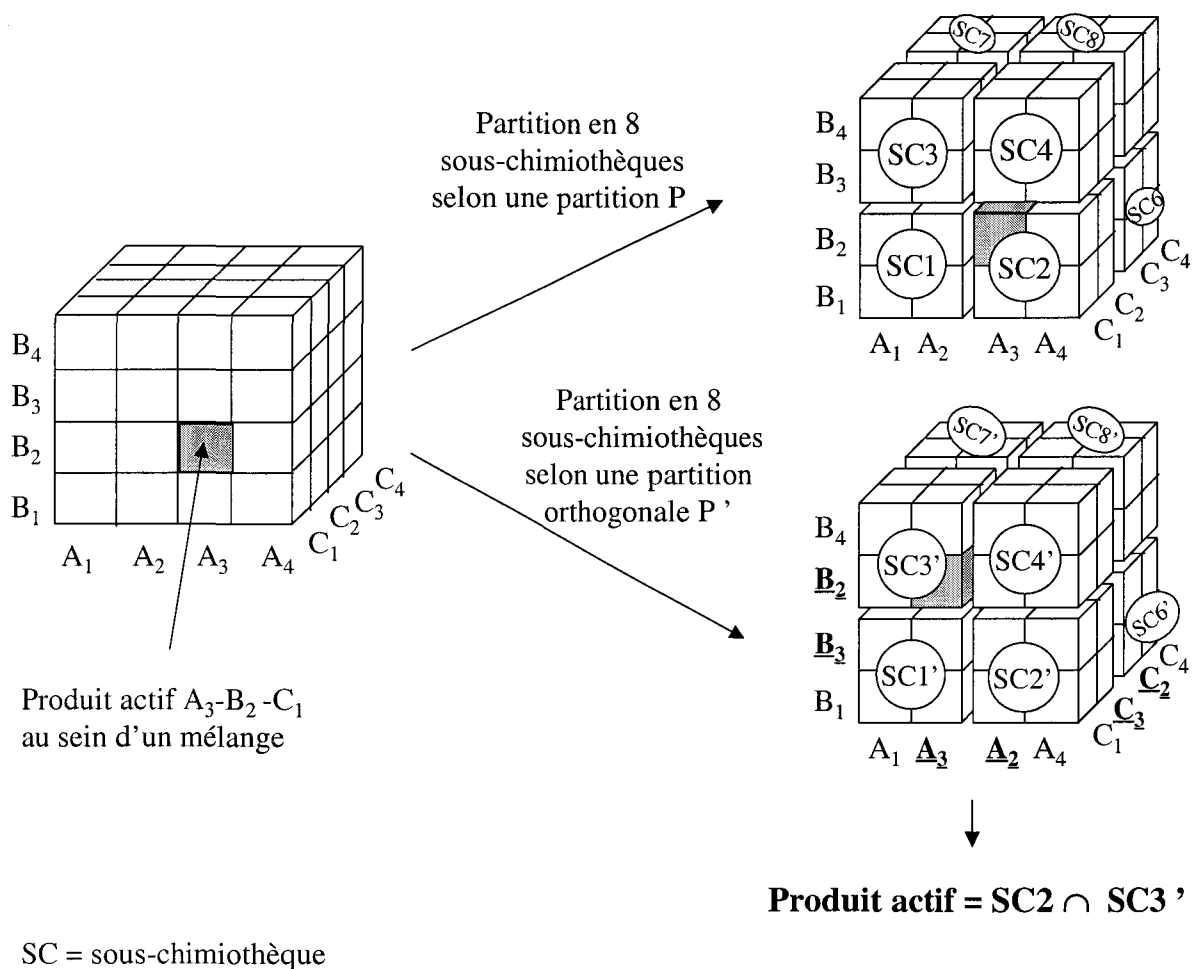


Figure 12 : Composition de deux chimiothèques orthogonales de formule générale $A_n-B_n-C_n$ ($n = 4$)

Discussion :

Le nombre et la durée des étapes permettant l'identification du produit actif

Comme pour la technique du balayage positionnel, la synthèse suivi du criblage de deux chimiothèques orthogonales permettent une **identification directe** et rapide d'un composé actif (lorsqu'il est le seul à être actif). En théorie, une seule synthèse sera nécessaire à l'issue du criblage de 2 chimiothèques orthogonales : la synthèse du produit supposé actif afin de confirmer les résultats biologiques.

Le nombre et la taille des sous-chimiothèques nécessaires à la déconvolution par des chimiothèques orthogonales sera directement fonction du nombre de positions combinatoires

ainsi que du nombre de groupes de monomères définis lors de leur partition. Le tableau 6 illustre à la fois l'organisation d'une chimiothèque de tripeptides et d'hexapeptides impliquant 15 ou 30 monomères (partitionnés en groupes de 3) pour chaque position combinatoire ainsi qu'une chimiothèque d'hexapeptides impliquant 30 monomères (partitionnés en groupes de 6).

Deux informations principales peuvent être déduites de ce tableau :

- il est impératif de **limiter le nombre de groupes** constitués lors de la partition des monomères occupant les différentes positions combinatoires, le nombre de sous-chimiothèques préparées croissant de manière multiplicative avec le nombre de groupes constitué pour chaque position

- la conséquence directe de préférer un nombre limité de groupes de taille importante plutôt qu'un grand nombre de petits groupes sera la synthèse de sous-chimiothèques de grandes tailles pouvant, comme nous l'avons décrit précédemment, impliquer un effet d'activités cumulées (cf. § III-B-1). Or cet effet est à éviter tout particulièrement pour une déconvolution de ce type car il conduit à augmenter le nombre de sous-chimiothèques présentant une activité, ce qui multiplie les intersections possibles avec les sous-chimiothèques actives qui leurs sont orthogonales. Dans le cas où 2 sous-chimiothèques (SC 1 et SC 2) d'une chimiothèque I et 2 sous-chimiothèques (SC 1' et SC 2') d'une chimiothèque II (orthogonale à la première) présentent une activité, 4 composés devront être resynthétisés :

SC 1 \cap SC 1' SC 2 \cap SC 1'

SC 1 \cap SC 2' SC 2 \cap SC 2'

Le nombre de possibilités croissant de façon multiplicative avec le nombre de sous-chimiothèques actives conduit rapidement à un grand nombre de combinaisons à synthétiser. Afin de limiter au maximum cet effet cumulatif, il est indispensable à la fois de limiter la taille des sous-chimiothèques, et surtout d'éviter la présence concomitante de structures apparentées. Il faudra par conséquent, **limiter la taille des groupes** constitués lors de la partition et **assurer une grande diversité structurale des monomères** au sein de ces groupes.

		Tripeptides	Hexapeptides
A partir de 15 acides aminés partitionnés en 5 groupes de 3	Nombre total de peptides	3.375	11.390.625
	Nombre de sous-chimiothèques	5³ soit 125	5⁶ soit 15625
A partir de 30 acides aminés partitionnés en 10 groupes de 3	Nombre total de peptides	27.000	729.000.000
	Nombre de sous-chimiothèques	10³	10⁶
A partir de 30 acides aminés partitionnés en 5 groupes de 6	Nombre total de peptides	27.000	729.000.000
	Nombre de sous-chimiothèques	5³ soit 125	5⁶ soit 15625

Rq : les chiffres indiqués concernent la synthèse d'une des deux chimiothèques orthogonales

Tableau 6 : comparaison du nombre de sous-chimiothèques (préparées pour une des deux chimiothèques orthogonales) nécessaires à la déconvolution de chimiothèques de tripeptides et d'hexapeptides impliquant 15 ou 30 acides aminés

Le nombre de composés synthétisés à l'issue des étapes de déconvolution.

Quelque soit le type de chimiothèque, il sera nécessaire de synthétiser deux fois tous les composés, soit un excès de synthèse de **100 %**, ce qui constitue un faible effort de synthèse si l'on prend en compte le fait que ces chimiothèques orthogonales pourront ensuite être utilisées plusieurs fois pour différents tests biologiques.

La fiabilité de la méthode sera fortement dépendante de l'organisation de la chimiothèque.

Comme nous l'avons décrit dans le paragraphe « Le nombre et la durée des étapes permettant l'identification du produit actif », il est indispensable de garantir une diversité structurale importante au sein de chaque sous-chimiothèque. D'autre part, cette méthode offre l'avantage, par rapport aux méthodes de déconvolution décrites précédemment, de valider toute activité détectée dans une sous-chimiothèque par une activité dans une sous-chimiothèque orthogonale dont la composition du mélange est totalement différente. Ce « contrôle interne » permet de détecter la présence de « faux positifs », pouvant notamment apparaître par un effet d'activités cumulées. Cette méthode mise au point et utilisée dans notre laboratoire nous a permis la découverte d'un ligand du récepteur V₂ à la vasopressine (cf. § V).

La génération d'informations complémentaires

En dehors de la possibilité de corréler les résultats obtenus sur chacune des chimiothèques, pour établir un « contrôle interne », aucune information complémentaire ne peut être utilisée sans manipulation supplémentaire.

Conclusion :

La méthode de déconvolution par chimiothèques orthogonales est une méthode pouvant à la fois s'appliquer au criblage de chimiothèques combinatoires, qu'elles soient polymériques ou non, mais aussi au criblage de composés provenant d'origines diverses, mais organisés en mélanges « orthogonaux », cette méthode est actuellement utilisée par d'autres groupes de recherche pour ces deux applications. L'identification du (des) produit(s) actif(s) se fait très rapidement et ne nécessite, en général, que la synthèse d'un nombre limité de composés pour confirmer les résultats biologiques. La fiabilité de cette méthode restant très dépendante des problèmes d'activité par effet cumulatif, limitera cette technique à la déconvolution de chimiothèques, de petite taille (quelques milliers de composés), à forte diversité structurale.

III-C Comment assurer l'équimolarité des composés synthétisés en mélange ?

Les méthodes de déconvolution décrites ci-dessus, reposent toutes sur les valeurs d'activité biologique obtenues pour différentes sous-chimiothèques afin d'orienter la démarche d'identification du produit actif au sein d'un mélange. L'hypothèse de départ est que, à chaque étape de sélection, la sous-chimiothèque la plus active contient le produit le plus actif. Comme nous avons pu le voir précédemment, l'influence de ligands suboptimaux (cf. § III-B-1) fausse ce raisonnement. Un autre paramètre pouvant biaiser ces méthodes est la présence des différents composés à des concentrations variables, au sein des différentes sous-chimiothèques testées. Dans ce cas, un produit très actif, mais présent à une faible concentration dans une sous-chimiothèque, lui conférera une activité sous-estimée. Dans les divers processus de déconvolution, cette sous-chimiothèque risque donc d'être écartée, au profit d'une autre sous-chimiothèque contenant un produit moins actif mais dont la concentration est supérieure. Ainsi seules les procédures de synthèse permettant d'aboutir à la synthèse de mélanges de composés présents dans les mêmes proportions devront être retenues. Les deux approches suivantes peuvent être utilisées. Nous illustrerons leur description en faisant intervenir la synthèse en phase solide, car des excès de réactifs doivent souvent être utilisés.

- **Le couplage de mélanges de monomères**

Cette approche, techniquement simple à réaliser, fût la première utilisée pour la construction de chimiothèques. Elle implique de constituer le mélange de monomères avant l'étape de couplage. Deux solutions ont été proposées :

La première (cf. figure 13) consiste à effectuer un premier couplage du mélange de monomères à incorporer en assurant non seulement une **concentration identique en chaque monomère** dans le mélange à incorporer mais aussi une **stœchiométrie du nombre d'équivalents de monomères à coupler par rapport au nombre d'équivalents de fonctions réactives présentes sur la résine**. En appliquant un temps de couplage prolongé, les monomères les moins réactifs pourront réagir avec les fonctions encore disponibles sur la résine. Un second couplage sera ensuite effectué avec un mélange identique de monomères (par conséquent en excès par rapport aux fonctions réactives résiduelles de la résine) afin d'aboutir à une réaction totale. Cette méthode a été utilisée avec succès pour le couplage d'acides aminés. Nous avons d'ailleurs contrôlé, après une hydrolyse acide totale de mélanges de 125 tripeptides, que les taux d'incorporation obtenus variaient en moyenne de +/- 20 % par rapport aux taux d'incorporation théoriques.

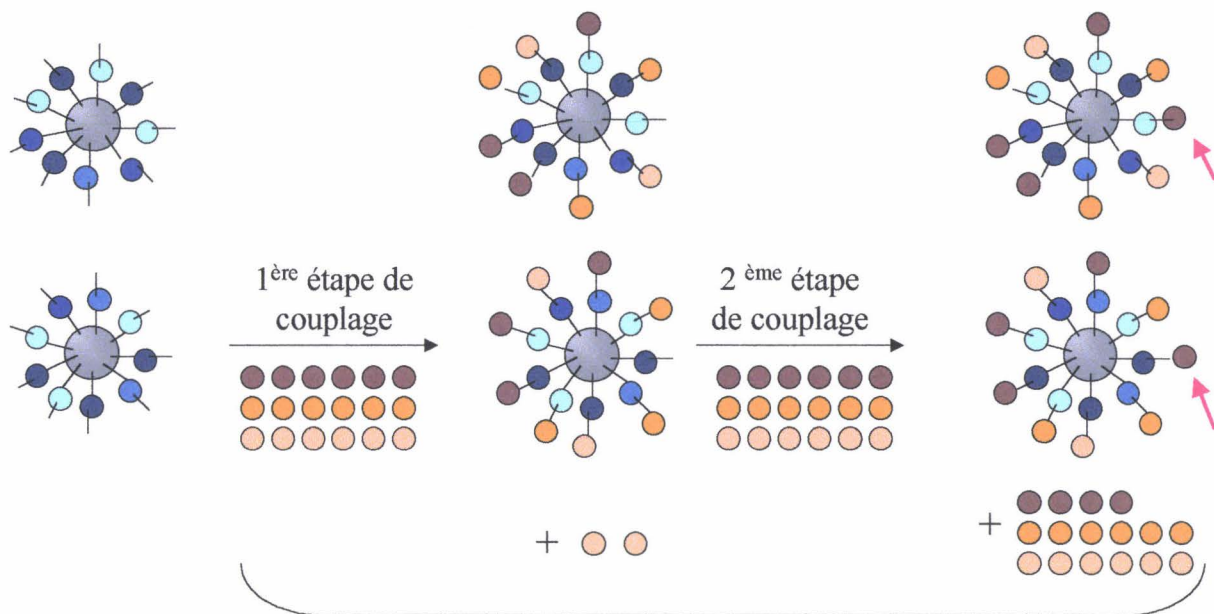
La seconde solution proposée par Ostresh et al (Ostresh 1994 b) consiste à coupler un mélange d'acides aminés, présents à des concentrations différentes, afin de niveler les cinétiques de couplage de tous les acides aminés. Un acide aminé peu réactif sera donc introduit en plus large excès qu'un acide aminé réactif. Le tableau 7 décrit les ratios déterminés pour chacun des 18 acides aminés d'un mélange introduit en excès (10 équivalents). Ce ratio devra être redéfini si la composition du mélange est changée.

Ala	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Tyr
3,58	3,7	3,84	2,66	3,04	3,76	18,3	6,56	5,23	2,42	5,64	4,56	5,61	6,87	2,93	5,04	11,9	4,36

Rq : chaque nombre représente le pourcentage molaire d'acide aminé nécessaire pour le couplage du mélange de ces 18 Boc-acides aminés introduits en excès (10 équivalents) activés sous forme d'esters d'HOBt (en utilisant la diisopropylcarbodiimide), le solvant utilisé étant le DMF.

Tableau 7 : ratios de 18 acides aminés permettant un couplage équimolaire

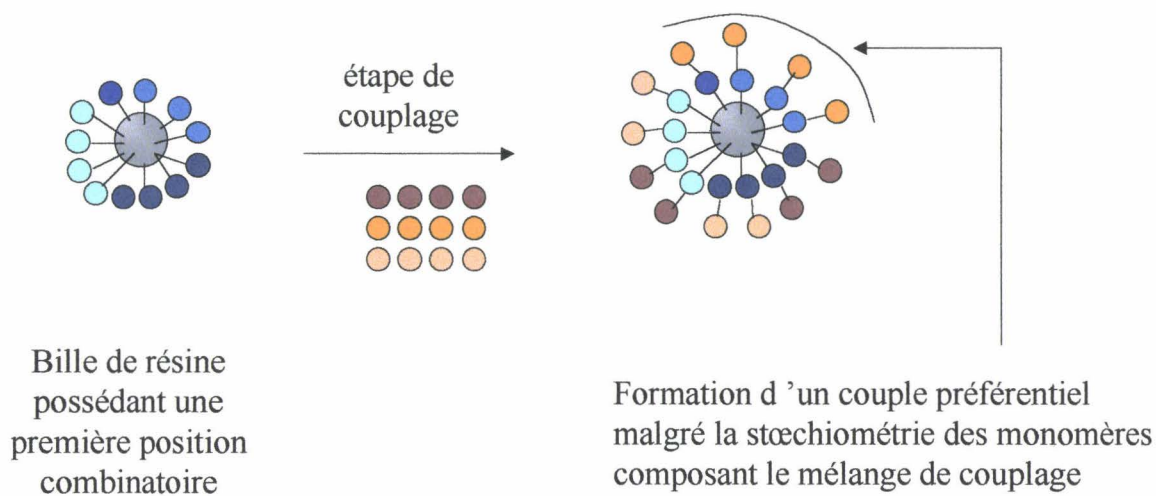
Il faudra cependant remarquer que la formation de couples préférentiels d'acides aminés (cf. figure 14) correspondant au couplage préférentiel d'un des monomères du mélange introduit, avec un des monomères fixés sur la résine, ne peut être évitée pour l'une ou l'autre des deux solutions proposées précédemment.



Billes de résine possédant une **première position** combinatoire

Couplage d'un mélange de monomères sur une **deuxième position** combinatoire.

Figure 13 : technique de couplage de monomères en mélange permettant de tendre vers l'équimolarité de tous les composés synthétisés dans un mélange



Bille de résine possédant une première position combinatoire

Formation d'un couple préférentiel malgré la stœchiométrie des monomères composant le mélange de couplage

Figure 14 : formation d'un couple préférentiel après le couplage d'un mélange de monomère

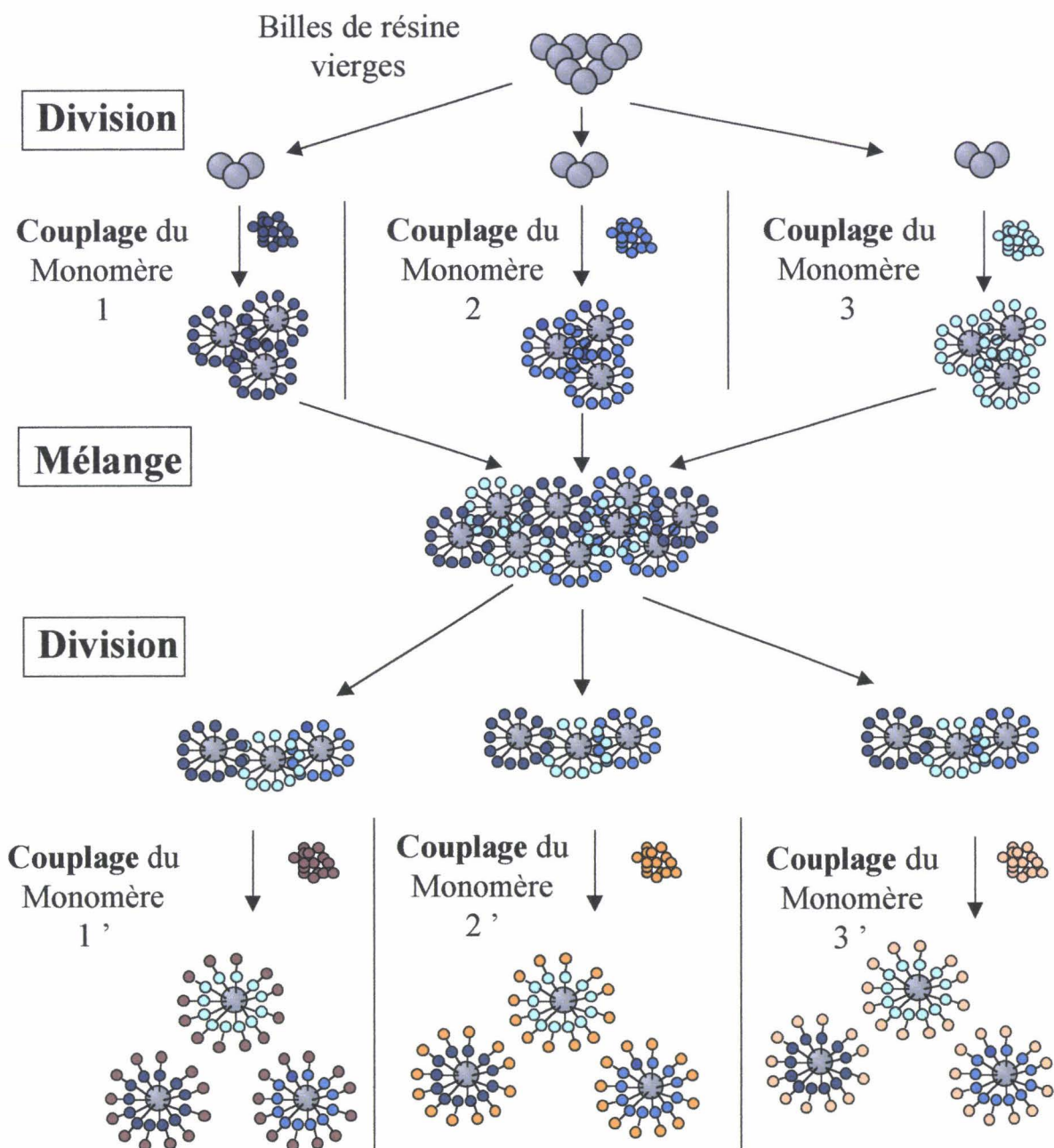
- *Le couplage selon le mode « division / mélange » ou « split and mix »*

Encore appelé méthode DCR (Divide-Couple-Recombine) ou méthode « one bead one peptide ».

Principe : soit un mélange de n monomères à coupler sur un lot de résine de taille T , le lot de résine sera d'abord divisé en n fractions identiques de taille T / n , chaque monomère sera couplé individuellement à une des n fractions (le monomère 1 avec la fraction 1, le monomère 2 avec la fraction 2 ...le monomère n avec la fraction n). Les fractions seront ensuite réunies et la présence équimolaire de chaque monomère sera ainsi garantie. Une propriété inhérente à cette méthode est la présence d'un seul type de composé en surface de chaque bille. En effet, à aucun moment, un mélange de monomères ne sera couplé. Cette caractéristique est schématisée figure 15.

Cette méthode reste cependant soumise à une contrainte majeure : l'existence d'un nombre de billes de résine suffisant pour aboutir à une équimolarité de tous les composés synthétisés.

En effet, pour que cette méthode soit fiable, il faudra qu'à chaque étape de division de résine, les lots formés soient identiques et qu'ils contiennent toutes les combinaisons synthétisées antérieurement, avec une répartition équimolaire. Zhao et al (Zhao 1995, Zhao 1996) ont analysé la déviation de la distribution des produits par rapport à la distribution équimolaire idéale d'une chimiothèque de tripeptides impliquant 20 acides aminés à chaque position combinatoire, soit $20 \times 20 \times 20 = 8000$ composés. Les billes prises en considération ont un diamètre de $90 \mu\text{m}$ et un gramme de résine en contient environ $2,86 \cdot 10^6$.



En surface de chaque bille n'est présent qu'un seul type de combinaison, d'où l'expression « one bead, one peptide »

Figure 15 : illustration de la méthode de synthèse « split and mix »

L'évaluation de cette déviation s'effectue en prenant en compte les abondances relatives obtenues pour chaque composé, au sein de toutes les paires de composés possibles. Pour un intervalle de confiance $[(1 - \alpha) \times 100 \text{ \%}]$ défini, la quantité de résine nécessaire sera calculée de telle sorte que le ratio des concentrations du composé le plus représenté sur le moins représenté, soit inférieur ou égal au ratio $(1 + L_2) / (1 - L_2)$ où L_2 compris entre 0 et 1

représente la valeur de la limite tolérable. Le tableau 8 permet d'obtenir, pour un intervalle de confiance fixé, la quantité de billes nécessaires pour la synthèse des 8000 tripeptides possibles en fonction de limites L_2 définies. Dans le cas, par exemple, où le ratio maximum acceptable du composé le plus représenté sur le moins représenté doit être $\leq 1,5$, alors $(1 + L_2) / (1 - L_2)$ sera $\leq 1,5$, d'où $L_2 \leq 0,2$, la quantité de résine nécessaire, pour un intervalle de confiance 95 % sera de 1,417 g. Si le ratio acceptable du composé le plus représenté sur le moins représenté devient 1,1 ($L_2 = 0,05$) la quantité de résine nécessaire sera plus importante (22,669 g).

Limite tolérable L_2	Quantité de résine* nécessaire (g)	
	$(1 - \alpha) = 0,95$	$(1 - \alpha) = 0,99$
0,01	566,734	652,502
0,05	22,669	26,100
0,1	5,667	6,525
0,2	1,417	1,631
0,3	0,630	0,725
0,5	0,227	0,261
0,8	0,089	0,102
0,99	0,058	0,067

* : les billes de résines ont un diamètre de 90 μm et 2,86 10^6 billes pèsent 1 gramme et les quantités de résine nécessaire sont calculés sur la base de la synthèse d'une chimiothèque de 8000 tripeptides issus de 20 acides aminés (20 x 20 x 20)

Tableau 8 : quantités de résine nécessaires pour permettre un couplage équimolaire en utilisant une méthode de type « split and mix »

Conclusion :

La méthode « split and mix » offre de nombreux avantages. Elle permet :

- d'assurer l'équimolarité des produits présents dans une chimiothèque, lorsque la quantité de résine de départ est suffisante
- de contrôler le couplage de chaque acide aminé incorporé
- d'éviter des réactions secondaires pouvant subvenir lors de couplages d'acides aminés en mélange
- d'éviter la formation de couples préférentiels

- d'obtenir un seul type de composé par bille, caractéristique pouvant s'avérer intéressante lors d'identification de composés actifs par séquençage direct sur une bille de résine (cf. § IV).

Cependant cette méthode restera limitée à la synthèse de chimiothèques de quelques milliers de composés, car au-delà, pour assurer l'équimolarité des produits synthétisés, elle nécessiterait d'utiliser des lots de résine de taille trop importante, de plus celle-ci pourrait s'avérer particulièrement fastidieuse du fait du nombre important d'étapes de division et de mélange de résine à effectuer.

En revanche, même si la méthode faisant appel au couplage de mélanges de monomères reste moins fiable sur le plan de la composition équimolaire en composés présents dans la chimiothèque que la méthode précédente, celle-là permettra d'accélérer les cadences de production et sera beaucoup mieux adapté à la synthèse de chimiothèques de grande taille (plusieurs dizaines de milliers de composés).

IV Les méthodes d'élucidations structurales, des produits conférant l'activité aux mélanges, ne faisant pas appel à la déconvolution (Cortese 1996, Williard 1996)

L'identification directe de la molécule active : cette technique exige la possibilité de repérer, d'isoler puis d'identifier le produit actif par séquençage (peptides (Lam 1991)) ou par spectrométrie de masse (utilisation de la fragmentation MS/MS (Stankova 1994)). La synthèse en phase solide de type « split and mix » de la chimiothèque ainsi que des techniques de micromanipulation sont donc nécessaires.

L'étiquetage ou « tagging » : cette technique consiste à associer un agent identifiant à chaque molécule synthétisée. L'étiquette peut être une structure chimique ou composant électronique émetteur d'une radiofréquence (cf. Seconde partie, Chapitre A). Dans le cas d'une étiquette chimique, celle-ci pourra être liée directement aux produits combinatoires soit liée à la bille de résine. La technique de l'étiquetage nécessite donc en général de repérer la bille de polymère portant le produit actif, d'isoler l'étiquette correspondante puis de l'identifier par PCR puis séquençage (Brenner 1992) ou par chromatographie phase gaz (Ohlmeyer 1993, Borchardt, 1994). La synthèse en phase solide de type « split and mix » de la chimiothèque ainsi que des techniques de micromanipulation sont donc nécessaires.

L'inconvénient majeur des deux techniques précédentes, reste la présence de la bille de résine (ou éventuellement des étiquettes, liées aux molécules testées) lors du test biologique, empêchant l'ensemble de pénétrer au sein de macromolécules, de cellules ..., ce qui limite les tests biologiques utilisables. Cet obstacle a pu être contourné en utilisant le **principe du clivage séquentiel** (Salmon 1993) basé sur le clivage orthogonal de composés présents sur les billes de résine.

Principe : les peptides combinatoires, issus d'une synthèse de type split and mix, fixés sur les billes de résine sont répartis dans différents récipients (à raison de 500 billes par récipient). Le clivage du premier tiers des composés (sensibles au premier agent de clivage), présents sur chaque bille, est réalisé et les produits sont testés en solution. Le récipient ayant libéré le mélange actif est identifié et les 500 billes de ce mélange sont ensuite réparties dans 500 récipients différents. Le clivage du deuxième tiers des composés (sensibles au deuxième agent de clivage) est réalisé, la bille ayant libéré le produit actif est isolée puis le dernier tiers des composés, restant sur la bille, est séquencé.

La séparation « spatiale » des composés synthétisés : tous les composés résultant de la matrice combinatoire seront synthétisés mais de façon isolée. Les premières synthèses de ce type ont fait appel au système des pins (cf. Seconde partie, Chapitre A) développé par Geysen dès 1984 (Geysen 1984) puis à des méthodes d'adressage spatial par irradiations lumineuses (Fodor 1991) utilisées par la société Affymax (techniques développées antérieurement pour la fabrication de circuits imprimés faisant appel à des systèmes de caches). Cette approche a été poursuivie par l'utilisation massive des techniques de synthèse parallèle robotisée faisant l'objet de la seconde partie de cette thèse.

Deconvolution Tools in Solution-Phase Synthesis

Xavier Williard and André Tartar

When natural products are screened for new pharmacological functions, the detection of an activity must be followed by purification and structure determination. Combinatorial chemistry, however, offers the possibility of generating mixtures of structurally diverse compounds using a predetermined format, greatly simplifying the identification of a biologically active compound within the mixture. There are two basic approaches to identifying active compounds in combinatorial chemistry. Tagging involves linking an identifier (either directly or, more often, indirectly) to each type of molecule present in the library; unfortunately the chemistry of the tag may prevent the use of certain types of assays or modify the recognition by target biological macromolecules. Here, we concentrate on the second approach, deconvolution, in which the library is partitioned into a series of sublibraries that are organized in such a way as to allow the identification of an active compound without the necessity for a tag.

Combinatorial Libraries are N -Dimensional Matrices

The basis for deconvolution strategies and the reason they work is that, as outlined by Pirrung (1), any combinatorial library can be conceptually represented as an N -dimensional matrix. Each axis of the matrix has as many elements as the number (n) of building blocks present in each set used in the synthesis, and N is the number of steps of combinatorial synthesis involved in the preparation of the library. Two examples are given in Figure 1. The first matrix (Figure 1a) corresponds to two steps of combinatorial synthesis using four different building blocks in position A ($n_A = 4$) and four different building blocks in position B ($n_B = 4$). This library con-

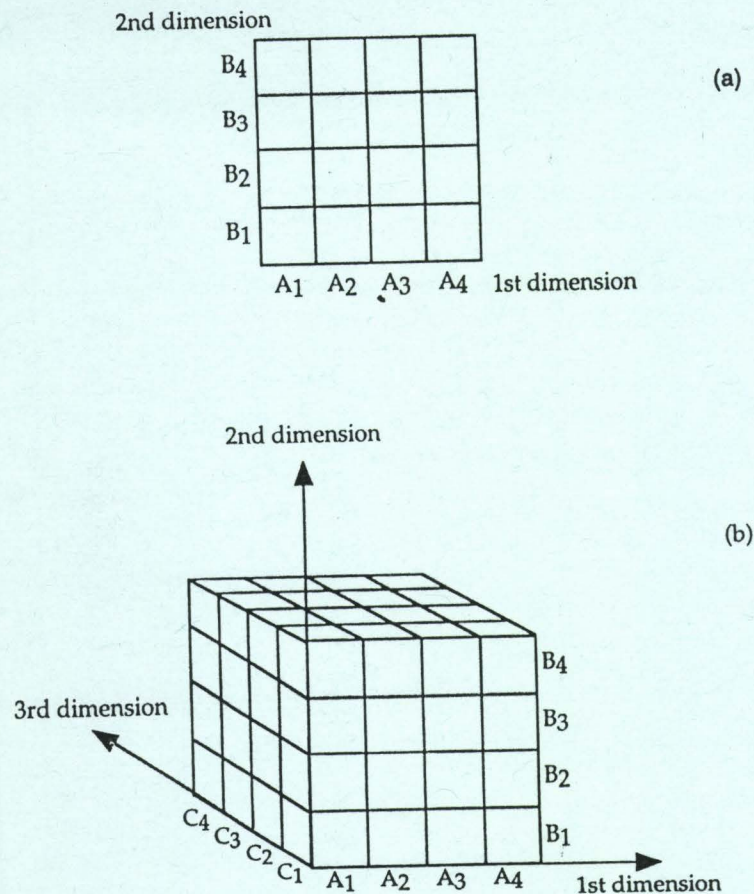


FIGURE 1. Matrix representation of combinatorial libraries. (a) A library constructed in two steps ($N = 2$) with four building blocks used at each step ($n_A = n_B = 4$) has 16 compounds ($A_n - B_n$). (b) A library constructed in three steps ($N = 3$) with four building blocks used at each step ($n_A = n_B = n_C = 4$) has 64 compounds ($A_n - B_n - C_n$).

tains $n_A \times n_B = 16$ compounds, and each compound can be localized in a different cell of the matrix (A_x, B_y). The second matrix (Figure 1b) illustrates a library constructed with three steps ($N = 3$) of combinatorial synthesis, each involving four building blocks ($n = 4$). This library contains 64 compounds ($n^N = 4^3$), and each compound can be localized in a different cell of the matrix defined by the combination A_x, B_y, C_z . Although graphical representation is more difficult for $N > 3$, any combinatorial library can be conceptually represented as an N -dimensional matrix in which each compound of the library is localized in a different cell.

To obtain all the information available from a library, in our case the biological activity of each compound, this value would have to be assessed for each of the

cells by determining the activity of each pure compound (and this is the procedure that is followed for spatially addressable combinatorial libraries (2) prepared by parallel synthesis). In general, however, the purpose of random screening of compound libraries is not to obtain complete information on all compounds present in these libraries. For example, the simplest question that can be asked is whether there is at least one active compound present in the library. Provided that the activity of the active compound can be detected above the background of the mixture, answering this question requires that only one activity determination be performed on a pool containing all the compounds. The next step is to determine the structure of the active compound present in this library. This determination requires additional information from additional activity-determination experiments, but this information still represents a small subset of the complete information that can theoretically be obtained from the library (the activity of each individual compound). Thus the principle of all deconvolution methods is to limit the number of biological experiments, so that only those that will allow the determination of the structure of the most active compound are performed.

Several different deconvolution methods have been proposed. All of them are based on the partition of the complete library into a series of sublibraries, which are then tested individually. The three major methods, which are described in the following sections, are iterative deconvolution, positional scanning, and orthogonal libraries.

Iterative Deconvolution

Iterative deconvolution involves the detection of an activity in the full library, followed by the synthesis and screening of a series of sublibraries containing a decreasing number of compounds until a unique structure is identified. The basic concept is illustrated in Figure 2, which shows as an example a small library constructed using three coupling steps with a mixture of four building blocks used at each step. This yields $4 \times 4 \times 4 = 64$ trimers. According to the general convention used by Houghten (3), in which X represents the mixture of building blocks in a position (four building blocks in this case) and O defines one individual building block (among the four), the complete initial library can thus be represented as XXX.

If a biological activity is detected in this mixture, three steps of deconvolution will be sufficient to identify the structure of the active compound. The first step is to determine the building block that is required in the first position. Four sublibraries (represented as OXX) are prepared, with the first position occupied by a defined building block and the last two positions still containing a mixture of the four building blocks. These four different sublibraries, each containing 16 ($1 \times 4 \times 4$) trimers, are then screened to determine which of the building blocks at the first position confers the biological activity. If the most active sublibrary contains A_3 at this position, four new sublibraries (represented by A_3OX) will be prepared in which the first position is defined as A_3 , the second position is one of four defined building

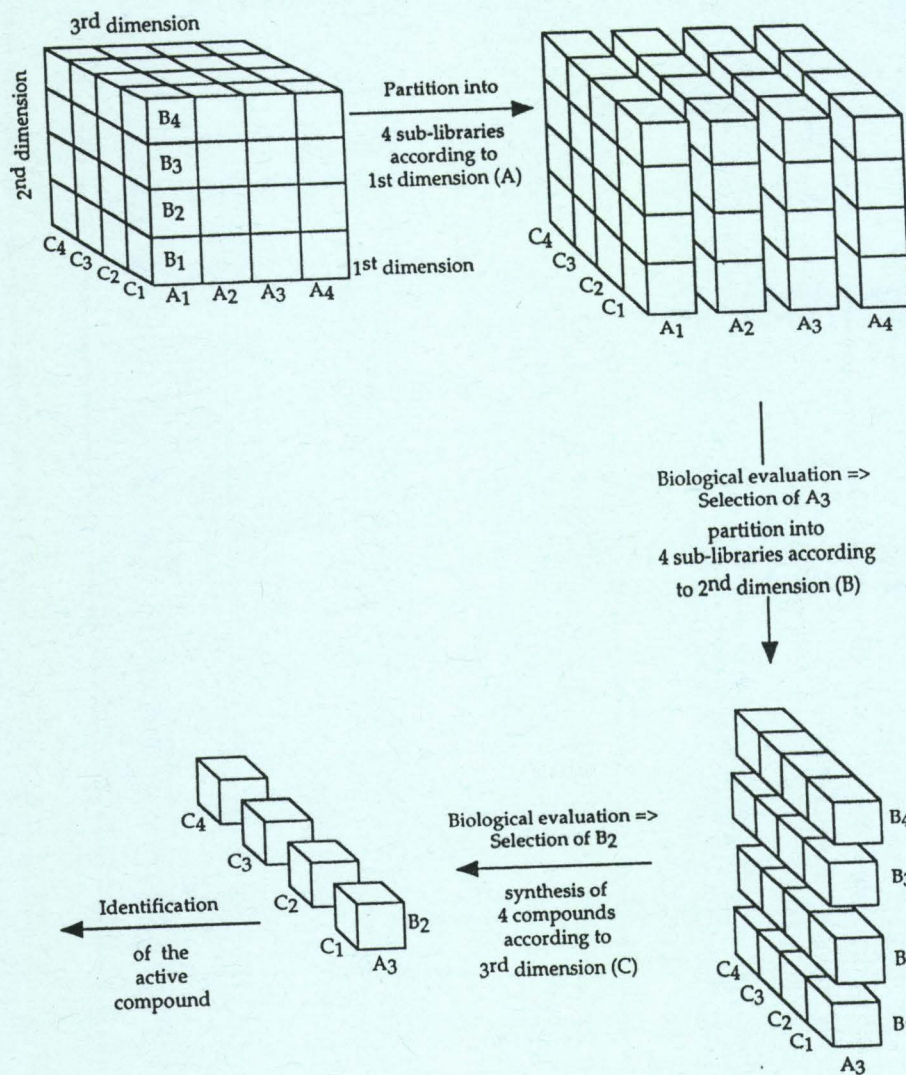


FIGURE 2. Three-step iterative deconvolution of a three-position ($N = 3$), 64-compound combinatorial library $A_n B_n C_n$ with $n_A = n_B = n_C = 4$.

blocks, and the third position still consists of the mixture of four building blocks. The determination of the biological activity of these four new sublibraries will indicate which of the building blocks in the second position confers the activity. If the most active sublibrary contains B₂ in the second position, the third deconvolution step will involve the synthesis of four additional sublibraries (of one compound each, and represented as A₃B₂O), and the evaluation of their biological activity will identify the structure of the active compound. The matrix representation of these three steps of deconvolution is shown in Figure 2. This method generates less complex mixtures at each deconvolution step, so the active sublibraries are enriched relative to their predecessors.

In this example, with three deconvolution steps each involving the preparation of four sublibraries, 12 sublibraries are required to identify an active compound from the 64 possibilities. The efficiency of the method increases sharply as the size of the library increases. The iterative deconvolution of a library that was synthesized in N steps with n building blocks used at each step (and thus containing n^N compounds) requires the preparation and biological evaluation of $N \times n$ sublibraries (over N iterations). Thus a hexapeptide library built from 20 natural amino acids contains 20^6 or 64,000,000 peptides, but its deconvolution requires the synthesis and evaluation of only $6 \times 20 = 120$ sublibraries.

The main limitation in cases such as this is the difficulty in detecting a biologically active compound in a mixture containing so many compounds. In general, however, such libraries contain compounds of limited diversity, and any active compound is generally associated with several suboptimal binders. It is crucial for the chemist to realize that the combined effect of the active and partially active compounds will generate a detectable biological activity, although it may make selecting the best binder difficult. The impact of this effect can be monitored by following the evolution of biological activities during the deconvolution. At each step, an additional position is defined, thus increasing the ratio of the concentration of the active compound to the concentration of inactive compounds. In an ideal library composed of totally unrelated compounds and with n building blocks in one position, the concentration of the active component, and so the biological activity of the mixture, should increase by a factor of n after the deconvolution of this position. In most cases reported to date, biological activities did not improve this much.

Freier et al. (4) have analyzed the evolution of activity when deconvoluting oligomer libraries for 15 cases reported in the literature (Table I). Using these values, they define a suboptimal binding factor (SBF) value, which represents the ratio of the theoretical activity for the initial sublibrary (based on the assumption that the winner is the only active compound among the pool of compounds present in the initial sublibrary) to the observed activity for the initial sublibrary. A high SBF indicates a high incidence of suboptimal binders in the initial sublibrary, whereas a value of 1 would indicate that the winner was the only active compound in the initial sublibrary.

The presence of large numbers of suboptimal binders must be taken into account during deconvolution as each step involves an irreversible choice, yielding more and more restricted arrays of compounds. The simplest approach in the decon-

TABLE I. Suboptimal Binding Factor (SBF) is a Measurement of the Apparent Number of Active Compounds in a Sublibrary

Target for Assay ^a	Library Composition	Number of Monomers	Oligomer Length	Number of Fixed Positions ^b	Subset Complexity ^c	Subset Activity ^d	Winner Activity ^d	SBF ^d
Cellular Screens								
HIV	Phosphorothioate DNA	4	8	2	4096	20 μ M	0.3 μ M	61
HSV	Phosphorothioate DNA	4	8	1	16,384	70 μ M	0.4 μ M	94
<i>S. aureus</i>	Peptide	18	6	2	104,976	450 μ g mL ⁻¹	3.4 μ g mL ⁻¹	793
<i>S. aureus</i>	Peptide	19	6	2	130,321	1730 μ g mL ⁻¹	1.1 μ g mL ⁻¹	829
Antibody Binding								
pAB-FMRF	Peptide	15	4	0	50,625	1400 μ g mL ⁻¹	0.5 μ g mL ⁻¹	18
pAB-pep3	Peptide	16	6	0	16,777,216	6500 μ g mL ⁻¹	0.08 μ g mL ⁻¹	206
mAB-19B10	Peptide	18	6	2	104,976	250 μ M	0.03 μ M	13
mAB-125-10F3	Peptide	19	6	2	130,321	20 μ M	0.004 μ M	26
Nucleic Acid Binding (Experimental)								
<i>Ha-ras</i> RNA	2'-O-methyl RNA	4	9	1	65,536	10 μ M	0.01 μ M	66
Protein Binding or Activity								
HIV-protease	Peptide	22	4	1	11,132	4400 μ M	1.4 μ M	3.5
Opioid receptor	Peptide	19	6	2	130,321	3.452 μ M	0.025 μ M	1057
Opioid receptor	Peptide	19	6	2	130,321	2.1 μ M	0.005 μ M	310
Nucleic Acid Binding								
18-mer target	RNA	4	9	0	262,144	0.020 μ M	5.5×10^{-7} μ M	7.1
9-mer target	RNA	4	9	0	262,144	0.034 μ M	1.1×10^{-6} μ M	8.0
6-mer target	RNA	4	9	0	262,144	7.9 μ M	0.047 μ M	1561

^aAbbreviations: HIV, human immunodeficiency virus; HSV, herpes simplex virus.

^bNumber of fixed positions in the initial round of deconvolution.

^cThe number of unique molecules in each initial subset (the complexity) is N^{L-F} , where N is the number of monomers, L is the oligomer length, and F is the number of fixed positions.

^dThe concentration of the most active subset in the initial round necessary for half-maximal activity (subset activity), the activity of the final compound (winner activity), and the complexity were used to calculate the SBF of the most active sublibrary during the first step of deconvolution. (SBF = winner activity \times subset complexity/subset activity.)

SOURCE: Adapted from Freier et al. (4).

volution strategy assumes that, when comparing several mixtures, the mixture with the greatest activity contains the most active compound. If, however, the building block selected in one position is not the one in the corresponding position of the most active compound, but the one that is shared by the largest number of suboptimal binders, the next steps of deconvolution will lead to a suboptimal binder. To avoid this problem it is important, at each step, to examine which building blocks are found in the three or four most active sublibraries. If two of these building blocks are structurally very different, then it is likely that two different paths exist for iterative deconvolution, and that these paths will eventually lead to totally different compounds. In this case, two independent deconvolutions must be performed using a different building block in each case (see the first of the examples in the section "Selected Examples", later in the chapter). Iterative deconvolution has proven its efficiency in a large number of cases, such as the isolation of nanomolar inhibitors of seven-transmembrane receptors (see example 2).

A major drawback of iterative deconvolution is the requirement for the preparation and biological evaluation of a new set of sublibraries at each step. (If the library is used again with a different assay, it is unlikely that the same sublibraries will be needed.) The number of steps can only be reduced by synthesizing more sublibraries. For example, the step in which the last randomized position of a library is deconvoluted can be eliminated by directly preparing an individual sublibrary for each different building block (avoiding the last recombine step of the divide-couple-recombine (DCR) method), and so generating sublibraries with formats such as XXO or XXXO. The dual positional format described by Houghten (3) is an extension of this strategy to the last two randomized positions. For example, a hexapeptide library can be prepared using four combinatorial steps in which mixtures of 20 amino acids are used, followed by two steps in which individual amino acids are used. This yields $20 \times 20 = 400$ different sublibraries with the general structure OOOXXX, which are then screened individually. Two advantages of this method are that there are only 4 instead of 6 steps at which sublibraries have to be resynthesized, and there are also 400-fold fewer compounds ($20^4 = 160,000$ rather than $20^6 = 64,000,000$) present in the pools that are submitted to the initial screening.

Erb et al. (5) have suggested that recursive deconvolution could be used to accelerate the iterative deconvolution process. Although this method was developed for use with solid-phase chemistry, it could be applied to solution-phase problems equally well. In this strategy, a portion of each resin sample is set aside after each coupling step and catalogued. The resins are then recombined before the next coupling. When an activity is detected in a pool of compounds, the partial libraries that have been set aside are used to speed up the deconvolution procedure.

Positional Scanning

This method was pioneered by Houghten with peptide libraries and has been the subject of several reviews (6, 7). It involves the initial synthesis of series of subli-

braries in which one position is defined with a single building block while the remaining positions are composed of mixtures of building blocks. Thus, the number of sublibraries synthesized for each position i is n_i , the number of building blocks used in that position. The determination of the biological activity of these sublibraries yields information about the most important building block for each position. The same number of sublibraries is produced as in iterative deconvolution, but no resynthesis of sublibraries is required and no increase in biological activity is expected during positional scanning. According to the rules proposed by Houghten, a hexapeptide library containing the 20 natural amino acids at each position will be formatted for positional scanning as 120 (6×20) sublibraries corresponding to:

OXXXXX
 XOXXXX
 XXOXXX
 XXXOXX
 XXXXOX
 XXXXXO

A simpler positional scan is shown as a matrix representation in Figure 3. In this case the library is constructed in two combinatorial steps ($N = 2$) from two sets of four building blocks (A at the first position and B at the second position; $n_A = n_B = 4$). The resulting library is represented as a two-dimensional (2D) matrix with 16 cells. Positional scanning of this library involves the division of the matrix into rows and columns. In each of the first four sublibraries, a single A building block is reacted with a mixture of the four B building blocks. In each of the second series of four sublibraries, a single B building block is reacted with a mixture of the four A building blocks. This method has been used by Smith et al. (8) for the synthesis and biological evaluation of amides and esters from a single coupling step. A similar approach, developed by Pirrung (1) to evaluate a library of 54 dimers (carbamates), is described in detail in the examples at the end of this chapter (example 3).

This representation can be extended to multidimensional matrices, as shown in Figure 4 for a three-position ($N = 3$) combinatorial library containing four different building blocks at each position ($n_A = n_B = n_C = 4$). The main advantage of positional scanning over iterative deconvolution is that the different positions of the library are evaluated simultaneously, and these results, found for each position, can be combined to directly identify active compounds. There is, however, no enrichment of the active species during the selection procedure with this method. This can be a problem if there are several active compounds in the library, as the screening does not provide any information on their connectivity. Thus, if several building blocks are found to be active at each position, the number of compounds that have to be individually prepared for biological testing corresponds to all possible combinations and so grows exponentially (e.g., three active amino acids at each position of a hexapeptide library would require the preparation of $3^6 = 729$ peptides).

Dooley and Houghten (7) proposed a number of alternative solutions to the synthesis of all possible peptides, including the synthesis of positional scanning

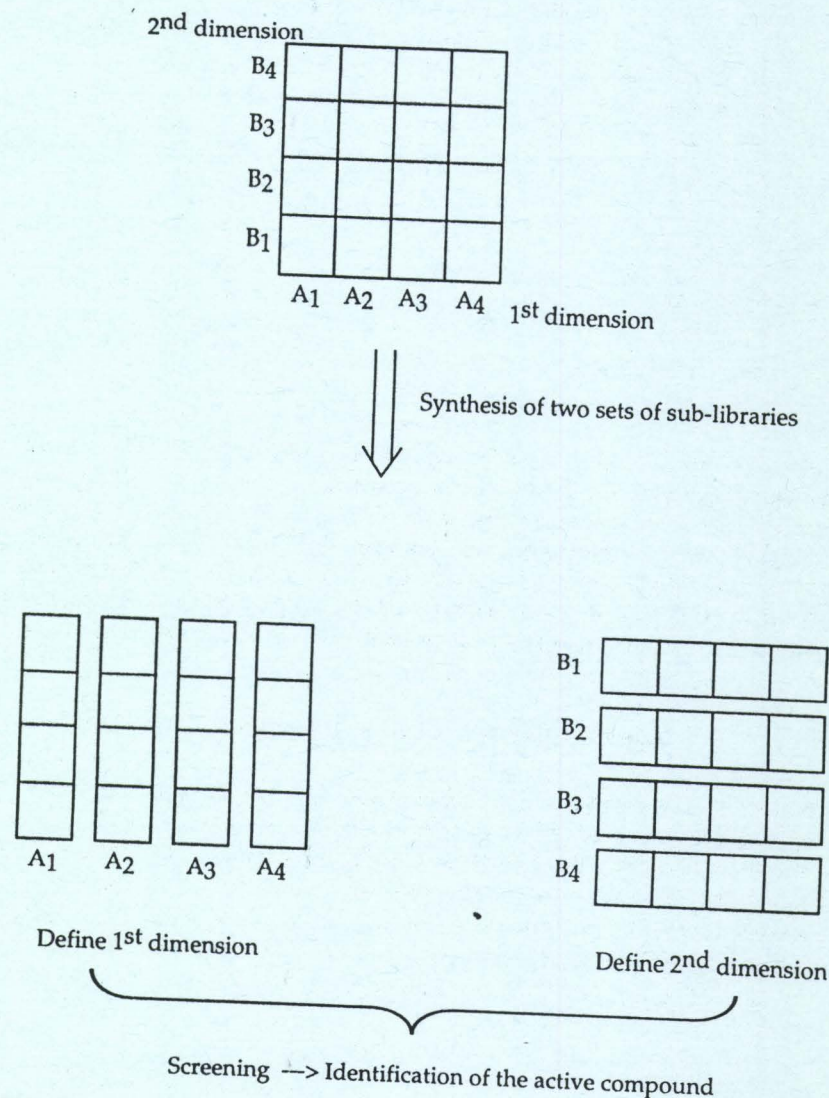


FIGURE 3. Matrix representation of the eight sublibraries that are required for positional scanning of a 4×4 library.

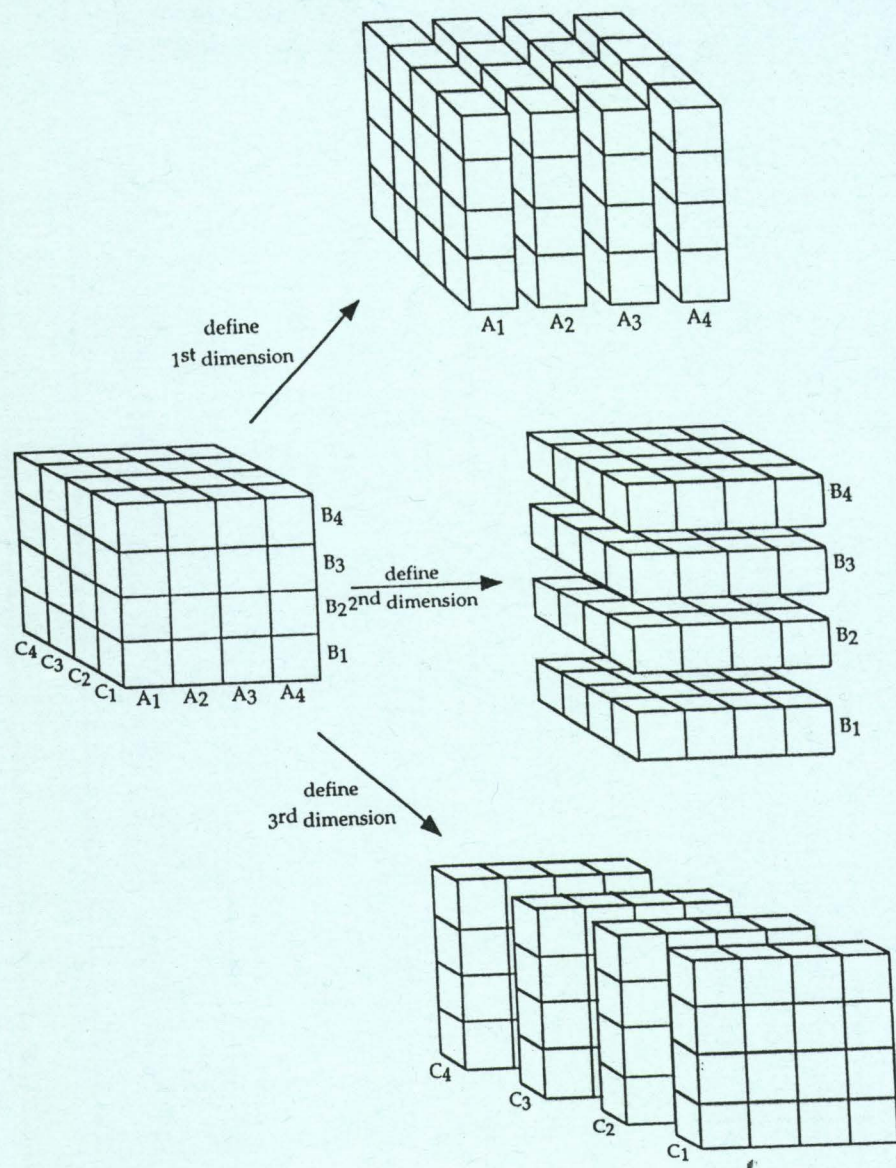


FIGURE 4. Matrix representation of the 12 sublibraries that are required for positional scanning of a $4 \times 4 \times 4$ library. The complete set of 12 sublibraries is submitted to evaluation.

sublibraries such as $X_1X_2X_3X_4O$, in which X_n is limited to the subset of amino acids which proved to be effective for position n in the initial positional scan. This will increase the concentration of active compounds but requires additional synthesis. Another possibility is to pursue a classical iterative deconvolution using only the most active amino acids for each position. In this case, the best choice will be to start the iterative deconvolution from the position that gave the most variation in the biological assay when it was occupied with the different building blocks.

Orthogonal Libraries

Unlike iterative deconvolution but as with positional scanning, orthogonal libraries are screened simultaneously and allow the direct identification of active compounds without iterative resynthesis and selection steps.

This strategy involves the synthesis of two libraries, I and II, which contain the same compounds. The libraries are divided into sublibraries in such a way that the number of compounds in each sublibrary is equal to (or less than) the number of sublibraries in each library. For example, in Figure 5 a library containing 16 compounds is divided into four sublibraries of four compounds each. The principle of orthogonal partition is that all compounds that are present in the same sublibrary in library I are partitioned into different sublibraries in library II. Thus in Figure 5, each of the four compounds that are pooled in a sublibrary of library I are found in the four different sublibraries of library II. Initially, the biological activity of the four sublibraries of library I are tested. If an activity is detected in one of these sublibraries, then the four sublibraries of library II are screened. The fact that each sublibrary of II contains only one compound from a given sublibrary of I allows a straightforward identification of the active compound. This approach can be extended to matrices with more than two dimensions [see example 4 below for a three-dimensional matrix of 15,625 compounds (9)] provided that the number of sublibraries in each of libraries I and II is at least equal to the number of compounds present in each sublibrary. (For example, libraries containing 10,000 compounds can be organized into 2×100 sublibraries containing 100 compounds each.)

Although orthogonal libraries can be prepared by mixing individual compounds, they can more easily be assembled using classical methods of combinatorial chemistry. To obtain an orthogonal partition of compounds within the sublibraries, the mixtures of building blocks are partitioned orthogonally, so that any group of building blocks used to prepare library I shares only one building block with any group of building blocks used to prepare library II. Thus in Figure 5, library I was synthesized by pooling building blocks $A_1 + A_2$, $A_3 + A_4$, $B_1 + B_2$, and $B_3 + B_4$ while library II was synthesized by pooling building blocks orthogonally ($A_1 + A_3$, $A_2 + A_4$, $B_1 + B_3$, and $B_2 + B_4$). The orthogonal partition of building blocks leads directly to an orthogonal partition of compounds within the libraries. For larger libraries there are generally many possible ways to orthogonally partition groups of building blocks. However, to maximize the diversity within each subli-

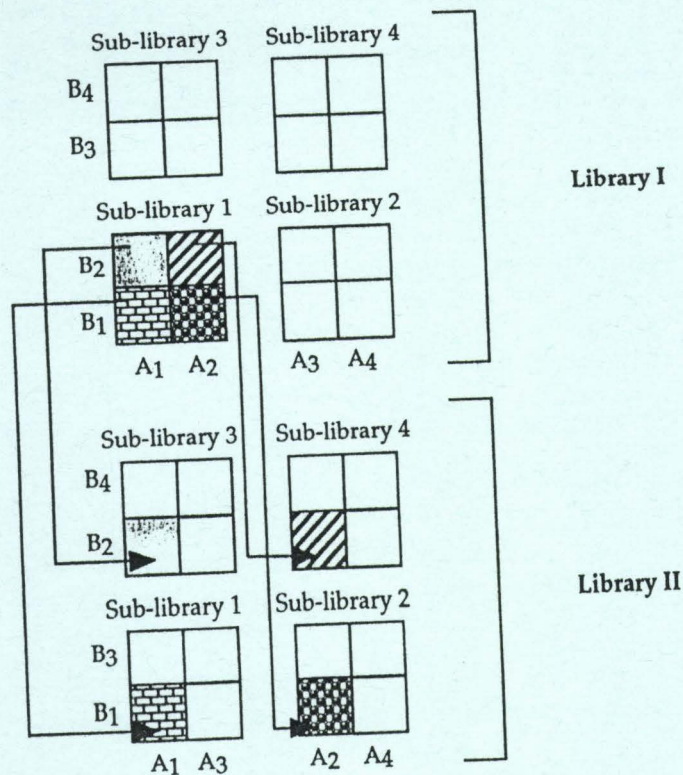


FIGURE 5. Composition of two orthogonal libraries corresponding to the general structure A_n-B_n with $n = 4$. Each of the four compounds pooled in library A1 ($A_1B_1, A_1B_2, A_2B_1, A_2B_2$) is partitioned into one of four different sublibraries of library B. As an example, detection of an activity in sublibrary 1 of library I and in sublibrary 3 of library II directly identifies the compound represented by the gray square as the active compound.

library, it is important to select those partitions that maximize structural diversity within each group of building blocks.

The use of orthogonal libraries is not limited to polypeptides or other polymers that can be sequenced, and these libraries do not restrict the chemical strategies used to generate diversity. This method can even be applied to the rapid screening of pools of existing compound libraries of diverse origins.

Moreover, orthogonal libraries offer an internal validation during the screening, as any positive result in one library must be confirmed by a positive result in the orthogonal library, thus providing an internal control for the detection of false positive results. Another advantage is that an active compound will be pooled with different compounds in library I and library II, thus allowing detection of possible "suboptimal binding" effects (4). Unlike the two previous strategies, which can be applied to large libraries of low diversity (such as peptide or oligonucleotide

libraries) in which the presence of several compounds of similar activity does not preclude the identification of an active species, orthogonal libraries are better suited to groups of compounds with high structural diversity. If several compounds with similar activities are present in the library, the number of ways to decode the readout will rapidly increase. For example, if two compounds with the same activity exist in the library, they will generally be responsible for two hits in both library I and library II. These can be decoded as any of four possible structures, of which only two will be active. The number of possibilities increases with the square of the number of equally active compounds in each library. In such cases, iterative deconvolution is required.

Synthetic Strategies

Although solution screening offers a clear advantage in terms of the broad variety of assays that can be used, this strategy requires that compounds be available in a soluble form. After polymer-supported synthesis, some compounds are tested while still linked to their support, but for solution screening, a suitable linker between the compounds and the polymer must be designed. This linker, which has to be stable enough to withstand the reaction conditions used during synthesis, must be cleaved to release the products in a soluble form. This may restrict the molecular diversity of the products, as only a limited number of functional groups may be compatible with this last chemical step. On the contrary, solution-phase combinatorial chemistry yields compounds that can be directly screened. Moreover, with solution chemistry there is no requirement that at least one of the building blocks carry a functional group that can be linked to the support. A typical example is for a library of carbamates (example 3) prepared in solution using monofunctional isocyanates and alcohols. This synthesis yields several relatively unstable compounds, which would probably not have resisted the chemical reagents generally used to cleave linkers.

Purification from reagents used for the synthesis (e.g., coupling agents) is a problem, however. Given the structural heterogeneity of the compounds present in the library, any purification runs the risk of losing at least some compounds. Moreover, after solution synthesis each combinatorial compound is more or less diluted (depending on the number of members in the pool), but the reagents do not undergo such a dilution. Thus, a library of 100 amides synthesized using a carbodiimide as coupling agent between a set of 10 amines and 10 acids will contain 0.01 equiv of each amide but 1 equiv of the urea generated by the carbodiimide (if the reaction has reached completion).

To minimize these problems, solution synthesis should favor combinatorial chemistries using reagents that can be easily eliminated. The classical strategies involve reagents that generate volatile products, or polymer-supported reagents. In all cases, the compatibility of the different byproducts with the biological assays should be tested in a blank assay as a basic precaution before building any new

different compounds in library I and library II, thus allowing detection of possible "suboptimal binding" effects (4). Unlike the two previous strategies, which can be applied to large libraries of low diversity (such as peptide or oligonucleotide

The deconvolution strategies described above rely on the determination of the biological activity (A) of mixtures of compounds. The result of such a determination depends not only on the biological activity of each individual compound but also on its molar representation in the mixture. As an example, a higher than expected concentration of a suboptimal binder can lead to an overestimation of the biological activity of the corresponding pool of compounds and introduce bias in the deconvolution. Two approaches to this problem are described below.

The Divide-Couple-Recombine (DCR) Method

In this method (10), polymer beads are divided equally into several batches. Each batch is then coupled with one of the building blocks before the different batches are pooled. If the division of the resin has been performed accurately, this method results in an equimolar distribution of the building blocks. This method will not work, however, unless the number of beads at the beginning of the synthesis is large enough. The magnitude of this problem increases dramatically with the size of the library.

Deviation from the ideal equimolar distribution of final products has been rigorously examined by Zhao et al. (11, 12). They analyzed the deviation from the ideal equimolar distribution of products in the final mixture of a DCR synthesis of a tripeptide library ($N = 3$), synthesized using 20 amino acids at each position ($n_A = n_B = n_C = 20$) and so containing $20 \times 20 \times 20 = 8000$ compounds. The synthesis used 90- μm beads, and the resin contained 2.86×10^6 beads g^{-1} . The deviation assessment takes into account the difference in abundance of every possible pair of compounds. The required number of beads is determined so that with a particular level $[(1 - \alpha) \times 100\%]$ of confidence, each individual relative error is controlled to be less than a predetermined tolerable limit, L_2 . The maximum ratio of the concentrations of any compound in the final mixture is less than $(1 + L_2)/(1 - L_2)$, where $0 < L_2 < 1$. Table II lists the required number of beads for different confidence levels and several different tolerable limits of $(1 - \alpha) \times 100\%$. For example, if we want the concentration ratio between the most and the least prevalent compounds to be < 1.5 , then $(1 + L_2)/(1 - L_2) = 1.5$ and $L_2 = 0.2$. Depending on the required confidence level, this would require 1.417 or 1.631 g of resin. If we want the ratio of all concentrations to be < 1.1 , this corresponds to an L_2 of 0.05, and the synthesis would require 22–26 g of resin. In contrast, if L_2 values are close to 1, only small quantities of resin are required in the synthesis, but this choice of L_2 will only ensure that all compounds are present (with the corresponding confidence level). Such ratios are acceptable when compounds are screened still linked to beads, as the requirement in this case is that each compound should at least be pres-

Chemistries using reagents that can be easily eliminated. The classical strategies involve reagents that generate volatile products, or polymer-supported reagents. In all cases, the compatibility of the different byproducts with the biological assays should be tested in a blank assay as a basic precaution before building any new

0.1	5.667	6.525
0.2	1.417	1.631
0.3	0.630	0.725
0.5	0.227	0.261
0.8	0.089	0.102
0.99	0.058	0.067

NOTE: The split-recombine synthesis involves $20 \times 20 \times 20$ reactions (i.e., three coupling steps generating a total of 8000 compounds).

*The beads have a diameter of 90 μm , and there are 2.86×10^6 beads in 1 g.

SOURCE: Reproduced with permission from reference 11.

ent on one bead, in order to be selected. For the solution phase, however, too high a ratio of the concentrations of different compounds (an L_2 value of 0.8 leads to a maximum ratio of 9 between the most and the least prevalent compound) precludes any accurate evaluation of the contribution of each individual compound to the biological activity of a pool, and so prevents the use of deconvolution strategies with such libraries.

Thus, considering the need to generate approximately equimolar mixtures and the amounts of resin that can easily be manipulated (10 to 100 g), the use of the DCR method in strategies involving deconvolutions will be limited to libraries containing between 10,000 and 100,000 compounds.

Coupling Mixtures of Building Blocks

The alternative to individual coupling reactions is to have mixtures of compounds in coupling reactions, yielding beads that carry mixtures of compounds. With this method the number of different compounds that can be prepared is limited only by the number of functional groups available on the beads, but, once again, achieving equimolar distribution of the compounds obtained after cleavage from the polymer is a problem. In this case, the factors controlling the abundance of different compounds are quite different. In the DCR method, excesses of reagents are used to force reactions to completion, but when coupling mixtures of building blocks the most reactive building blocks will be preferentially incorporated, leading to an unsatisfactory distribution of compounds in the final mixtures.

Solutions to this problem have been proposed for use with building blocks of similar reactivities, such as amino acids or nucleotides. One possibility is to perform a first coupling using only one equivalent of the mixture of building blocks per equivalent of reactive groups on the resin. Using a prolonged coupling time, even

TABLE III. Ratios of Amino Acids Necessary for Equimolar Coupling

Ala	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Tyr
3.58	3.70	3.84	2.66	3.04	3.76	18.3	6.56	5.23	2.42	5.64	4.56	5.61	6.87	2.93	5.04	11.9	4.36

NOTE: Each number represents the mole percentage of amino acid derivative necessary for equimolar coupling when using a 10-fold excess of mixtures of Boc-protected amino acids and 1-hydroxybenzotriazole ester derivatives in dimethylformamide (prepared in situ using diisopropylcarbodiimide).

SOURCE: Reproduced with permission from reference 13.

the less reactive building blocks will be able to react with reactive groups available on the resin. A second coupling step is then performed using an excess of the same building blocks (or a capping reagent) to force the reaction to completion. If a kinetically favored reaction occurs at this stage, it will only occur at the reactive groups that were left free after the first step. An improvement of this method, used when one of the building blocks displays too large a difference in reactivity, is to isolate part of the resin and to couple this building block separately from the mixture (partial DCR). This method has been used successfully with amino acid mixtures (amino acid analysis of the resulting peptide mixtures showed reasonable agreement with the expected values corresponding to an equimolecular representation of each amino acid).

A more accurate method for peptides proposed by Houghten (13) involves coupling mixtures of side-chain-protected Boc amino acids in a predetermined molecular ratio that compensates for the different coupling rates of the various amino acid derivatives. Table III gives an example of the molecular ratio for a mixture of 18 Boc-protected amino acids coupled in a 10-fold excess.

Although these methods offer the advantage over DCR of ease of synthesis and unlimited library size, they are less efficient as soon as building blocks of larger structural diversity are involved, as these building blocks generally display a wider range of reactivity than amino acids. Moreover, none of these methods can prevent the preferential "pairing" of two building blocks during couplings, corresponding to the preferential coupling of one of the building blocks of the mixture to one of the building blocks present on the polymer.

Selected Examples

To illustrate the different approaches outlined in this chapter, five examples have been selected from the literature.

Example 1

In this example (14), iterative deconvolution is used to select peptides selective for the μ opioid receptor from a hexapeptide library.

Composition of the Library

Six-residue peptides with free *N*-termini and amidated *C*-termini were prepared according to the dual positional format, so that the first two residues (at the *N*-termini) of each sublibrary were individually defined as one of the 20 natural amino acids and the last four positions consisted of equimolar mixtures of 19 of the 20 natural L-amino acids (with cysteine omitted). This generated 400 sublibraries, each represented by the general formula $O_1O_2XXXX-NH_2$, and each containing a total of 130,321 (19^4) different peptides. Thus the complete library contained $20 \times 20 \times 19^4 = 52,128,400$ hexapeptides.

Synthesis of the Library

The library was prepared by classical solid-phase peptide synthesis methods on a methylbenzhydrylamine polystyrene resin using *t*-Boc-protected amino acids. The XXXX-resin mixture was prepared as a unique batch using the DCR method, with each amino acid coupled individually.

The XXXX-resin was then divided into 400 equal portions, and two defined amino acids (O_1 and O_2) were coupled to each batch. Cleaved peptide mixtures were extracted with water, lyophilized, and reconstituted in water to yield a final concentration of 1–5 mg mL⁻¹, and used in standard in vitro receptor assays.

Biological Assay

Sublibraries were tested in a radioreceptor assay using an analog of methionine enkephalin ($[^3H]$ -[D-Ala², MePhe⁴, Gly-ol⁵]enkephalin; DAMGO), which binds specifically and with high affinity to μ binding sites from rat brain membranes. IC_{50} and ID_{50} (effective concentration of peptide mixture required to achieve 50% inhibition of binding of the radioligand) values were determined in each case using serial dilutions of peptide mixtures starting from an initial concentration of 0.08 mg mL⁻¹ of peptide mixture.

Design of the Sublibraries

The complete library was designed according to the dual positional format (O_1O_2XXXX). According to this strategy, the initial library consists of 400 sublibraries in which the first two amino acids are individually and specifically defined.

Identification of Active Compounds

The initial step of deconvolution involved the screening of the 400 O_1O_2XXXX sublibraries to determine their ability to inhibit the binding of 3-nM $[^3H]$ -DAMGO to rat brain membranes. IC_{50} values were then determined for each of the most effective peptide mixtures. The three most effective inhibiting peptide mixtures were YGXXXX-NH₂ ($IC_{50} = 3450$ nM), WWXXXX-NH₂ ($IC_{50} = 3500$ nM), and YPXXXX-NH₂ ($IC_{50} = 4050$ nM), which were separately deconvoluted. Figure 6 illustrates the different deconvolution steps in the case of the YPXXXX-NH₂ sublibrary in which the four X positions were sequentially defined. This clearly illustrates that the most active compound (YPFGFR-NH₂, $IC_{50} = 13$ nM) would not have

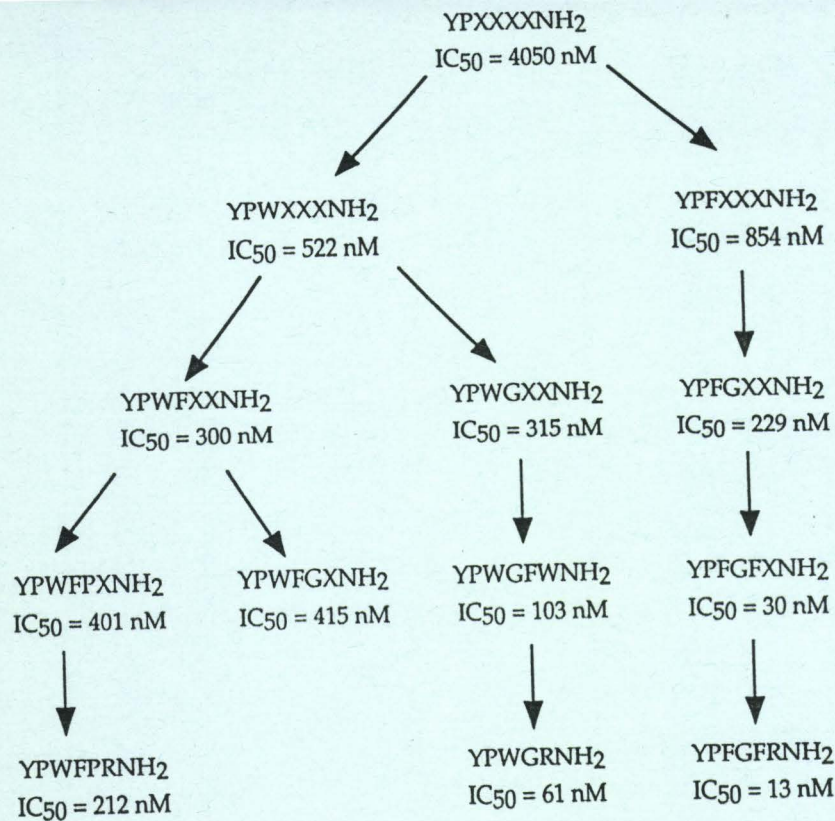


FIGURE 6. The iterative deconvolution of YPXXX-NH₂. IC₅₀ values (shown below each compound or compound mixture) were determined at each step, and the most potent mixtures were deconvoluted further. Note that the most potent of the final compounds comes from the sublibrary that was not the most potent after the first round of deconvolution.

been found if, at the first step of deconvolution, only the most active sublibrary (YPWXXX-NH₂) had been considered.

Example 2

This example (15) involves the discovery of nanomolar ligands for the α 1-adrenergic receptor from a ~5000-component *N*-substituted glycine peptoid library using iterative deconvolution.

Composition of the Library

The general structure of the library is shown in Figure 7. It consists of trimers assembled from three sets of *N*-substituted glycine monomers and capped with one of three different *N*-terminal caps. For the synthesis of the *N*-substituted glycines,

the total set of amines was partitioned into three groups that had either 3 hydroxylic building blocks, 4 aromatic building blocks, or 17 diverse building blocks. These sets of building blocks were designated O, A, and D, respectively. An additional source of diversity was to leave a free amine at the terminus or cap it as an acetamide or cyclohexylurea. Building blocks in the aromatic and hydroxylic sets were selected on the basis of their occurrence in known inhibitors of seven transmembrane receptors, and building blocks in the diverse set were chosen for their structural diversity. Mixtures generated from the 6 possible permutations of the 3 sets of building blocks further diversified by 3 different *N*-terminal cappings generated the 18 sublibraries shown in Table IV. Each combinatorial mixture included $30 \times 4A \times 17D = 204$ peptoid trimers and was capped in one of three possible ways. Each trimer in the library therefore contained at least one aromatic and one hydroxylic side chain (in addition, all possible dimers were included, adding 12 ($30 \times 4A$), 51 ($30 \times 17D$) or 68 ($4A \times 17D$) additional compounds to pools with D, A, and O groups respectively). Thus, each of the 18 sublibraries contained either 216, 255, or 272 compounds.

Synthesis of the Library

The peptoid backbone was assembled in two steps from two readily available reagents, as shown in Scheme 1. The first step consists of an acylation with bromoacetic acid followed by nucleophilic displacement of bromide by a wide variety of primary amines. The repetition of these two steps allows the extension of the peptoid backbone. Only aliphatic hydroxyl, amine, and acid side chains required protection. The synthesis was performed according to the DCR strategy on an automated synthesizer using the isopycnic slurry method. Dichloroethane:DMF at 3:2 was used to split the resin. Each building block was coupled individually, after which all portions were combined. The complete library was generated through six independent syntheses, corresponding to different permutations of O, A, and D sets. After completion of the trimeric peptoids, resins were split into three equal portions, which had one of three different *N*-terminal caps added. These portions were not recombined.

TABLE IV. Composition of the 18 Sublibraries Representing the Complete Peptoid Library

Identity of Cap		
H- = free amine	Ac- = acetamido	Chu- = cyclohexylureido
H-AOD-NH ₂	Ac-AOD-NH ₂	Chu-AOD-NH ₂
H-ADO-NH ₂	Ac-ADO-NH ₂	Chu-ADO-NH ₂
H-DAO-NH ₂	Ac-DAO-NH ₂	Chu-DAO-NH ₂
H-DOA-NH ₂	Ac-DOA-NH ₂	Chu-DOA-NH ₂
H-ODA-NH ₂	Ac-ODA-NH ₂	Chu-ODA-NH ₂
H-OAD-NH ₂	Ac-OAD-NH ₂	Chu-OAD-NH ₂

NOTE: Each sublibrary was made from all combinations drawn from the six permutations of three sets of building blocks (amines) and three *N*-terminal endings. Thus each compound contains at least one aromatic and one hydroxylic group.

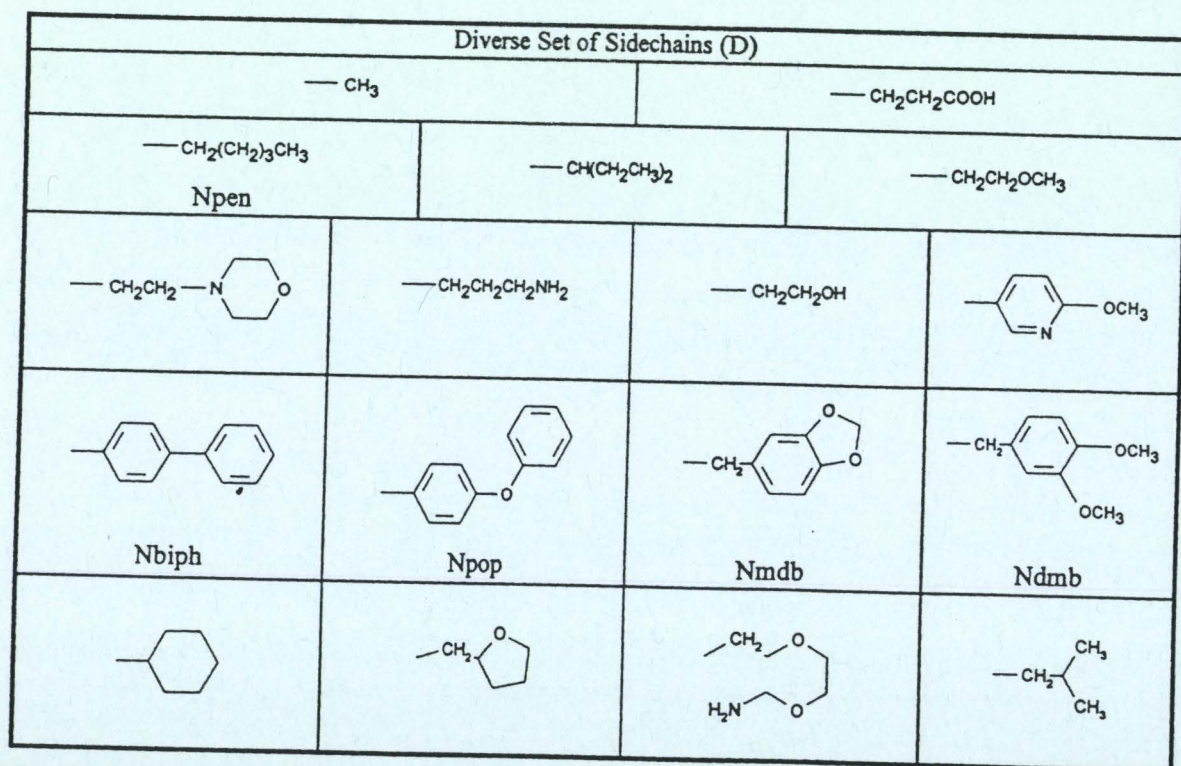
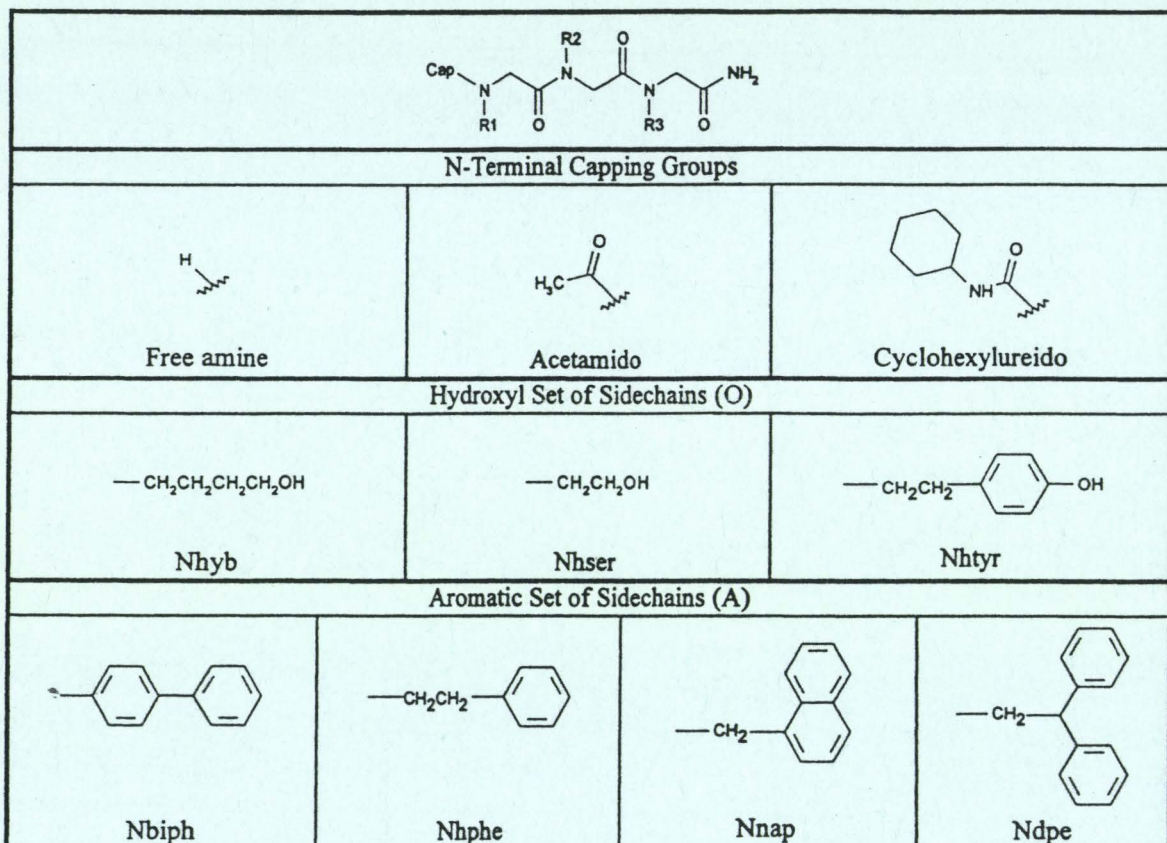
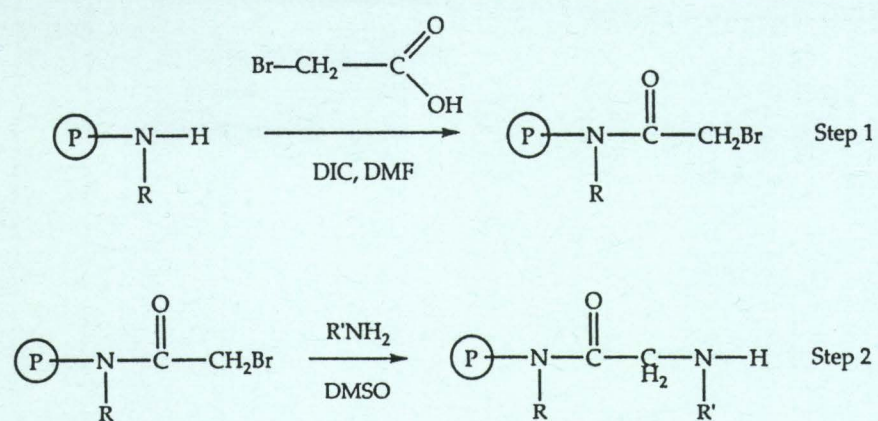


FIGURE 7. Building blocks for the construction of a peptid library. The overall structure of the library is shown in the top panel, and the three caps are shown in the next set of panels. Below this are the side chains of the amino building blocks used, which were divided into three sets. O is a set of 3 hydroxylic building blocks, A is a set of 4 aromatic building blocks, and D is a set of 17 diverse monomers.



SCHEME I. Solid-phase synthesis of *N*-substituted glycine peptoids using a two-step addition cycle for each building block.

Biological Assay

Binding to the $\alpha 1$ -adrenergic receptor was measured by competitive displacement of [^3H]-prazosin from a rat brain membrane preparation. The original eighteen sublibraries were tested at 100 nM per peptoid (21.6–27.2 μM total peptoids), and the smaller sublibraries were tested at 1 μM per peptoid.

Results

Individual active compounds were identified by iterative deconvolution (Figure 8). The organization of the library allowed the direct identification of the cap and the order in which the different types of side chains appear (although not the exact identity of any one side chain) in the most active peptoid by screening the 18 initial sublibraries. The H-ODA-NH₂ sublibrary was identified as the most inhibitory, indicating that the preferred peptoid had H as an N-Cap and a member of the hydroxyl set in the first peptoid position. This sublibrary was then resynthesized as four smaller sublibraries, each containing 68 compounds of the form H-XDA-NH₂, where X represents one of the three hydroxylic side chains of the O set or a solvent blank. A single sublibrary with a homotyrosine (NhTyr) side chain was responsible for most of the activity of the initial sublibrary (Figure 8). The next deconvolution step involved the preparation of 17 sublibraries of 4 compounds each. The compounds were of the format (H-NhTyr)-X-A-NH₂, with X now representing each of the 17 side chains from the diverse set. Two mixtures showed significant activities, and their side chains were fixed for the last round, which involved the synthesis of 8 individual peptoids each with a different side chain from the aromatic set. The most active compound was CHIR 2279, with the inhibition constant (K_i) of 5 ± 3 nM (Structure I).

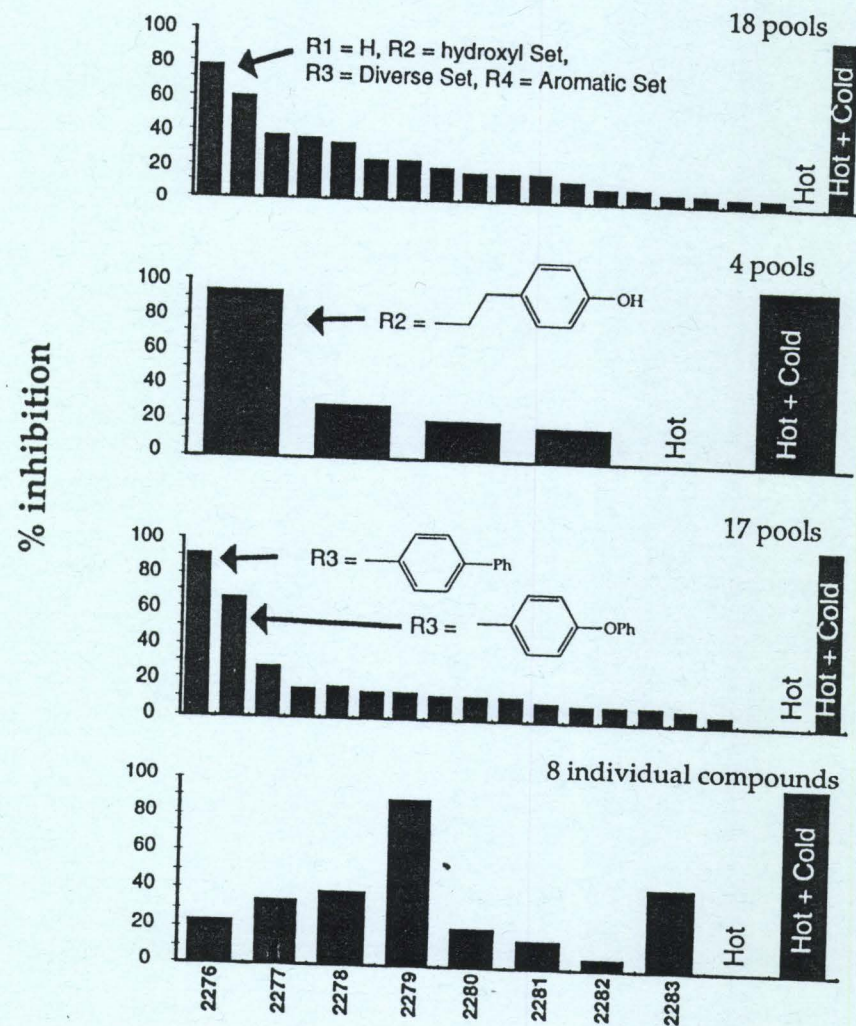
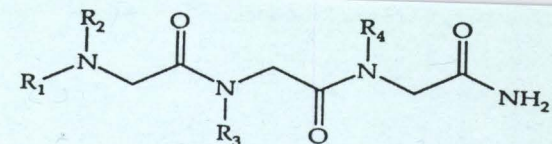
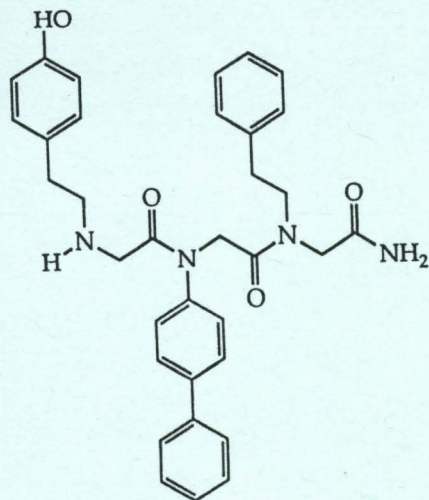


FIGURE 8. Iterative steps of deconvolution for the identification of high-affinity ligands for the $\alpha 1$ -adrenergic receptor. At the top is the generic structure of the library members. In the first step, the identity of the cap and the preferred order of the types of side chains were identified. The subsequent steps deciphered the identity of the side chains one by one.



STRUCTURE I. Structure of CHIR 2279, the most potent inhibitor of the α 1-adrenergic receptor, discovered from the library described in Figure 7. The compound has a K_i of 5 ± 3 nM.

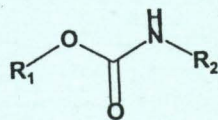
A similar deconvolution performed with this library using an opiate receptor assay resulted in the identification of a potent μ -specific opiate receptor inhibitor ($K_i = 6$ nM).

Example 3

This example (1) covers the solution synthesis and screening of a nonpeptide "indexed" combinatorial library.

Composition of the Library

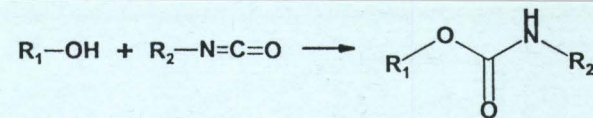
The complete library consisted of 54 carbamates corresponding to the general structure:



and synthesized using nine alcohols (R_1OH) and six isocyanates ($R_2-N=C=O$).

Synthesis of the Library

Compounds were prepared by solution synthesis by directly combining isocyanates and alcohols in the following reaction.



To eliminate kinetic effects, reactions were conducted with stoichiometric quantities of reagents or mixtures in tetrahydrofuran (THF) solution at 60–80 °C for 4 h in sealed pressure tubes. The solvent was then removed by a rotary evaporator.

Biological Assay

Mixtures were tested for their ability to inhibit acetylcholinesterase activity. An IC_{50} was determined for each mixture.

Design of the Sublibraries

Sublibraries were constructed according to the concept of an indexed library, which is similar to a two-dimensional positional scanning library (Figure 2). Two sets of sublibraries were assembled. In the first dimension, six sublibraries were prepared by reacting each individual isocyanate with the mixture of nine alcohols, while in the second dimension, nine sublibraries were prepared by reacting each individual alcohol with the mixture of six isocyanates.

Identification of Active Compounds

The 15 sublibraries described in the preceding paragraph were tested and their IC_{50} values determined. In the isocyanate dimension, potency decreases in the order $Me > iPr > Et > tBu$, while in the alcohol dimension, potency decreases in the order succinimide > benzotriazole > 4-hydroxybenzaldehyde. The most active compound was identified as *O*-succinimidyl *N*-methylcarbamate with an individual IC_{50} of 0.7 mM.

Example 4

This example (9) describes the discovery of a potent vasopressin V2 receptor antagonist from two orthogonal libraries of 15,625 trimers.

Composition of the Library

According to the general principle of orthogonal libraries described above, two libraries of the same soluble 15,625 trimers ($N = 3$) resulting from the combination of 25 building blocks at each position ($n_A = n_B = n_C = 25$) were independently prepared.

The building blocks were 23 D-amino acids and two nonchiral amino acids, glycine and isonipecotic acid. For the first library (library I), the 25 building blocks were partitioned into five sets, I_1 to I_5 (Figure 9). Three combinatorial reactions of these five sets yields $5^3 = 125$ sublibraries, each resulting from the incorporation of 1 set of 5 building blocks at each of the 3 variable positions and so containing $5^3 =$

	I ₁	I ₂	I ₃	I ₄	I ₅
II ₁	D-Leucine	D-Proline	D-Serine	(4-nitro)-D-phenylalanine	D-Isoglutamic acid
II ₂	D-Arginine	D-Isoleucine	Glycine	D-Threonine	D-Isoglutamine
II ₃	D-Glutamine	D-Tyrosine	D-Valine	D-e-Nicotinoyl-lysine	D-Histidine
II ₄	D-Tryptophane	D-Asparagine	D-Glutamic acid	D-Methionine sulfoxide	D-Alanine
II ₅	Isonipecotic acid	D-Lysine	D-Tetrahydroisoinoleic acid	D-Aspartic acid	D-Phenylalanine

FIGURE 9. Orthogonal partition of the 25 building blocks used to construct a tripeptide library into five sets of five compounds (I₁ to I₅ for library I and II₁ to II₅ for library II).

125 different compounds. For the second library (library II), the same 25 building blocks were used, also yielding 125 sublibraries and the same 15,625 trimers, but the building blocks were partitioned orthogonally into 5 different groups, II₁ to II₅, as shown in Figure 9. As any set of 5 building blocks used for library I has only 1 building block in common with the sets of building blocks used for library II, any sublibrary from library I and any sublibrary from library II share only one compound.

Synthesis of the Library

The two libraries were synthesized by solid-phase synthesis according to the Boc/benzyl strategy. At each step, each of the resulting peptidyl-modified resins was split into five portions, each of which was coupled to a different set of five building blocks. To avoid preferential incorporation of amino acids having the most favorable coupling rates, a first coupling was performed with only 1.1 equiv of acylating species for 1 equiv of amino groups available on the resin. A second coupling step was then performed to force the reaction to completion. Amino acid analysis on 10 sublibraries showed that the levels of most of the amino acids were within 10% of the expected value, the most divergent being within 25% of this value. After completion of the synthesis, peptides were cleaved from the resin by treatment with HF and lyophilized.

Biological Assay

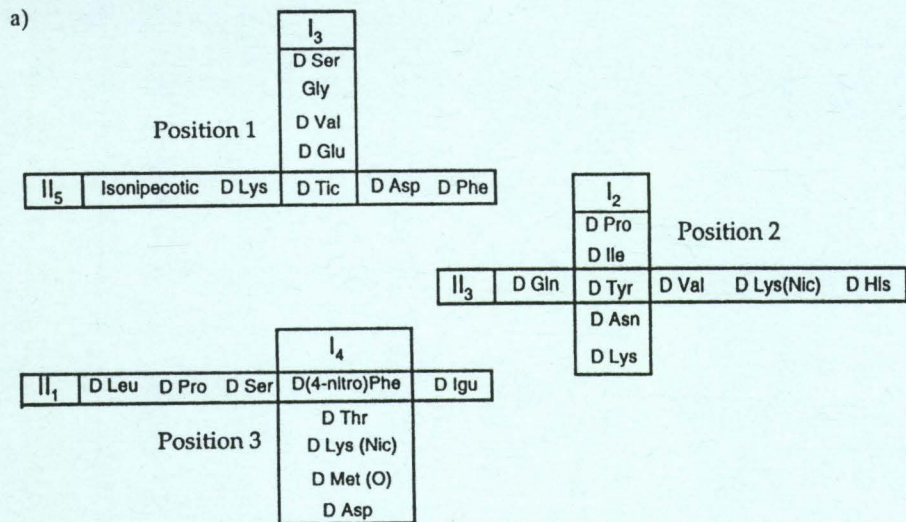
Sublibraries were tested at a concentration of 25 $\mu\text{g ml}^{-1}$, which corresponds to $\sim 0.2 \mu\text{g ml}^{-1}$ for each individual compound. The assay measured inhibition of the binding of 3 nM ³H-AVP (tritiated arginine vasopressin) to porcine kidney epithelial cells.

Identification of Active Compounds

In the initial screening, three sublibraries with high inhibiting potency were identified in each library. These sublibraries were further tested at a concentration of 2.5 $\mu\text{g ml}^{-1}$ and, on the basis of the potency of inhibition, the selection was narrowed to one sublibrary in I (I₃-I₂-I₄) and two sublibraries in II (II₅-II₃-II₁ and II₄-II₃-II₁). The trimer common to I₃-I₂-I₄ and II₅-II₃-II₁ was easily identified, as shown in Figure 10, as H-D-Tic-D-Tyr-(4-nitro)-D-Phe-OH, and turned out to be a potent inhibitor (IC₅₀ = 63 nM). No inhibition activity could be detected with the candidate shared by I₃-I₂-I₄ and II₄-II₃-II₁, however. It was thus suspected that this activity arose as a by-product only when the reactions were performed in pools.

References

- Pirrung, M. C.; Chen, J. *J. Am. Chem. Soc.* **1995**, *117*, 1240-1245.
- Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science (Washington, D.C.)* **1991**, *251*, 767-773.



b)

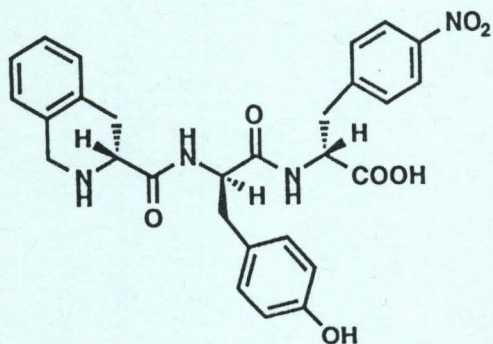


FIGURE 10. Screening of orthogonal libraries. (a) Identification of the trimer shared by I₃-I₂-I₄ and II₅-II₃-II₁. (b) The structure of this compound, H-D-Tic-D-Tyr-(4-nitro)-D-Phe-OH.

9. Déprez, B.; Williard, X.; Bourel, L.; Coste, H.; Hyafil, F.; Tartar, A. *J. Am. Chem. Soc.* **1995**, *117*, 5405-5406.
10. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature (London)* **1991**, *354*, 82-84.
11. Zhao, P. L.; Zambias, R.; Bolognese, J. A.; Boulton, D.; Chapman, K. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10212-10216.
12. Zhao, P. L.; Nachbar, R. B.; Bolognese, J. A.; Chapman, K. *J. Med. Chem.* **1996**, *39*, 350-352.
13. Ostresh, J. M.; Winkle, J. H.; Hamashin, V. T.; Houghten, R. A. *Biopolymers* **1994**, *34*, 1681-1689.
14. Dooley, C. T.; Kaplan, R. A.; Chung, N. N.; Schiller, P. W.; Bidlack, J. M.; Houghten, R. A. *Peptide Res.* **1995**, *8*, 124-137.
15. Zuckermann, R. N.; Martin, E. J.; Spellmeyer, D. C.; Stauber, G. B.; Shoemaker, K. R.; Kerr, J. M.; Figliozzi, G. M.; Goff, D. A.; Siani, M. A.; Simon, R. J.; Banville, S. C.; Brown, E. G.; Wang, L.; Richter, L. S.; Moos, W. H. *J. Med. Chem.* **1994**, *37*, 2678-2685.

Reprinted from A Practical Guide to Combinatorial Chemistry
 Anthony W. Czarnik and Sheila H. DeWitt, Editors
 Published 1997 by the American Chemical Society

3. Houghten, R. A.; Pinilla, C.; Blondelle, S.; Appel, R. J.; Dooley, C. T.; Cuervo, J. H. *Nature (London)* **1991**, *354*, 84-86.
4. Freier, S. M.; Konings, D. A. M.; Wyatt, J. R.; Ecker, D. J. *J. Med. Chem.* **1995**, *38*, 344-352.
5. Erb, E.; Janda, K. D.; Brenner, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11422-11426.
6. Pinilla, C.; Appel, R. J.; Blondelle, S.; Dooley, C. T.; Eichler, J.; Ostresh, J. M.; Houghten, R. A. *Drug Dev. Res.* **1994**, *33*, 133-145.
7. Dooley, C. T.; Houghten, R. A. *Life Sci.* **1993**, *52*, 1509-1517.
8. Smith, P. W.; Lai, J. Y. Q.; Whittington, A. R.; Cox, B.; Houston, J. G.; Stylli, C. H.; Banks, M. N.; Tiller, P. R. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2821-2824.

V Découverte d'un ligand du récepteur V₂ à la vasopressine, au moyen de chimiothèques orthogonales

« Orthogonal combinatorial libraries » : B. Déprez, X. Williard, L. Bourel, H. Coste, F. Hyafil and A. Tartar, Journal of the American Chemical Society, 1995, 117, 5405-5406

L'ensemble du travail suivant a été effectué avec la participation et sous la direction de Benoît Déprez.

V-A Conception de la chimiothèque généraliste

V-A-1 L'objectif

L'objectif principal de ce travail était de montrer, en utilisant des techniques de chimie combinatoire, qu'il était possible de trouver, **rapidement et à moindre coût**, un ou plusieurs composés doués d'une **activité biologique inconnue** (obtention de têtes de séries ou hits).

V-A-2 Les moyens

V-A-2-1 Choix des molécules à synthétiser

Le choix des molécules à synthétiser constitue le facteur principal du succès futur d'une chimiothèque. Les structures désirées devront posséder des caractéristiques physicochimiques trouvées dans les médicaments actuellement sur le marché (être « drug-like ») mais devront aussi être accessibles par des voies de synthèses permises par la chimie combinatoire. Deux type de contraintes peuvent donc être distinguées :

- les contraintes pharmacologiques

Celles-ci doivent non seulement être prise en compte pour optimiser les chances de trouver une nouvelle tête de série, mais aussi, au-delà des essais in vitro, pour anticiper des problèmes de biodisponibilité liés à des caractéristiques structurales connues. W. Curatolo (Curatolo 1998) a d'ailleurs récemment récapitulé les caractéristiques physico-chimiques les plus rencontrées chez les médicaments administrables par voie orale. Il mentionne notamment la

difficulté d'améliorer les caractéristiques pharmacocinétiques d'un composé dont l'affinité pour sa cible a déjà été optimisée (risque de perte d'affinité).

L'absorption intestinale passive des molécules est fonction : de la solubilité de ces molécules dans le milieu gastrointestinal ainsi que de la perméabilité intestinale. Ces deux facteurs dépendent de la lipophilie et de la masse moléculaire des molécules, du nombre de groupements donneur ou accepteur de liaisons hydrogènes ainsi que du nombre de charges présentes sur ces molécules.

Remarque : En général, l'objectif en thérapeutique est d'administrer un composé à une dose de 1 mg / kg avec une solubilité $\geq 0,05$ mg / mL. Le point de départ d'un programme de découverte de médicament sera généralement un composé possédant une solubilité de l'ordre de 1 mg / mL qui devrait par la suite assurer une absorption suffisante, si son activité et sa constante de vitesse d'absorption au travers de la paroi intestinale se situent dans une gamme acceptable.

Sur la base de 2245 composés, étant pour la plupart entrés en phase clinique II et biodisponibles par voie orale mais n'appartenant pas aux classes suivantes : polymères, peptides, ammoniums quaternaires et phosphates, Lipinski et al (Lipinski 1997) ont tenté d'établir une listes de caractéristiques physico-chimiques communes à la majorité de ces composés :

la masse moléculaire : seulement 11 % des composés possèdent une masse moléculaire supérieure à 500 Da. Les hautes masses moléculaires sont souvent associées à de faibles perméabilités.

le ClogP : seulement 10 % des composés ont un ClogP supérieur à 5. Les ClogP importants sont souvent associés à de faibles solubilités en milieu aqueux.

la présence de groupes donneurs de liaisons hydrogène : seulement 8 % des composés ont plus de 5 groupes OH ou NH confondus. L'augmentation du nombre de groupes donneurs de liaisons hydrogène est souvent associée à une diminution de perméabilité.

la présence de groupes accepteurs de liaisons hydrogène : seulement 12 % des composés ont plus de 10 atomes d'oxygène ou d'azote confondus. L'augmentation du nombre de groupes accepteurs de liaisons hydrogène est souvent associée à une diminution de perméabilité.

Cette analyse se résume en une règle mnémotechnique appelée « règle des cinq » car les valeurs limites obtenues sont des multiples de cinq : « le risque d'obtenir une mauvaise biodisponibilité pour un composé administré par voie orale devient important s'il possède :

- une masse moléculaire supérieure à 500 Da
- un ClogP supérieur à 5
- plus de 5 groupes donneurs de liaisons hydrogène
- plus de 10 groupes accepteurs de liaisons hydrogène ».

Cette règle est actuellement en place au sein de la société Pfizer, un message informatique alertant les chimistes médicaux lorsqu'ils enregistrent des produits dont les caractéristiques physico-chimiques dépassent ces limites.

La *stabilité des composés* dans le milieu gastrique, mais aussi la demi-vie plasmatique de ces mêmes composés doivent être prises en compte afin d'éviter que ceux-ci ne se dégradent trop rapidement dans l'organisme.

- les contraintes chimiques

Outre la nécessité de disposer d'une chimie compatible avec la synthèse de composés en mélange, le problème majeur rencontré dans toute synthèse de chimiothèque combinatoire reste la capacité analytique du laboratoire. Le nombre de composés synthétisés étant considérable, le goulot d'étranglement sera la capacité à analyser les produits de synthèse puis à interpréter les résultats obtenus. Il est pour cette raison préférable d'acquérir une **connaissance approfondie** de la chimie utilisée afin d'**anticiper** le plus possible tout événement pouvant nuire à la qualité de la chimiothèque. Cette connaissance nous permettra de plus, d'assurer une **reproductibilité** de toute synthèse le cas échéant.

- Bilan

La synthèse peptidique en phase solide nous a semblé adaptée à notre projet.

Celle-ci est :

→ particulièrement bien **documentée** et **utilisée au laboratoire**. La synthèse de peptides sur supports solides est décrite depuis 1963 (Merrifield 1963) et a été largement utilisée par la suite. De nombreux travaux concernant, les réactions secondaires rencontrées, l'optimisation des réactifs de couplage, des groupements protecteurs ainsi que des réactifs de

clivage ont été publiés et ont permis de rendre cette méthode particulièrement efficace. Une assurance qualité garantie tout au long de la production devait nous permettre de compenser un contrôle qualité difficile à mettre en œuvre.

→ compatible avec la synthèse d'un **grand nombre de produits**, de nombreux acides aminés étant commercialement disponibles (environ 800 N-tBoc-acides aminés figurent dans ACD).

→ réalisée **sans** exiger de **conditions opératoires particulières** (conditions ambiantes)

→ pourvue des divers **avantages** qu'offre la **synthèse sur support solide** : absence de phases d'évaporations de solvants et de purifications d'intermédiaires réactionnels, facilité de manipulation des composés liés à un support solide (tous les composés se trouvent dans un même état physique au niveau macroscopique)

→ compatible avec un **contrôle** systématique de chaque étape de couplage, indiquant l'état d'avancement des réactions (ex. test de Kaiser), mais compatible aussi avec une analyse d'acides aminés après hydrolyse acide totale permettant d'en déduire la composition du mélange.

→ compatible avec la synthèse de produits en **mélange**, cette caractéristique ayant été décrite par Geysen en 1986.

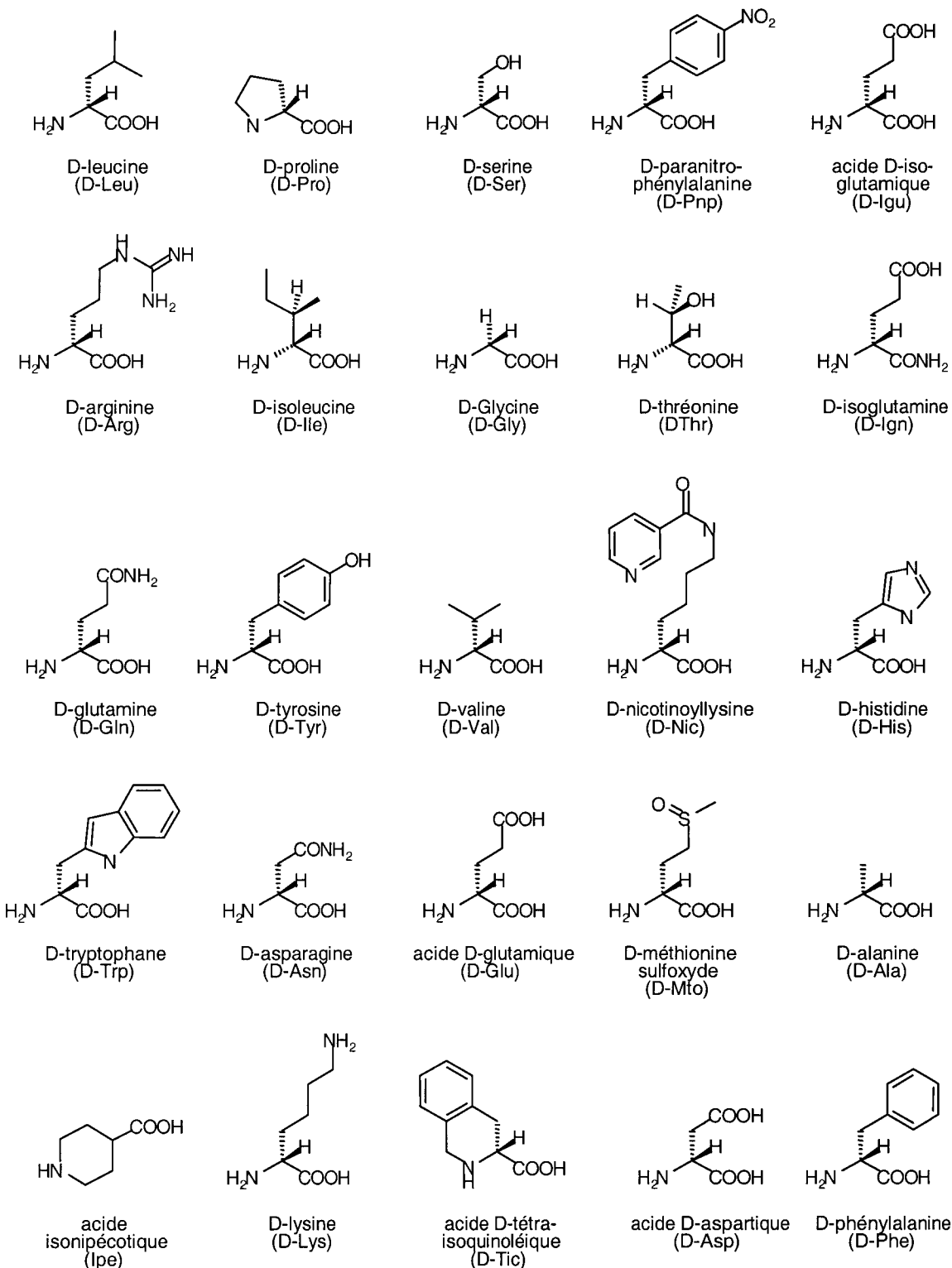
Nous avons donc décidé de synthétiser grâce à la synthèse peptidique en phase solide, une chimiothèque :

1) de **tripeptides** dont la masse moléculaire moyenne des composés serait, en nous basant sur les 20 acides aminés naturels, de **375 g / mole** (les deux extrêmes étant le tripeptide de la glycine : 189 g / mole et le tripeptide du tryptophane : 577 g / mole).

2) en utilisant comme monomères des acides aminés de **série D ou achiraux** pour garantir une stabilité vis à vis des protéases de l'organisme.

3) à partir des monomères pourvus de **diverses fonctions** (cf. figure 16) susceptibles d'interagir avec différentes cibles biologiques (par liaisons hydrogène, de Van de Waals, ioniques ou hydrophobes). Parmi les 25 acides aminés choisis, 19 portent des chaînes latérales trouvées dans les aminés naturels et 6 portent des chaînes latérales « exotiques ».

4) dont tous les composés possèdent une fonction amine protonable (amine terminale), afin que ceux-ci aient une **solubilité supérieure ou égale à 0,1 mg / mL** (soit 0,25 mM).



Rq : masse moléculaire moyenne des tripeptides : 407 g / mole (les deux extrêmes étant le tripeptide de la glycine : 189 g / mole et le tripeptide de la D-nicotinoyllysine : 718 g / mole).

Figure 16 : monomères utilisés pour la synthèse de la chimiothèque de 15625 tripeptides

V-A-2-2 Utilisation d'une nouvelle méthode de déconvolution développée au laboratoire

Comme nous l'avons décrit au paragraphe III B, les techniques de déconvolution alors utilisées souffraient de divers inconvénients. Schématiquement, la déconvolution itérative nécessite des « aller-retours » entre les équipes de chimie et de biologie, ce qui ralentit considérablement le processus, d'autre part, la déconvolution par balayage positionnel demande un effort de synthèse souvent important et peut mener à l'identification de ligands suboptimaux. Nous avons donc apporté une nouvelle réponse à la question : comment identifier le plus vite possible un composé actif se trouvant dans un mélange ? (cf. Première partie, chapitre III-B-3).

V-A-2-3 Définition de l'architecture de la chimiothèque

→ la taille des mélanges testés

La chimiothèque initiale sera partitionnée en plusieurs sous-chimiothèques qui seront ensuite testées sur diverses cibles biologiques. Afin de détecter facilement le produit responsable d'une éventuelle activité biologique au sein d'une sous-chimiothèque, nous avons limité la taille des mélanges testés à une centaine de composés (faible bruit de fond et risque limité d'activités par effet cumulatif). Le couplage de 5 acides aminés à chaque étape combinatoire permet d'aboutir à des sous-chimiothèques contenant chacune **125 tripeptides**.

Des groupes de 5 acides aminés ont donc été constitués.

→ nombre de monomères à utiliser

Afin de pouvoir réaliser une partition orthogonale des groupes de 5 monomères, il fallait pouvoir répartir les monomères constituant les groupes de 5 acides aminés, formés pour la partition A, dans des groupes différents pour la partition orthogonale B. Le choix d'une organisation « carrée », où 5 groupes de 5 monomères sont constitués (soit $5 \times 5 = 25$ monomères différents), permet une réorganisation aisée de tous ces groupes de façon orthogonale en schématisant par un carré de 5×5 , les **25 monomères** choisis puis en regroupant dans une première partition, les monomères par lignes et dans une seconde

partition, par colonnes (cf. figure 17). Ainsi lors de la synthèse de la chimiothèque selon la partition A, les acides aminés seront toujours couplés par 5 en respectant les groupes A1 à A5. Pour la synthèse selon la partition B, les groupes utilisés seront B1 à B5.

→ nombre total de composés synthétisés

$$25 \times 25 \times 25 = \mathbf{15625 \text{ tripeptides}}$$

→ constitution des groupes de 5 monomères

La composition de ces groupes influe de façon importante sur la fiabilité de la déconvolution. En effet, l'exigence d'une forte diversité structurale des monomères constituant chacun des groupes va permettre d'engendrer des tripeptides eux-mêmes très divers au sein de chaque sous-chimiothèque. Ainsi, l'activité observée au sein d'une sous-chimiothèque sera probablement portée par un seul composé, les autres constituants du mélange étant structurellement éloignés. La détection d'activité par un effet cumulatif est donc limitée. Le groupe A1 par exemple, comporte un acide aminé hydrophobe (la leucine), un acide aminé chargé positivement (l'arginine), un acide aminé polaire non chargé (la glutamine), un acide aminé aromatique (le tryptophane) et un acide aminé non naturel (l'acide isonipécotique).

A1	A2	A3	A4	A5
D-Leu	D-Pro	D-Ser	D-Pnp	D-Igu
D-Arg	D-Ile	Gly	D-Thr	D-Ign
D-Gln	D-Tyr	D-Val	D-Nic	D-His
D-Trp	D-Asn	D-Glu	D-Mto	D-Ala
Ipe	D-Lys	D-Tic	D-Asp	D-Phe

Partition A

B1	D-Leu	D-Pro	D-Ser	D-Pnp	D-Igu
B2	D-Arg	D-Ile	Gly	D-Thr	D-Ign
B3	D-Gln	D-Tyr	D-Val	D-Nic	D-His
B4	D-Trp	D-Asn	D-Glu	D-Mto	D-Ala
B5	Ipe	D-Lys	D-Tic	D-Asp	D-Phe

Partition B

Figure 17 : constitution des groupes de monomères pour la réalisation de 2 partitions orthogonales A et B d'une même chimiothèque.

V-B Préparation des chimiothèques orthogonales

La synthèse en phase en solide a été utilisée pour les différentes raisons évoquées au paragraphe V-A-2-1. La stratégie la mieux connue à l'époque était la stratégie tBoc-benzyle après ancrage d'un premier acide aminé sur une résine chlorométhylée (dite de Merrifield). Actuellement il existe une variété importante de supports solides dont l'utilisation permettrait la simplification de l'étape de clivage des molécules finales, de leur support.

V-B-1 Protections choisies pour les différents acides aminés

La protection temporaire de l'amine des aminoacides utilisés étant le tBoc, les groupes protecteurs des chaînes latérales devaient donc lui être orthogonaux (c'est à dire résistants au TFA, réactif de déprotection du tBoc).

→ Les monomères commerciaux

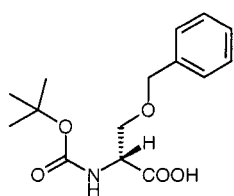
Les alcools et les acides carboxyliques:

Protégés respectivement sous forme d'éther ou d'ester benzyles.

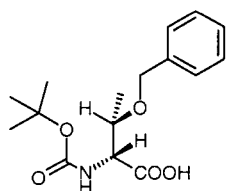
Ces protections sont : clivables par l'acide fluorhydrique anhydre

résistantes au TFA à 50 % dans le DCM

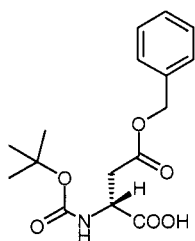
résistantes à la pipéridine à 10 % dans la NMP



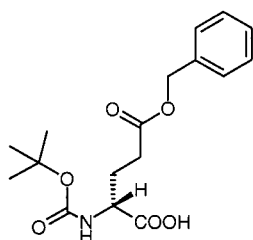
tBoc-D-Ser(Bzl)-OH



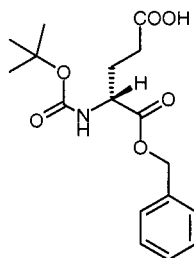
tBoc-D-Thr(Bzl)-OH



tBoc-D-Asp(OBzl)-OH



tBoc-D-Glu(OBzl)-OH

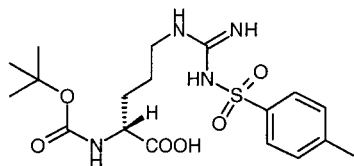


tBoc-D-Glu-OBzl
(D-Igu)

L'arginine:

Protégée par un groupement tosyle.

Cette protection est : clivable par l'acide fluorhydrique anhydre
résistante au TFA à 50 % dans le DCM
résistante à la pipéridine à 10 % dans la NMP

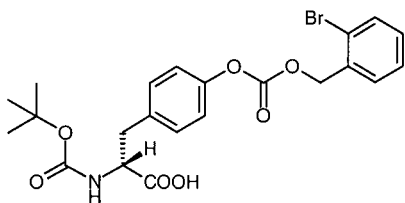


tBoc-D-Arg(Tos)-OH

La tyrosine :

Protégée par un groupement 2-Bromo-benzyloxycarbonyl.

Cette protection est : clivable par l'acide fluorhydrique anhydre
résistante au TFA à 50 % dans le DCM
résistante à la pipéridine à 10 % dans la NMP



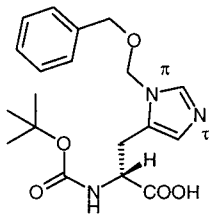
tBoc-D-Tyr(2-Br-Z)-OH

La protection de la chaîne latérale permet d'éviter l'alkylation du noyau phénolique par une transposition de Fries.

L'histidine :

Protégée sur l'atome d'azote π du noyau imidazole, par un groupement benzyloxyméthyl.

Cette protection est : clivable par l'acide fluorhydrique anhydre
résistante au TFA à 50 % dans le DCM
résistante à la pipéridine à 10 % dans la NMP



tBoc-D-His(π -Bom)-OH

La protection de l'atome d'azote π , permet de bloquer tout risque de racémisation de l'histidine, lié à un réarrangement en milieu basique au cours du couplage de l'histidine.

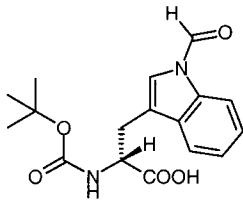
Le tryptophane :

Protégé sur l'atome d'azote du noyau indole, par un groupement formyl.

Cette protection est : clivable par la pipéridine à 10 % dans la NMP

résistante au TFA à 50 % dans le DCM

résistante l'acide fluorhydrique anhydre



tBoc-D-Trp(For)-OH

La protection du noyau indole, permet à la fois d'éviter son oxydation mais aussi son alkylation par des carbocations lors des étapes de clivage.

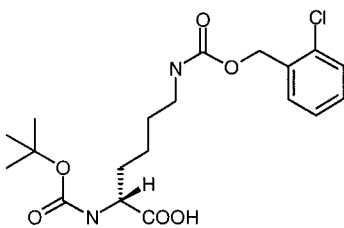
La lysine :

Protégée par un groupement 2-Chloro-benzyloxycarbonyl.

Cette protection est : clivable par l'acide fluorhydrique anhydre

résistante au TFA à 50 % dans le DCM

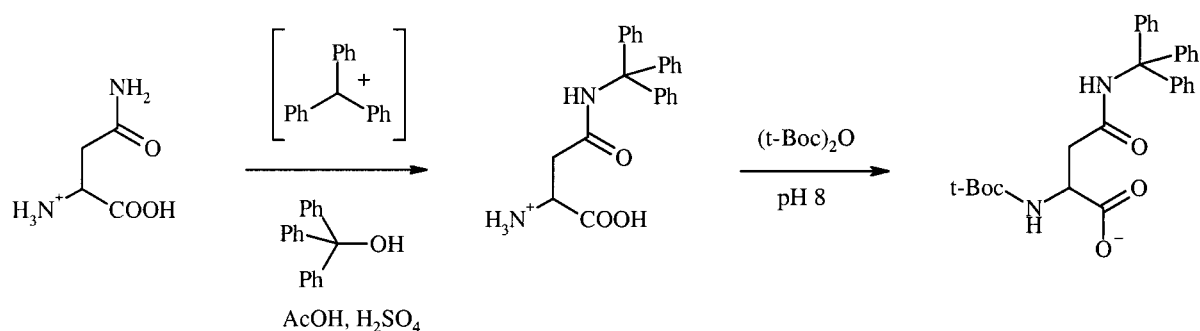
résistante à la pipéridine à 10 % dans la NMP



tBoc-D-Lys(2-Cl-Z)-OH

→ Les monomères préparés au laboratoire.

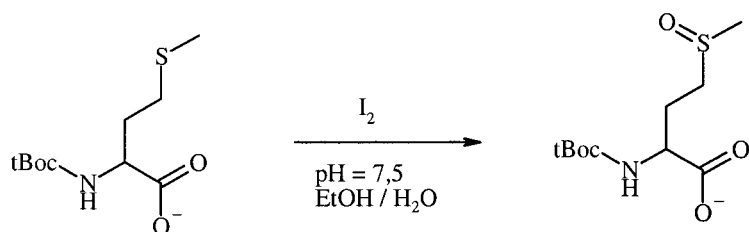
La tBoc- D-Asparagine protégée par un groupement trityle sur sa chaîne latérale : tBoc-D-Asn(trityl)-OH



La faible solubilité de la tBoc-D-Asn-OH sous forme de sel de césium (forme impliquée lors du couplage du premier acide aminé sur une résine chlorométhylée, cf. § VI-B-2) nécessite de réaliser la substitution d'un atome d'hydrogène du carboxamide, par un groupement trityle pour augmenter la solubilité de ce composé. Ce groupe protecteur permet aussi d'éviter la réaction de déshydratation de la fonction carboxamide conduisant à la fonction nitrile observée lorsque des agents de couplage tels que les carbodiimides, le BOP, le PyBop et l'HBTU sont utilisés (Gausepohl 1989). Cette réaction secondaire est minimisée en présence d'HOBT.

Remarque : le groupement trityle est un groupement clivable en milieu acide (TFA).

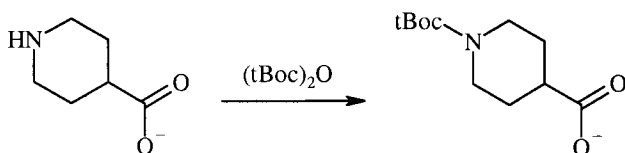
La tBoc D-Méthionine S-sulfoxyde : tBoc-D-Mto-OH



La principale réaction secondaire concernant la méthionine reste l'oxydation conduisant à la méthionine sulfoxyde, pouvant se produire lors des étapes d'élongation de chaîne, de clivage en milieu acide fort ou lors d'un stockage prolongé. Nous avons donc décidé d'oxyder

préalablement cet acide aminé, afin d'en maîtriser la stabilité. Cette protection permet aussi d'éviter toute alkylation du soufre lors des étapes de clivage.

L'acide N-tBoc Isonipécotique : tBoc-Ipe-OH



Ce composé fait partie des 2 aminoacides (avec l'acide isoglutamique) non α -aminé, présents dans nos chimiothèques orthogonales.

V-B-2 Ancrage des premiers acides aminés au support solide : accès aux aminoacyl-résines protégées

Cette étape consiste en la substitution nucléophile de l'atome de chlore d'une résine chlorométhylée, par un anion carboxylate. Afin de renforcer la nucléophilie de l'anion carboxylate, celui-ci sera introduit sous la forme de sel de césium, car dans ce cas les espèces chargées sont particulièrement dissociées (méthode dite « aux ions nus »). L'ancrage obtenu, de type ester benzylique, ne pourra être clivé qu'en présence d'un acide fort (acide fluorhydrique).

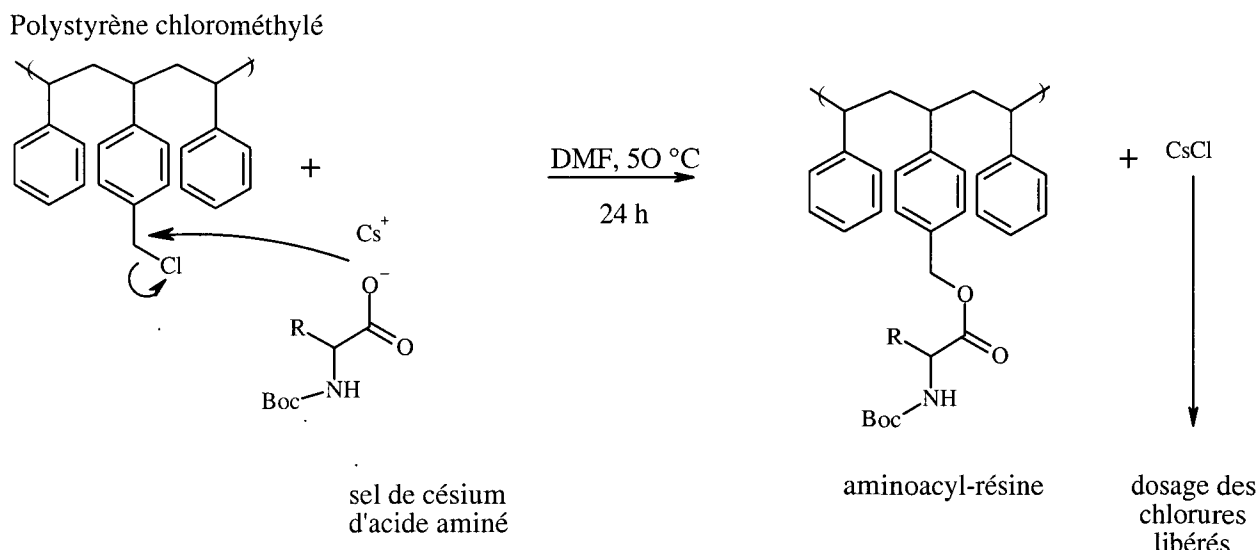


Figure 18 : ancrage d'un acide aminé sur une résine chlorométhylée par la méthode aux ions nus

Les 25 aminoacyl-résines nécessaires à la synthèse des 2 chimiothèques orthogonales ont été préparées individuellement (à l'échelle de 4 mmoles par aminoacyl-résine) et les rendements de couplage calculés par dosage des ions chlorures libérés (par la méthode de Charpentier-Volhard, cf. partie expérimentale). Les rendements obtenus s'échelonnent de 66 % pour la tBoc-D-Arg (Tos)-OH à 99 % pour la tBoc-D-Mto-OH. Dans 20 cas sur 25 le rendement obtenu était supérieur à 90 %.

Chaque aminoacyl-résine obtenue a été divisée en deux parties. Les premières moitiés ont été réunies en 5 groupes de 5 résines selon la partition A et les secondes moitiés, en 5 groupes de 5 résines selon la partition B (cf. figure 17).

Les lots d'aminoacyl-résines obtenues étaient donc :

A₁-P, A₂-P, A₃-P, A₄-P, A₅-P

B₁-P, B₂-P, B₃-P, B₄-P, B₅-P

(**P** représente le polymère polystyrène)

Les quantités alors manipulées étaient de l'ordre de 11 grammes par mélange de 5 aminoacyl-résines (5 x 2,1 g en moyenne).

Afin de limiter toute réaction secondaire pouvant faire intervenir l'amine présente sur une peptidyl-résine, celle-ci ne sera déprotégée (clivage du t-Boc) que juste avant l'étape de couplage de l'acide aminé suivant. La réaction de cyclisation des glutamines N-terminales en résidus pyrroglutamiques (cf. figure 19) sera notamment évitée. En effet, les conditions de

clivage du groupement t-Boc, accélèrent cette réaction favorisée en milieu acide. La formation du résidu pyrroglutamique a alors lieu pendant tout stockage prolongé de la peptidyl-résine déprotégée.

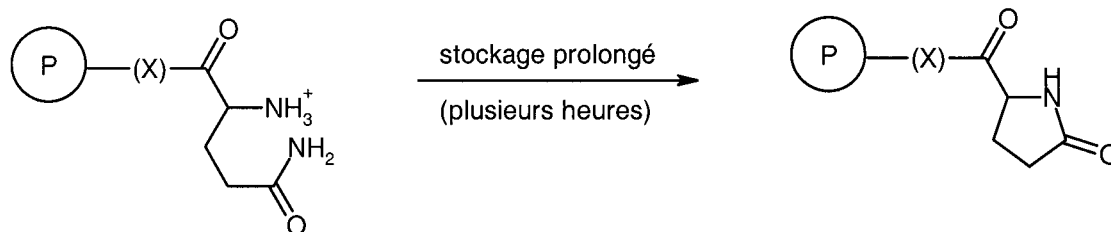


Figure 19 : formation de résidus pyrroglutamiques à partir de résidus glutamine N-terminaux, déprotégés.

Remarque : dans la suite du paragraphe, nous ne détaillerons que la synthèse d'une des 2 chimiothèques orthogonales, la seconde étant préparée selon la même procédure.

V-B-3 Couplage des acides aminés de la position 2 : accès aux dipeptidyl-résines

Les tBoc-aminoacyl-résines devant être déprotégées juste avant la réaction de couplage (cf. paragraphe précédent), les aminoacyl-résines (A_1-P à A_5-P) ont été successivement divisées en 5 parties égales (division pondérale des résines essorées du DMF) puis traitées par du TFA à 50 % dans le DCM pour cliver le groupement protecteur t-Boc et libérer la fonction amine. Une fois les aminoacyl-résines lavées et neutralisées, les acides aminés (5 x 0.22 Eq) à coupler ont été activés sous forme d'esters d'HOBT, grâce à l'HBTU (1,1 Eq) dans le DMF, en présence d'HOBT (1,1 Eq) et de Diisopropyléthylamine (4,4 Eq). Les équivalents sont calculés par rapport à la quantité théorique d'amine réactive présente sur la résine (1 Eq).

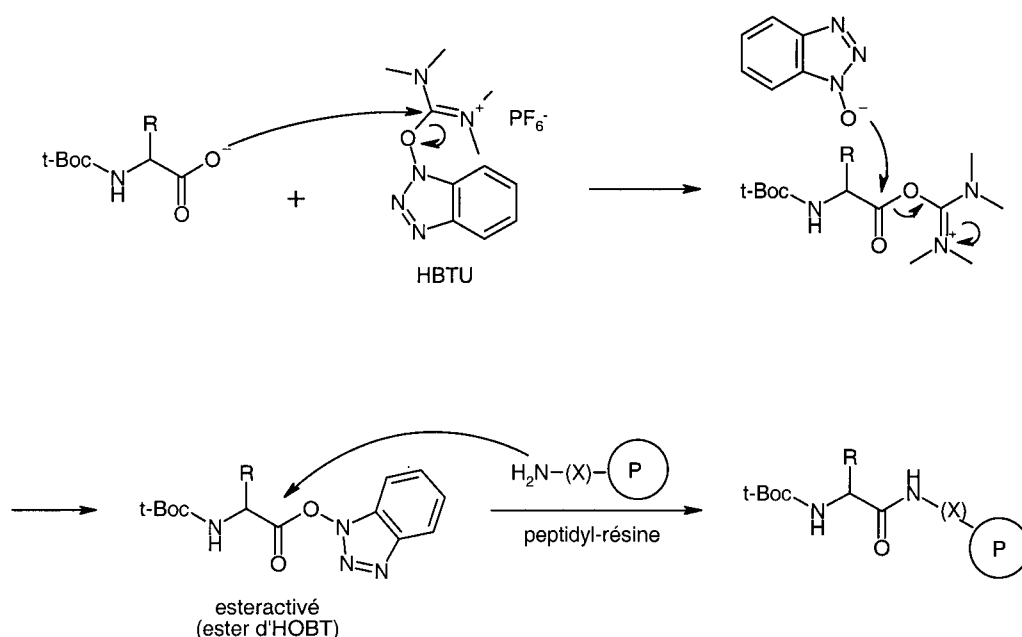


Figure 20 : formation d'une liaison amide grâce à un acide activé par l'HBTU

- Le groupe de 5 acides aminés A₁ a été couplé aux premières fractions de A₁-P à A₅-P.
- Le groupe de 5 acides aminés A₂ a été couplé aux deuxièmes fractions de A₁-P à A₅-P.
- Le groupe de 5 acides aminés A₃ a été couplé aux troisièmes fractions de A₁-P à A₅-P.
- Le groupe de 5 acides aminés A₄ a été couplé aux quatrièmes fractions de A₁-P à A₅-P.
- Le groupe de 5 acides aminés A₅ a été couplé aux cinquièmes fractions de A₁-P à A₅-P.

Les acides aminés ont été couplés, selon la méthode du couplage de monomères en mélange (cf. première partie, chapitre III-B). Un excès de 10 % en acides activés par rapport aux fonctions amines à acyler sur la résine a été utilisé. Des tests de Kaiser quantitatif (mise en évidence d'amines primaires) ou de Christensen (mise en évidence d'amines secondaires) ont été effectués après chaque couplage pour vérifier l'absence de fonctions amines non acylées. Dans 95 % des cas, aucune fonction amine libre n'a pu être détectée. Cependant, un second couplage, selon les mêmes conditions a systématiquement été effectué.

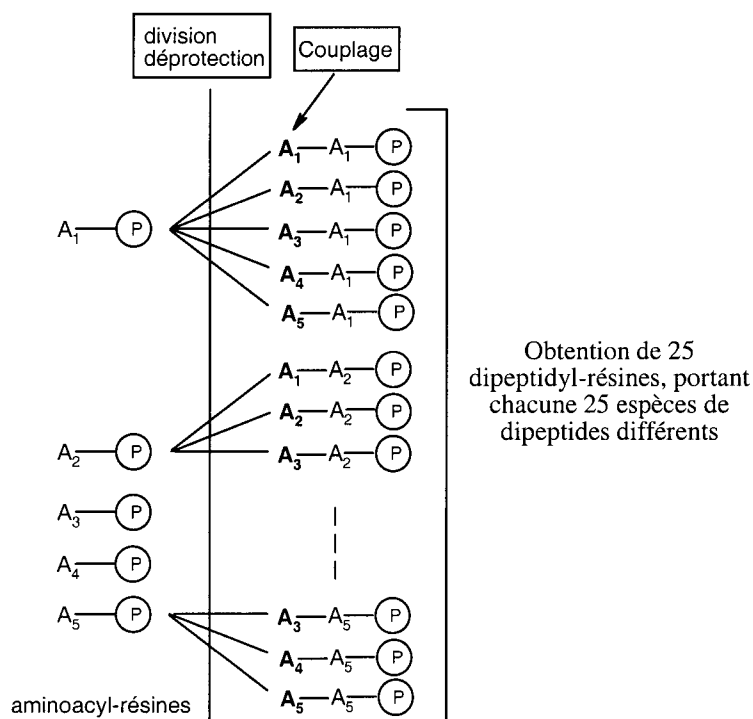


Figure 21 : composés obtenus après le couplage des acides aminés en position 2

V-B-4 Couplage des acides aminés de la position 3 : accès aux tripeptidyl-résines

La technique est identique à celle décrite au paragraphe précédent. Les quantités de résine manipulées sont ici de l'ordre de 400 mg.

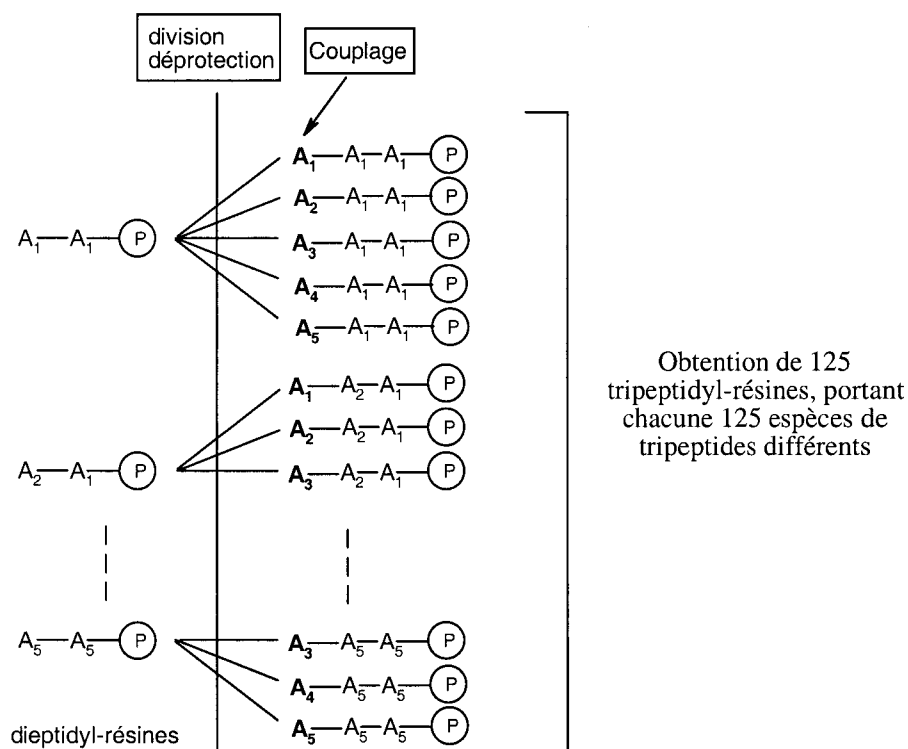


Figure 22 : composés obtenus après le couplage des acides aminés en position 3

V-B-5 Déprotection et clivage des tripeptides du support solide

Après avoir terminé la synthèse sur support solide, 100 à 200 mg de toutes les tripeptidyl-résines ont été traitées au TFA à 50 % dans le DCM, afin d'éliminer la fonction t-Boc de l'amine terminale puis à la pipéridine à 10 % dans la NMP afin d'éliminer le groupement formyl protégeant la chaîne latérale du tryptophane. Ce traitement a aussi été appliqué aux tripeptidyl-résines ne contenant pas de tryptophane afin que le traitement soit homogène pour tous les composés, quel que soit la composition du mélange dans lequel ils se trouvent. Ainsi une réaction secondaire liée à ce traitement, apparaissant pour un composé présent dans une sous-chimiothèque contenant le résidu tryptophane, sera reproduite pour ce même composé dans une sous-chimiothèque orthogonale ne contenant pas de tryptophane. Les tripeptidyl-résines ont ensuite été traitées par l'acide fluorhydrique anhydre (5 mL) en présence de pièges à carbocations (paracrésol -7,5 %- et parathiocrésol -2,5%-) 1 heure à 0°C. Les peptides sont alors clivés de leur support et les protections de chaînes latérales des acides aminés éliminées par un mécanisme de type SN1 (les carbocations libérés étant piégés par le paracrésol et le parathiocrésol). Les peptides ont ensuite été précipités dans un mélange éther/pentane 50/50 v/v, centrifugés, redissous dans l'acide acétique puis filtrés sur filtre 5 µm afin d'éliminer les résines. Les solutions obtenues ont ensuite été congelées puis lyophilisées.

Nous avons ensuite contrôlé la composition de 7 sous-chimiothèques prises au hasard, afin de vérifier la représentation de acides aminés. Celles-ci ont été soumises à une hydrolyse acide totale (HCl 6N, 10 µL de phénol aqueux à 10 %, 24 h à 110 °, sous vide) puis ont été analysées par HPLC pour déterminer de façon quantitative les proportions en chaque acide aminé. Les résultats ont montré que la plupart des acides aminés étaient présents dans une fourchette de 10 % autour de la valeur attendue et que les plus divergents étaient à moins de 30 % de la valeur attendue (cf. tableau 9). Ces résultats nous ont permis à la fois de valider le protocole de couplage des acides aminés en mélange, mais aussi la technique de purification des peptides par précipitation dans l'éther/pentane.

<i>Lots analysés</i>													
		<i>Leu</i>	<i>Arg</i>	<i>Glu</i>	<i>Trp</i>	<i>Lys</i>	<i>Asp</i>	<i>His</i>	<i>Ala</i>	<i>Phe</i>			
A₁A₄A₅	Théo.	1	1	3	1	1	1	1	1	1			
	Obs.	1,3	0,9	2,9	1,3	1,1	1,3	0,7	1,2	0,9			
		<i>Leu</i>	<i>Arg</i>	<i>Glu</i>	<i>His</i>	<i>Ala</i>	<i>Phe</i>						
A₁A₅A₁	Théo.	2	2	4	1	1	1						
	Obs.	2,5	1,9	3,6	1,1	1	1						
		<i>Pro</i>	<i>Ile</i>	<i>Tyr</i>	<i>Asp</i>	<i>Lys</i>	<i>Thr</i>	<i>Ser</i>	<i>Gly</i>	<i>Val</i>	<i>Glu</i>		
A₂A₄A₃	Théo.	1	1	1	2	2	1	1	1	1	1		
	Obs.	1	1	0,7	2	2,3	0,8	0,9	1	0,9	1		
		<i>Pro</i>	<i>Ile</i>	<i>Tyr</i>	<i>Asp</i>	<i>Lys</i>	<i>Thr</i>						
A₂A₄A₄	Théo.	1	1	1	3	3	2						
	Obs.	0,9	1	0,6	3	3,5	1,9						
		<i>Pro</i>	<i>Ile</i>	<i>Tyr</i>	<i>Asp</i>	<i>Lys</i>	<i>Glu</i>	<i>His</i>	<i>Ala</i>	<i>Phe</i>	<i>Leu</i>	<i>Arg</i>	
A₂A₅A₁	Théo.	1	1	1	1	1	3	1	1	1	1	1	
	Obs.	1	0,9	0,9	0,9	1,2	3,3	1	1	0,9	1,1	0,8	
		<i>Ser</i>	<i>Gly</i>	<i>Val</i>	<i>Glu</i>	<i>Thr</i>	<i>Lys</i>	<i>Asp</i>					
A₃A₄A₄	Théo.	1	1	1	1	2	2	2					
	Obs.	1	1	1,2	1,1	1,5	2,1	1,8					
		<i>Thr</i>	<i>Lys</i>	<i>Asp</i>	<i>Glu</i>	<i>His</i>	<i>Ala</i>	<i>Phe</i>	<i>Leu</i>	<i>Arg</i>			
A₄A₅A₁	Théo.	1	1	1	3	1	1	1	1	1			
	Obs.	0,9	1,1	0,8	3,3	1	1,1	0,9	1,2	0,8			

Tableau 9 : évaluation par HPLC de la composition en acides aminés (par rapport à la composition théorique) de 7 sous-chimiothèques ayant subi une hydrolyse acide totale

V-C Résultats biologiques

Les cibles biologiques sur lesquelles ont été testés les composés des chimiothèques orthogonales ont été choisies par la société Glaxo. Celles-ci ont concerné par exemple, des cibles virales, des canaux ioniques et des récepteurs 7-TM tels que CCKB ou V2. Les tests ont été réalisés dans les laboratoires Glaxo ou bien sous-traités à des sociétés comme CEREP, spécialisée notamment dans le criblage biologique.

V-C-1 Conditions générales des tests biologiques

→ Mise en solution des sous-chimiothèques

Les sous-chimiothèques ont d'abord été solubilisées dans le DMSO puis diluées dans le tampon d'expérimentation pour atteindre une concentration de 100 µg / mL soit $2,5 \cdot 10^{-4}$ M (MM moyenne des peptides = 407 g/mole). Cette concentration rapportée à chacun des 125 composés revient à une concentration de $2 \cdot 10^{-6}$ M pour chaque composé du mélange.

→ Détermination de la concentration à laquelle seront testées les sous-chimiothèques

Celle-ci sera déterminée afin d'obtenir un ratio signal sur bruit compatible avec le test intéressé. Cette mise au point sera systématique pour chaque nouveau test. Pour cela, 10 sous-chimiothèques prises au hasard seront testées à 3 concentrations différentes : 6,25 ; 25 et 100 mg / mL afin de définir la concentration du criblage de l'ensemble de la chimiothèque.

→ Test des sous-chimiothèques orthogonales

Une des deux chimiothèques orthogonales est d'abord testée. La seconde ne le sera, que lorsqu'une activité intéressante a pu être observée dans la première.

V-C-2 Résultats obtenus sur le récepteur à la vasopressine V₂

V-C-2-1 Intérêt thérapeutique de la cible biologique

La vasopressine ou ADH (antidiuretic hormone) est un neuropeptide cyclique de 9 acides aminés. Sa formule est apparentée à celle de l'ocytocine car ces deux nonapeptides ont le même précurseur et sont activés de la même façon. Il existe plusieurs sous-types de récepteurs à la vasopressine. Notre criblage a concerné le type V₂, présent au niveau du tube collecteur du rein. La fixation de la vasopressine sur ce récepteur entraîne une réabsorption d'eau par le rein. Des antagonistes du ligand naturel, pourraient être utilisés pour la lutte contre l'hypertension artérielle (par effet diurétique) ou contre le syndrome de sécrétion inappropriée d'ADH.

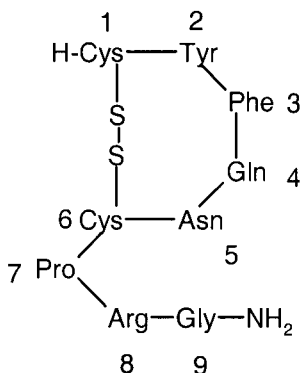


Figure 23 : structure de la vasopressine

V-C-2-2 Identification d'un produit actif au cours des tests de déplacement de la vasopressine tritiée du récepteur V₂ porcin

→ Détermination de la concentration à laquelle seront testées les sous-chimiothèques

Le tableau ci-dessous donne les résultats obtenus après criblage de 10 sous-chimiothèques aux trois concentrations : 6,25 ; 25 et 100 µg / mL. Les résultats sont exprimés pourcentage d'inhibition de liaison spécifique du ligand de référence : Arginine-Vasopressine (IC₅₀ = 4,7-5,2 nM).

<u>Sous-chimiothèques</u>	<u>6,25 µg/mL</u>	<u>25 µg/mL</u>	<u>100 µg/mL</u>
A ₄ A ₁ A ₁	11	14	49
A ₅ A ₁ A ₁	-	-	21
A ₄ A ₁ A ₂	-	-	13
A ₁ A ₂ A ₂	-	17	33
A ₅ A ₁ A ₃	-	-	-
A ₁ A ₂ A ₃	22	-	14
A ₃ A ₂ A ₃	-	19	-
A ₃ A ₁ A ₄	22	73	91
A ₃ A ₂ A ₄	74	90	82
A ₃ A ₁ A ₅	15	23	58

Tableau 10 : résultats du test de criblage de 10 sous-chimiothèques sur le récepteur V₂ (le signe "-" indique une inhibition inférieure à 10%)

La concentration de 100 µg / mL était trop peu discriminante. Nous avons donc choisi la concentration de 25 µg / mL, celle-ci permettant de conserver un bruit de fond relativement faible tout en permettant la mise en évidence de composés actifs possédant des IC₅₀ de l'ordre du micromolaire et inférieur (concentration en chacun des composés lors du test : 0,2 µg / mL soit 0,5 µM).

→ Résultats obtenus après criblage des 2 chimiothèques orthogonales

Après avoir criblé les sous-chimiothèques de la **chimiothèque A** (la première des deux chimiothèques orthogonales), nous avons obtenu **4 sous-chimiothèques** capables d'inhiber à plus de 60 % la liaison de la vasopressine tritiée à son récepteur : **A₃-A₂-A₄** (92 %), **A₃-A₂-A₂** (76 %), **A₃-A₁-A₄** (73 %), (76 %), **A₃-A₂-A₅** (73 %). Les acides aminés en position

intermédiaire et C-terminale semblent pouvoir être plus facilement substitués que ceux placés en position N-terminale. Les 125 sous-chimiothèques de la **chimiothèque B** ont alors été testées. Les sous-chimiothèques **B₅-B₃-B₁**, **B₅-B₅-B₁** et **B₄-B₃-B₁** présentaient respectivement des activités inhibitrices de 73 %, 59 % et 68 %. Afin de réduire le nombre de combinaisons à synthétiser (3 x 3 = 9 combinaisons possibles), les 7 sous-chimiothèques ont été testées après une dilution au dixième (soit 2,5 µg / mL). Dans ces conditions, **seules A₃-A₂-A₄ (30 %), B₅-B₃-B₁ (23 %) et B₄-B₃-B₁ (18 %) restent actives.**

Les intersections des sous-chimiothèques actives de A avec celles de B conduisent à deux tripeptides actifs possibles (cf. figure 25) : H-D-Tic-D-Tyr-D-Pnp-OH et H-D-Glu-D-Tyr-D-Pnp-OH. Ces deux peptides ont donc été synthétisés, mais seul le premier a révélé une activité inhibitrice de la liaison de l'arginylvasopressine tritiée (3 nM) à son récepteur avec une IC₅₀ 63 nm (cf. figure 26).

Nous avons ensuite déterminé grâce à un test fonctionnel cellulaire, si le composé actif était agoniste ou antagoniste. Pour cela, le produit à tester est incubé en présence de doses croissantes de vasopressine. Dans les conditions physiologiques, lorsque la vasopressine se fixe sur le récepteur V2, celle-ci déclenche un signal post membranaire : la libération d'AMPc. En l'absence d'inhibiteur, la quantité d'AMPc est mesurée et la valeur 100 % lui est attribuée. Notre composé a induit une inhibition de l'activation de l'adénylate avec une IC₅₀ de 210 nm, en présence de vasopressine (2 nM), il s'agissait donc d'un ligand antagoniste. L'étape ultérieure a été de tester ce composé sur un récepteur V2 humain. Nous avons alors constaté, malgré une analogie de 85 % entre les deux récepteurs, une perte totale d'activité.

Des travaux d'optimisation d'activité ont aussi été réalisés (thèse de Line Bourel soutenue le 13 Janvier 1997 à l'université des Sciences et Techniques de Lille). Trente analogues ont été synthétisés : les modifications effectuées ont concerné l'acétylation de l'amine N-terminale, la synthèse du dérivé carboxamide en position C-terminale, la synthèse de diastéréoisomères et le remplacement de noyaux aromatiques par des cycles isostères. Parmi ces composés, seul le dérivé dans lequel la fonction a été substituée par un atome de chlore, s'est montré plus actif (IC₅₀ = 100 nm dans le test fonctionnel alors que le composé issu de la chimiothèque, possédait une IC₅₀ = 210 nm). Dans de nombreux cas, une perte totale d'activité a été observée.

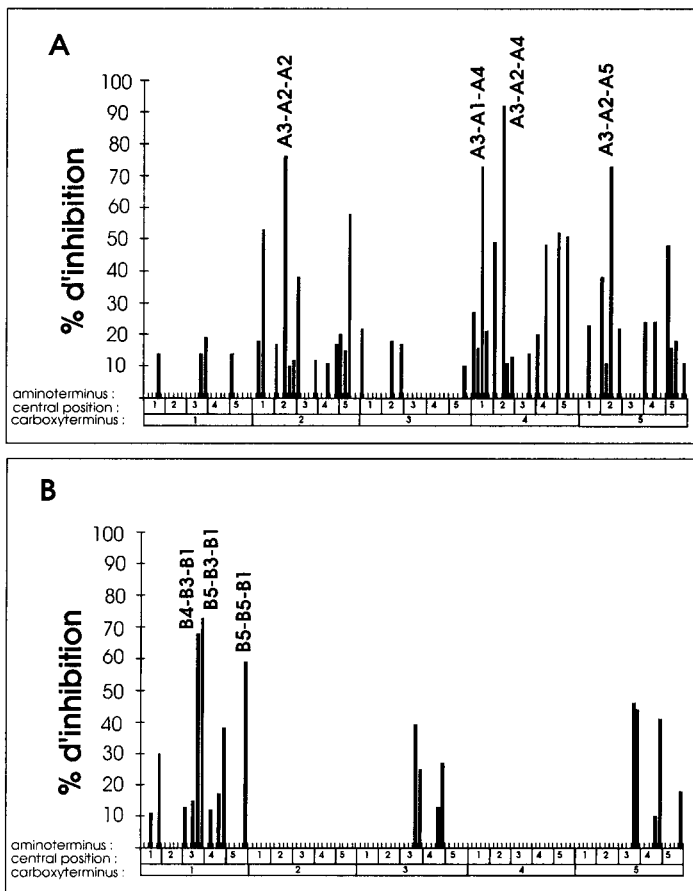


Figure 24 : Screenogramme des 2 chimiothèques orthogonales sur le récepteur V₂ (sous-chimiothèques testées à une concentration de 25 µg / mL).

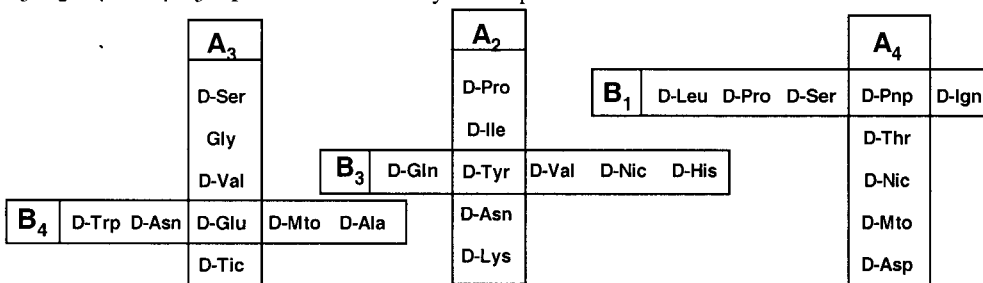
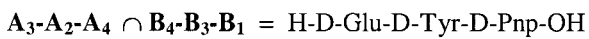
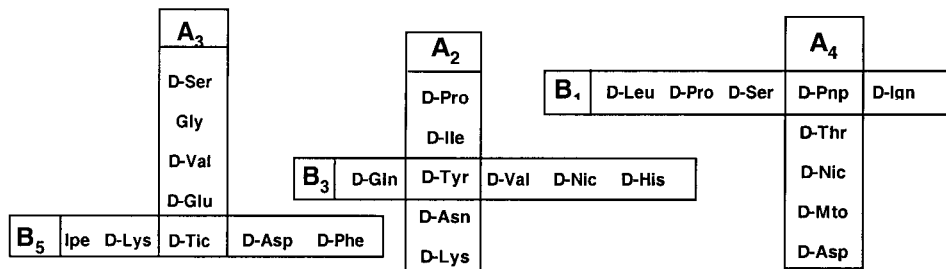
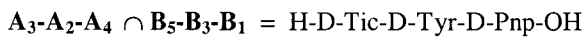
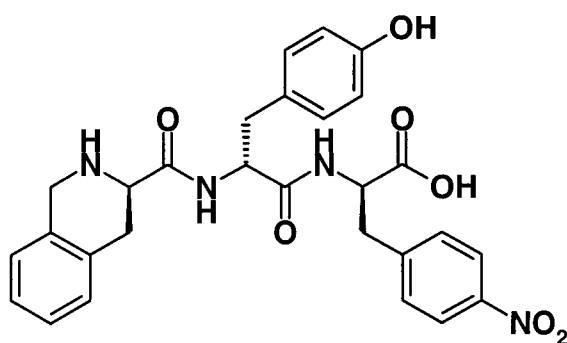


Figure 25 : identification des composés partagés par les sous-chimiothèques actives



H-D-Tic-D-Tyr-D-Pnp-OH

Figure 26 : structure du tripeptide antagoniste V2 porcine, trouvé dans les chimiothèques orthogonales de 15625 tripeptides

V-C-2-3 Discussion

L'efficacité de la méthode

Notre objectif principal consistait à prouver qu'il était possible d'accéder à des composés biologiquement actifs au sein de chimiothèques combinatoires, sans connaissance préalable de données relatives aux cibles biologiques et à leurs ligands. Des tests réalisés sur des virus, des canaux ioniques ou des récepteurs à 7 domaines transmembranaires ont montré que nos chimiothèques répondaient à nos attentes.

Nous avons parallèlement cherché à mettre au point une nouvelle méthode de déconvolution pour accélérer l'identification de composés actifs au sein de mélanges. La découverte de l'antagoniste V2 nous a permis d'illustrer parfaitement son efficacité, l'identité d'un produit actif présent dans une sous-chimiothèque appartenant à une première chimiothèque, ayant pu être déduite par l'identification d'une sous-chimiothèque active appartenant à une chimiothèque orthogonale à la première.

Les limites de la méthode

Comme nous l'avons indiqué au paragraphe concernant les résultats biologiques obtenus sur le récepteur V2, cette méthode nous a également conduit à un composé « faux positif ». En admettant que les résultats biologiques révélant une activité pour la sous-chimiothèque B₄-B₃-

B₁ à deux reprises (lors des tests à 25 puis à 2,5 µg / mL) confirmaient la présence d'au moins un composé actif, nous avons trouvé trois explications possibles à l'obtention d'un faux positif.

Explication 1 : l'activité de la sous-chimiothèque B₄-B₃-B₁ pourrait résulter du cumul d'activité de plusieurs composés. Il devient alors très difficile d'identifier ces composés par le biais d'une chimiothèque orthogonale, ces composés ayant des activités individuelles souvent faibles, et ceux-ci étant répartis dans plusieurs sous-chimiothèques orthogonales à celle présentant l'activité biologique.

Explication 2 : la présence d'un produit secondaire n'apparaissant que dans la sous-chimiothèque B₄-B₃-B₁.

Explication 3 (la plus probable) : le composé actif présent dans B₄-B₃-B₁ n'est pas partagé par A₃-A₂-A₄, mais par une autre sous-chimiothèque de A n'ayant pas été sélectionnée lors du criblage à 2,5 µg / mL. Par analogie structurale avec le composé actif H-D-Tic-D-Tyr-D-Pnp-OH, nous avons supposé que le composé H-D-Trp-D-Tyr-D-Pnp-OH pouvait porter l'activité, la sous-chimiothèque A₁-A₂-A₄ (sous chimiothèque de A contenant ce composé) devait donc elle aussi être active. En analysant plus en détail le screenogramme (à 25 µg / mL), nous avons d'ailleurs constaté que cette sous-chimiothèque figurait parmi les 10 sous-chimiothèques de A les plus actives (pourcentage d'inhibition de l'ordre de 50 %). Ce peptide a donc été resynthétisé et son IC₅₀ sur test fonctionnel a été évaluée à 5 µM.

Cet exemple nous a montré que dans le cas où plusieurs sous-chimiothèque était trouvées actives, une sélection drastique des sous-chimiothèques actives basées sur des constantes d'inhibitions pouvait engendrer des résultats complexes. En effet, des données telles que l'équimolarité effective des composés dans un mélange, la potentialisation ou l'inhibition de composés les uns par rapport aux autres et la présence de produits secondaires ne peuvent être appréhendées. En revanche quand cette sélection ne souffre d'aucune ambiguïté (un seul composé actif ou composé nettement plus actif que les autres) la méthode démontre tout son intérêt. Dans ce contexte, l'identification sans équivoque du composé actif H-D-Tic-D-Tyr-D-Pnp-OH est probablement liée à la difficulté de l'optimisation du composé actif H-D-Tic-D-Tyr-D-Pnp-OH et donc à l'absence de ligands suboptimaux dans la chimiothèque.

Il faut cependant toujours garder à l'esprit qu'une chimiothèque combinatoire est conçue pour conduire à l'obtention de composés actifs identifiés et non pour apporter une série de données exhaustives sur chaque composé théoriquement présent dans cette chimiothèque.

VI Conclusion de la première partie

La chimie combinatoire, bien qu'apparaissant maintenant comme une technique de transition, s'est révélée être un passage obligé vers la production de molécules organiques à haut débit. Parmi les nombreuses expériences acquises au cours de cette période, l'enseignement majeur reste la possibilité offerte aux chimistes médicaux de synthétiser rapidement, à faible coût, un grand nombre de composés et de ne purifier et caractériser que ceux justifiant d'une activité biologique. Le second enseignement est la possibilité d'utiliser « un hasard maîtrisé » pour la découverte de composés pouvant ultérieurement conduire à des médicaments originaux. Les conséquences directes de la maîtrise des éléments constitutifs (monomères) d'une chimiothèque sont : la connaissance des composés théoriquement présent dans cette chimiothèque, ainsi que la possibilité d'utiliser des éléments constitutifs stables, à réapprovisionnement aisé et dont l'accès à des éléments analogues est tel que tout composé trouvé actif pourra devenir le point de départ d'une famille de composés dérivés. Le troisième enseignement est la possibilité offerte aux biologistes de tester des composés en mélange (ces composés ayant pu être synthétisés en mélange ou de façon individuelle puis remélangés), un gain de temps considérable ainsi qu'une diminution des coûts en réactifs, sont alors induits lors des phases de criblage.

La chimie combinatoire présente cependant quelques limites que les chimistes tentent de contourner : le suivi complexe des composés en cours de synthèse ou dans le mélange final, la difficulté d'identifier le(s) composé(s) conférant l'activité à un mélange, le format de synthèse inadapté aux tests de criblage (synthèse en tube, criblage en microplaques 96 puits), les chimies utilisées peu perfectionnées et comprenant un nombre d'étape limité. Nous assistons actuellement à la synthèse de composés individuels, en parallèle, mais de façon automatisée, afin de maintenir des cadences élevées de production. Les installations nécessaires demeurent très coûteuses et sophistiquées, mais permettent d'éviter les problèmes cités précédemment. Un exemple de développement d'une chimie automatisée est décrit dans la seconde partie de cette thèse.

Orthogonal Combinatorial Chemical Libraries

Benoît Déprez, Xavier Williard, Line Bourel, Hervé Coste, François Hyafil, and André Tartar*

Department of Chemistry, Institut Pasteur de Lille
 Rue Calmette, 59045 Lille, France
 GLAXO Centre de recherches
 Avenue du Québec, 91951 Les Ulis, France

Received November 2, 1994

Although knowledge-based *de novo* design of high-affinity ligands for receptors or enzymes appears to be the most rational approach to the discovery of new pharmacologically active substances, high-throughput random screening of compounds having no known structural similarity with the natural ligand has proved to be a highly efficient method. To increase the input of new molecules, methods of combinatorial synthesis have been considerably developed in recent years.¹⁻⁴ Much work has been devoted to the development of various strategies designed for the identification of the active compounds in a mixture.⁵ These strategies often reduce the scope of the chemical combinations and the biological assays. To circumvent these difficulties, we have developed a general concept which we name "orthogonal libraries". Herein we describe the design, synthesis, and first use of a pair of such orthogonal libraries,⁶ leading to the discovery of a nanomolar ligand to the V2 vasopressin receptor.

Using a classical solid phase procedure, we prepared two libraries⁷ of the same soluble 15 625 trimers resulting from the combination of 25 synthons at each position, linked by two amide bonds (Figure 1a). The synthons were 23 D-amino acids and 2 nonchiral amino acids, glycine and isonipecotic acid (Figure 1b). Each library was partitioned into 125 sublibraries of 125 trimers. For the first library (A), the 25 synthons were partitioned in five groups, A₁–A₅. Each sublibrary resulted from the incorporation of one group, A_n (*n* = 1–5), at each of the three variable positions of the synthesis ($5^3 = 125$). For the second library (B), the same 25 synthons were used, yielding the same 15 625 trimers. They were, however, partitioned *orthogonally* in five different groups, B₁–B₅ (Figure 1b).

The important property of the couple of libraries A and B is that, for any given *n* and *n'*, A_{*n*} and B_{*n'*} share only one synthon. Thus, any sublibrary from A and any sublibrary from B share only one trimer. In a given screen, an active trimer will confer

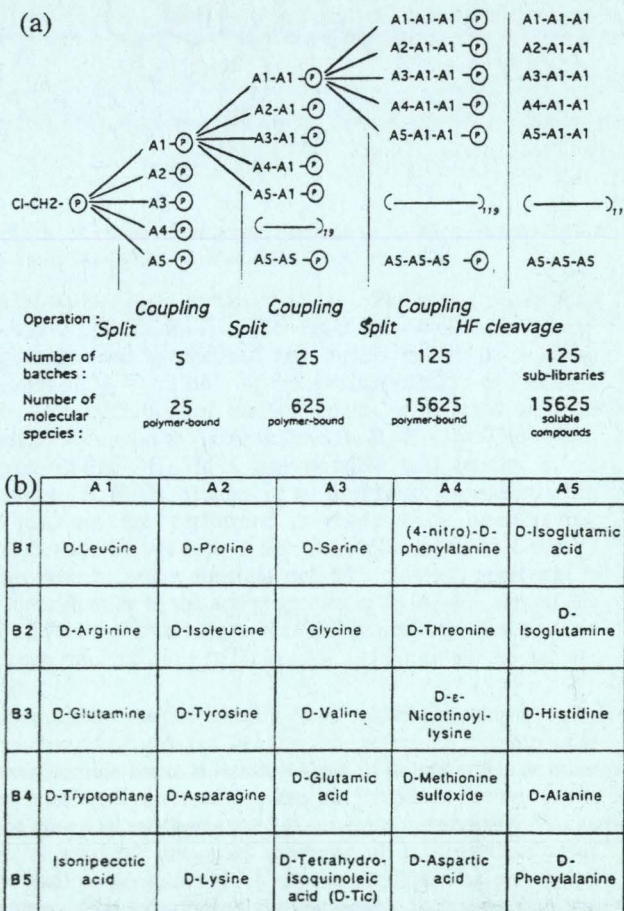


Figure 1. Synthesis of the two libraries. (a) The two libraries were synthesized by solid phase synthesis.⁷ (b) The 25 amino acids were partitioned into five groups of five amino acids, each named A₁–A₅ for library A and B₁–B₅ for library B. In each group, amino acids were selected to maximize the structural diversity. For the synthesis of library A, each group (A₁–A₅) was coupled to a chloromethylated Merrifield resin (Cl-CH₂-P), and the five mixtures of aminoacyl resins were split into five parts. To each part, one different group of amino acids were coupled. The resulting 25 dipeptidylresins were split again into five parts each and coupled with the mixtures of amino acids, leading, after deprotection and cleavage from the resin, to 125 sublibraries in a soluble form. Library B was prepared according to the same procedure, except that the orthogonal groups of amino acids, B₁–B₅, were used in this case.

activity to one single sublibrary of A and one single sublibrary of B. This active trimer can be unequivocally identified as the unique trimer shared between these sublibraries. The distribution of the 25 synthons in the groups A_{*n*} or B_{*n'*} has been made in such a way that the structural diversity within a group of synthons is as large as possible. This avoids the synthesis of sublibraries containing structurally related compounds which could generate cumulative biological activities. Typically, a group consisted of one hydrophilic, one aromatic, one charged, and one small-chain α-amino acid in addition to a nonnatural side chain amino acid. Thus, each sublibrary of A or B explores a wide structural space, and if one trimer confers an activity to its sublibrary, the 124 other trimers of this sublibrary are less likely to interfere with the biological assays. This structural diversity within the mixtures submitted to the screening procedure, together with the limited number of chemical entities present in each sample (125), allows a low background and an easy detection of biological activity.

According to our strategy, library A was screened initially in a series of assays. When a positive result was obtained in

(1) Houghten, R. A.; *et al. Nature* **1991**, *54*, 84–86.
 (2) Lam, K. S.; *et al. Nature* **1991**, *354*, 82–84.
 (3) Gallop, M. A.; *et al. J. Med. Chem.* **1994**, *37*, 1235–1251.
 (4) Gordon, M. E.; *et al. J. Med. Chem.* **1994**, *37*, 1385–1401.
 (5) Janda, K. D.; *et al. Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10779–10785.

(6) A related concept, named "indexed libraries" has recently been reported for the deconvolution of a mixture of 54 carbamates: Pírrung, M. C.; *et al. J. Am. Chem. Soc.* **1995**, *117*, 1240–1245. Smith, P. W.; *et al. BioMed Chem. Lett.* **1994**, *4*, 2821–2824. This group describes a similar strategy for the screening of 1600 esters or amides.

(7) Boc-protected amino acids were esterified to a chloromethylated resin. Benzyl-based protections were used for the side chains of trifunctional amino acids. 1-Hydroxybenzotriazolyltetramethylisouronium hexafluorophosphate (HBTu) was used for the couplings of Boc-protected amino acids mixtures. To avoid the preferential incorporation of amino acids having the most favorable coupling rates, a first coupling was performed with only 1.1 equiv of acylating species for 1 equiv of amino groups available on the resin. A second coupling step was then performed to force the reaction to completion. Amino acid analysis of 10 sublibraries showed that most amino acids were within 10% of the mean value, the most divergent ratios being within 25% of this value. At the end of the synthesis, the 250 polymer-bound sublibraries were treated for 1.5 h at 0 °C with anhydrous hydrogen fluoride in the presence of *p*-cresol and *p*-thiocresol. The resulting mixtures of peptides were precipitated with a mixture of diethyl ether and pentane, dissolved in trifluoroacetic acid, reprecipitated, and lyophilized from acetic acid. (a) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154. (b) Tam, J. P.; Merrifield, R. B. *The peptides*; Academic Press: New York, 1987; pp 185–249.

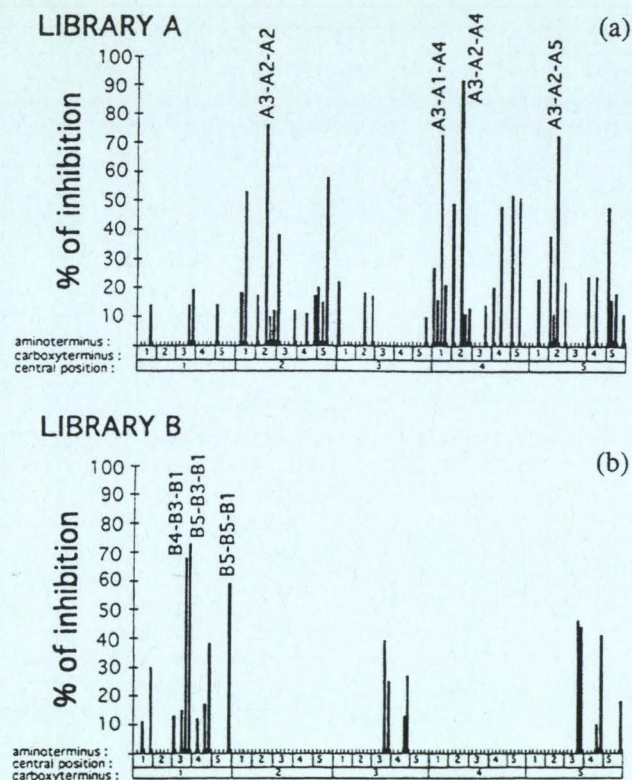


Figure 2. Biological activity of the two libraries in the V2 receptor binding assay. Cells were incubated in the presence of sublibraries from library A (a) or library B (b) at 25 $\mu\text{g/mL}$. Each bar represents the percentage of inhibition observed for one sublibrary.

one assay, library B was screened under the same conditions. In addition to allowing a direct identification of the active compound, this approach provides an internal control, as this compound is associated with a different set of 124 trimers in library A and in library B. As an example, we report the discovery of a new ligand for the V2 vasopressin receptor. All sublibraries of library A were tested at a concentration of 25 $\mu\text{g/mL}$, corresponding roughly to 0.2 $\mu\text{g/mL}$ of each individual compound. Activities were determined by the inhibition of binding of radiolabeled vasopressin to renal LLCPK1 cells⁸ (Figure 2a). The most active sublibrary was A₃-A₂-A₄, with 92% inhibition; three other sublibraries gave inhibitions higher than 70%: A₃-A₂-A₂, 76%; A₃-A₁-A₄, 73%; and A₃-A₂-A₅, 73%, indicating that some replacements could be tolerated in the middle and carboxy-terminus positions, while in the amino-terminus position, A₃ contains a more stringent synthon. When the orthogonal library B was tested (Figure 2b), three sublibraries exhibited high activities, the most potent being B₄-B₃-B₁ with 73% inhibition, B₅-B₃-B₁ with 69%, and B₅-B₅-B₁ with 59%. This selection yielded 12 (4 × 3) possible inhibitors. To narrow the choice, we selected sublibraries on the basis of the dose-dependence of the inhibition and the potency at lower concentration. At 7.5 $\mu\text{g/mL}$, while only A₃-A₂-A₄ remained active in A (31% inhibition), in B, two sublibraries, B₅-B₃-B₁ (37%)

(8) Porcine kidney epithelial LLC-PK1 cells (American Type Culture Collection) were incubated with 3 nM [³H]AVP in the presence of inhibitors. Following incubation at 4 °C for 180 min, the cells were filtered and washed. Radioactivity was determined by liquid scintillation counting. Lutz, W.; Londowski, J. M.; Sanders, M.; Salisbury, J.; Kumar, R. *J. Biol. Chem.* **1992**, *267*, 1109–1115.

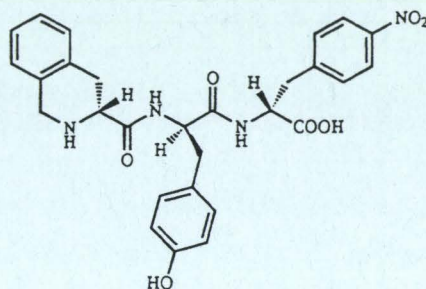


Figure 3. Structure of the trimer common to A₃-A₂-A₄ and B₅-B₃-B₁, H-D-Tic-D-Tyr-(4-nitro)-D-Phe-OH.

and B₄-B₃-B₁ (31%) remained active. The trimer common to A₃-A₂-A₄ and B₅-B₃-B₁, H-D-Tic-D-Tyr-(4-nitro)-D-Phe-OH (Figure 3), was synthesized and turned out to be a potent inhibitor (IC₅₀ = 63 nM) of vasopressin binding to LLCPK1 cells. No inhibition of binding could be detected with the candidate common to A₃-A₂-A₄ and B₄-B₃-B₁, H-D-Glu-D-Tyr-(4-nitro)-D-Phe-OH. It is thus possible that activity in the sublibrary B₄-B₃-B₁ resides in an unwanted byproduct when the reactions are performed in pools. The unambiguous identification of D-Tic as a necessary pharmacophore for V2 antagonism is further substantiated by a classical approach¹ to the identification of the active species in A₃-A₂-A₄: among the five mixtures of 25 trimers where ambiguity at the first position has been resolved, only D-Tic-A₂-A₄ (and not D-Glu-A₂-A₄) was active.

In summary, the use of orthogonal self-deciphering libraries offers several advantages: it allows the screening of compounds in their soluble form, it is not limited to polypeptides or other sequenceable polymers, and it does not introduce any restriction to the chemical strategies used to generate the diversity. It can even be used for the rapid screening of existing libraries of individual compounds which can be organized as orthogonal mixtures. The interest of this approach increases with the number of screens to which the libraries are submitted, as once the libraries are prepared, no further step of deconvolution or analysis is required.⁹ Moreover, orthogonal libraries also offer an internal validation, as any positive result detected in one library must be confirmed by a positive result in the orthogonal library, thus allowing the detection of false positive results.

Acknowledgment. We thank M. G. Néliat (CEREP, Poitiers, France) for performing the binding assays and Ms. Florence Lorient for technical assistance. B.D. and X.W. are supported by a Region Nord-Pas-de-Calais grant.

JA943561X

(9) Although also based on partitioning the library into sublibraries, orthogonal libraries differ from the positional scanning described by R. Houghten (Pinilla, C.; *et al. Biotechniques* **1992**, *13*, 901–905) in several aspects. A positional scan in this case would have required the synthesis of 75 sublibraries (OXX, XOX, and XXO) instead of 250. However, each sublibrary consists of 25 × 25 = 625 compounds instead of 5 × 5 × 5 = 125, introducing a 5-fold dilution factor (although in each case, the libraries contain 15 625 different compounds, in the positional scanning, each compound is synthesized three times). Moreover, the 625 compounds share the same defined position, introducing the risk that among the different possibilities, the determination of the most active sublibrary will favor the defined residue (O) which is compatible with the greatest number of combinations in the two variable positions (X) and not the residue belonging to the most active molecular entity. This risk is greatly diminished with orthogonal libraries, which are organized to maximize the diversity within each sublibrary. Moreover, each compound is tested in the presence of two different sets of 124 compounds in libraries A and B.

Seconde partie

La synthèse parallèle à haut débit
l'outil de choix pour la découverte de nouvelles
molécules biologiquement actives

Seconde partie

La synthèse parallèle à haut débit

**l'outil de choix pour la découverte de nouvelles
molécules biologiquement actives**

Chapitre A : L'automatisation en chimie organique

Le criblage, systématique, de bibliothèques de composés organiques sur des cibles biologiques, s'est imposé ces cinq dernières années comme la méthode de choix pour la découverte de futurs médicaments. L'efficacité de cette méthode est étroitement liée au nombre de composés testés : en supposant que les composés soumis au criblage soient variés, plus le nombre de composés testés est important, plus la chance de trouver un composé actif augmente. Afin de maintenir une croissance commerciale suffisante, les dix plus grandes sociétés pharmaceutiques mondiales devraient chacune mettre sur le marché 5 molécules innovantes par an, alors qu'en moyenne, entre 1990 et 1994, seulement 0,45 molécule innovante était lancée, par an, par les dix plus grandes sociétés pharmaceutiques. C'est dans ce contexte que s'est manifestée la volonté d'automatiser l'étape de découverte de structures actives originales, au niveau des tests biologiques, mais aussi au niveau de la synthèse des composés testés. En février 1996, le consortium " Myriad " (regroupant en partenariat BASF, Chiroscience, Merck & Co., Novartis, Pfizer, SmithKline Beecham, Takeda Chemical Industries et The Technology Partnership, sur le thème de la chimie automatisée) a révélé la volonté de l'industrie pharmaceutique d'accentuer ses efforts dans ce domaine.

Les stratégies d'automatisation :

L'automatisation de différentes étapes de production a permis à la chimie combinatoire d'atteindre des cadences de synthèse en parallèle de plusieurs milliers de composés par jour, mais aussi de limiter les risques d'erreur de manipulation. Cependant ce type de production, nécessite un équipement particulièrement sophistiqué et ne tolère aucune faille dans la gestion des données relatives à la synthèse de chaque composé.

Deux stratégies peuvent être adoptées pour fiabiliser, automatiser et " enregistrer " les étapes clé d'une production : la **stratégie modulaire** et la **stratégie intégrée**.

La stratégie modulaire consiste à automatiser différentes tâches nécessaires à une production en les considérant indépendantes les unes par rapport aux autres. Plusieurs automates seront donc utilisés, les paramètres à gérer sur chacun de ces automates étant relativement simples. Des tâches spécifiques à chaque production peuvent être facilement ajoutées. Le principal inconvénient de cette approche reste la nécessité d'interventions humaines tant au niveau de l'organisation d'une production, qu'au niveau de la gestion des produits (et des données associées) issus de cette production. Ce type de stratégie est plutôt adapté à la synthèse de chimiothèques de taille moyenne (1000 à 10.000 produits).

La stratégie intégrée consiste à automatiser toutes les tâches nécessaires à une production en les réunissant dans une macro-tâche gérée par un système unique. Les avantages liés à cette approche sont : l'autonomie importante du système, l'homogénéité des données informatiques recueillies et l'excellente traçabilité des événements. Cependant le peu de flexibilité du système ainsi que l'importante qualification technique nécessaire à sa programmation, en font un outil coûteux, réservé à quelques sociétés spécialisées dans la production de chimiothèques de grande taille (plusieurs dizaines de milliers de composés) synthétisées à très haut débit (plusieurs milliers de produits par jour).

	Automatisation modulaire	Automatisation intégrée
Principe	Automatiser différentes tâches par des systèmes indépendants	Automatiser différentes tâches et les réunir dans une même macro-tâche pilotée par un système unique
Flexibilité	+++	
Complexité de programmation	+	+++
Interventions humaines en cours de production	++	(+)
Traçabilité	+	+++
Cadences de production	Plusieurs centaines de produits par jour	Plusieurs milliers de produits par jour
Frais d'équipement	+	+++
Maintenance	+	++
Fournisseurs	Advanced ChemTech, Argonaut, Beckman, Bioblock, Bodhan, Gilson, Roucaire, Stemm, Tecan, Zymark ...	Bodhan, Scitec, Zymark ...

Les tâches automatisées :

- La pesée des matières premières

Une des étapes manuelles de la synthèse à haut débit, difficilement automatisable, reste la pesée des matières premières, le problème majeur étant la variété des états physiques des produits à peser (liquide, solide, huileux). Nous avons donc développé au laboratoire, un logiciel permettant d'assister les pesées. L'opérateur doit d'abord identifier grâce à un lecteur code barres le produit à peser, s'affiche alors la quantité à peser, avec une certaine tolérance. Le produit est ensuite pesé, cette pesée ne pouvant être validée par l'opérateur que si elle se situe dans la fourchette indiquée par le logiciel.

Un autre système développé par la société Bodhan, permet de tarer automatiquement jusqu'à 350 tubes (à un rythme de 240 / heure). Un opérateur ajoute ensuite dans chacun des tubes le produit attendu. La pesée est finalement réalisée par l'automate, prenant en compte de la tare de chacun des tubes.

- La mise en solution des matières premières

A partir d'une masse de poudre pesée (renseignée dans un fichier informatique), un volume calculé de solvant peut être distribué pour atteindre la concentration en produit souhaitée. Une agitation de l'échantillon permet ensuite sa mise en solution. (Combiprep™ de la société Gilson)

- Le prélèvement et la distribution des réactifs en solution

Les préparateurs automatiques d'échantillons sont de deux types : les monoaiguille et multiaiguilles, ces derniers permettant d'accroître considérablement les cadences de productions. Le temps pour pipeter 96 x 100 µL d'échantillon dans une plaque 96 puits est inférieur à 4 minutes sur un préparateur tel que le Genesis™ pourvu de 8 aiguilles (Société Tecan) alors qu'il serait d'au minimum 30 minutes pour un système monoaiguille.

Les systèmes peuvent être pourvus de détecteurs de niveau à l'extrémité de chaque aiguille. Les volumes traités s'échelonnent de quelques microlitres à plusieurs millilitres.

- L'agitation des milieux réactionnels

De nombreuses méthodes peuvent être utilisées, celles-ci étant étroitement liées au nombre ainsi qu'au type de réacteurs à agiter. Il est par exemple possible de distinguer les systèmes d'agitation collective (agitation par vibrations, agitation par vortex, agitation par retournement), des systèmes d'agitation individuelle (distribution d'un gaz inerte au sein du réacteur, agitation par vortex, agitation magnétique, agitation par hélice).

- Le maintien en atmosphère inerte des milieux réactionnels

Les systèmes pourvus de cette fonctionnalité sont conçus de deux manières : soit l'atmosphère inerte est assurée localement, dans le réacteur où s'effectue la synthèse, soit l'ensemble du système automatisé est placé sous atmosphère inerte.

Dans le premier cas les réacteurs peuvent être traités individuellement ou en groupe. Dans le cas de réacteurs non étanches, l'apport du gaz inerte se fait de façon continue. Au contraire, si l'étanchéité des réacteurs est garantie, une simple mise sous gaz inerte, au début de la réaction peut suffire. L'avantage de ce type d'installation réside dans la faible consommation de gaz inerte ainsi que dans la possibilité d'intervention sur les autres composants du système robotique sans influencer la composition de l'atmosphère présente au sein de chaque réacteur. Cependant cette installation limite les opérations effectuées sous atmosphère inerte à celles s'opérant au niveau des réacteurs.

Dans le second cas, le système automatisé est généralement placé dans une enceinte en légère surpression de gaz inerte. Un apport continu de ce gaz est donc nécessaire. L'avantage majeur de ce type d'installation est sa flexibilité, aussi bien en variété de réacteurs utilisables, qu'en diversité des opérations pouvant s'effectuer sous atmosphère inerte (mise en solution, distribution des réactifs, manipulation des intermédiaires de synthèse ...). Cependant, au cours des phases de synthèse, le système doit rester clos, ce qui limite d'éventuelles interventions humaines à l'intérieur de l'enceinte.

- Le chauffage et le refroidissement des milieux réactionnels

Des systèmes de chauffage permettant d'atteindre des températures de 150°C, sont maintenant adaptés à une grande variété de réacteurs, allant du récipient en verre à la microplaque. Les calories sont généralement apportées par un support métallique chauffé par une résistance

électrique. Un contrôle individuel de température est possible sur certains systèmes (ex. le Nautilus™ de la société Argonaut) essentiellement utilisés pour la mise au point de conditions opératoires. La condensation des vapeurs issues des milieux réactionnels chauffés, offre la possibilité d'effectuer des synthèses à reflux. Le plus souvent la circulation d'un fluide thermostaté à proximité de la partie supérieure des réacteurs permet la condensation des vapeurs. Pour des synthèses effectuées en microplaques, des systèmes tels que des aiguilles refroidies, placées à l'aplomb de chaque puits, doivent être développés (un système de ce type, exclusif à CEREP, nous a été fourni par la société Scitec).

Le refroidissement des milieux réactionnels jusqu'à des températures avoisinant les $-70\text{ }^{\circ}\text{C}$ est actuellement possible sur de nombreux systèmes. Les technologies utilisées sont l'effet Pelletier et la circulation de fluides réfrigérants. (Comme précédemment, un contrôle individuel de température est possible sur certains systèmes). Cependant, pour des systèmes au contact de l'air ambiant, un problème fréquemment rencontré reste la condensation de l'humidité atmosphérique.

- L'évaporation des milieux réactionnels

Les systèmes rencontrés font appel soit à la technique de l'évaporation sous pression réduite, soit à la technique de l'évaporation sous flux gazeux.

Les évaporations sous pression réduite permettent d'éliminer des solvants à point d'ébullition $> 100^{\circ}\text{C}$ avec des débits rapides. Ces appareils sont généralement pourvus d'un système de chauffage, d'un système de centrifugation intégré, ainsi que de pièges permettant de recueillir les solvants. La résistance à des solvants très corrosifs, comme le TFA, est maintenant assurée par des constructions entièrement téflonnées (systèmes commercialisés par les sociétés Savant ou Genevac). Un problème souvent rencontré lors de l'évaporation de solvants à faible point d'ébullition (comme le DCM) est le dégazage brutal de ces solvants, conduisant à des projections de solutions pouvant contaminer les tubes voisins.

Les évaporations sous flux gazeux sont particulièrement adaptées à l'évaporation de solvants à faible point d'ébullition. Celles-ci nécessitent un bloc chauffant, une source de gaz ainsi qu'un système permettant d'isoler l'environnement extérieur des vapeurs toxiques ou corrosives. Des exemples de vitesses d'évaporation, sous une pression gazeuse de 20 psi, sont donnés ci-dessous :

Solvant (2 mL)	
DCM	5 min. / 40°C
Hexane / Acétate d'éthyle (50 : 50)	5 min. / 40°C
Méthanol	8 min. / 40°C
DMF	40 min. / 55°C

- La purification des milieux réactionnels

Lorsqu'elle est nécessaire, cette étape constitue le goulot d'étranglement majeur de toute production robotisée. Les techniques d'extraction liquide / liquide (système Combitec™ de la société Técan) et solide / liquide (système Aspec™ de la société Gilson) peuvent permettre d'éliminer des produits secondaires attendus, ayant des propriétés physicochimiques proches voire identiques. Cependant la forte diversité caractérisant les composés présents dans une chimiothèque généraliste semble peu compatible avec des méthodes de purification aussi peu universelles que les méthodes d'extraction automatisées. Les recherches semblent maintenant s'orienter vers l'utilisation de l'HPLC couplée à la spectrométrie de masse, permettant d'isoler les produits ayant les caractéristiques de masse atomique attendues (système FractionLynx™ de la société Micromass). Il faut cependant reconnaître, qu'à ce jour, cette méthode se limite à la purification de faibles quantités de produits (1 à 20 mg).

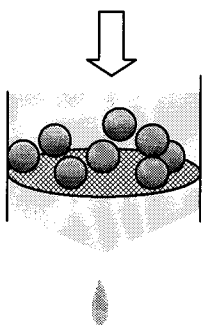
- La filtration de résines pour la synthèse en phase solide

L'utilisation de la chimie en phase solide constitue une technique alternative aux méthodes de purification décrites précédemment. Elle peut être utilisée, à la fois pour permettre la synthèse du composé en utilisant le support solide, mais aussi pour introduire dans les différents milieux réactionnels, des composés, immobilisés sur ce support solide, permettant de piéger certaines impuretés puis de les éliminer par simple filtration (Kaldor 1996) ou d'activer des réactifs (Pop 1997).

La chimie combinatoire ayant pour point de départ la chimie peptidique, les premiers automates apparus permettaient la manipulation de résines de type polystyrène, utilisées pour

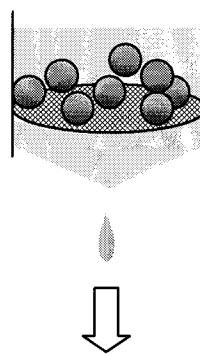
la synthèse en phase solide. Il faut dans ce cas utiliser des systèmes permettant d'isoler le support solide des solutions. Trois techniques sont principalement utilisées :

La filtration sur fritté



Sous pression

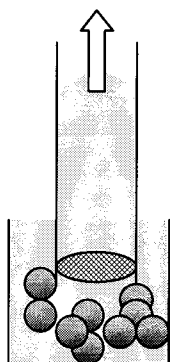
ex : le système ReACTECH™
(Société Advanced ChemTech)



Sous vide

ex : le système Gilson Personal
Synthesizer System™
(Société Gilson)

La filtration au moyen d' "aiguilles"



Par aspiration

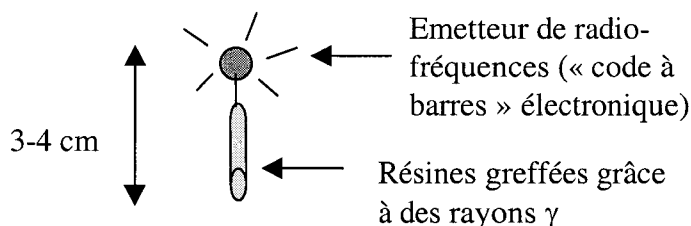
ex : le système Combitec™
(Société Tecan)

La société CEREP, en partenariat avec la société Scitec, a développé un système basé sur ce principe.

La technique des couronnes, des pins et des résines encapsulées

Le but de ces techniques est d'éviter les étapes de filtration au cours des synthèses en phase solide. Les résines seront, soit soudées à un support (couronnes SynPhase™ et pins et de la société Chiron), soit placées dans une capsule perméable aux solvants (MicroKans™ de la société Irori). La séparation résine / solution, ne se faisant plus par filtration, permet d'éviter les problèmes de colmatage de filtres ou frittés.

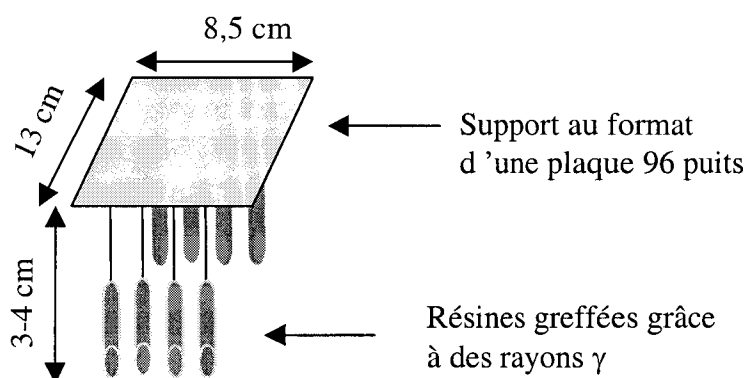
Les couronnes



Ces couronnes sont particulièrement adaptées à la synthèse combinatoire de type "split and mix". Le transpondeur (émetteur de radiofréquence), permet d'identifier, en fin de synthèse, le produit théoriquement présent sur la couronne.

(Les charges disponibles s'étalent, selon le type de couronne, de 4 à 40 μ moles par couronne).

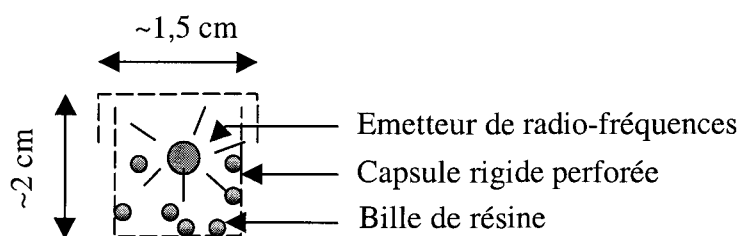
Les pins



Les pins permettent la synthèse de chimiothèques dans un format du type plaque 96 puits. L'identité théorique du produit obtenu après clivage, est définie par les coordonnées du puits qu'il occupe dans la plaque 96 puits.

(Les charges disponibles s'étalent, selon le type de couronne, de 4 à 40 μ moles par couronne).

Les résines encapsulées
(*MicroKans*)



La technique des résines encapsulées s'apparente à celle des couronnes. Elle est adaptée à la synthèse combinatoire de type "split and mix". Le transpondeur (émetteur de radiofréquence), permet d'identifier, en fin de synthèse, le produit théoriquement présent dans la capsule.

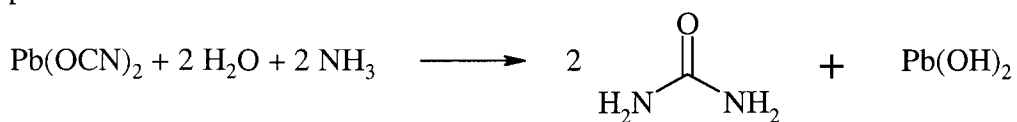
(Il est possible d'introduire 30 mg de résine soit une charge 15 à 30 μ moles par capsule)

Conclusion :

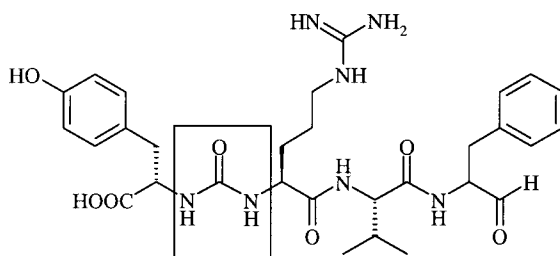
La chimie automatisée, ayant maintenant à sa disposition de nombreux systèmes capables d'exercer des tâches très variées, permet d'engendrer rapidement, un grand nombre de composés. Les données relatives à ces composés, qu'elles soient issues des tests biologiques ou de la modélisation moléculaire, doivent être utilisées au mieux, pour accélérer l'étape ultérieure de sélection de la structure susceptible d'entrer en développement. Des relations structures activités précoces doivent être établies afin de déterminer les meilleurs candidats en terme d'activité, de toxicité estimée, de caractéristiques pharmacocinétiques anticipées ainsi qu'en terme de protection industrielle.

Chapitre B : La synthèse robotisée d'urées asymétriques

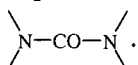
Les urées constituent une classe très importante de molécules organiques qui compte de nombreux médicaments. D'un point de vue historique, l'urée fût la première molécule organique synthétisée à partir de « substances inanimées » grâce à Friedrich Wöhler qui en 1828 réussit à transformer le cyanate de plomb en urée par action de l'ammoniaque, en présence d'eau.



Des dérivés substitués de l'urée constituent un motif de base retrouvé dans de nombreux médicaments, notamment dans la diéthylacétylurée (chef de file de la famille des barbituriques), qui possède des propriétés hypnotiques, dans les pyridylurées ayant des effets anti-inflammatoires, ou encore dans les 3-ureidopyrrolidines utilisés comme analgésiques. Plus récemment, ce motif est apparu dans plusieurs inhibiteurs de la protéase du VIH (Zhang 1997).



Ce lien est d'autre part souvent utilisé comme analogue de la liaison peptidique. De nombreuses synthèses sont décrites pour aboutir à l'enchaînement :



Parmi celles-ci, compte tenu de notre objectif combinatoire, nous nous sommes surtout intéressés aux synthèses conduisant à des urées asymétriques. Un rappel des grandes voies de synthèses d'urées sera donc effectué, où les réactions de condensation d'amines sur des isocyanates ou leurs dérivés seront particulièrement détaillées.

I Etude Bibliographique

A Les différentes voies d'accès aux urées

A-1 Réactions des amines avec l'urée



Conditions opératoires (Kurzer 1963 a) : Ces réactions se font généralement à des températures supérieures ou égales à 100°C, dans des solvants comme l'eau, le xylène ou l'*o*-dichlorobenzène. Les amines primaires et secondaires sont utilisables. Les rendements obtenus sont très bons. L'urée disubstituée (symétrique) constitue un sous-produit classiquement retrouvé.

Mécanisme supposé : l'urée serait la source d'isocyanate d'ammonium réagissant avec l'amine.

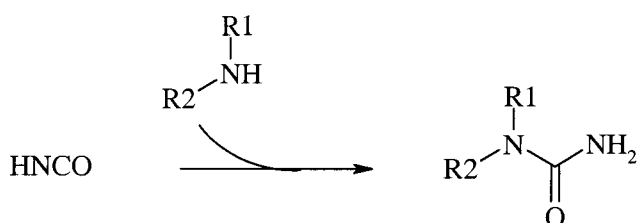
Intérêts :

- Matières premières facilement accessibles
- Grande diversité de la famille des amines utilisées

Inconvénients :

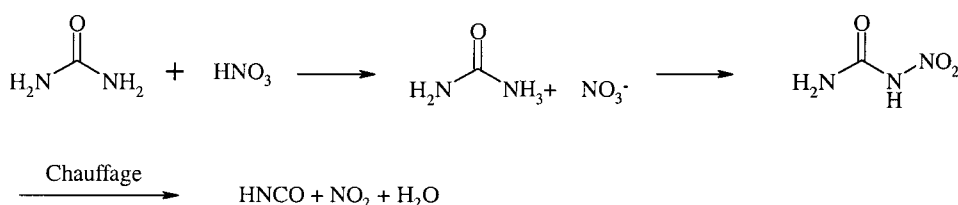
- Conditions opératoires drastiques
- Faible réactivité des amines secondaires
- Un seul site de diversité sur la molécule finale (la synthèse de la 1-chlorophenyl-3-diméthylurée a pourtant été décrite (Thompson 1954) en appliquant un protocole similaire, mais cette voie de synthèse n'a jamais été généralisée)

A-2 Réactions des amines avec l'acide cyanique

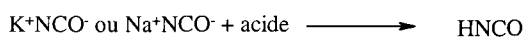


Formation in situ de l'acide cyanique

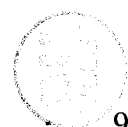
Méthode 1



Méthode 2



Conditions opératoires (Kurzer 1963 b) : Ces réactions se font généralement à des températures proches de la température ambiante, dans un mélange eau / acide acétique, l'acide isocyanique étant produit in situ. Les rendements obtenus sont très bons. L'urée disubstituée (symétrique) constitue un sous-produit classiquement retrouvé.

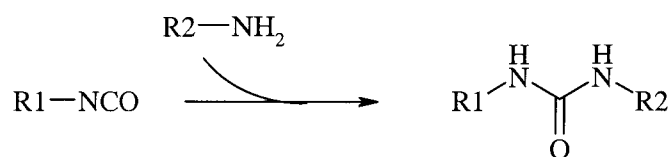


Intérêts :

- Matières premières facilement accessibles
- Les amines I^{aires} et II^{aires} peuvent être utilisées
- Forte diversité de la famille des amines utilisées

Inconvénients :

- Nécessité de produire l'HNCO in situ
- Toxicité de l'acide cyanique
- Un seul site de diversité sur la molécule finale

A-3 Réactions des amines avec les isocyanates

Conditions opératoires (French 1926, Shachat 1963, Ozaki 1972, Datta 1975) : Réactions réalisées à des températures proches de 0°C. Des amines tertiaires peuvent être introduites comme catalyseurs. Les rendements observés sont souvent très bons.

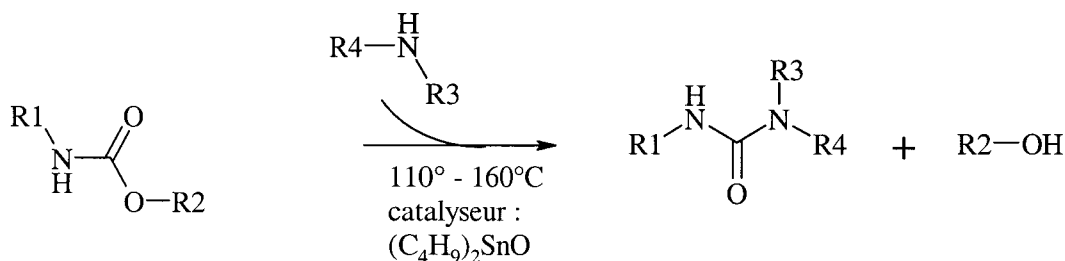
Exemple de solvant utilisé : l'éther éthylique

Intérêts :

- Amines I^{aires} et II^{aires} réactives
- Grande diversité de la famille des amines utilisées
- 2 sites de diversité sur la molécule finale (assemblage combinatoire)

Inconvénients :

- Nombre réduit d'isocyanates commercialement disponibles, ce qui limite la diversité au niveau de R1
- Instabilité des isocyanates qui nécessite d'effectuer les synthèses en atmosphère inerte afin d'éviter leur hydrolyse en acide carbamique puis en amine, elle-même susceptible de réagir sur l'isocyanate restant pour former une urée symétrique

A-4 Réactions des amines avec les carbamates

Conditions opératoires (Adams 1964) : Réactions réalisées à des températures supérieures à 100°C. L'oxyde de dibutyl étain (acide de Lewis) est utilisé comme catalyseur. Les rendements observés sont souvent très bons. Les carbamates utilisés sont essentiellement des dérivés méthoxy ou éthoxy.

Exemple de solvant utilisé : le xylène.

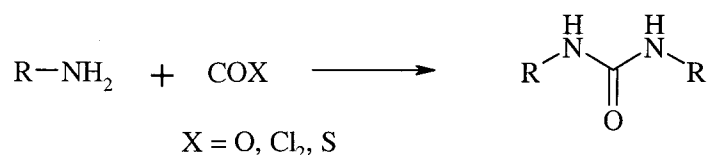
Intérêts :

- Les amines I^{aires} et II^{aires} non encombrées peuvent être utilisées
- Grande diversité de la famille des amines utilisées
- 2 sites de diversité sur la molécule finale (assemblage combinatoire)

Inconvénients :

- Faible diversité des carbamates commerciaux
- Nécessité d'utiliser des acides de Lewis comme catalyseurs (0.05 à 1% par rapport aux amines) : l'acétate cuivrique, le chlorure stannique, l'oxalate stanneux, les alcoxydes d'aluminium et les oxydes d'alkyl étain.

A-5 Réactions des amines avec les dérivés carbonylés activés

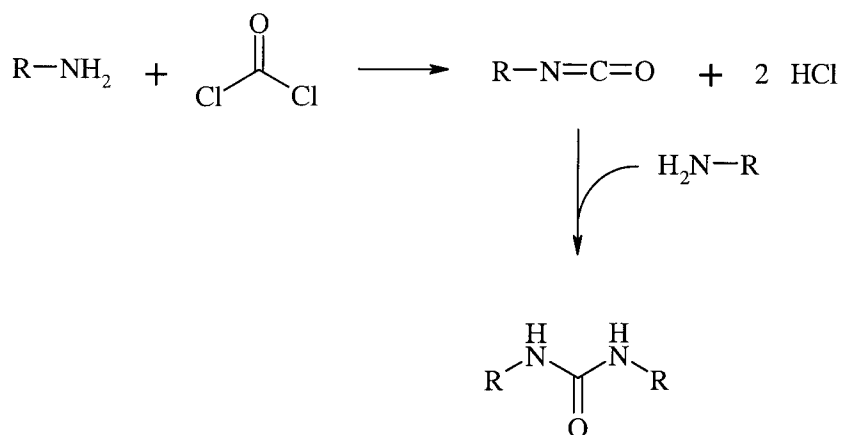


A-5-1 Utilisation du Phosgène et de ses succédanés

A-5-1-1 Utilisation du Phosgène

- avec des amines I^{aires} :

La substitution nucléophile du phosgène par les amines primaires permet, après déshydratation du chlorure de carbamoyl intermédiaire, d'accéder aux isocyanates.

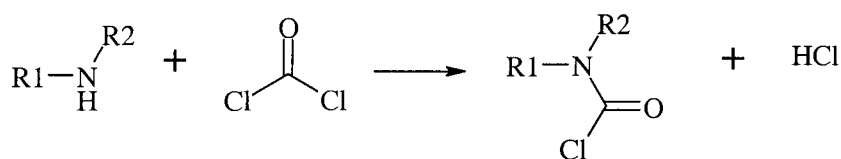


Conditions opératoires (Crain 1984): Les amines primaires réagissent avec le phosgène pour donner des isocyanates puis en chauffant le milieu réactionnel, les urées symétriques peuvent être obtenues (Farbenindustrie 1937, Slocombe 1949, Shingu 1957).

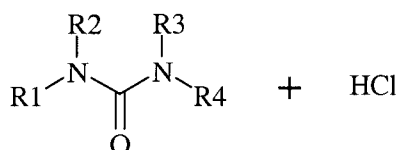
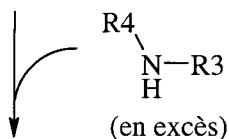
Exemples de solvants utilisés : le benzène, le toluène, le tétrachlorure de carbone et le THF.

- avec des amines II^{aires}:

La substitution nucléophile du phosgène par les amines secondaires permet d'accéder aux chlorures de carbamoyles intermédiaires qui, contrairement aux chlorures de carbamoyles d'amines primaires ne peuvent pas se déshydrater pour engendrer des isocyanates.



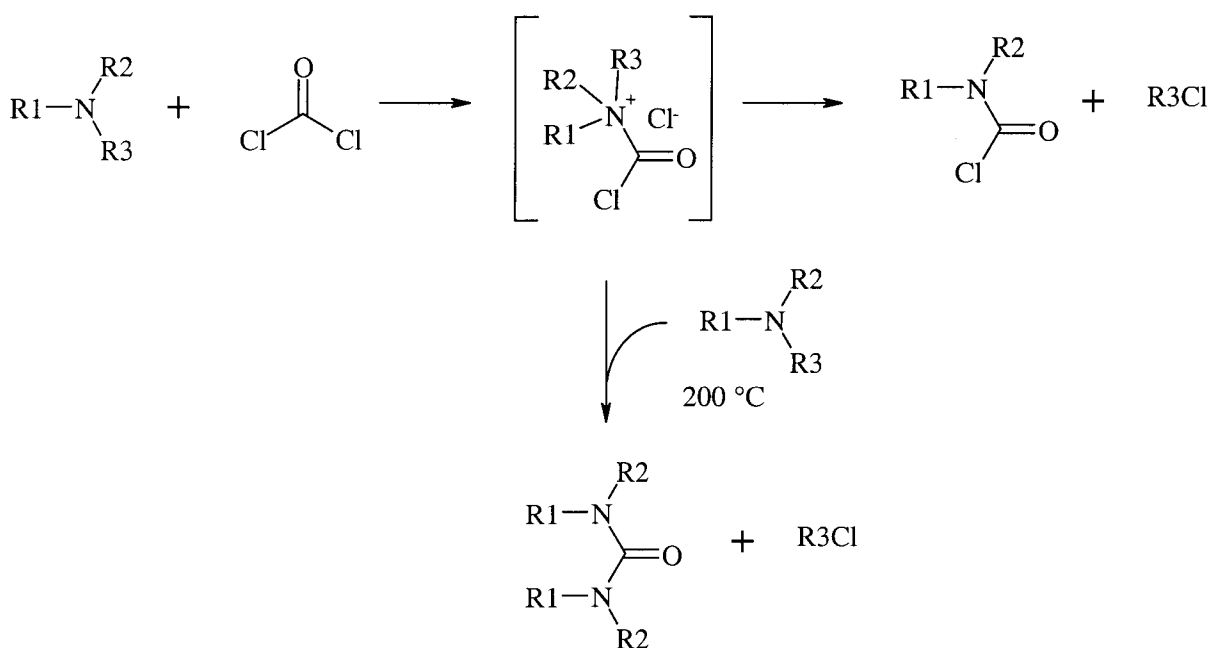
Chlorure de carbamoyle



Conditions opératoires (Abrahart 1936, Price 1926) : Les amines secondaires réagissent avec le phosgène pour donner des chlorures de carbamoyle qui conduisent à des urées asymétriques après réaction avec des amines primaires ou secondaires. Cette deuxième étape nécessite souvent de chauffer le milieu réactionnel à des températures de 30 à 80°C. Exemples de solvants utilisés : le dioxane, le benzène, l'éthanol absolu.

- avec des amines III^{aires} :

Un des trois substituants portés par l'azote doit être éliminé (par β élimination) pour permettre de créer le lien urée, les conditions opératoires font donc appel à des températures élevées.



Conditions opératoires (Lastovskii 1946) : Les amines tertiaires réagissent avec le phosgène pour donner des urées tétrasubstituées symétriques. Cette réaction pourrait impliquer la

formation d'un intermédiaire chlorure de carbamoyle car il est connu que ceux-ci réagissent avec les amines tertiaires à 200°C pour donner des urées tétrasubstituées.

Intérêts :

- Forte diversité de la famille des amines utilisées et des urées résultantes
- Accès aux urées asymétriques en faisant réagir les chlorures de carbamoyl avec différentes amines primaires ou secondaires
- Les amines tertiaires ne réagissant pas dans des conditions classiques, pourront être utilisées pour piéger des protons si nécessaire

Inconvénients :

- L'action du phosgène sur les amines primaires est surtout utilisée pour la synthèse d'urées symétriques. La synthèse d'urées asymétriques peut cependant s'effectuer en synthétisant dans un premier temps l'isocyanate, par action du phosgène sur une solution d'amine (Shriner 1944, Norwich 1992), mais du fait de l'instabilité des isocyanates, leur purification et leur conservation restent délicates. Il est aussi possible de produire in situ l'isocyanate et de le faire réagir directement avec une autre amine, cependant la contamination du produit attendu par les urées symétriques des deux amines introduites (à cause de phosgène résiduel) est souvent peu maîtrisable, et conduit à un milieu réactionnel difficile à purifier.

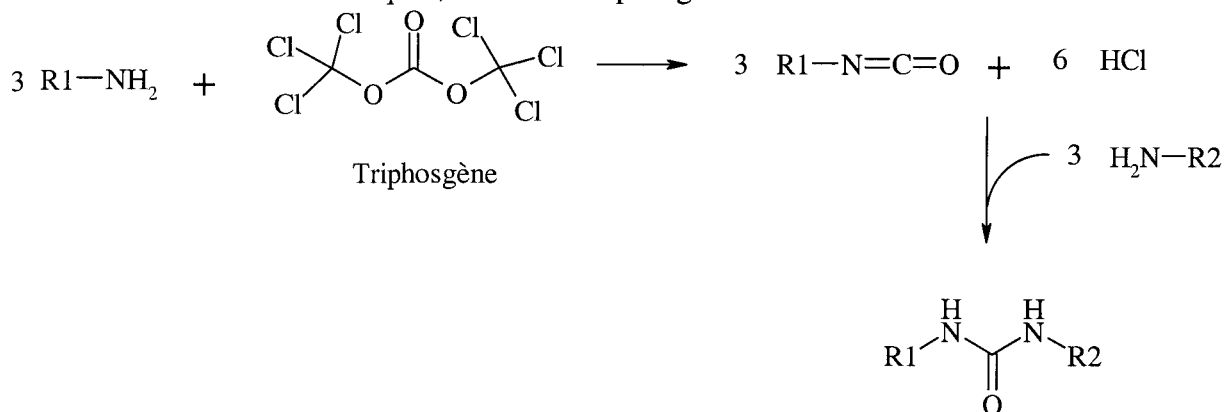
- L'utilisation des amines tertiaires dans ce type de synthèses reste délicate car d'une part, les conditions opératoires étant relativement drastiques, peu de chaînes latérales d'amines pourraient résister, d'autre part, dans le cas où deux bons groupes partants sont présents sur l'azote, la compétition entre le départ de l'un ou l'autre groupe conduit alors à un mélange de produits

- La forte toxicité du phosgène impose une ventilation efficace du lieu de synthèse et se prête mal à la robotisation

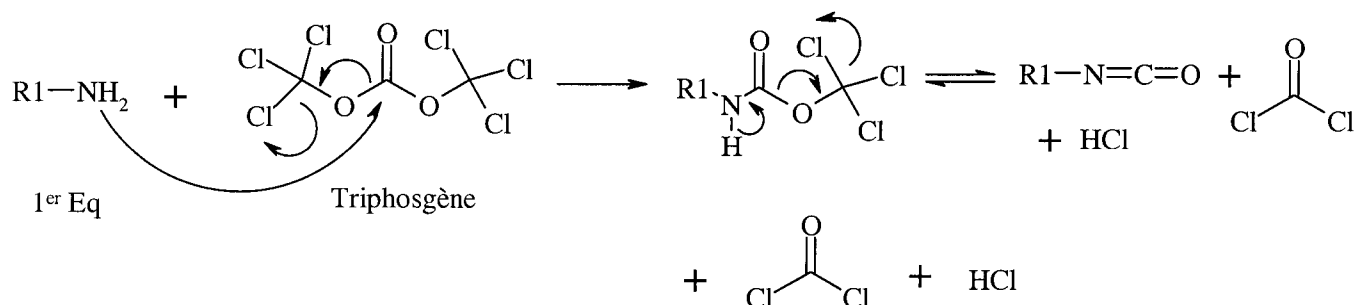
- La forte réactivité du phosgène nécessite de travailler sous atmosphère inerte

A-5-1-2 Utilisation du Triphosgène

La toxicité du phosgène a suscité la mise au point de dérivés plus facilement utilisables, étant moins volatiles et moins toxiques, comme le triphosgène.



Mécanisme de production de phosgène après la condensation d'une amine sur le triphosgène :



Conditions opératoires (Majer 1994) : Cette synthèse se fait de manière séquentielle, dans un seul réacteur (« one pot »), sous argon. Au triphosgène est ajouté progressivement un mélange amine 1 / DIEA, puis l'amine 2 est introduite. Il est important d'introduire en première position l'amine la moins réactive des deux, afin de minimiser la formation d'urée symétrique de la première amine. Solvant utilisé : le dichlorométhane

Intérêts :

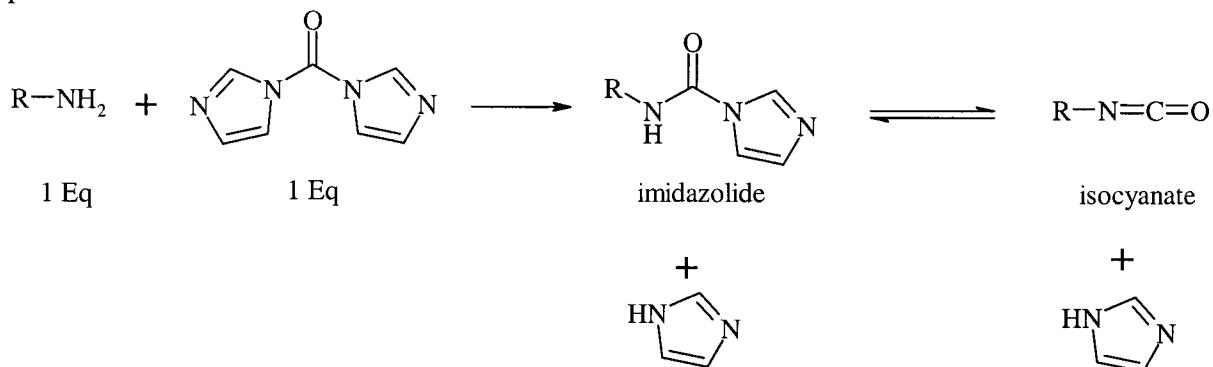
- Forte diversité de la famille des amines utilisées et des urées résultantes
- Accès aisé aux urées asymétriques
- Utilisation d'un réactif plus facilement manipulable que le phosgène
- Conditions opératoires préservant les fonctions alcools primaires et secondaires
- Les 2 urées symétriques sont largement minoritaires : 2 à 5 % du produit final (évaluation par HPLC)

Inconvénients :

- Nécessité de travailler en atmosphère inerte dans du DCM anhydre
- L'utilisation de cette réaction avec des anilines n'est pas décrite

A-5-1-3 Utilisation du Carbonyldiimidazole

L'utilisation de ce réactif est détaillée au paragraphe B car cette méthode sera retenue pour la production robotisée d'urées.



Conditions opératoires (Walsh 1990, Staab 1962): Cette synthèse se fait de manière séquentielle, dans un seul réacteur (« one pot »). A 35 μ moles de CDI sont ajoutées 34 μ moles de l'amine 1. Le milieu réactionnel est agité 1h30 à température ambiante, puis 30 μ moles de l'amine 2 sont introduites et le milieu réactionnel chauffé 18h au reflux du THF. Des réactions faisant intervenir des amines primaires aliphatiques et aromatiques avec le CDI pour accéder aux urées ont été décrites (Staab 1957). Ces réactions passent par un

intermédiaire isocyanate qui a été mis en évidence par infrarouge (bande caractéristique des isocyanates à 2250 cm^{-1}) (Staab 1957, W. Otting 1959). Solvants utilisés : solvants anhydres tels que le THF, le chloroforme ou le DMF.

Remarque : l'isocyanate ne peut se former qu'à partir d'imidazolides d'amines primaires, les imidazolides d'amines secondaires ne pouvant pas se décomposer en isocyanate par perte d'imidazole. Dans ce cas les conditions opératoires permettant de faire ces composés doivent être plus vigoureuses car cet intermédiaire est moins réactif que l'isocyanate.

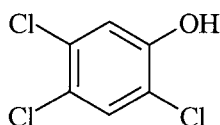
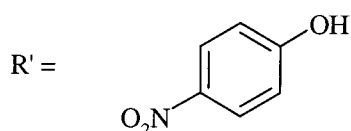
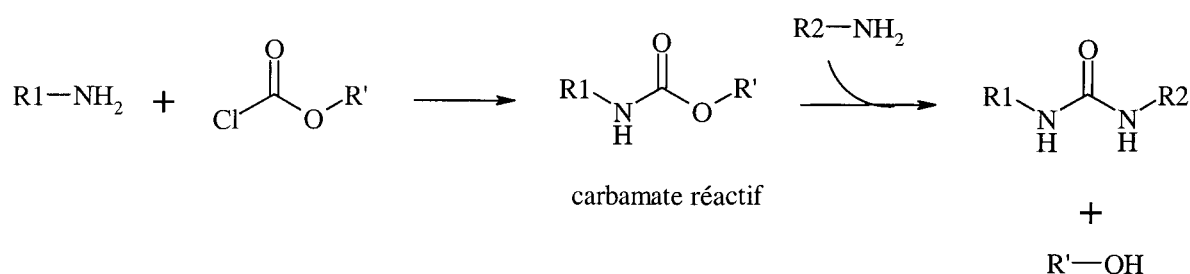
Intérêts :

- Forte diversité de la famille des amines utilisées et des urées résultantes
- Accès aisé aux urées asymétriques
- Utilisation d'un réactif plus facilement manipulable que le phosgène
- Le CDI est non toxique
- Les réactions se font en milieu neutre (sans formation d'HCl)

Inconvénient :

- Conditions particulières pour utiliser les amines secondaires

A-5-1-4 Utilisation des Chloroformates



Conditions opératoires : (Gante 1965, Lipkowski 1986) Les chloroformates seront choisis de manière à ce que, après une première attaque de l'amine substituant l'atome de chlore, la fonction ester restante ait une réactivité suffisante pour permettre le couplage avec une seconde amine. Pour cela, il sera généralement fait appel à des groupes partant de type phénol porteur de groupements électroattracteurs comme le *p*-nitrophenyl chloroformate (Gante 1965) ou le 2,4,5-trichlorophenyl chloroformate (Lipkowski 1986). Cette réaction a été décrite au paragraphe I-A-4.

Cette méthode nécessite des temps de réaction prolongés et l'utilisation d'excès de réactifs. De plus, une étape de purification de l'intermédiaire afin d'éliminer les réactifs introduits en excès ainsi que les contaminants libérés par l'activateur (nitrophenol, trichlorophenol) est indispensable.

Intérêts :

- Grande diversité de la famille des amines utilisées et des urées résultantes
- Accès aisé aux urées asymétriques

Inconvénient :

- Utilisation d'excès de réactifs
- Nécessité de purifier les milieux réactionnels

A-5-2 Utilisation du dioxyde de carbone et de ses dérivés**Utilisation du dioxyde de carbone**

Conditions opératoires : Ces réactions se font généralement à haute température et sous des pressions élevées (> 100 atm), ces conditions opératoires sont donc difficiles à mettre en œuvre sur un robot (Buckley 1951, Cook 1967, Luetzow 1979, Mitchelitsch 1973 a, b)

D'autres méthodes faisant appel au sulfure de carbonyle, le monoxyde de carbone ou à certains carbonates ont aussi été décrites.

Remarque : toutes les voies de synthèses d'urées citées au paragraphe A font appel à des réactions de condensation. Il est cependant possible d'utiliser des réactions d'oxydation comme l'oxydation d'amines en présence cyanure de sodium et de peroxyde d'hydrogène ou l'oxydation de carbazides et de semicarbazides, mais leur application pour la synthèse de chimiothèques ne peut être envisagée.

B L'utilisation du carbonyldiimidazole pour la synthèse d'urées

Les principaux critères retenus pour la synthèse parallèle robotisée d'une chimiothèque généraliste étant :

- l'obtention de milieux réactionnels propres (chemin réactionnel unique, absence de sous-produits pouvant gêner les tests, bons rendements...)
- la robotisation aisée de la chimie
- la possibilité d'engendrer un grand nombre de composés très divers (nécessité d'utiliser des monomères comme les amines, les alcools, les acides carboxyliques, les aldéhydes) dont une grande variété est disponible commercialement.

Deux méthodes parmi celles décrites dans le chapitre précédent, ont retenu notre attention : la méthode au triphosgène et la méthode au carbonyldiimidazole. En effet, celles-ci paraissent robotisables et adaptées à la synthèse d'un grand nombre de produits, les conditions opératoires permettant l'accès aux urées asymétriques par condensation de deux amines différentes sur ces succédanés du phosgène.

Une étude expérimentale comparative (cf. § II Mise au point de la chimie) a montré que le carbonyldiimidazole était mieux adapté à nos conditions de synthèse, nous avons donc approfondi notre étude bibliographique sur ce réactif.

B-1 Premières mises en évidence de la réactivité des imidazolides

Dès 1928, Bergmann et Zervas (Bergmann 1928) ont montré que les dérivés de l'histidine acylée sur un des azotes de l'imidazole, réagissaient avec d'autres aminoacides pour former des dipeptides. Ils s'interrogeaient alors sur le rôle éventuel de cette réaction dans la biosynthèse des protéines. Vingt cinq ans plus tard Wieland et Schneider (Wieland 1953) ont publié que la liaison peptidique pouvait être formée en activant un acide sous forme d'un acylimidazole au moyen du N-benzoyl-L-histidinate de méthyle, mais leur protocole n'était pas suffisamment général pour s'adapter à la synthèse de peptides. C'est grâce à la découverte du carbonyldiimidazole (CDI) par Heinz A. Staab en 1957 (Staab 1957 a,b) qu'un intérêt particulier fût porté aux azolides (amides dont l'azote fait partie d'un hétérocycle aromatique à 5 sommets, fusionné ou non avec un autre cycle aromatique, contenant au moins deux atomes d'azote, comme dans les N-acylimidazoles, N-acyltriazoles, N-acylbenzimidazoles...). C'est également à Staab que l'on doit la démonstration de la réactivité du CDI vis à vis d'amines et d'alcools pour mener à différents azolides. En 1958 l'utilisation du CDI pour la synthèse peptidique fût proposée par Anderson et Paul (Anderson 1958) dans une publication décrivant l'activation d'acides carboxyliques en N-acylimidazoles, par le CDI, suivie par la réaction avec différentes amines conduisant aux amides correspondants.

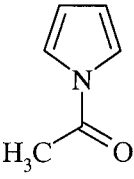
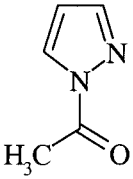
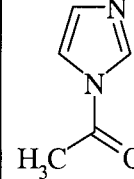
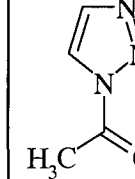
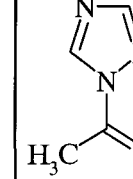
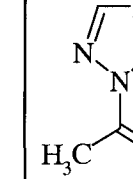
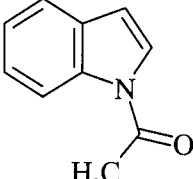
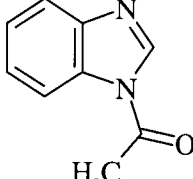
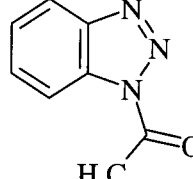
Le CDI reste actuellement utilisé, quoique de manière marginale, pour la synthèse d'amides (Ming 1988) et d'esters (Ohta 1982, Morton 1988), par l'activation d'acides, la synthèse d'urées (Walsh 1990) par l'activation d'amines, et la synthèse de carbamates par l'activation d'alcools. Les intérêts principaux de ce réactif résident dans sa faible toxicité et dans l'élimination aisée de ses sous-produits.

Staab a décrit en 1962 (Staab 1962) les propriétés de divers azolides ainsi que plusieurs voies de synthèses utilisant ces composés. Dans notre travail, nous nous sommes intéressés plus spécialement aux réactions de transacylation conduisant aux urées et aux carbamates

B-2 Réactivité des azolides

- Réactivités comparées de différents azolides

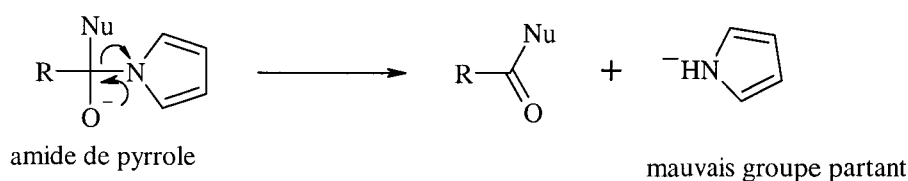
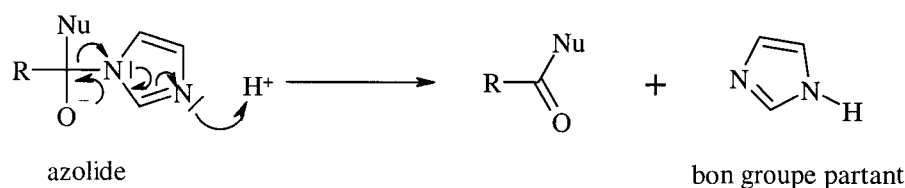
Les amides sont normalement caractérisés par une faible réactivité vis à vis des nucléophiles (réactions d'hydrolyse, d'alcoolyse...), pourtant les azolides possèdent une réactivité proche des halogénures d'acides ou des anhydrides. Celle-ci est illustrée par les cinétiques d'hydrolyse comparée de différents acétamides figurant dans le tableau ci-dessous.

						
	I	II	III	IV	V	VI
$k' \cdot 10^5 \text{ (s}^{-1}\text{)}$	$\rightarrow 0$	1,27	28,1	43,5	180	> 2000
$\tau_{1/2} \text{ (min)}$	$\rightarrow \text{infini}$	908	41	26,6	6,4	< 0,5
						
	VII		VIII		IX	
$k' \cdot 10^5 \text{ (s}^{-1}\text{)}$	$\rightarrow 0$		1,75		10,1	
$\tau_{1/2} \text{ (min)}$	$\rightarrow \text{infini}$		760		115	

Hydrolyse de différents acétamides placés dans l'eau à pH 7, à 25°C (Staab 1962)

La forte réactivité des azolides ainsi que les différences de réactivité observées entre les azolides sont en relation avec le caractère aromatique de ces composés, plus ou moins appauvris en électrons au niveau des différents cycles azotés.

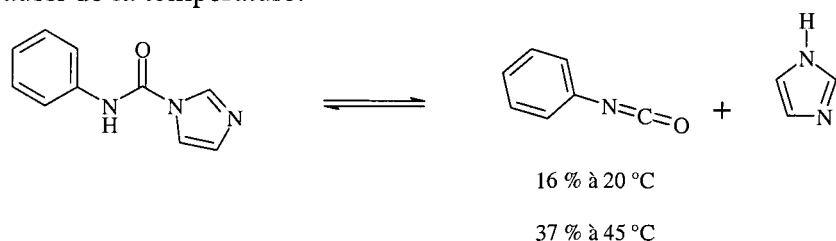
Contrairement au N-acétyl-pyrrole ou au N-acétyl-indole, les N-acétyl-azoles possèdent au moins deux azotes dans le cycle aromatique. Le remplacement d'un carbone par un azote (plus électronégatif), contribue à l'augmentation de l'attraction des électrons par le cycle aromatique et en conséquence à l'augmentation de la polarisation de la liaison N-C exocyclique. Mais le principal élément à prendre en compte est la capacité des azoles à s'éliminer de l'intermédiaire à 4 centres (formé après l'attaque d'un nucléophile) grâce à la protonation relais d'un atome d'azote, qui permet d'accélérer les réactions de substitutions nucléophiles.



Comparaison de la stabilité des groupes partants impliqués lors d'attaques nucléophiles d'azolides ou d'amides du pyrrole

Afin d'éviter toute formation d'urée symétrique à partir de l'imidazolide ou de l'isocyanate avec un reste d'amine n'ayant pas réagi, il est important d'ajouter l'amine en quantité stoechiométrique par rapport au CDI.

La dissociation du dérivé N-imidazolide en isocyanate et imidazole est fonction de l'amine impliquée dans la réaction (l'isocyanate correspondant étant plus ou moins favorisé), mais aussi de la température.

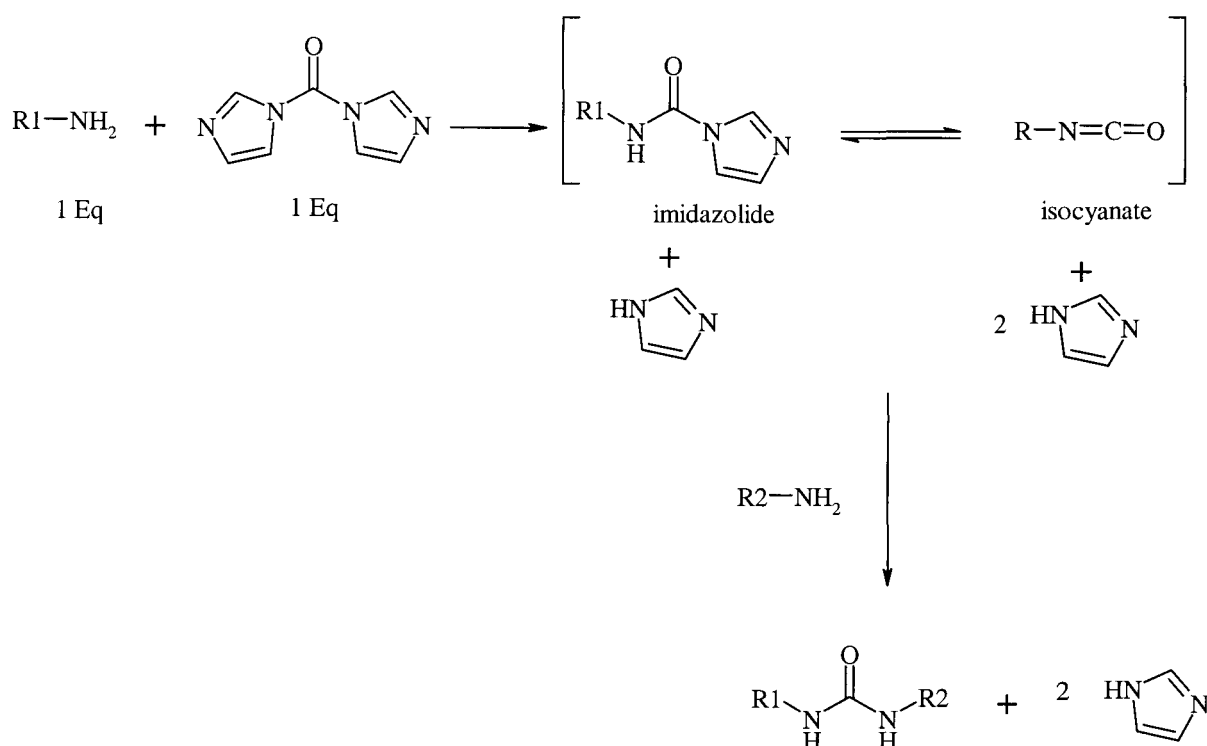


Cette dissociation permet donc la synthèse d'isocyanates d'après le protocole suivant (Staab 1959) :

- 1 mole d'amine est ajoutée goutte à goutte à une mole de CDI en solution (le THF étant le solvant le plus utilisé)

- la distillation du milieu réactionnel permet d'isoler l'isocyanate

La synthèse d'urées asymétriques peut s'effectuer en ajoutant, directement au milieu réactionnel, une seconde amine.

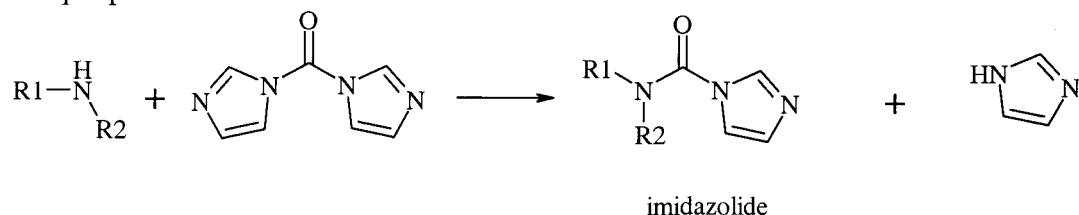


Les imidazolides, synthétisés à partir d'amines primaires peuvent être considérés comme des isocyanates masqués, du fait de leur tendance à la dissociation. En solution leur réactivité est similaire à celle des isocyanates, cependant, à l'état solide, les imidazolides sont des composés cristallins, stables et beaucoup moins toxiques que les isocyanates.

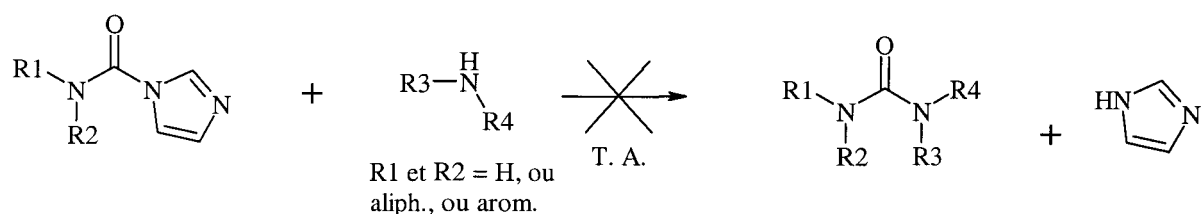
Remarque : les rendements en urées asymétriques (synthétisées grâce à un protocole similaire à celui ci-dessus) décrits dans une publication récente (Walsh 1990), restent cependant modérés (12 à 67%).

- Cas des amines secondaires

Une monosubstitution du CDI a lieu avec les amines secondaires conduisant aux imidazolides. A température ambiante, ceux-ci sont stables et ne subissent pas une seconde attaque par d'autres amines.



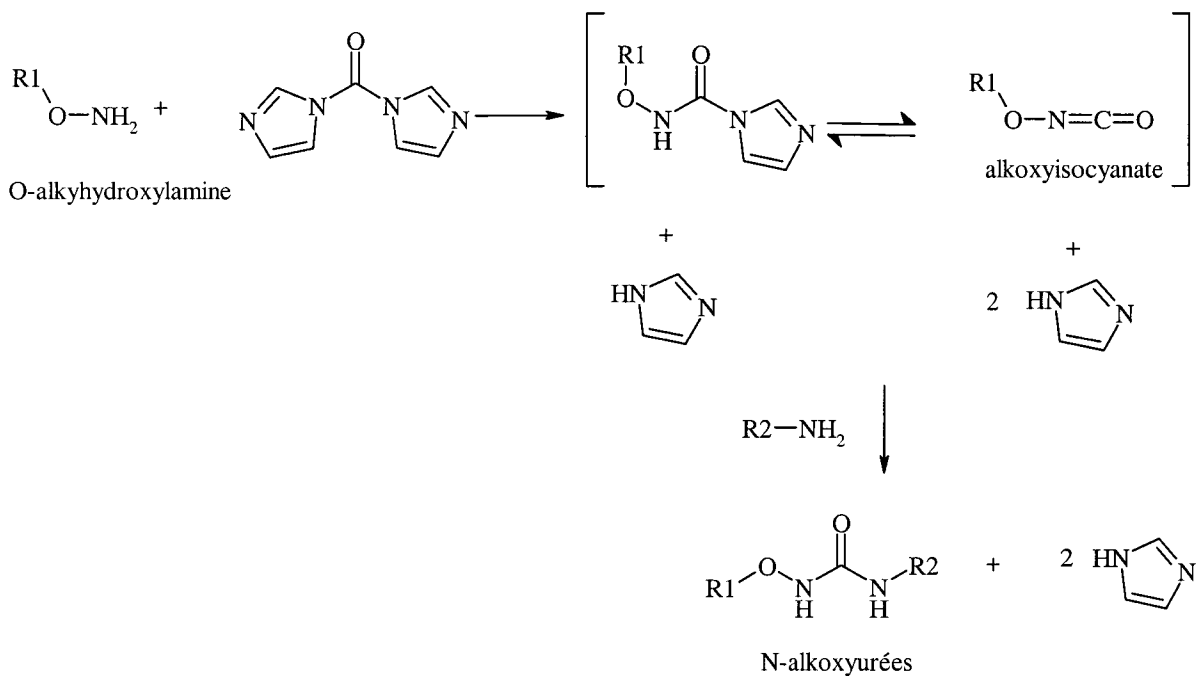
R1 et R2 ≠ H
 R1 et R2 = aliph. ou arom.



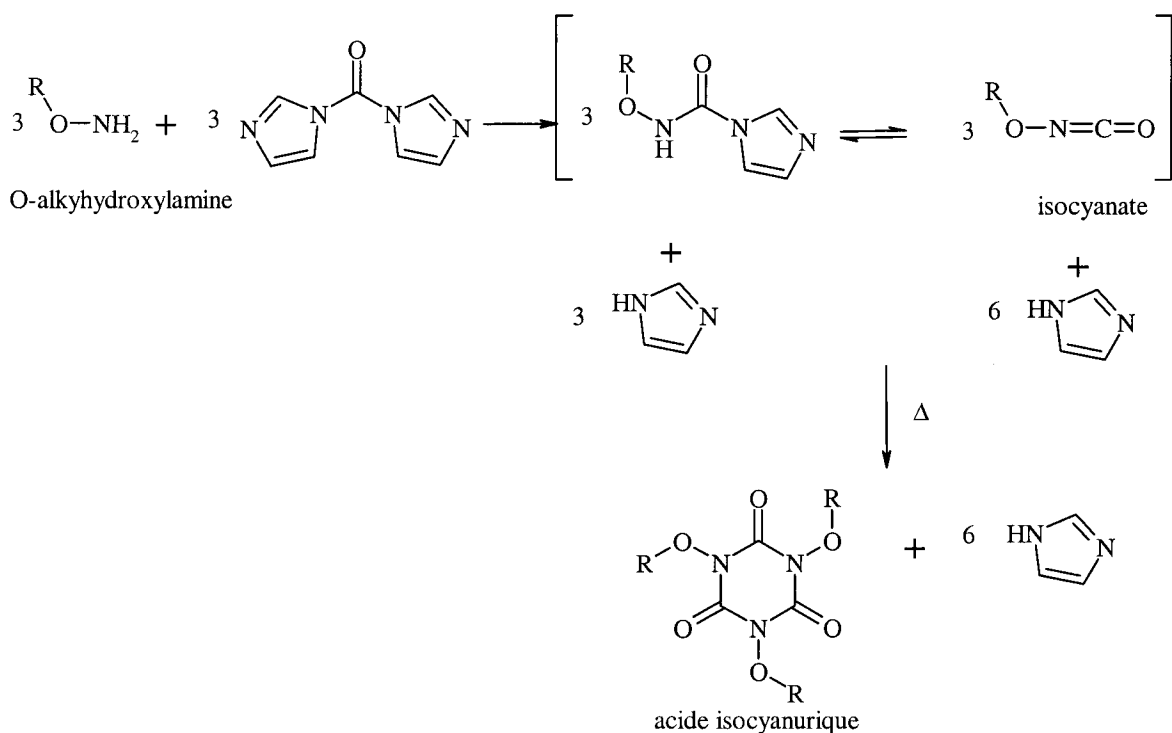
En effet, l'absence d'hydrogène sur l'azote de l'amine secondaire engagé dans l'imidazole, empêche toute formation d'isocyanate par départ d'imidazole. L'attaque de la seconde amine doit se faire sur le CO de l'imidazole. Cette réaction est plus difficile que l'attaque de l'isocyanate, du fait d'un encombrement stérique accru ainsi que de la diminution de l'électrophilie du CO. Ce phénomène confirme le fait qu'à température ambiante, l'intermédiaire isocyanate est l'intermédiaire réactif lors de la formation des urées au CDI (cf. paragraphe précédent).

- Cas des hydroxylamines (Staab 1962)

En présence de CDI, les hydroxylamines conduisent à un intermédiaire du type N-carboxamide-imidazole pouvant subir une aminolyse pour donner les N-alkoxyurées, en quelques minutes à température ambiante. Cette réaction semble également faire intervenir un intermédiaire de type alkoxyisocyanate car si des hydroxylamines N-substituées sont utilisées, l'aminolyse ne se produit plus.



Dans ce cas, l'intermédiaire alkoxyisocyanate semble particulièrement réactif car si après l'introduction de l'hydroxylamine sur le CDI, le milieu réactionnel est chauffé, une trimérisation à lieu et constitue une voie d'accès aux acides trialkoxyisocyanuriques.



L'accès aux N-alkoxyurées est possible en faisant réagir l'intermédiaire type N-carboxamide imidazole avec une amine. Cette réaction ferait aussi intervenir la formation d'un isocyanate, car comme précédemment, si l'atome d'azote de l'O-alkylhydroxylamine utilisée (à la première étape) est secondaire, ni la formation de l'urée, ni la formation de l'acide isocyanurique ne sont réalisables, du fait de l'impossible accès à un isocyanate.

II Mise au point de la synthèse robotisée d'urées

A la suite de l'étude bibliographique précédente, peu de méthodes nous sont apparues utilisables pour la préparation d'une chimiothèque généraliste d'urées asymétriques. Le choix devait se faire en fonction de **deux contraintes principales**:

- la première d'ordre **méthodologique** : permettre la synthèse rapide d'un grand nombre de produits et réaliser ces synthèses au moyen d'un robot. Idéalement la réaction doit se faire à température ambiante, sous pression atmosphérique, en présence d'air (pouvant être climatisé et partiellement séché). Quoique possible techniquement, toute contrainte supplémentaire, une étape de chauffage par exemple, se traduit par un ralentissement sensible des cadences de production.

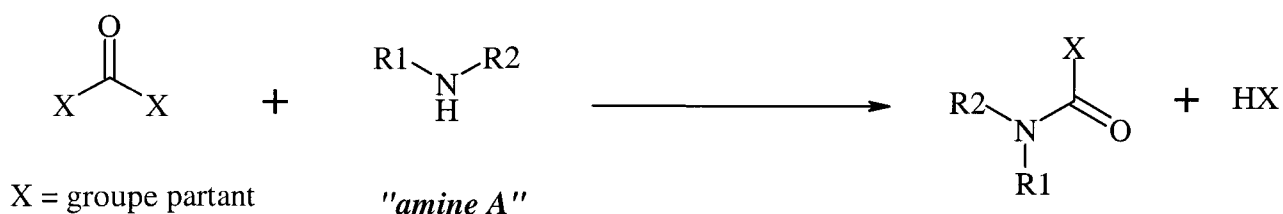
- la seconde d'ordre **structural** : la méthode choisie doit permettre de recruter le plus grand nombre de monomères très divers, et ce à chacune des étapes combinatoires.

Ces contraintes nous ont amenés à écarter les méthodes suivantes :

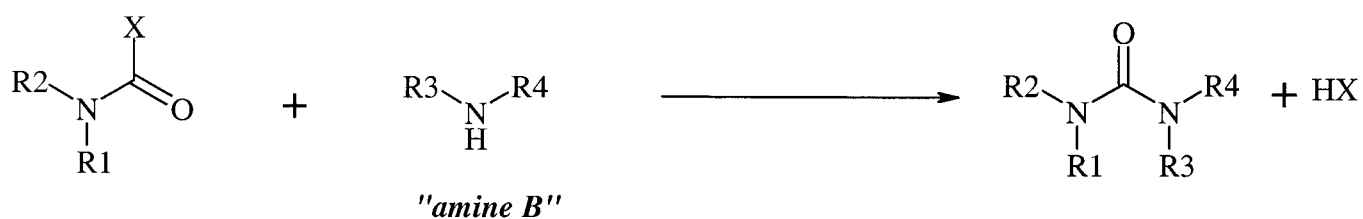
- la condensation d'amines sur les isocyanates commerciaux, de diversité trop limitée
- la préparation et l'isolement d'un intermédiaire réactionnel, qui constitue une difficulté robotique (une méthode alternative peut être l'utilisation de la phase solide)
- l'utilisation de solvants ou réactifs difficiles à manipuler par un robot, en particulier les réactifs dangereux
- l'utilisation de solvants imposant des restrictions trop importantes en termes de solubilité des monomères
- des conditions opératoires pouvant ralentir les cadences de production (chauffage, filtration...)

Deux réactifs ont retenu notre attention : le triphosgène et le carbonyldiimidazole. Ces deux réactifs de couplage symétriques, permettent d'introduire, en deux étapes, deux amines différentes. Par souci de simplification, dans la suite du document, il sera fait mention d'*amine A* ou d'*amine B*, pour désigner l'amine introduite lors de la première étape (A) ou l'amine introduite lors de la seconde étape (B) (cf. schéma ci-dessous)

1^{ère} étape



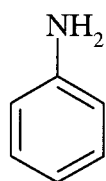
2^{nde} étape



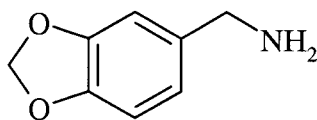
A Etude comparative du triphosgène et du carbonyldiimidazole

Sélection des urées à synthétiser pour valider les méthodes de synthèse

Nous avons choisi d'utiliser comme monomères modèles l'aniline (**1**) et la pipéronylamine (**2**) pour synthétiser l'urée asymétrique de référence : la N-Phényl-N'-(3,4-méthylènedioxybenzyl)urée (**3**), ainsi que les deux urées symétriques correspondantes : la N, N'-Bis(phényl)urée (**4**) et la N, N'-Bis(3,4-méthylènedioxybenzyl)urée (**5**).



aniline (**1**)



pipéronylamine (**2**)

Ces deux amines ont été choisies sur la base des critères suivants :

- évaluation de 2 classes de réactivité d'amines primaires : l'aniline, amine aromatique et la pipéronylamine, amine aliphatique
- possibilité de suivre la réaction par CCM et HPLC des monomères et des produits (détection par UV)
- possibilité d'analyser les 3 produits principaux par RMN, leurs protons respectifs ayant des déplacements chimiques bien différenciés.

A-1 Utilisation du triphosgène

A-1-1 Mise au point du protocole expérimental

Nous avons tout d'abord évalué les possibilités offertes par le triphosgène en adaptant la méthode proposée par Majer et al (Majer 1994).

Mode opératoire de référence :

Le triphosgène (**6**) (0,37 mmole, 110 mg) est dissous dans le CH_2Cl_2 (2 mL). Un mélange du chlorhydrate de l'amine A (1 mmole) et de DIEA (2,2 mmole) dans le CH_2Cl_2 (3,5 mL) est lentement ajouté avec une pompe seringue, sur 30 minutes, à la solution de triphosgène placée sous agitation. Le milieu réactionnel est ensuite agité 5 minutes puis un mélange du chlorhydrate de l'amine B (1 mmole) et de DIEA (2,2 mmole) dans le CH_2Cl_2 (2 mL) est ajouté en une seule fois. Le milieu réactionnel est agité 10 minutes, évaporé à sec, repris dans l'acétate d'éthyle, lavé avec une solution aqueuse à 10 % de KHSO_4 , puis avec une solution aqueuse à 5 % de NaHCO_3 , puis avec une solution aqueuse saturée en NaCl . La phase organique est alors séchée sur MgSO_4 puis évaporée. Les rendements en produits isolés obtenus pour les 7 urées décrites sont tous supérieurs à 85 %, avec une contamination en urées symétriques de 2 à 5 % (évaluée par HPLC).

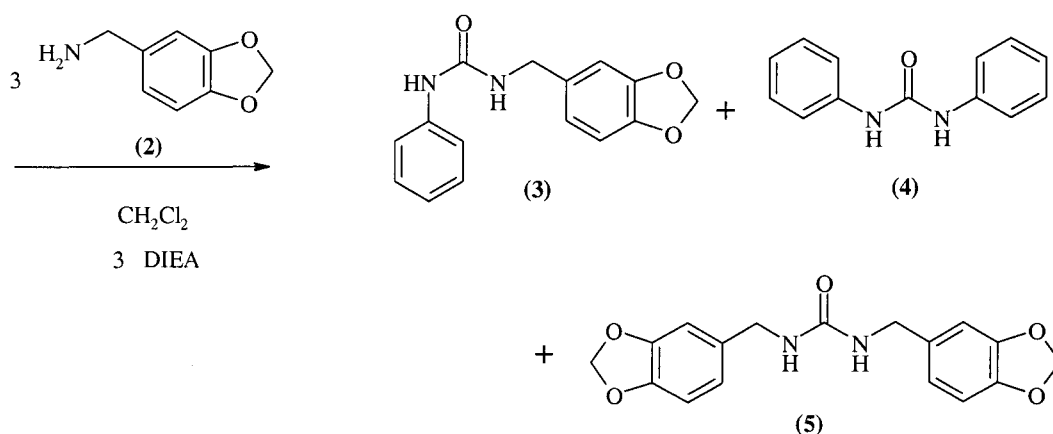
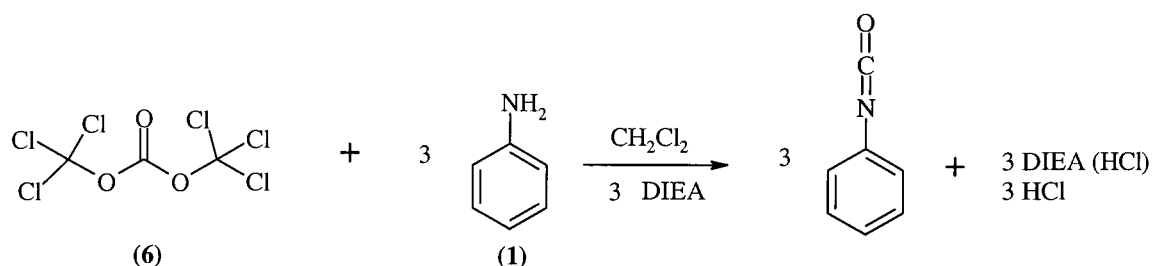
Conditions générales figurant dans la publication : les solvants utilisés doivent être anhydres et la synthèse réalisée sous argon.

A-1-2 Evaluation de l'importance des conditions anhydres

En nous inspirant du protocole expérimental décrit, nous avons synthétisé l'urée asymétrique (3) issue de l'aniline (1) et de la pipéronylamine (2) puis les deux urées symétriques correspondantes : la N,N'-Bis(phényl)urée (4) et la N,N'-Bis(3,4-méthylènedioxybenzyl)urée (5). La distribution de l'amine A s'est faite en une demi heure à raison de 7 ajouts de 500 μ L goutte à goutte, à 5 minutes d'intervalle. Le DCM utilisé n'avait pas été séché au préalable et la synthèse s'est faite dans les conditions ambiantes. L'utilisation de l'aniline nous permettait également de valider ce protocole pour la synthèse d'urées issues d'amines aromatiques, non décrites dans la publication de Majer et al. (1994).

- Préparation de l'urée asymétrique Aniline-Pipéronylamine

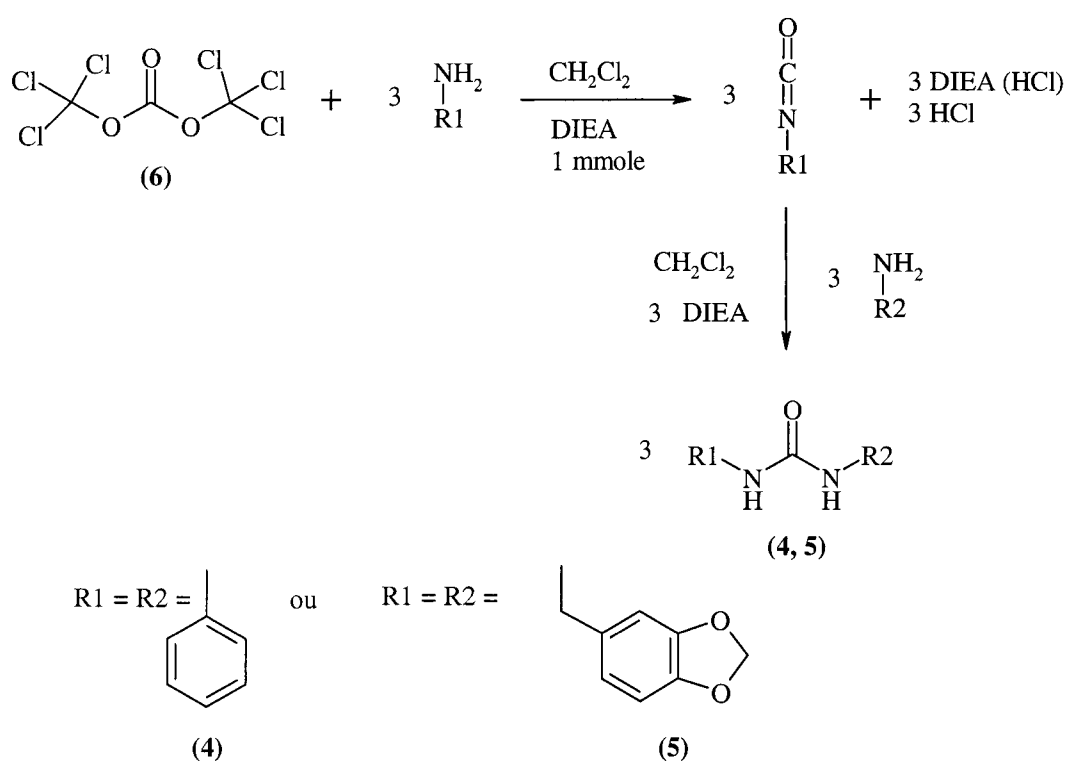
En suivant les observations effectuées par Majer, l'aniline a été introduite lors de la première étape car sa nucléophilie étant plus faible que celle de la pipéronylamine, le risque de formation d'urée symétrique (4) était dans ce cas diminué.



Résultats :

L'utilisation du protocole décrit par Majer et al. (Majer 1994), mais sans maintenir une atmosphère inerte durant la réaction nous a permis d'obtenir l'urée asymétrique attendue avec un rendement de 80 % en produit isolé. Cependant deux pics contaminants étaient détectés sur le chromatogramme de l'urée (3) purifiée par extractions. Ceux-ci, situés de part et d'autre du produit attendu, correspondaient aux urées symétriques (4) et (5). Cette contamination s'élevait à 6 % pour le composé (4) et 1 % pour le composé (5). Ces résultats étaient comparables aux données de la publication.

- Préparation des 2 urées symétriques Aniline-Aniline, Piperonylamine-Piperonylamine



Résultats :

- La formation de l'urée symétrique de l'aniline a montré un possible accès aux diarylurées grâce au protocole de Majer et al. (1994).
- Dans les deux cas, des restes d'amines de départ ont été détectés dans les milieux réactionnels avant leur purification. La conversion incomplète de la piperonylamine en urée symétrique correspondante s'expliquerait surtout par une hydrolyse partielle du triphosgène

(un défaut de nucléophilie ne pouvant être évoqué dans le cas de la pipéronylamine contrairement à l'aniline).

Conclusion du paragraphe A-1-2 :

Nous avons donc montré que le maintien de l'atmosphère inerte, recommandé dans le protocole de Majer et al. (1994), pour la formation d'urées asymétriques en utilisant le triphosgène n'était pas indispensable. En effet, il nous a été possible d'obtenir l'urée asymétrique désirée (**3**) en présence d'air ambiant. Le produit attendu n'était que faiblement contaminé par les urées symétriques (**4**) et (**5**). Cependant la conversion des amines de départ n'étant pas totale, nous avons supposé que le triphosgène s'hydrolysait partiellement. L'introduction d'un excès de triphosgène, dès le début de la réaction, devrait permettre de compenser cette hydrolyse.

D'autre part, en préparant les urées (**3**) et (**4**) à partir de l'aniline, nous avons étendu le protocole Majer et al. (1994) à la synthèse d'urées issues d'amines aromatiques.

Il nous fallait ensuite modifier le protocole de Majer et al. pour adapter au robot l'addition de l'amine A au triphosgène, celle-ci devant idéalement s'effectuer en une distribution rapide et non plus sur 30 minutes.

A-1-3 Evaluation de l'influence de la vitesse d'introduction de la première amine

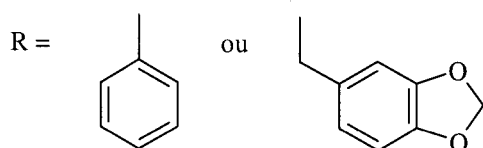
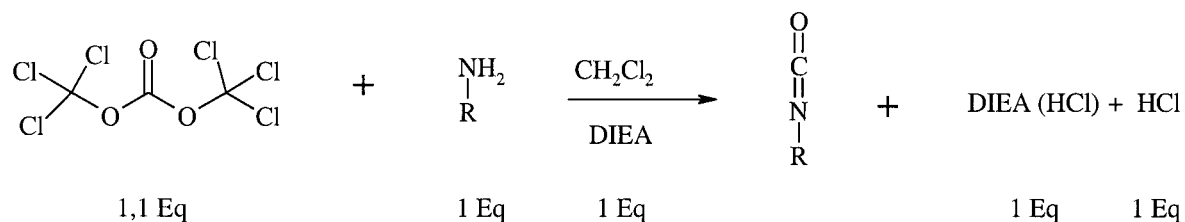
La distribution des réactifs par le robot est difficile à mettre en œuvre en utilisant le protocole de Majer et al.. En effet, pour la synthèse des 6 μ moles d'urée asymétrique que constituent notre échantillon standard, la distribution en 30 minutes de quelques dizaines de microlitres (60 μ L pour des solutions 0,1 M) pose les problèmes suivants :

- la distribution continue de l'amine A durant 30 min pose le problème du réglage de la vitesse des pistons et se traduit par un ralentissement des cadences de production
- la distribution de l'amine A en plusieurs ajouts successifs, répartis sur 30 min implique des volumes trop faibles compte tenu de la précision possible sur le robot (erreur de $\pm 10\%$ pour 10 μ L) et ici également, un ralentissement des cadences de production.

Nous avons donc étudié la possibilité d'introduire en une seule fois la première amine. La formation de l'isocyanate intermédiaire après l'addition de l'amine A sur le triphosgène a été

estimée pour deux amines : l'aniline et la pipéronylamine. Un équivalent de chaque amine a ainsi été ajouté à 1,1 équivalent de triphosgène. Les milieux réactionnels ont ensuite été analysés par CCM toutes les 5 minutes la première demi-heure de réaction puis une CCM a été effectuée après 1h30 puis 3h de réaction.

Schéma réactionnel :

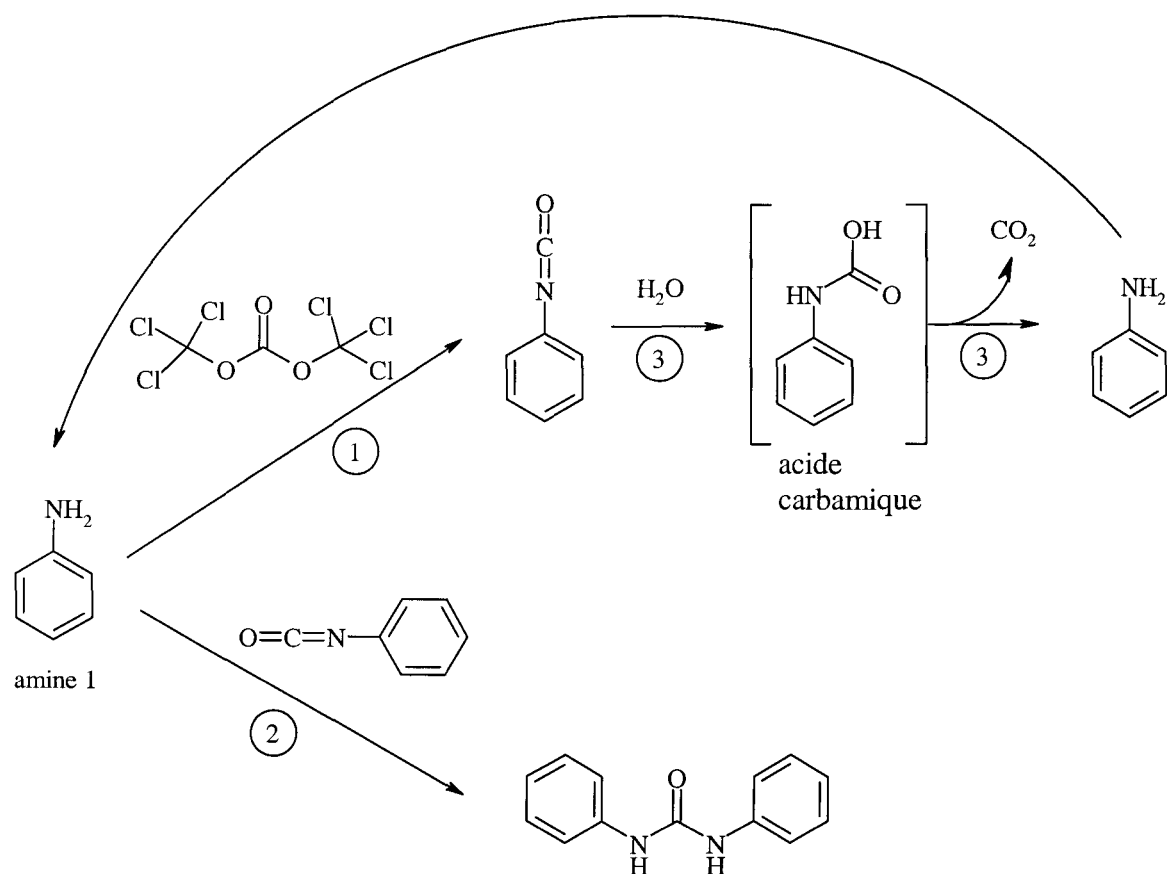


Résultats :

Nos analyses ont montré que 5 minutes après l'introduction de l'amine A, celle-ci était consommée pour conduire à l'urée symétrique correspondante. Cette réaction était très rapide car les analyses du milieu réactionnel effectuées toutes les 5 minutes, la première demi-heure de réaction, puis après 1h30 et 3h de réaction ne révélèrent aucune évolution dans le temps. L'introduction rapide de la première amine favorisait donc très nettement la formation de l'urée symétrique correspondante.

Interprétation :

La vitesse d'introduction de la première amine peut influencer sur les 3 réactions suivantes :



- ① formation de l'isocyanate
- ② formation de l'urée symétrique
- ③ hydrolyse de l'isocyanate

L'amine introduite lors de la première étape peut réagir avec deux électrophiles : le triphosgène présent initialement ou l'isocyanate qui se forme progressivement.

La réaction recherchée est la formation de l'isocyanate (réaction 1), celui-ci devant réagir ultérieurement avec l'amine B pour former l'urée. Afin de favoriser la réaction 1, Majer et al. (1994) ont ajouté l'amine A par petites fractions à la solution de triphosgène, ce qui permettait à chaque molécule d'amine d'être en présence d'un excès de triphosgène. Cet excès de triphosgène diminuait progressivement au cours des ajouts successifs mais le triphosgène ayant été introduit en excès (+10%) et la réaction se faisant sous atmosphère inerte, le triphosgène restait en excès par rapport à l'amine A jusqu'au dernier ajout d'amine. La probabilité de rencontre entre une molécule d'amine et une molécule de triphosgène était supérieure à la probabilité de rencontre entre une molécule d'amine et une molécule d'isocyanate jusqu'à ce que 60 % de la solution d'amine soit ajoutée. Cependant si la réaction

1 est cinétiquement favorisée par rapport à la réaction 2, **la formation de l'isocyanate reste prépondérante.**

Dans notre cas, l'amine A ayant été introduite en une seule fois, la probabilité de rencontre entre une molécule d'amine et une molécule de triphosgène devenait dépendante de l'agitation. Si celle-ci ne permettait pas aux molécules d'amines de rencontrer suffisamment rapidement le triphosgène, le seul électrophile proche de l'amine devenait l'isocyanate. La réaction 2 étant alors favorisée, **l'urée symétrique devient le produit majoritaire.**

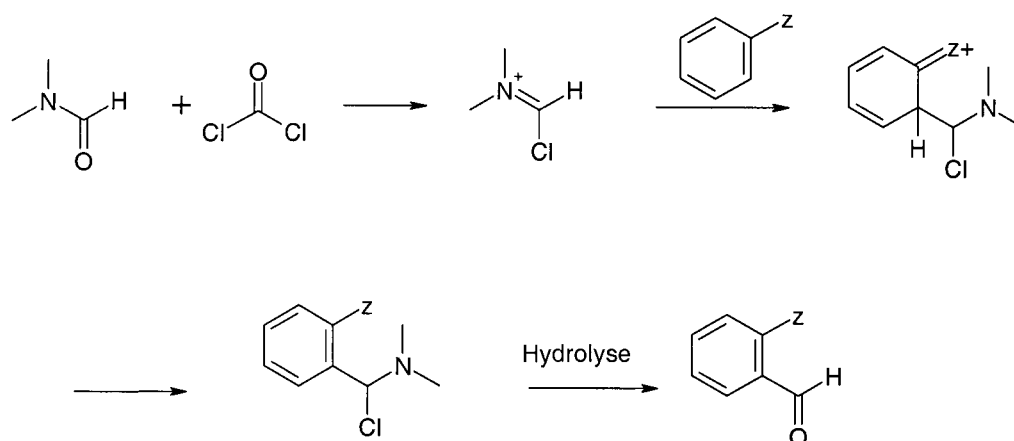
Conclusion du paragraphe A-1-3 :

La distribution rapide de la solution de l'amine A sur le triphosgène s'est traduite par la formation d'une quantité importante de l'urée symétrique de l'amine A. Ce protocole de synthèse n'était donc pas adapté à la synthèse parallèle d'urée.

A-1-4 Utilisation du DMF comme solvant de réaction

Le DMF étant un excellent solvant des petites molécules organiques, et tout particulièrement des amines, nous avons tenté de reproduire le mode opératoire de Majer et al. (1994) mais en substituant le DCM par le DMF. Une réaction très violente, apparentée à la réaction de Vielsmeier, ayant été observée lors de la tentative de mise en solution du triphosgène dans le DMF, a compromis l'utilisation de ce solvant pour la synthèse d'urées au triphosgène.

Réaction de Vielsmeier :



Conclusion du paragraphe A-1 :

Cette étude concernant l'utilisation du triphosgène comme agent de couplage d'amines pour la synthèse d'urées asymétriques a mis en évidence les deux contraintes suivantes :

- 1- la nécessité de l'ajout progressif de la première amine sur le triphosgène afin de limiter la formation de l'urée symétrique de la première amine
- 2- la nécessité d'un solvant de réaction ne dégradant pas le triphosgène tout en permettant la mise en solution du plus grand nombre d'amines.

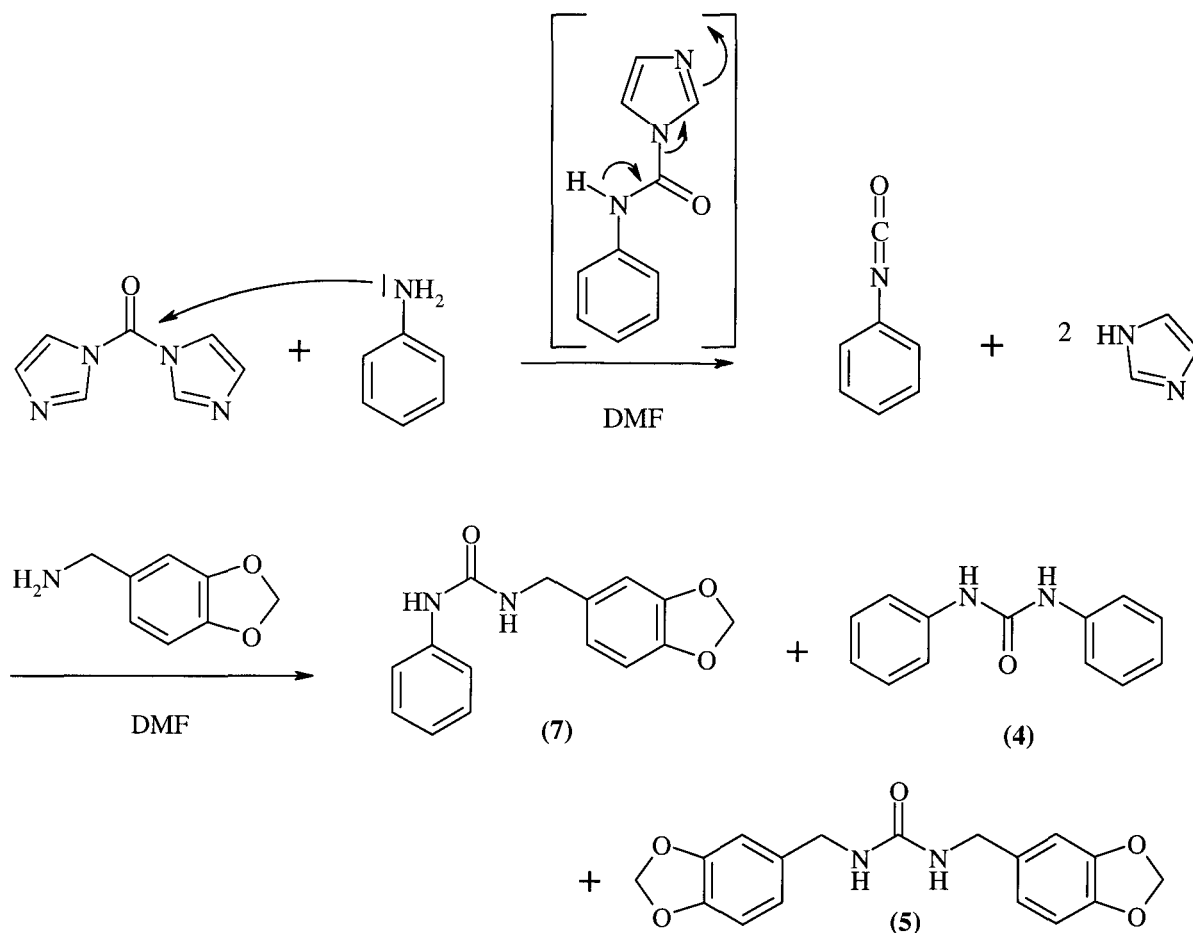
Avant de poursuivre nos recherches sur ce réactif, nous avons préféré évaluer le carbonyldiimidazole afin d'orienter nos travaux vers la voie de synthèse la plus facilement transposable au robot.

A-2 Utilisation du carbonyldiimidazole

A-2-1 Evaluation de l'importance des conditions anhydres

Comme lors de l'évaluation du triphosgène pour la synthèse d'urées, nous avons d'abord synthétisé l'urée asymétrique issue de l'aniline et de la pipéronylamine. Nous avons décidé d'utiliser, dès le départ, le DMF comme unique solvant de réaction compte tenu de nos connaissances sur la mise en solution des amines. La réaction ne s'accompagnant pas de formation d'acide, aucun ajout de DIEA n'a été nécessaire. La distribution de l'amine A (0,5 M) à la solution de CDI (1,25 M), placée sous agitation, s'est faite sur une demi heure en distribuant en goutte à goutte, 7 aliquotes de la solution, à 5 minutes d'intervalle. Le DMF utilisé n'a pas été déshydraté et la synthèse s'est faite dans les conditions ambiantes. Une fois la première amine ajoutée, le milieu réactionnel a été laissé sous agitation pendant 5 minutes puis la solution de l'amine B (0,5 M) a été ajoutée en une seule fois. Après 4 heures de réaction, le milieu réactionnel a été contrôlé par HPLC et le DMF a été évaporé.

Schéma réactionnel :



Résultats :

Le milieu réactionnel analysé par CCM à différent temps a montré que 10 minutes après l'introduction de la deuxième amine, la réaction était terminée. L'HPLC du milieu réactionnel après 4 heures de réaction nous a permis d'estimer un taux de conversion des amines de départ en urées, supérieur à 90% (conversion déduite de la surface des pics d'amines résiduelles). L'urée asymétrique (7) constituait le produit principal de la réaction, la contamination en urée symétrique de l'aniline (4) ne représentant que 10 % du pic de l'urée (7).

Conclusion du paragraphe A-2-1 :

Contrairement au triphosgène, le CDI s'est montré compatible avec le DMF, les amines de départ étant presque totalement consommées pour former l'urée attendue. La meilleure

stabilité du CDI nous a permis d'obtenir l'urée asymétrique attendue en utilisant le DMF comme solvant de réaction et dans des conditions ambiantes. Ce protocole autorisait la solubilisation d'un plus grand nombre d'amines que dans le cas du protocole avec le triphosgène.

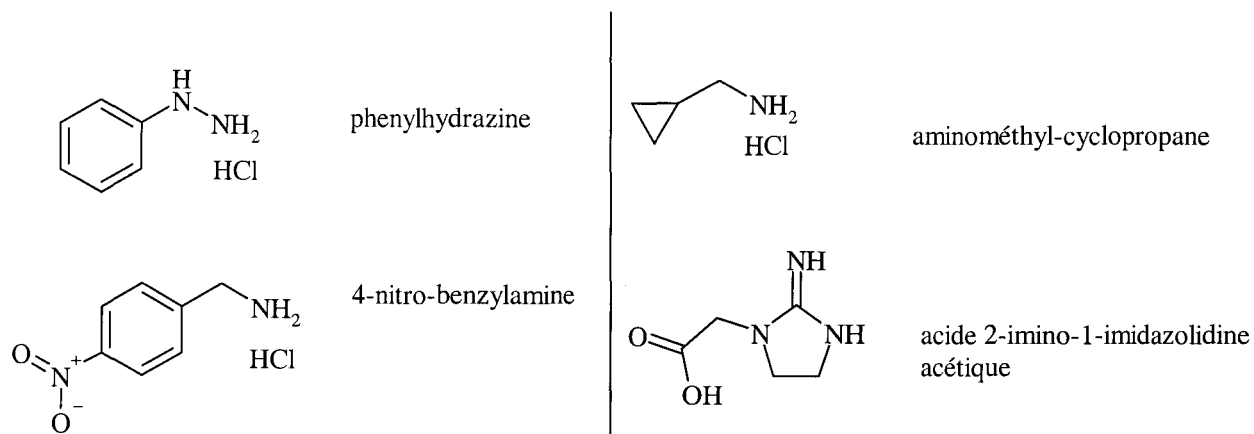
Nous avons ensuite voulu optimiser les vitesses d'introduction des réactifs ainsi que les temps de réaction.

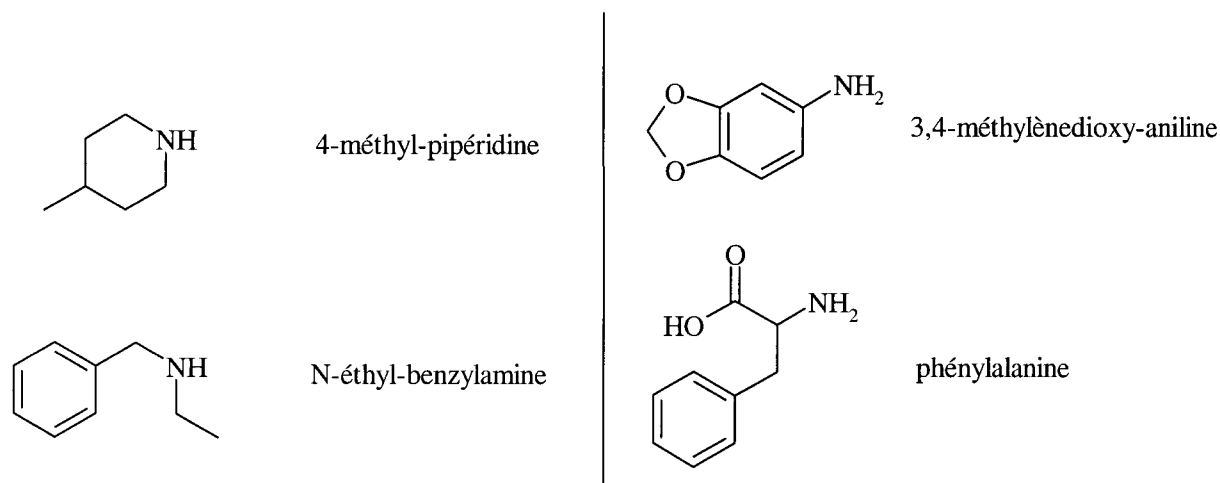
A-2-2 Optimisation des temps de réaction et des vitesses d'introduction des réactifs

A-2-2-1 Mise au point de l'étape d'addition de l'amine A sur le carbonyldiimidazole

Nous avons tout d'abord étudié l'influence d'un temps de contact prolongé (4h30) entre l'amine A et le CDI au cours de la préparation de 8 urées asymétriques correspondant à 8 amines B différentes et à une seule amine A : l'aniline.

Les 8 amines B :





Les analyses effectuées 5 jours après l'introduction des différentes amines B ont montré que le produit principal obtenu était dans tous les cas l'urée symétrique de l'aniline.

La très faible formation d'urées asymétriques peut s'expliquer par le fait qu'en l'absence d'amine B, il se produit une hydrolyse partielle de l'imidazolide d'aniline (ou de l'isocyanate correspondant) en aniline, réagissant en partie avec l'imidazolide d'aniline résiduel pour donner l'urée symétrique. L'intermédiaire activé (imidazolide d'aniline) étant alors consommé avant l'ajout des amines B.

Une nucléophilie insuffisante des différentes amines B pourrait être évoquée, cependant il est difficilement concevable que ni la 4-méthyle-pipéridine, ni la 4-nitro-benzylamine ne permettent d'accéder aux urées asymétriques correspondantes, alors que la pipéronylamine avait permis la formation de l'urée (7) lors de l'essai préliminaire (§ A-2-1).

Interprétation :

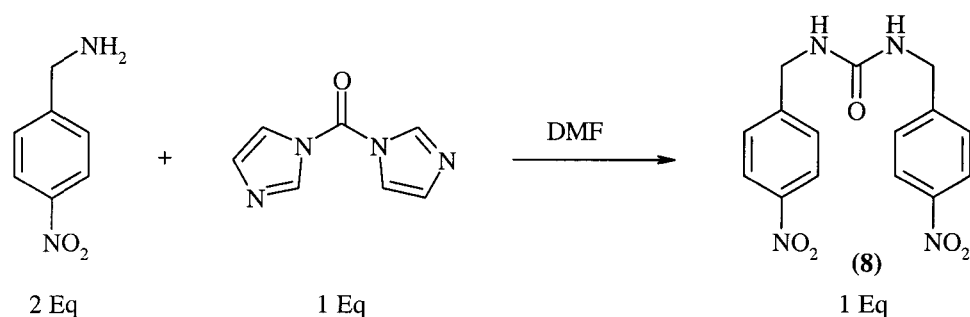
Un temps de contact prolongé entre l'amine A et le CDI (1^{ère} étape) s'est montré défavorable à la formation d'urées asymétriques. Cet effet est vraisemblablement lié à l'instabilité de l'intermédiaire activé.

Il nous fallait donc déterminer le temps de contact optimum entre l'amine A et le CDI. Le contrôle par HPLC ou CCM de la formation de l'imidazolide de l'amine A étant peu fiable du fait de son instabilité dans les conditions d'analyse, nous avons évalué sa formation en quantifiant la formation de l'urée asymétrique issue de la réaction de cet intermédiaire avec une amine très nucléophile : la 4-nitro-benzylamine.

- Détermination du temps de contact optimum entre l'amine A et le CDI après une distribution de l'amine A pendant 30 minutes

Nous avons tout d'abord synthétisé l'urée symétrique de la 4-nitro-benzylamine, afin de nous assurer que son temps de rétention HPLC était suffisamment différent de celui de l'urée symétrique de l'aniline pour pouvoir les identifier sans ambiguïté lors de l'étude de formation de l'urée asymétrique.

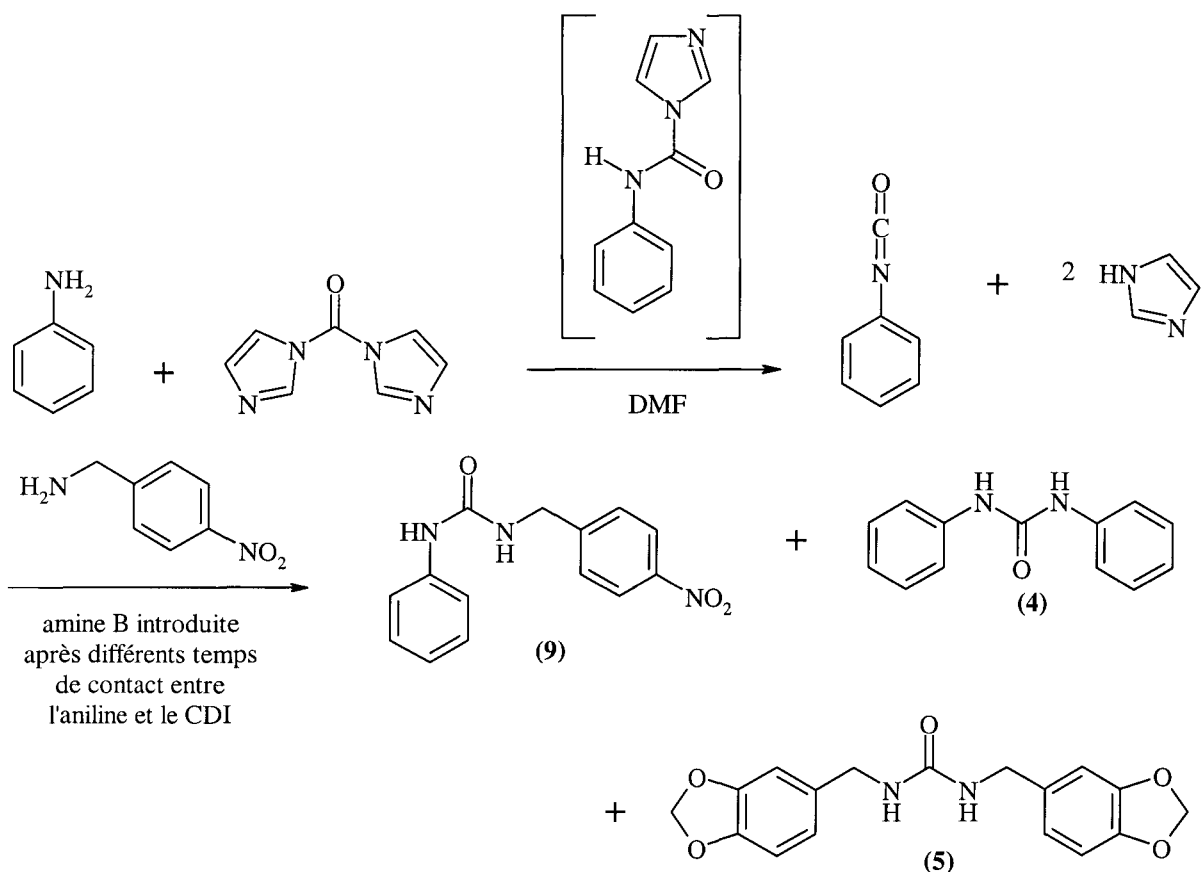
Schéma réactionnel :



L'urée symétrique de la 4-nitro-benzylamine a montré un temps de rétention HPLC, sur colonne cyano, de 22,3 minutes contre 20,2 minutes pour l'urée symétrique de l'aniline. Le chromatogramme obtenu après coinjection des 2 urées confirme que la différence de temps de rétention est suffisante pour pouvoir étudier la formation de l'urée asymétrique.

Nous avons alors fait varier le temps de réaction entre l'amine A et le CDI avant l'introduction de l'amine B.

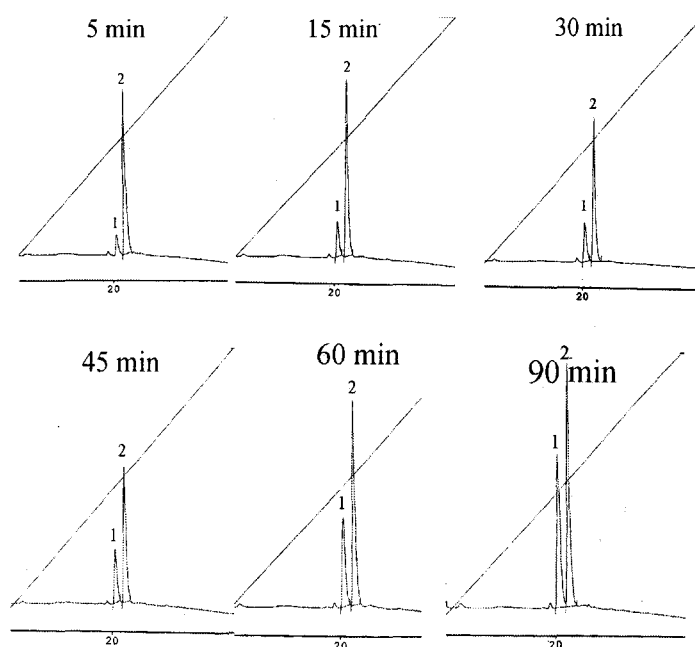
Schéma réactionnel :



Résultats :

Les CCM réalisées 15 minutes ou 1 heure après l'introduction de l'amine B se sont avérées similaires : dans tous les cas, la présence d'amines résiduelles (amine A et B) était mise en évidence, mais pour des temps de contact courts entre l'amine A et le CDI, les quantités d'amines résiduelles étaient plus faibles.

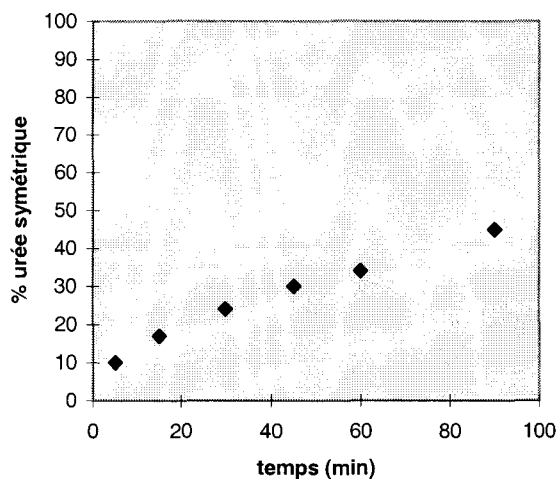
D'autre part, les HPLC ont clairement montré la relation directe entre la formation d'urée symétrique de l'amine A (4) et la durée de l'étape d'activation de l'amine A avec le CDI. Pour des temps de contact (sans tenir compte de la distribution de l'amine A faite sur 30 minutes) entre l'amine A (aniline) et le CDI supérieurs à 30 minutes, les rendements en urée asymétrique (9), obtenue après l'introduction de l'amine B, diminuaient fortement au profit de la formation de l'urée symétrique (4).



(1) = urée symétrique de l'amine A, (2) = urée asymétrique

En nous basant sur les aires des pics HPLC (sur colonne cyano, après évaporation des milieux réactionnels, et injection d'un aliquote) obtenues nous avons pu déterminer les proportions de formation d'urée symétrique de l'amine A par rapport à l'urée asymétrique, en fonction de la durée de l'étape d'activation de l'amine A (étape 1).

% d'urée symétrique de l'amine 1 formée en fonction du temps de contact amine 1 / CDI



Temps de contact *	% d'urée symétrique de l'amine A	% d'urée asymétrique
5 min	10	90
15 min	17	83
30 min	24	76
45 min	30	70
60 min	34	66
90 min	45	55

*Temps de contact entre l'amine A et le CDI après distribution de l'amine A (distribuée sur 30 minutes)

Tableau 1 : influence du temps de contact entre l'amine A et le CDI sur la formation d'urée symétrique de l'amine A

Interprétation :

Dans ce cas, où l'amine B introduite était la 4-nitro-benzylamine, l'étape 2 pouvait se limiter à 15 minutes du fait de la forte nucléophilie des amines benzyliques.

La présence d'amine A résiduelle, quelque soit le temps de contact entre l'amine A et le CDI pouvait s'expliquer de deux façons. La première impliquait un défaut de CDI par rapport à l'amine A, lié à une **hydrolyse partielle du CDI au cours de sa conservation** (en poudre ou en solution cf. § B-3 de ce chapitre), ou bien **au cours de la réaction**. La seconde impliquait une **hydrolyse de l'intermédiaire activé de l'amine A** avant l'introduction de l'amine B. Une troisième hypothèse selon laquelle le temps de réaction de l'amine A sur le CDI était insuffisant a été écartée car dans ce cas le CDI résiduel aurait pu réagir avec l'amine B pour conduire à l'intermédiaire activé ou à l'urée symétrique correspondants, espèces qui n'ont pu être mises en évidence par HPLC.

Pour modéliser le phénomène de contamination croissante en urée symétrique de l'amine A, pour des temps de contact croissants entre l'amine A et le CDI, nous avons supposé que l'urée symétrique de l'aniline (**4**) était formée avant l'introduction de la 4-nitro-benzylamine. Cette hypothèse nous est apparue raisonnable car celle-ci étant plus nucléophile, a pu réagir plus vite que l'aniline sur l'intermédiaire activé pour conduire à l'urée asymétrique. Nous avons alors envisagé deux effets conjugués :

- **l'effet de régénération** de l'amine A : l'instabilité de l'intermédiaire activé permet de régénérer progressivement, au cours du temps, une partie de l'amine A de départ qui peut ensuite réagir avec l'intermédiaire activé pour conduire à l'urée symétrique
- **l'effet d'incubation** : le temps de contact entre l'amine A et le CDI augmentant, la formation d'urée symétrique entre l'amine A résiduelle et l'intermédiaire activé peut s'opérer, mais dès l'introduction de l'amine B, celle-ci étant plus nucléophile, l'intermédiaire activé a été consommé pour conduire à l'urée asymétrique.

Conclusion du paragraphe A-2-2-1 :

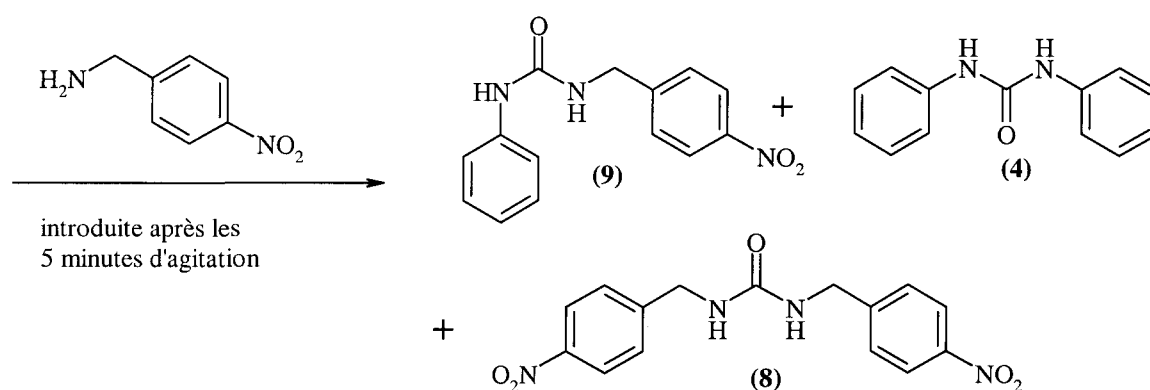
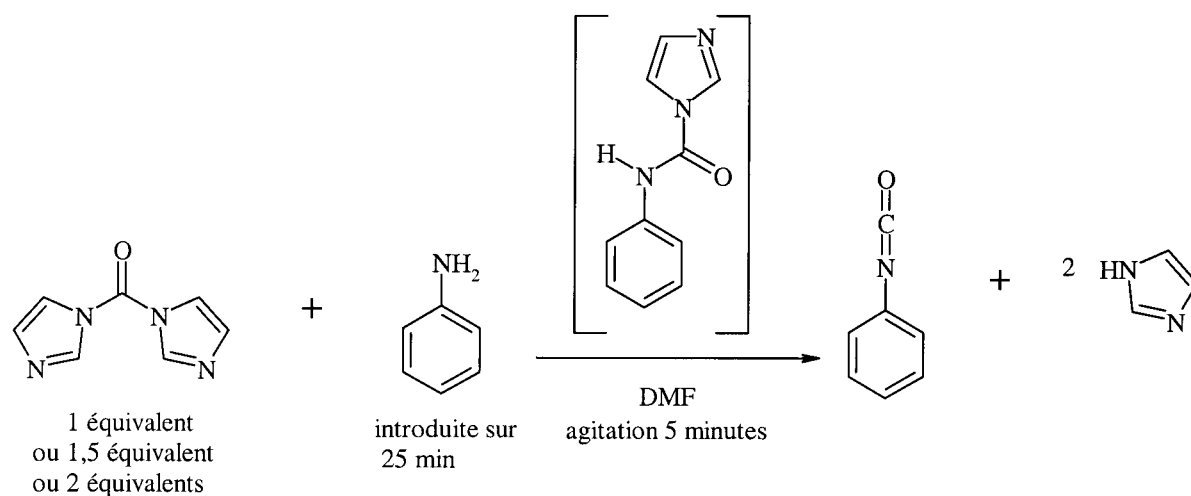
Les résultats précédents ont mis en évidence l'effet néfaste d'un temps de contact prolongé entre l'amine A et le CDI, favorisant la contamination en urée symétrique de l'amine A. Ce temps de contact a été par la suite limité au temps de distribution de l'amine A sur le CDI (25 minutes) suivi de 5 minutes d'agitation.

Il nous fallait cependant améliorer le rendement de la réaction afin de limiter les quantités d'amines résiduelles. De ce fait, le nombre d'équivalents de CDI à introduire nous semblait être un paramètre important à définir.

A-2-2-2 Détermination du nombre d'équivalents de CDI à utiliser

Afin de garantir une consommation totale de l'amine A lors de l'étape 1, nous avons étudié l'influence d'un excès de CDI (1, 1,5 et 2 équivalents) sur la réaction. Cette étude devait nous permettre de compenser, à la fois, une éventuelle dégradation du CDI avant la réaction ou au cours de la réaction ainsi que la possible formation d'urée symétrique de l'amine B, si au moment de son introduction, un excès de CDI était présent.

Schéma réactionnel :



Résultats :

Les HPLC des milieux réactionnels évaporés ont permis d'estimer la formation des urées symétriques en fonction du nombre d'équivalents de CDI introduits au départ.

	1 Eq	1,5 Eq	2 Eq
urée symétrique de l'amine A (4)	10 %	6 %	3 %
urée symétrique de l'amine B (5)	0 %	1 %	7 %
urée asymétrique (9)	90 %	93 %	90 %

Tableau 2 : influence du nombre d'équivalent de CDI (en solution dans le DMF) sur la formation de l'urée asymétrique

Interprétation :

Lorsque 1 équivalent de CDI était utilisé, l'urée symétrique (**4**) représentait 10 % des urées formées, l'urée asymétrique (**9**) en représentant 90 %. En revanche, pour le milieu réactionnel contenant 1,5 équivalent de CDI, la formation d'urée symétrique (**4**) a été réduite à 6 % des urées formées. L'excès de CDI introduit a pu compenser l'hydrolyse du CDI lors du stockage, ou bien, a permis de consommer l'amine A (résultant de l'hydrolyse de l'intermédiaire activé), pour régénérer l'intermédiaire activé et donc de limiter la formation d'urée symétrique (**4**). Quand 2 équivalents de CDI étaient introduits, l'urée symétrique de l'amine B est apparue, traduisant la présence de CDI résiduel après l'étape d'activation de l'amine A de 30 minutes. Dans ce cas, la proportion d'urée symétrique de l'amine B formée nous a paru relativement faible, au vu de l'excès de CDI introduit. La formation de faibles quantités d'urée symétrique de l'amine B, même en présence de 2 équivalents de CDI, montre clairement que le CDI s'hydrolyse très rapidement lors de la réaction. Ce phénomène trouve toute son importance pour une production robotisée à grande échelle.

Remarque : nous n'avons pu mettre en évidence, l'intermédiaire activé de l'amine B, celui-ci ayant vraisemblablement engendré l'urée symétrique correspondante après évaporation.

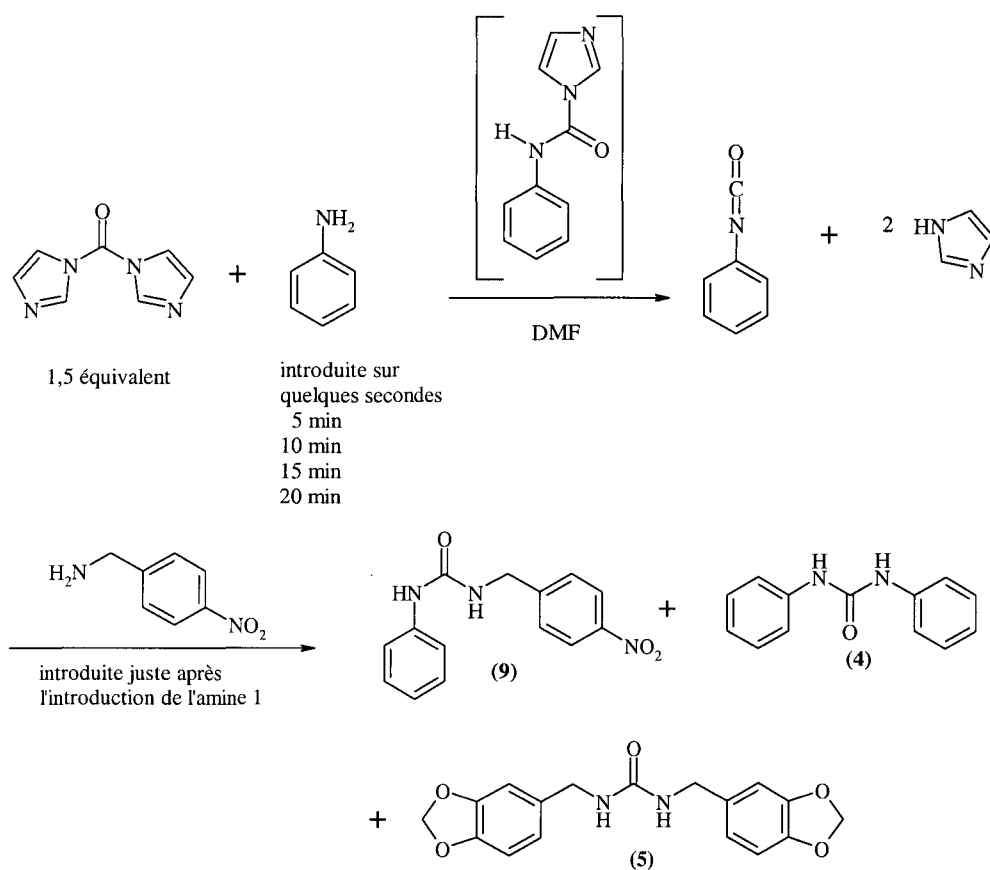
Conclusion du paragraphe A-2-2-2 :

L'étude du nombre d'équivalents de CDI à introduire pour minimiser la formation des urées symétriques a permis de mettre en évidence une hydrolyse du CDI à l'état solide lors du stockage, ou en solution dans le DMF (avant ou pendant la distribution de l'amine A sur 30 minutes). Afin de compenser cette hydrolyse et donc de favoriser la formation de l'urée asymétrique, il était nécessaire de placer 1,5 équivalent de CDI pour 1 équivalent d'amine A et 1 équivalent d'amine B. L'introduction d'un excès supérieur de CDI (2 équivalents) s'est traduit par la formation accrue, quoique limitée, de l'urée symétrique de l'amine B.

A-2-2-3 Etude de l'influence de la vitesse d'addition de l'amine A au CDI, sur la formation des urées symétriques ou asymétriques

Pour des raisons robotiques (cf. § A-1-3 de ce chapitre), et compte tenu de l'instabilité du CDI en solution, il nous semblait important de réduire la durée d'introduction de l'amine A. Nous avons donc contrôlé par HPLC les milieux réactionnels obtenus quand l'amine A (aniline) était ajoutée en quelques secondes, 5, 10, 15 et 20 minutes.

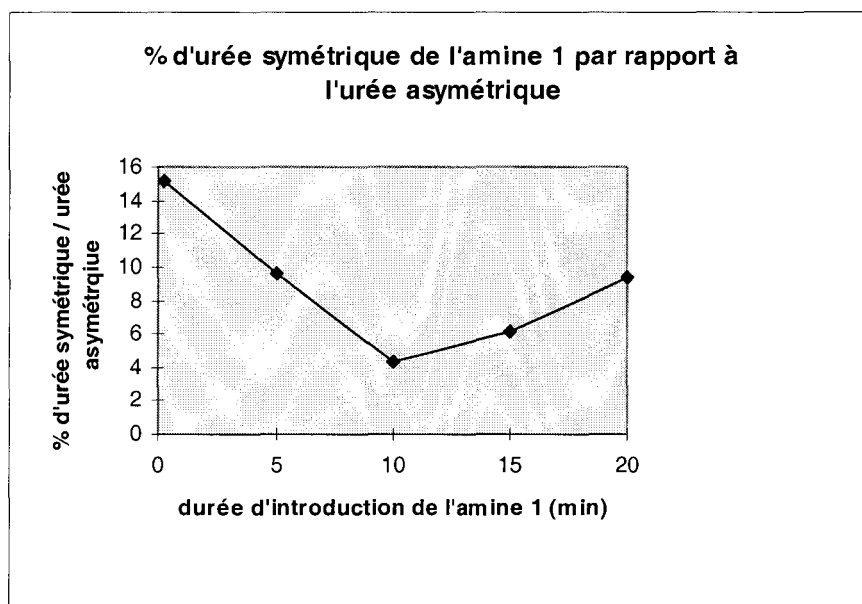
Schéma réactionnel :



Résultats :

Les HPLC ont montré que lorsque l'amine A était introduite en **quelques secondes** ou en **5 minutes**, plusieurs **pics secondaires** s'ajoutaient aux urées symétriques. (Nous verrons par la suite au paragraphe A-2-3-2 que ce phénomène n'est pas généralisable.)

Nous avons aussi constaté que l'**urée symétrique** de l'amine A se formait **peu** pour des distributions de l'amine A réalisées sur **10 à 15 minutes**.



Interprétation :

Pour des distributions rapides d'amine A (• 5 minutes), l'amine A se trouve localement en excès par rapport au CDI et peut conduire à l'urée symétrique. Le CDI résiduel est alors disponible pour réagir avec l'amine B et conduit à l'intermédiaire activé correspondant (cf. pic HPLC à 17,3 min). Contrairement à l'imidazolide de l'aniline, l'imidazolide de la 4-nitro-benzylamine a pu être détecté par HPLC, celui-ci étant vraisemblablement plus stable. Les imidazolides évoluant vers les formes isocyanates correspondantes (espèces plus réactives), il est important d'associer la réactivité de l'intermédiaire avec la facilité d'accès à l'isocyanate. Dans le cas de l'aniline, la formation du phénylisocyanate est favorisée par la délocalisation des électrons sur l'ensemble de la molécule car tous les atomes sont dans un état d'hybridation sp^2 . Dans le cas de la 4-nitro-benzylamine, l'isocyanate n'est plus conjugué avec le cycle aromatique, sa formation est donc plus difficile.

Pour des durées de distribution de l'amine A supérieures à 20 minutes, l'intermédiaire activé s'hydrolyse pour régénérer l'amine A, puis réagit sur l'intermédiaire activé résiduel pour conduire à l'urée symétrique (4).

Remarque :

Dans la synthèse impliquant l'utilisation de 2 équivalents de CDI pour un équivalent d'amines (A et B) (cf. § A-2-2-2), nous n'avons pas pu détecter l'imidazolide de la 4-nitro-benzylamine bien que du CDI résiduel devait être présent au moment de l'addition de l'amine B. Ce

phénomène pouvait s'expliquer par le fait que, l'analyse du milieu réactionnel avait dans ce cas été effectuée après évaporation du solvant. La conversion de l'imidazolide de la 4-nitro-benzylamine en urée symétrique correspondante, espèce thermodynamiquement favorisée avait alors pu s'opérer.

Conclusion du paragraphe A-2-2-3 :

La vitesse d'ajout de l'amine A au CDI est donc un paramètre important à contrôler. L'introduction de l'amine A sur une durée inférieure ou égale à 10 minutes permet de limiter l'hydrolyse du CDI et de l'intermédiaire activé de l'amine A (réactions se produisant pour des distributions plus lentes), sans qu'une contamination importante en urée symétrique de l'amine A ne soit observée (contrairement au triphosgène cf. § A-1-3).

A-2-2-4 Détermination du temps de réaction de l'amine B

Les résultats précédents ont montré que cette réaction est rapide. Nous avons même mis en évidence, lors de la détermination du temps de contact optimum entre l'amine A et le CDI (§ A-2-2-1), que les conversions obtenues (avec la 4-nitro-benzylamine) pour une 2^{ème} étape de 15 minutes ou d'une heure, étaient similaires (conversions > 90 %). Cependant, l'état de nos connaissances à ce moment n'était pas suffisant pour extrapoler les résultats obtenus avec la 4-nitro-benzylamine à d'autres amines B. Par précaution, nous avons étendu la durée de l'étape 2 à 12 heures pour permettre à des amines moins réactives de réagir.

Conclusion du § A-2-2

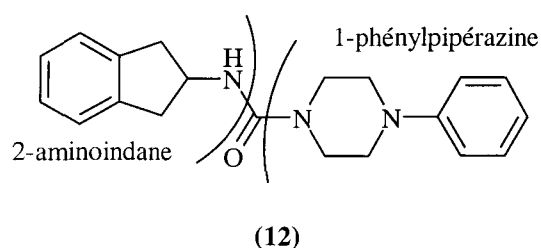
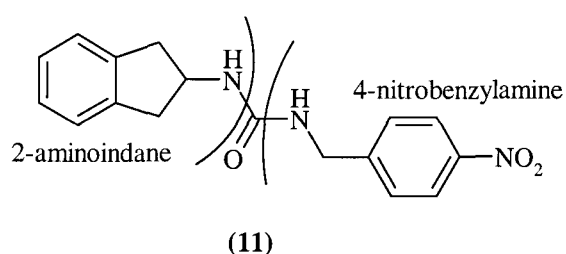
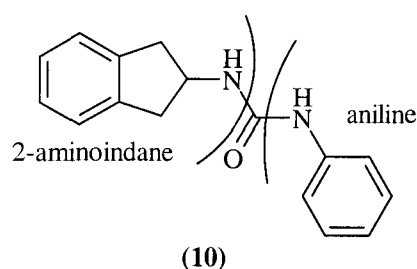
Les différentes études précédentes nous ont donc permis d'établir un protocole préliminaire pour la synthèse d'urées asymétriques en utilisant le CDI :

A 1,5 équivalent de CDI en solution (1,2 M) dans le DMF était ajoutée sur 10 minutes (en 6 aliquotes distribués goutte à goutte à 2 minutes d'intervalle), une solution 0,2 M de l'amine A dans le DMF. Après 5 minutes d'agitation, la solution d'amine B, 0,2 M dans le DMF était introduite en une seule fois puis le milieu placé sous agitation pendant au minimum 12 heures.

A-2-3 Evaluation des possibilités d'extension du protocole de synthèse à une amine A aliphatique

La mise au point du protocole de synthèse définie au § A-2-2 avait uniquement fait appel à l'aniline comme amine A, il fallait donc vérifier la possibilité d'utiliser des amines aliphatiques lors de la première étape. Cependant, du fait de la plus grande nucléophilie des amines aliphatiques, nous redoutions une contamination importante en urées symétriques formées lors de la première étape.

Nous avons synthétisé les trois urées asymétriques suivantes :



Les amines de départ ont été choisies car elles permettent :

- l'évaluation de la réactivité d'une amine aliphatique primaire pour l'amine A (2-aminoindane)
- l'évaluation de classes de réactivité différentes pour les amines B : une amine aromatique (l'aniline), une amine primaire aliphatique (la 4-nitrobenzylamine), une amine secondaire aliphatique (la N-phénylpipérazine)
- un suivi CCM et HPLC des monomères et des produits
- un suivi RMN des produits, leurs protons respectifs ayant des déplacements chimiques facilement attribuables.

Afin d'identifier facilement les possibles contaminants présents dans les milieux réactionnels, nous avons d'abord préparé les urées symétriques correspondantes.

porté à 100°C, ne révèle plus qu'un seul produit, le moins polaire et la disparition presque totale de l'amine de départ. Un résultat similaire a été obtenu lorsque le milieu réactionnel est placé à température ambiante pendant 24 h. La RMN et le spectre de masse du produit peu polaire, purifié par extraction, ont montré qu'il s'agissait de l'urée symétrique.

- les analyses des milieux réactionnels concernant la formation de l'urée symétrique de la 1-phénylpipérazine, après 4 heures de réaction (avec ou sans chauffage) ou après 24 heures, se sont avérées similaires. Des restes importants en amine de départ ont été mis en évidence. Un seul nouveau produit a été formé, très polaire au vu du Rf obtenu en CCM et du temps de rétention HPLC. La RMN et le spectre de masse de ce produit, purifié par chromatographie basse pression, ont montré qu'il s'agissait de l'imidazolide de l'amine de départ.

Interprétation :

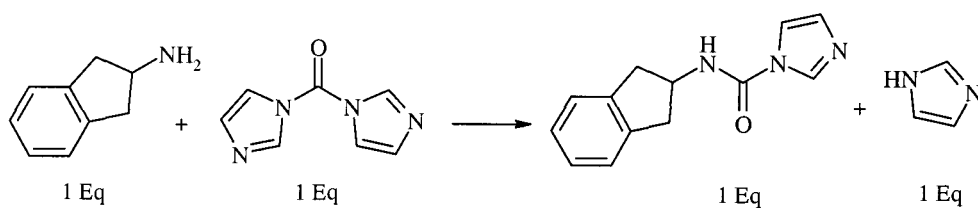
Les résultats précédents ont mis en évidence la formation de l'intermédiaire imidazolide après réaction d'une amine sur le CDI. Cet intermédiaire, particulièrement stable lorsqu'il était formé à partir d'une amine secondaire, ne permet pas d'accéder à l'urée symétrique correspondante. Après 5 minutes de chauffage à 100°C ou après 48 heures de réaction, en présence de l'équivalent d'amine n'ayant pas réagi (l'amine étant en excès par rapport au CDI), aucune trace d'urée n'a été détectée. De plus l'intermédiaire a pu être purifié, sans précaution particulière, par une chromatographie sur silice, à basse pression. Dans le cas de l'amine primaire, la formation de cet intermédiaire semblait également avoir lieu (cf. produit polaire détecté par CCM et HPLC). Cependant, l'urée symétrique de l'amine était détectée après 4 heures de réactions, ce qui révélait une réactivité supérieure. Après 24 heures de réaction, l'intermédiaire et l'amine de départ avaient presque totalement disparu, pour laisser place à l'urée symétrique. Cette réaction a d'autre part pu être nettement accélérée par chauffage, à 100°C pendant 5 minutes, du milieu réactionnel. La différence essentielle est dans le cas de l'amine primaire, la possibilité de former l'intermédiaire isocyanate (très réactif), après l'élimination de l'imidazole (Staab 1962).

Conclusion du § A-2-3-1

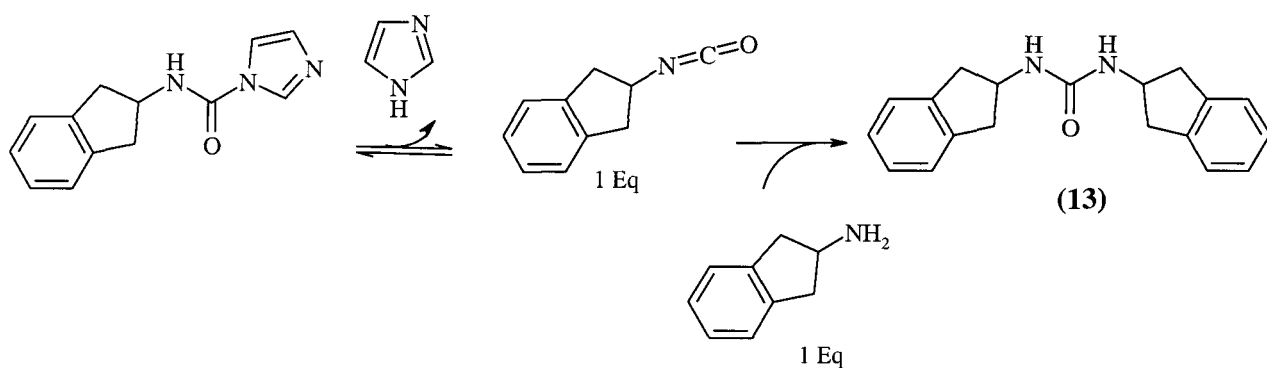
Les réactions des amines aliphatiques primaires ou secondaires avec le CDI se sont avérées très rapides et ont permis d'accéder aux imidazolides correspondants. L'obtention de l'urée

après l'attaque nucléophile du deuxième équivalent d'amine sur l'imidazolide de l'amine primaire, a nécessité un temps de réaction de 24 heures (ou un chauffage de 5 minutes à 100°C). En revanche, l'accès à l'urée symétrique de l'amine secondaire, s'est montré impossible dans les conditions utilisées (24 heures à température ambiante ou 5 minutes à 100°C), du fait de la stabilité de l'imidazolide, qui ne peut former l'isocyanate correspondant par perte d'imidazole. Ces deux cas sont schématisés ci-dessous.

Cas des amines primaires :

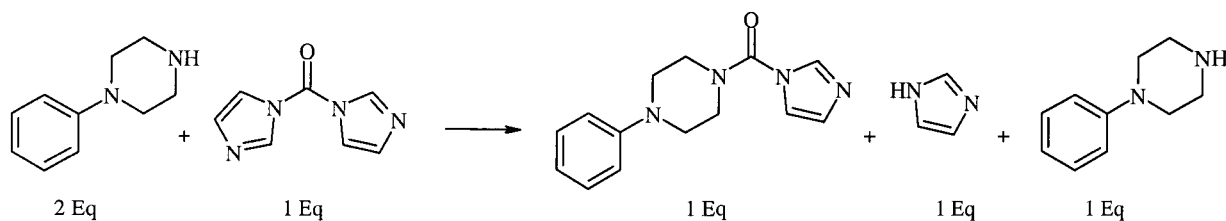


étape 1 : rapide

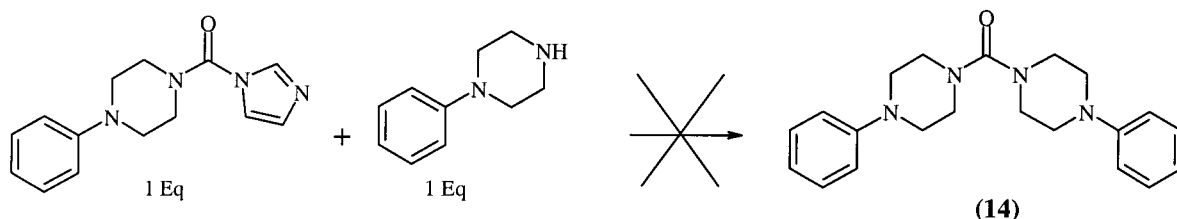


étape 2 : lente, accélérée par chauffage

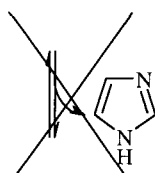
Cas des amines secondaires :



étape 1 : rapide



étape 2 : impossible à température ambiante pendant 24h ou à 100°C pendant 5 min

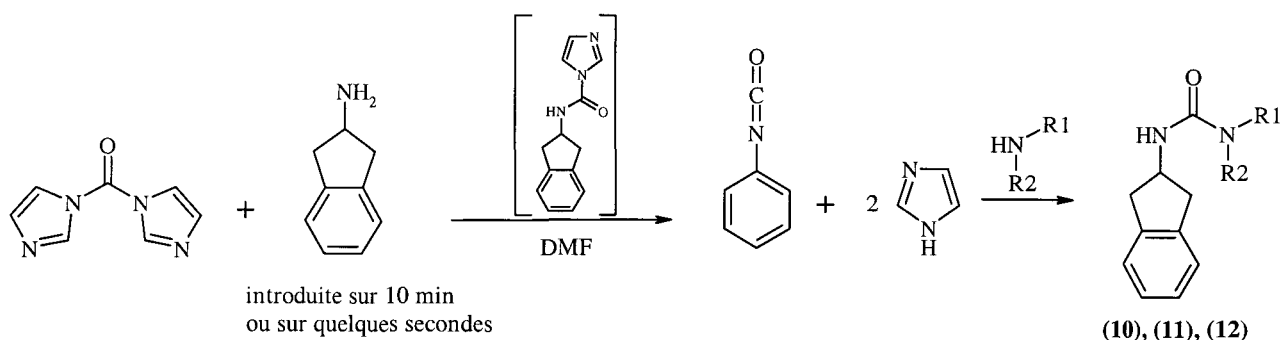


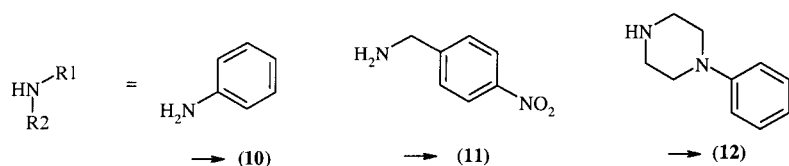
A-2-3-2 Préparation des urées asymétriques

Conditions opératoires :

Les trois urées asymétriques ont été synthétisées en utilisant le protocole préliminaire défini dans la conclusion du paragraphe A-2-2, l'amine A étant dans ce cas distribuée sur 10 minutes. Un autre protocole impliquant la distribution de l'amine A sur quelques secondes, a été utilisé à titre comparatif. Les milieux réactionnels ont été contrôlés par HPLC et CCM, 3h et 72h après l'introduction des amines 2.

Schéma réactionnel :





Résultats :

L'effet de la vitesse d'introduction de l'amine A s'est montré peu marqué, sauf pour la formation de l'urée (11), où lors de l'introduction rapide de l'aminoindane, des produits secondaires contaminaient le milieu réactionnel (résultats semblables à ceux du § A-2-2-3). Dans tous les cas les urées asymétriques étaient nettement majoritaires (conversions obtenues : 70% (10), 79% (11), 89% (12)). Les produits secondaires principaux sont l'imidazolide de l'amine A et l'urée symétrique de l'amine B (pour une amine B primaire), ou l'imidazolide de l'amine B (pour une amine B secondaire).

Conclusion du § A-2-3-2 :

Les problèmes de contamination du milieu réactionnel décrits dans le paragraphe A-2-2-3 pour une introduction rapide de l'amine A n'ont été ici confirmés que dans le cas où l'amine B était la 4-nitro-benzylamine. Ces résultats nous ont permis d'envisager une simplification du protocole préliminaire en remplaçant la distribution de l'amine A sur 10 minutes, par une distribution sur quelques secondes, beaucoup plus facile à réaliser par le robot.

Les contaminants observés dans les différents milieux réactionnels (urées symétriques de l'amine B ou l'imidazolide de l'amine B) reflétaient la présence de CDI résiduel au moment de l'introduction de l'amine B. Nous avons expliqué ce phénomène par la stabilité de l'imidazolide de l'aminoindane qui, contrairement à l'intermédiaire activé de l'aniline (cf. § A-2-2-2) ne régénèrait pas l'amine de départ et donc ne pouvait pas consommer le CDI en excès.

Conclusion du § A-2-3 :

L'utilisation d'une amine A de type aliphatique s'est montrée compatible avec le protocole de synthèse mis au point avec une amine A de type aromatique. Ces deux types d'amines pouvaient donc être utilisés indifféremment en position 1. L'introduction en une seule fois de l'amine A ne semblait pas compromettre la synthèse des urées asymétriques. L'effet défavorable de ce mode d'introduction décrit au § 2-2-3 et confirmé ci-dessus ne s'est manifesté que pour la synthèse d'urées faisant intervenir la 4-nitro-benzylamine comme

amine B. Il s'agissait vraisemblablement d'un effet propre à cette amine. D'autre part, les connaissances acquises sur la formation d'urées asymétriques au CDI, nous ont permis d'associer à la présence de divers produits secondaires, une série de causes potentielles :

Amine de départ A	<p>sa détection en fin de réaction témoigne :</p> <ul style="list-style-type: none"> - d'un défaut de réactivité vis à vis du CDI - d'un défaut en CDI introduit au départ (ou de son hydrolyse partielle) - d'une instabilité de l'imidazolide ou de l'urée issue de cette amine
Amine de départ B	<p>sa détection en fin de réaction témoigne :</p> <ul style="list-style-type: none"> - d'un défaut de réactivité vis à vis de l'isocyanate de l'amine A - d'un défaut en isocyanate de l'amine A - d'une instabilité de l'urée issue de cette amine
Imidazolide de l'amine A	<p>sa détection en fin de réaction témoigne :</p> <ul style="list-style-type: none"> - de la stabilité particulière de l'imidazolide de l'amine A - d'un défaut en amine B - de la faible réactivité de l'amine B
Imidazolide de l'amine B	<p>sa détection en fin de réaction témoigne de la présence de CDI au moment de l'introduction de l'amine B donc :</p> <ul style="list-style-type: none"> - d'un excès de CDI par rapport à l'amine A - d'un défaut de réactivité de l'amine A vis à vis du CDI
Urée symétrique de l'amine A	<p>sa détection en fin de réaction témoigne :</p> <ul style="list-style-type: none"> - de la réactivité particulièrement élevée de l'imidazolide ou de l'isocyanate de l'amine A - d'un défaut en CDI introduit au départ (ou de son hydrolyse partielle) - d'un défaut en amine B - de la faible réactivité de l'amine B
Urée symétrique de l'amine B	<p>sa détection en fin de réaction témoigne de la présence de CDI au moment de l'introduction de l'amine B donc :</p> <ul style="list-style-type: none"> - d'un excès de CDI par rapport à l'amine A - d'un défaut de réactivité de l'amine A vis à vis du CDI

Tableau 3 : produits secondaires possibles dans les milieux réactionnels issus de la synthèse d'urées au CDI

A-3 Conclusion

Les résultats décrits dans ce paragraphe nous ont amenés à choisir le CDI comme agent de couplage pour la synthèse d'urées asymétriques. Les quatre raisons majeures étant :

- l'obtention de milieux réactionnels peu complexes contenant principalement les produits de départ (minoritaires) et le produit attendu (majoritaire)
- la possibilité d'effectuer les synthèses en présence de DMF autorisant (contrairement au triphosgène) la mise en solution des amines dans ce solvant
- le transfert possible du protocole de synthèse à notre système robotisé
- la compatibilité de l'imidazole avec les tests biologiques.

B Mise au point du protocole final de la synthèse robotisée d'urées au carbonyldiimidazole

B-1 Objectifs et contraintes de la production robotisée

Objectifs : - atteindre des cadences de production de plusieurs milliers de produits par 24h afin de subvenir aux besoins des criblages biologiques à haut débit (HTS)

- produire les composés sous un format adapté aux futures campagnes de tests biologiques à haut débit (plaques de microtitration)

- disposer d'une traçabilité totale de chaque composé produit

- limiter le plus possible les interventions humaines.

Contraintes : - les temps de programmation informatique sont importants

- dans le cadre de distribution de composés variés (ex : les monomères), à l'échelle de la μ mole, le robot ne peut manipuler que des composés en solution

- les échantillons subissant tous la même suite d'opérations, il faut regrouper les monomères se comportant de manière homogène (ex : choisir les monomères d'une même production, de telle sorte qu'ils soient tous solubles à une même concentration)

- le robot devant fonctionner avec le minimum d'intervention humaine, il faut s'assurer de la stabilité des espèces réactives dans le temps (ex : stabilité du CDI en solution).

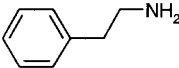
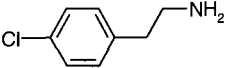
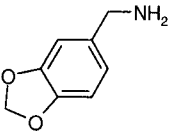
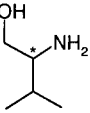
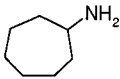
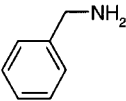
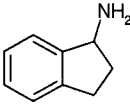
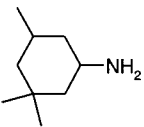
B-2 Etude de la mise en solution des amines

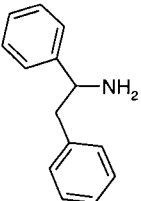
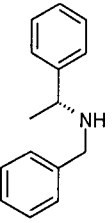
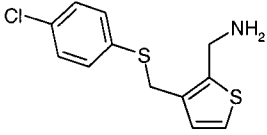
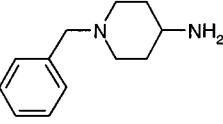
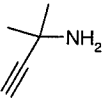
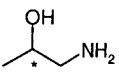
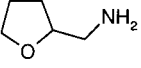
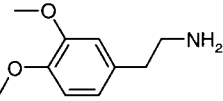
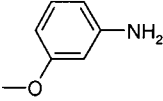
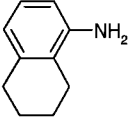
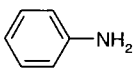
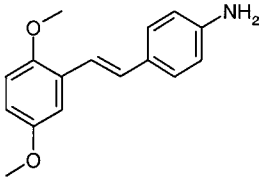
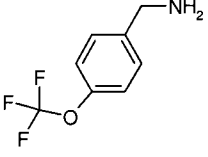
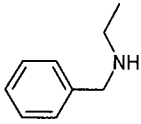
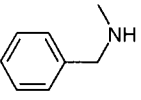
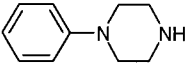
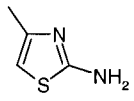
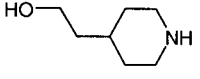
En chimie classique, la solubilité des réactifs est généralement traitée par le chimiste au cas par cas, en choisissant le solvant ou les mélanges de solvants appropriés. En chimie robotisée, par contre, il est indispensable de disposer d'une méthode générique permettant de traiter systématiquement de grandes séries de composés variés. Cette contrainte est accentuée par l'exigence de diversité structurale des différents monomères. La mise en solution constitue donc une étape cruciale pour toute production robotisée de chimiothèques, le choix d'une

méthode inappropriée pouvant se traduire, à terme, par une réduction inacceptable de la diversité de la chimiothèque si seuls les monomères parfaitement solubles sont retenus. En revanche, l'insolubilité d'un monomère, qui serait malgré tout retenu, entraîne une erreur de stœchiométrie lors de son introduction à tout milieu réactionnel. De plus, la distribution de ce réactif se faisant grâce à une aiguille, la fraction insoluble peut entraîner le colmatage de cette aiguille.

Le solvant retenu devra non seulement permettre la mise en solution du plus grand nombre de monomères (accès à une grande diversité de structure) mais aussi être compatible avec les conditions opératoires.

Une étude de solubilité dans trois solvants : le DCM, le DMF et le DMA, a été menée sur un lot de 66 amines, représentatif des 700 amines disponibles à l'époque (02/1995) au laboratoire. Parmi ces 66 amines, 37 étaient liquides. Nous avons supposé que celles-ci devaient être miscibles aux trois solvants à une concentration de 0,25 M. Les 29 autres amines étaient solides et ont été testées à deux concentrations (0,1 M et 0,25 M) dans les trois solvants.

Amines	Structures	Amines	Structures
Phénéthylamine		2-(4Chlorophényl)-éthylamine	
Pipéronylamine		DL-2-Amino-3-méthyl-1-butanol	
Cycloheptylamine		Benzylamine	
1-Aminoindane		3,3,5-Triméthylcyclohexylamine	

1,2-Diphényléthylamine		(R)-(+)-N-Benzyl- α -méthylbenzylamine	
S-(4-Chlorophényl)-2-aminométhyl-3-thiométhylthiophène		4-Amino-1-benzylpipéridine	
1,1-Diméthylpropargylamine		DL-1-Amino-2-propanol	
Tétrahydrofurfurylamine		3,4-Diméthoxyphénéthylamine	
m-Anisidine		1-Amino-5,6,7,8-tétrahydronaphtalène	
Aniline		2',5'-Diméthoxy-4-stibènamine	
4-(Trifluorométhoxy)benzylamine		N-Benzyléthylamine	
N-Benzylméthylamine		1-Phénylpipérazine	
2-Amino-4-méthylthiazole		4-(2-Hydroxyéthyl)pipéridine	

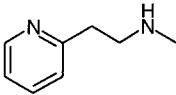
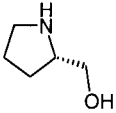
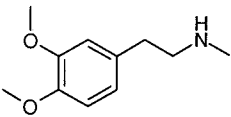
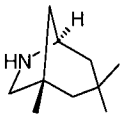


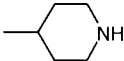
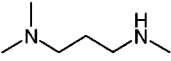

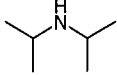
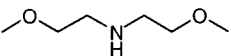
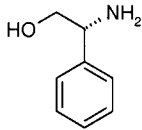
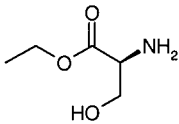
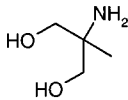
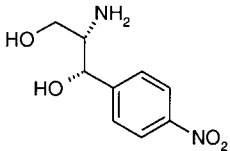
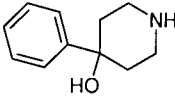
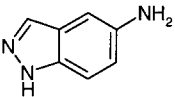
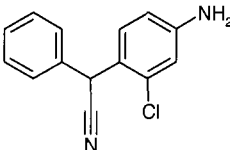
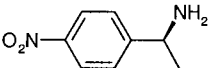
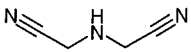
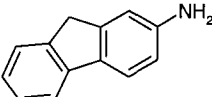
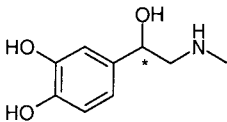
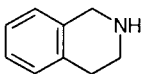
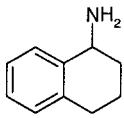
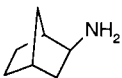
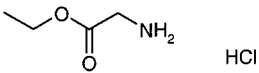
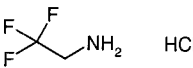
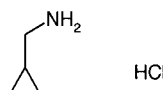
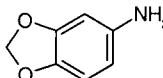
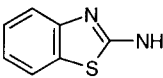
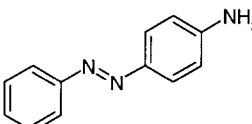
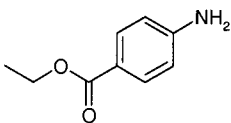
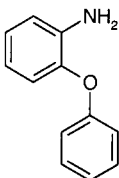
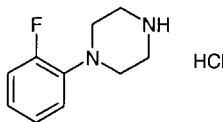
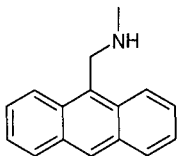
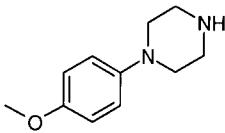
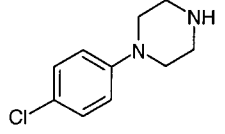
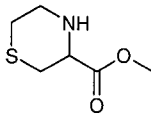
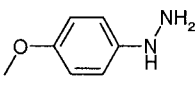
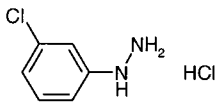
2-(2-(Méthylamino)éthyl)pyridine		(S)-(+)-2-Pyrrolidineméthanol	
N-Méthyl-3,4-diméthoxyphénéthylamine		1,3,3-Triméthyl-6-azabicyclo-(3,2,1)-octane	
Pyrrolidine		Morpholine	
4-Méthylpipéridine		N,N,N'-Triméthyl-1,3-propanediamine	
2-Méthyl-pyrroline		Diisopropylamine	
Bis(2-méthoxyéthyl)amine			

Tableau 4 : les 37 amines liquides présentes dans le lot des 66 amines représentatives des amines disponibles au laboratoire

Amines	Structures	DCM	DCM	DMF	DMF	DMA	DMA
		0,25 M	0,1 M	0,25 M	0,1 M	0,25 M	0,1 M
R-(-)-2-Phénylglycinol		0	0	1	1	1	1
Ester éthylique de la L-sérine		0	0	1	1	1	1
2-Amino-2-méthyl-1,3-propanediol		0	0	1	1	1	1

D-Thréo-2-amino-1-(4-nitrophényl)-1,3-propanediol		0	0	1	1	1	1
4-hydroxy-4-phénylpipéridine		0	0	1	1	1	1
5-Aminoindazole		0	0	1	1	1	1
4-Amino-2-chloro-diphénylacétonitrile		0	0	1	1	1	1
S-(-)-α-Méthyl-4-nitrobenzylamine		0	1	1	1	1	1
Iminodiacétonitrile		0	1	1	1	1	1
2-Aminofluorène		0	1	1	1	1	1
(+/-)-épinéphrine		0	0	0	0	0	0
1,2,3,4-Tétrahydroisoquinoline		1	1	0	1	0	0
1,2,3,4-Tétrahydro-1-naphttylamine		1	1	1	1	1	1
Endo-2-aminonorbornane		1	1	1	1	1	1

Ester éthylique de glycine (Chlorhydrate)		1	1	1	1	1	1
2,2,2-Trifluoroéthylamine (Chlorhydrate)		1	1	1	1	1	1
Aminométhylcyclopropane (Chlorhydrate)		1	1	1	1	1	1
3,4-(Méthylènedioxy)- aniline		1	1	1	1	1	1
2-Aminobenzothiazole		1	1	1	1	1	1
4-Phénylazoaniline		1	1	1	1	1	1
4-Aminobenzoate d'éthyle		1	1	1	1	1	1
2-Phénoxyaniline		1	1	1	1	1	1
1-(2-Fluorophényl) pipérazine (chlorhydrate)		1	1	1	1	1	1
9-(Méthylaminométhyl) anthracène		1	1	1	1	1	1

1-(4-Méthoxyphényl)- pipérazine		1	1	1	1	1	1
1-(4-Chlorophényl)- pipérazine		1	1	1	1	1	1
Thiomorpholino-4- carboxylate de méthyle		1	1	1	1	1	1
4-Méthoxyphénylhydrazine		1	1	1	1	1	1
3-Chlorophénylhydrazine (Chlorhydrate)		1	1	1	1	1	1

0 = insoluble, 1 = soluble

Tableau 5 : résultats des tests de solubilité effectués sur les 29 amines solides présentes dans le lot des 66 amines représentatives des amines disponibles au laboratoire

	DCM	DMF	DMA
0,25 M	62%	93%	93%
0,1 M	72%	97%	93%

Tableau 6 : récapitulatif des proportions d'amines solides, solubles dans le DCM, le DMF et le DMA à 0,25 et 0,1 M

Conclusion :

Pour des mises en solution d'amines à des concentrations de 0,1 et 0,25 M, le DMF et le DMA se sont avérés de bien meilleurs solvants que le DCM. La différence s'est surtout révélée lors de la mise en solution des aminoalcools, très peu solubles dans le DCM. Ce phénomène s'explique par la capacité du DMF et du DMA d'établir des liaisons hydrogène (en tant qu'accepteurs) avec les groupes hydroxyle des aminoalcools.

Le DMF, solvant aprotique polaire, est d'autre part reconnu comme un excellent solvant pour les réactions de type S_N2 (Vollhardt 1990) :

- en permettant la formation de réactifs nucléophiles « nus », dépourvus de couches de solvation isolantes et volumineuses de solvant (dues aux liaisons hydrogènes)
- en étant capable de dissoudre des molécules polaires (moment dipolaire important : 3,24D)
- en étant capable, en dissociant les paires d'ions, de maintenir les nucléophiles anioniques à l'écart de leurs contre-ions (grande constante diélectrique : 36,71)
- en permettant de stabiliser les états de transitions chargés (grande constante diélectrique : 36,71).

Par comparaison, le moment dipolaire du DCM est de 1,14 D et sa constante diélectrique est égale à 8,9.

Bien que le DMF et le DMA aient montré des pouvoirs solvant très similaires vis à vis des amines testées, nous avons préféré utiliser le DMF, celui-ci étant plus facilement évaporé que le DMA (point d'ébullition du DMF : 152-153° C / point d'ébullition du DMA : 164-166°C).

B-3 Etude de la stabilité du CDI en solution dans différents solvants

Afin de simplifier la production robotisée de la chimiothèque, nous avons recherché un solvant préservant la réactivité du CDI pendant une période minimale de 12 h. Le CDI, en solution, pouvait alors être facilement prélevé et distribué par le robot, de plus une stabilité garantie sur une période de 12 h permettait d'éviter une intervention humaine au cours des phases nocturnes de production.

En nous basant sur des données de la littérature (Staab 1962), le THF constituait le solvant classiquement utilisé pour les réactions faisant intervenir le CDI. Cependant deux autres solvants nous semblaient intéressants à étudier : le DMF et le DCM :

- le DMF permettait d'utiliser le même solvant pour toutes les étapes de la synthèse, les risques de précipitation liés à l'introduction d'un solvant différent étant alors écartés
- le DCM, du fait de son inertie, nous semblait compatible avec une conservation prolongée du CDI.

Afin d'évaluer la stabilité du CDI en solution, nous avons laissé évoluer dans le temps des solutions de CDI et synthétisé l'urée symétrique de la phénéthylamine en utilisant ces solutions de CDI plus ou moins récentes (voir partie expérimentale). Nous avons ensuite déterminé par HPLC, les rendements obtenus en urée pour ces différents temps.

Résultats :

Stabilité du CDI dans le THF : > 12 h

Les rendements en urée étaient quantitatifs pour les solutions de CDI récentes ou préparées depuis 13 heures.

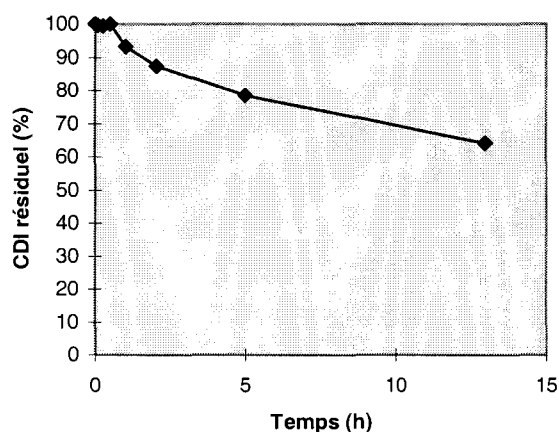
Stabilité du CDI dans le DCM : > 12 h

Les rendements en urée formée ont décliné de 100 % à 90 % pendant les 30 minutes suivant la mise en solution puis se sont stabilisés à 90 %. En nous appuyant sur cette cinétique biphasique, nous avons expliqué la diminution de réactivité du CDI pendant les 30 premières minutes par la présence d'eau dans le DCM. Le solvant ainsi déshydraté, a permis de préserver la réactivité du CDI pendant les 12 h 30 suivantes.

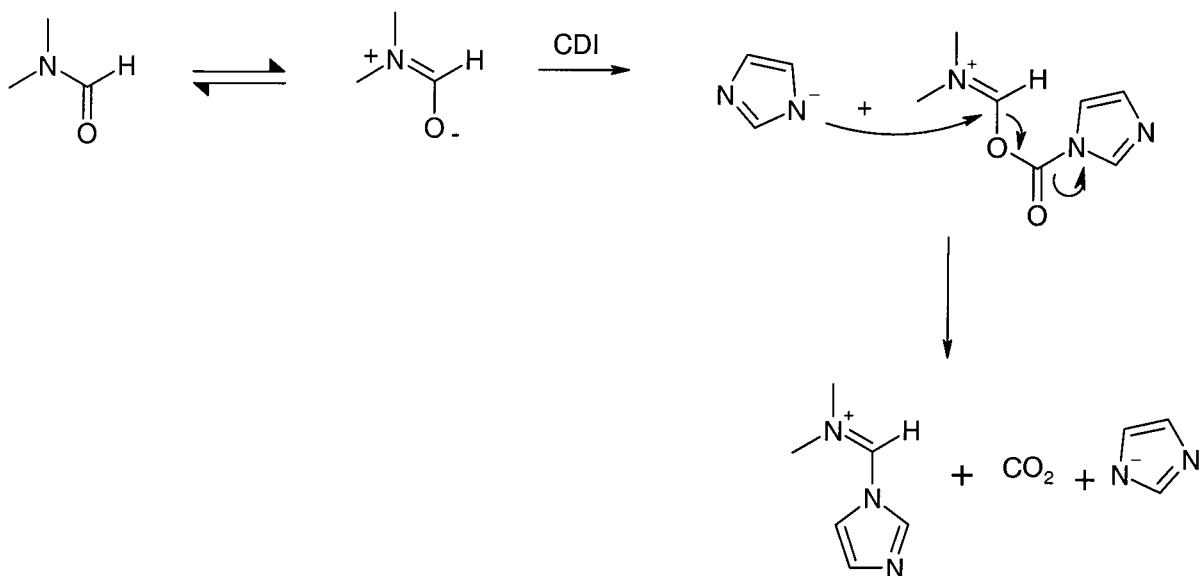
Stabilité du CDI dans le DMF : quelques heures

Le test de réactivité du CDI en solution dans le DMF effectué 13 h après la mise en solution, a permis d'évaluer la concentration en CDI résiduel à 64 % du CDI introduit au départ.

Cinétique de dégradation du CDI dans le DMF



Nous avons alors supposé l'attaque de la forme énolique du DMF sur le CDI et la formation du N-(N-diméthyliminium)imidazole par analogie avec la formation du chloroiminium lors de la première étape de la réaction de Vielsmeier.



Conclusion :

L'étude de stabilité du CDI dans le DMF, le DCM et le THF a montré que seules les solutions de CDI dans le DCM ou le THF pouvaient être conservées plus de 12 heures. La réactivité du CDI étant préservée durant cette période autorise des phases de production nocturnes sans intervention humaine. Notre choix s'est porté sur la mise en solution du CDI dans le THF plutôt que dans le DCM, afin de réduire les risques de concentration du CDI par évaporation du solvant (point d'ébullition du THF : 65-67°C / point d'ébullition du DCM : 39-40°C).

B-4 Mise au point des tests d'évaluation de la réactivité du CDI

Il était capital de vérifier avant chaque production robotisée, la réactivité du CDI pour garantir une distribution stœchiométrique de ce réactif d'activation. Nous désirions un test rapide et facile à mettre en œuvre.

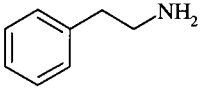
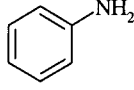
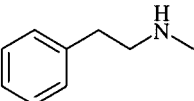
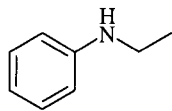
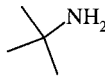
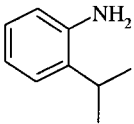
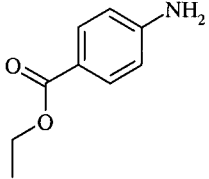
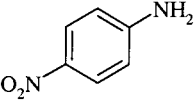
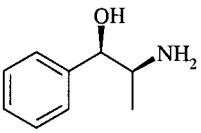
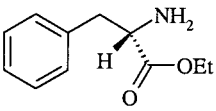
Nous avons d'abord pensé à la mise en place d'un **test colorimétrique** (cf. partie expérimentale).

Celui-ci consistait à peser une certaine quantité de CDI, à la placer en présence d'un excès de benzylamine puis à doser la quantité de benzylamine résiduelle. La quantité de CDI devant être pesée pour la préparation de la solution d'activateur utilisée lors de la production robotisée, dépendait du résultat obtenu précédemment. Nous avons alors mis en évidence l'hétérogénéité de la réactivité du CDI au sein du flacon de CDI commercial. Ce phénomène a été attribué à la présence de cristaux d'imidazole (correspondant à du CDI hydrolysé ou à de l'imidazole n'ayant pu être éliminé lors de la préparation du CDI), macroscopiquement indiscernables des cristaux de CDI. Afin de disposer de la réactivité réelle de la solution de CDI préparée pour la synthèse robotisée, nous avons mis au point un second test permettant cette fois la détermination de la réactivité du CDI en solution, par mesure de la formation d'une urée et non par mesure de la disparition d'une amine.

Le second test consiste en un **dosage HPLC de la réactivité d'une solution de CDI** (cf. partie expérimentale) : un volume précis d'une solution de CDI est prélevé, un excès de benzylamine lui est ajouté et la dibenzylurée formée est dosée par HPLC. Ce test sera systématiquement utilisé pour toute production faisant intervenir le CDI.

B-5 Evaluation de la réactivité de différentes familles de N-nucléophiles

Les différents N-nucléophiles ont été classés en familles sur la base de leurs caractéristiques stériques et électroniques, tant au niveau de l'atome d'azote lui-même (ex : amines primaires aliphatiques, hydrazides, benzothiazoles...) qu'au niveau de la structure prise dans son ensemble, celle-ci pouvant comporter d'autres groupes fonctionnels (ex : aminoalcools, acides aminés...). Chaque famille a été représentée par un composé modèle. Le protocole de synthèse des urées a ensuite été appliqué à chaque composé modèle afin d'en extrapoler la réactivité à chaque famille représentée. Le classement des N-nucléophiles en familles avait donc pour but de simplifier le choix des amines à incorporer dans la synthèse de la chimiothèque : si un composé modèle était réactif, l'ensemble de la famille qu'il représentait était supposé réactif.

Familles de N-Nucléophiles	Composés modèle	Formules
Amines primaires aliphatiques	Phénéthylamine	
Amines primaires aromatiques	Aniline	
Amines secondaires aliphatiques	N-méthylphénéthylamine	
Amines secondaires aromatiques	N-éthylaniline	
Amines primaires aliphatiques encombrées	<i>tert</i> -Butylamine	
Amines primaires aromatiques encombrées	2-isopropylaniline	
Amines primaires aromatiques faiblement désactivées	4-amino-benzoate d'éthyle	
Amines primaires aromatiques fortement désactivées	4-nitroaniline	
Aminoalcools	R, S, (-) noréphédrine	
Esters d'acides aminés*	L-phénylalanine-carboxylate d'éthyl	

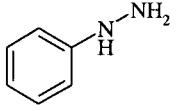
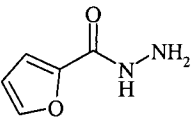
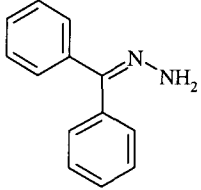
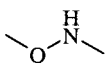
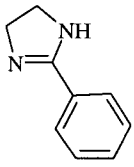
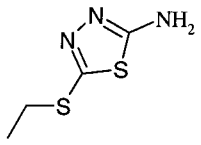
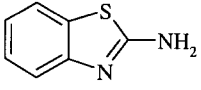
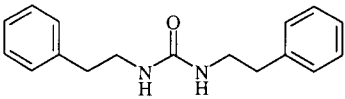
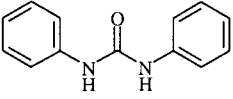
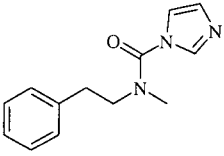
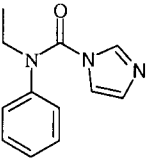
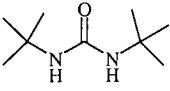
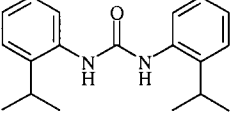
Hydrazines	Phénylhydrazine	
Hydrazides	2-furoic hydrazide	
Hydrazones	hydrazone de benzophénone	
Hydroxylamines	N,O-diméthyl- hydroxylamine	
Amidines	2-phényl-2-imidazoline	
1,3,4-thiadiazoles	2-amino-5-éthylthio-1,3,4- thiadiazole	
2-aminothiazoles	2-amino-benzothiazole	

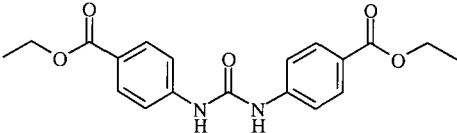
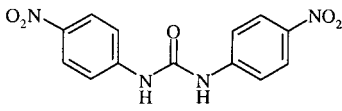
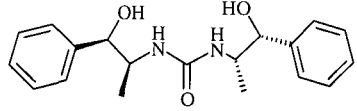
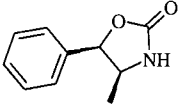
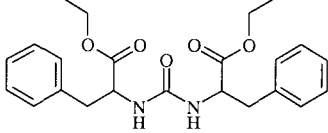
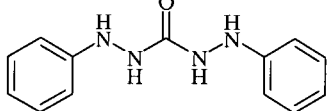
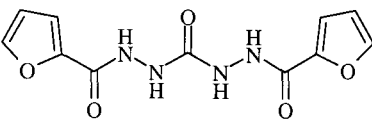
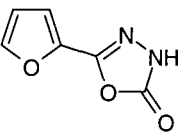
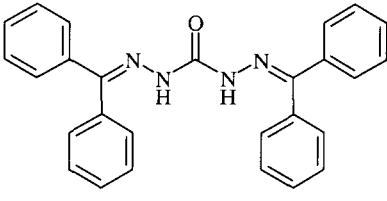
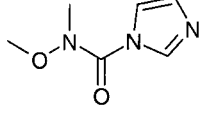
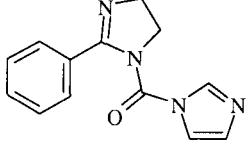
Tableau 7 : les différentes familles de N-nucléophiles et les composés modèle associés

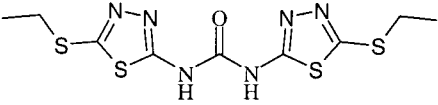
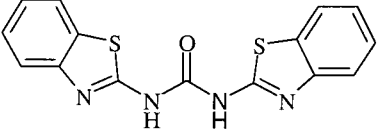
B-5-1 Préparation des urées symétriques ou des imidazolides issus des représentants des familles de N-nucléophiles

Les urées symétriques ont été préparées pour les amines primaires. Les imidazolides ont été préparés pour les amines secondaires (cf. conclusion du § A-2-3-1).

Les résultats obtenus devaient d'une part, nous permettre une première estimation de la réactivité de ces représentants dans la chimie des urées, d'autres part, les données chromatographiques recueillies allaient être utilisées pour l'étude de réactivité de ces représentants engagés dans la synthèse d'urées asymétriques (les urées symétriques et les imidazolides constituant les produits secondaires les plus probables).

Produits attendus	Produits obtenus		Rendements
	= produit attendu	(15)	> 95 % ¹
	= produit attendu	(16)	> 95 % ¹
	= produit attendu	(17)	> 95 % ¹
	= produits de départ	(18)	0 % ¹
	= produit attendu	(19)	72 % ²
	= produit attendu	(20)	> 95 % ¹

	= produit attendu	(21)	51 % ²
	= produits de départ	(22)	0 % ¹
		(23)	85 % ²
	= produit attendu	(24)	> 95 % ¹
	= produit attendu	(25)	> 95 % ¹
		(26)	> 95 % ¹
	= produit attendu	(27)	68 % ¹
	= produit attendu	(28)	76 % ²
	= produit attendu	(29)	48 % ²

	= produits de départ	(30)	0 % ¹
	= produits de départ	(31)	0 % ¹

¹ : rendement estimé par HPLC

² : rendement en produit isolé

Tableau 8 : rendements obtenus lors de la synthèse des urées symétriques ou des imidazolides de référence

Résultats :

Cette étude nous a permis d'identifier :

- les N-nucléophiles pouvant être utilisés indifféremment comme **amines A ou B**, les urées symétriques attendues ayant été obtenues. Il s'agit des **amines primaires aliphatiques** (encombrées ou non), des **amines primaires aromatiques** (encombrées ou non), des **amines primaires aromatiques faiblement désactivées**, des **hydrazines** et des **hydrazones**

- les N-nucléophiles ne pouvant être utilisés comme amines A : Les amines secondaires pour les raisons évoquées au paragraphe A-2-3-1, ne peuvent être utilisées comme amines A. De plus, les β-aminoalcools et les hydrazides. Dans ces deux cas, la formation de l'intermédiaire imidazolide, suivie d'une cyclisation intramoléculaire conduit à un composé stable, empêchant la formation ultérieure d'une urée (dans les conditions utilisées).

A l'exception de la N-éthylaniline, les amines secondaires ont réagi avec le CDI pour conduire aux imidazolides correspondants .

Les composés modèles n'ayant pas conduit aux urées attendues représentaient tous une famille de N-nucléophiles portant un atome d'azote très peu nucléophile : anilines fortement désactivées, 2-amino-1,3,4-thiadiazoles, 2-aminothiazoles. Nous devons cependant vérifier si ce défaut de réactivité compromettrait l'étape 1 ou l'étape 2.

B-5-2 Préparation d'urées asymétriques aux conditions de dilutions robotiques, à partir des représentants des familles de N-nucléophiles

Nous désirions d'une part, estimer les rendements en urées asymétriques (par rapport aux urées symétriques) obtenus avec des amines variées, d'autre part, compléter la sélection des familles de N-nucléophiles adaptées au protocole de synthèse robotisée.

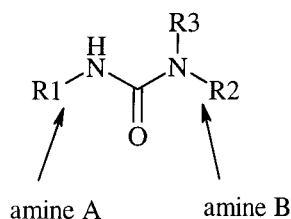
B-5-2-1 Adaptation du protocole pour la synthèse d'urées à partir d'anilines

Lorsque nous avons appliqué le protocole de synthèse d'urées asymétriques décrit ci-dessous à différents N-nucléophiles, les résultats obtenus se sont montrés décevants.

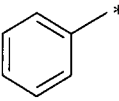
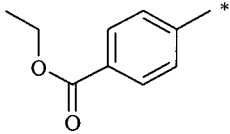
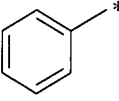
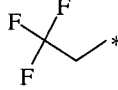
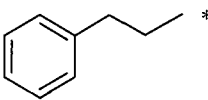
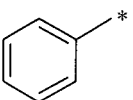
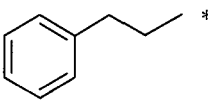
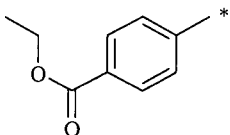
Protocole de synthèse :

A 2 mL d'une solution 0,25 M de CDI dans le THF, sont ajoutés 5 mL d'une solution 0,1 M d'amine A dans le DMF. Après 15 minutes de réaction sont ajoutés 5 mL d'une solution 0,1 M d'amine B dans le DMF. La réaction est prolongée 12 heures et les milieux réactionnels sont analysés.

Structure générale des composés synthétisés



	R1	R2	R3	ρ
(32)			H-	25 %
(33)			H ₃ C*	27 %
(34)			H-	65 %

(35)			H-	> 90 %
(36)			H-	70 %
(37)			H-	0 %
(38)			H-	0 %

ρ = rendement en urée asymétrique déduit de la conversion (estimée par HPLC) de l'amine A en urée asymétrique

Tableau 9 : rendements obtenus lors de la synthèse des urées asymétriques sans activateur de CDI

Les urées (32) et (33) ont été obtenues avec de bien plus faibles rendements que ceux déjà observés pour des urées similaires (cf. rendements de l'ordre de 90 %, paragraphe A-2), de plus les contaminants issus de la réaction des amines B avec le CDI (imidazolide ou urée symétrique) étaient présents dans de fortes proportions.

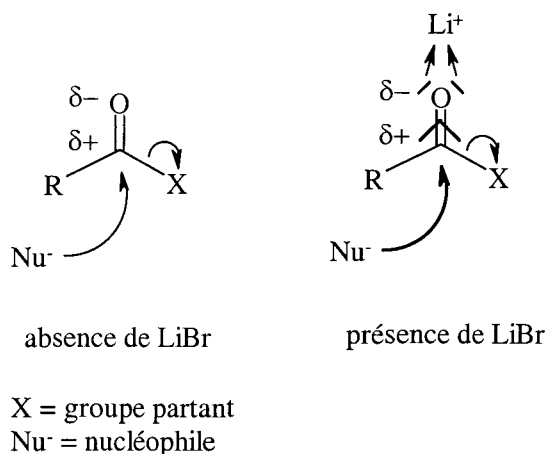
Le paramètre majeur ayant été modifié dans cette nouvelle étude, était la concentration des réactifs introduits. En effet, pour des raisons de solubilité, décrites au paragraphe B-2, les solutions d'amines ont été préparées à des concentrations de 0,1, et non plus de 0,5 M. Cette différence de concentration s'est traduite par une diminution de la vitesse de la réaction (ordre 2). Le CDI n'étant que partiellement consommé au moment de l'ajout de l'amine B, qui pour les urées (32) et (33), étant bien plus nucléophile que l'aniline, a conduit à l'imidazolide ou l'urée symétrique de l'amine B.

Dans le cas des urées (34) à (36), ce phénomène était moins marqué. En effet le différentiel de nucléophilie entre l'aniline et l'amine B étant plus faible, l'aniline continuait à réagir avec le CDI après l'introduction de la deuxième amine.

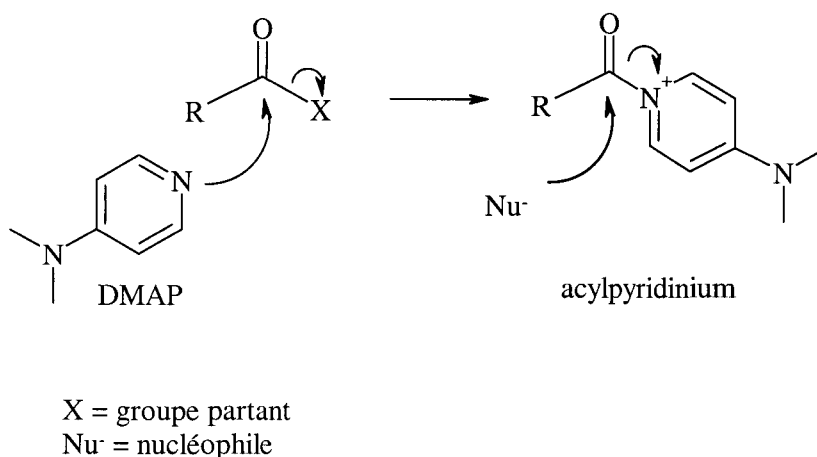
A ce stade de développement de la réaction, il nous était difficile de remplacer le CDI comme réactif de couplage. Pourtant, les anilines constituant une classe de N-nucléophiles fortement peuplée, nous devons les inclure dans notre future chimiothèque. L'utilisation des anilines comme amine B étant compromise (cf. produits (37) et (38)), nous avons donc cherché à optimiser la première étape en accentuant la réactivité du CDI vis à vis des anilines.

Deux activateurs potentiels de CDI ont donc été testés : le Bromure de Lithium (LiBr) et la N,N-diméthyl-aminopyridine (DMAP), ceux-ci ayant été décrits pour l'activation d'esters impliqués dans la synthèse de thioamides (Le 1995).

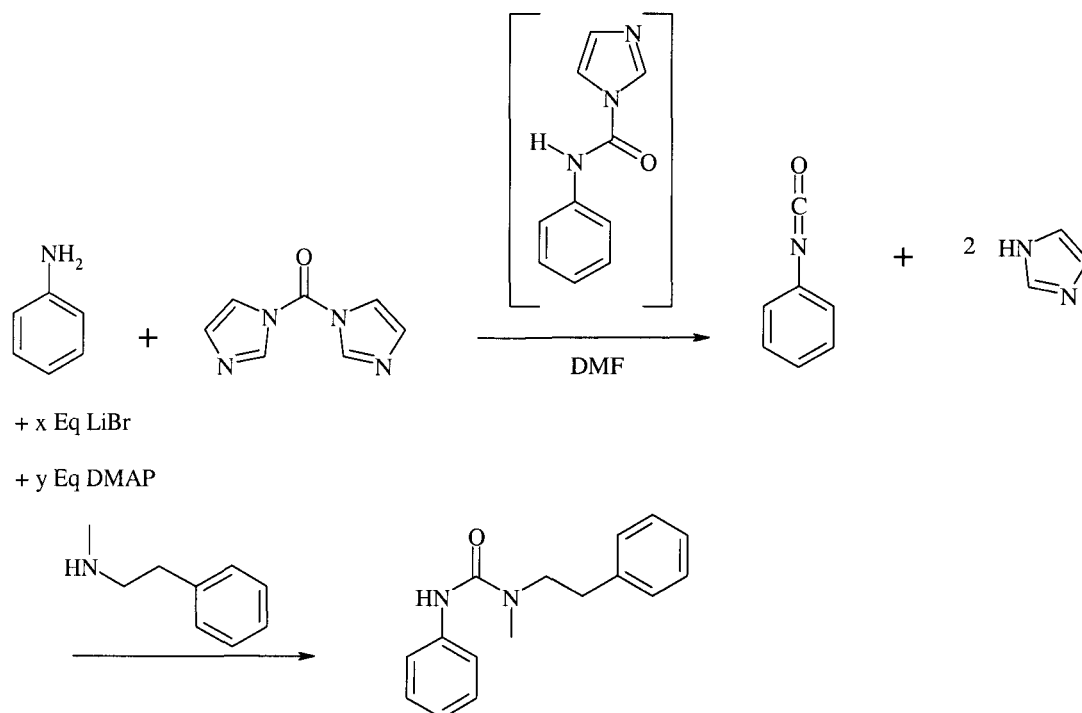
Le mécanisme d'action supposé du LiBr, repose sur une polarisation accentuée du groupement carbonyle, rendant l'atome de carbone correspondant plus sensible aux attaques nucléophiles.



Le mécanisme d'action supposé de la DMAP, implique la formation d'un intermédiaire très réactif du type acylpyridinium en présence d'esters activés.



Ces mécanismes appliqués au CDI (R = X= imidazole) laissaient espérer une activation de ce dernier. Nous avons alors étudié la formation de la N-Phényl-N'-méthyl-N'-(2-phényléthyl)urée en présence des catalyseurs précédents à différentes concentrations.



L'absence de l'imidazole de la N-méthylphénéthylamine (amine B) dans le milieu réactionnel final, devait nous permettre de vérifier la consommation totale du CDI avant l'addition de cette amine.

Résultats :

	DMAP (Eq)	LiBr (Eq)	ρ
Essai 1	0	1	74 %
Essai 2	0,1	0	36 %
Essai 3	0,1	0,1	54 %
Essai 4	0,1	1	84 %
Essai 5	1	0	66 %
Essai 6	1	0,1	93 %
Essai 7	1	1	> 95 %
Essai 8	1	5	> 95 %
Référence	0	0	23 %

ρ = rendement en urée asymétrique déduit de la surface du pic correspondant obtenu par HPLC (à 254 nm)

Tableau 10 : effets de la DMAP et du LiBr sur la formation d'une urée asymétrique issue de l'aniline en position A

Trois informations majeures sont déduites du tableau précédent :

- le LiBr comme la DMAP activent le CDI, quand ils sont utilisés seuls
- le couple LiBr / DMAP constitue un meilleur catalyseur que chaque réactif utilisé seul
- même si cette activation se manifeste pour de faibles concentrations (0,1 Eq) en ces activateurs, les rendements sont nettement améliorés lorsque ceux ci sont ajoutés de façon stoechiométrique par rapport aux amines de départ. Un effet activateur sensiblement supérieur est obtenu pour la DMAP, par rapport au LiBr (cf. essai 6 / essai 4).

Dans la majorité des cas, la présence de l'imidazolide de l'amine B a pu être détectée mais pour les essais 7 et 8, son absence, corrélée à la disparition des amines de départ, était cohérente avec les rendements obtenus.

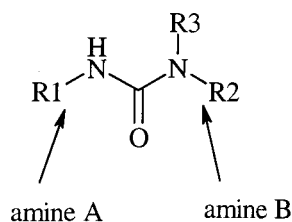
Conclusion du paragraphe B-5-2-1 :

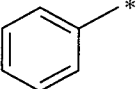
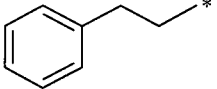
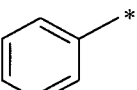
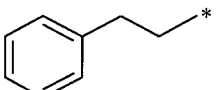
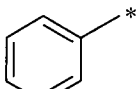
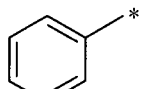

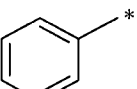

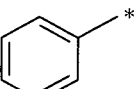
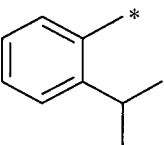
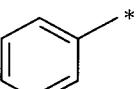
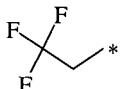
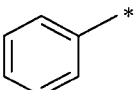
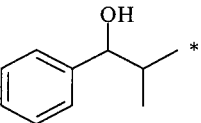
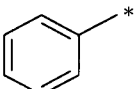
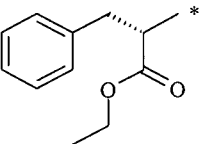
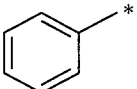
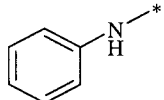
Le ralentissement de la cinétique réactionnelle impliquant l'aniline dans les conditions de dilutions indiquées pour la synthèse robotisée d'urées asymétriques (0,1 M dans le DMF) a pu être contourné par l'addition d'activateurs de CDI. Seul le protocole de mise en solution devait être modifié : une solution décimolaire de LiBr + DMAP dans le DMF allait être utilisée pour la mise en solution des anilines.

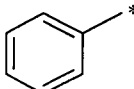
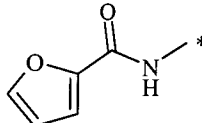
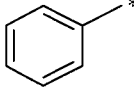
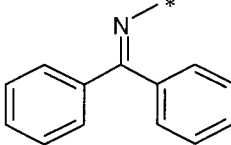
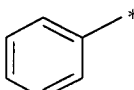
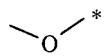
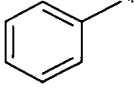
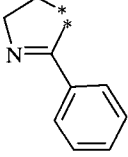
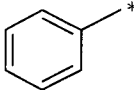
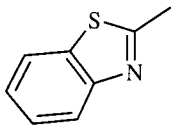
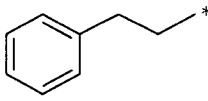
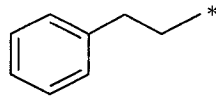
B-5-2-2 Détermination des N-nucléophiles réactifs

Le protocole de synthèse utilisé est identique à celui du § B-5-2-1 mais dans le cas où l'amine A est une aniline, celle-ci est préalablement dissoute par une solution LiBr + DMAP 0,1 M dans le DMF.

Structure générale des composés synthétisés



				Synthèse manuelle	Synthèse robotisée
	R1	R2	R3	ρ^1	ρ^1
(39)			H-	95 %	95 %
(40)			H_3C *	95 %	78 %
(41)				24 %	17 %
(42)			H-	49 %	52 %
(43)			H-	46 %	56 %
(44)			H-	46 %	52 %
(45)			H-	92 %	99 %
(46)			H-	42 %	53 %
(47)			H-	32 %	-

(48)			H-	60 %	100 %
(49)			H-	29 % (produit instable)	-
(50)			H ₃ C*	74 %	100 %
(51)				NC (instable)	-
(52)			H-	5 %	-
(53)			H ₃ C*	89 %	93 %

* : rendement déterminé par HPLC à partir d'une gamme étalon du produit attendu injecté en HPLC à différentes concentrations (seul le rendement du composé (49) correspond au taux de conversion de l'amine B en urée)

NC : non calculable

- : non synthétisée par le système robotisé

Tableau 11 : rendements obtenus lors de la synthèse des urées asymétriques (en présence de LiBr et de DMAP lorsque l'amine A est l'aniline)

Conclusion du paragraphe B-5-2-2 :

Le tableau ci-dessus révèle, comme nous pouvions l'attendre, l'influence de la nucléophilie et de l'encombrement stérique (au niveau de l'atome d'azote) sur la réactivité des différentes amines. Cependant, la composition des milieux réactionnels obtenus est restée très peu complexe (produits de départ et produit attendu), même lorsque les rendements obtenus

étaient médiocres. Cependant 3 familles de N-nucléophiles ont été écartées : les amines hétérocycliques (représentées par l'aminothiazole), la conversion obtenue étant beaucoup trop faible, ainsi que la famille des amidines et des hydrazones, du fait de l'instabilité des composés obtenus.

Sur la base de ces résultats nous avons établi des règles d'exclusion de N-nucléophiles inadaptés à notre protocole de synthèse robotisée d'urées :

- 1 - éviter les anilines (et amines hétérocycliques) désactivées en première et en deuxième position. Les anilines non désactivées peuvent être utilisées en première position mais 1 Eq de LiBr et 1 Eq de DMAP devront être ajoutés
- éviter les hydrazones en première et en deuxième position
- éviter les amines secondaires en première position (intermédiaire imidazolide stable ne pouvant réagir avec une amine)
- éviter les amidines et guanidines en première et en deuxième position
- 2 - éviter les hydrazides (et dérivés du type carbazides, semicarbazides...) en première position (cyclisation en oxadiazolinone)
- 3 - éviter les β et γ amino alcools en première position (cyclisation par carbamate interne)
- 4 - les esters méthyliques et éthyliques d'acides α aminés subissent une hydrolyse totale ou partielle (après évaporation du milieu réactionnel) par attaque de l'imidazole

B-6 Validation du protocole final par la synthèse robotisée de 11 urées de référence

Protocole :

Pour 5 μ moles de produit

- Placer 50 μ L de la solution de l'amine A 0.1M dans le DMF (ajouter 1 équivalent de DIEA par amine réactive protonnée)
- Ajouter 21 μ L de la solution de carbonyldiimidazole 0.247 M dans le THF, le titre de cette solution étant calculé par le test à la benzylamine (cf. partie expérimentale)
- Agiter 30 secondes
- Laisser réagir de 10 à 20 minutes
- Ajouter 50 μ L de la solution de l'amine B 0.1M dans le DMF (ajouter 1 équivalent de DIEA par amine réactive protonnée), agiter 1 minute puis laisser à température ambiante 10 h minimum
- Evaporer le DMF

Résultats : cf. tableau 11

Les résultats ont montré que notre protocole était transposable à notre système robotisé, les rendements obtenus lorsque les synthèses étaient effectuées avec le système robotisé étant comparables à ceux obtenus manuellement. La phase de production de chimiothèques pouvait débiter.

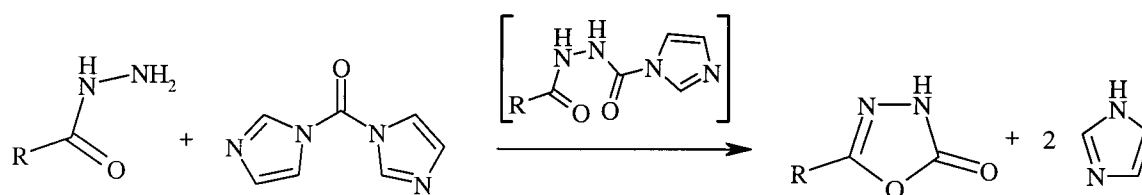
Remarque : dans la plupart des cas, les rendements obtenus pour des synthèses effectuées en plaques de microtitration (volume du milieu réactionnel 120 μL) sont supérieurs à ceux obtenus pour des synthèses à plus grande échelle (volume du milieu réactionnel 12 mL). Cet effet d'échelle est probablement lié à des échanges thermiques ainsi qu'à une évaporation des solvants s'opérant différemment.

Avant de décrire l'utilisation de notre protocole pour la synthèse de chimiothèques (voir paragraphe D), nous allons brièvement décrire l'intérêt potentiel d'une réaction secondaire observée lors de la mise au point de ce protocole.

C Intérêts potentiels d'une réaction secondaire conduisant aux oxadiazolinones (voir article à soumettre)

Lors de l'étape de détermination des classes d'amines réactives pouvant être incorporées dans la synthèse d'urées au CDI, une réaction secondaire est survenue lorsque des hydrazides étaient utilisées comme amines A. Des analyses HPLC réalisées sur des milieux réactionnels faisant intervenir différents hydrazides montraient, dans le cas d'hydrazides aromatiques, l'apparition d'un produit secondaire, une quantité importante d'amine B résiduelle et une consommation totale de l'hydrazide de départ (amine A). Dans le cas d'hydrazides aliphatiques, la réaction secondaire n'apparaissait pas systématiquement.

Le produit secondaire résultant d'une réaction consommant l'imidazolide de l'hydrazide (celui-ci n'étant plus disponible pour réagir avec l'amine B), nous avons tenté de reproduire cette réaction secondaire en faisant réagir l'hydrazide avec le CDI, sans ajout d'une seconde amine. La formation d'un produit correspondant au produit secondaire décrit précédemment a été observée. Celui-ci correspond à une cyclisation de l'intermédiaire imidazolide et conduit à l'hétérocycle oxadiazolinone après le départ d'imidazole :



Cette réaction est particulièrement favorisée lorsqu'elle fait intervenir des hydrazides aromatiques car elle aboutit à un système comportant deux cycles conjugués.

C-1 Intérêt biologique des oxadiazolinones

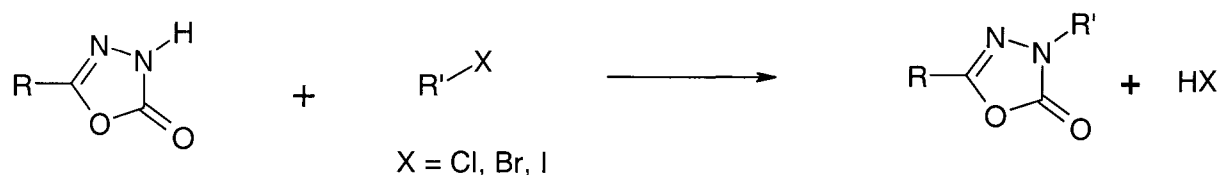
Après une étude bibliographique, nous avons remarqué que ces hétérocycles possédaient un intérêt biologique décrit dans le domaine de la chimie médicinale pour des activités tuberculostatiques (Stempel 1955), antibactériennes (Sherman 1961), ou inhibitrices de monoamines oxydases (Mazouz 1990), mais aussi dans le domaine de l'agrichimie pour des propriétés insecticides (Rüfenacht 1972).

C-2 Intérêt combinatoire des oxadiazolinones

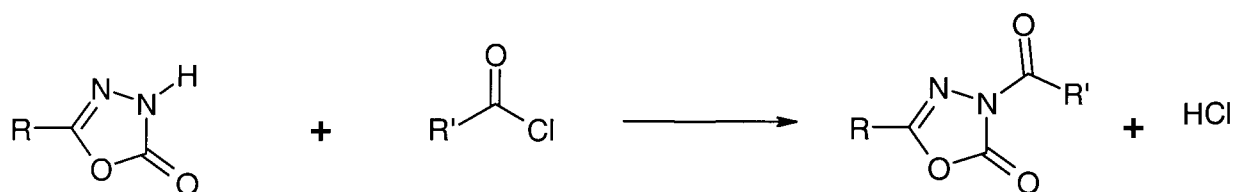
Une recherche effectuée dans ACD (available chemical directory) a révélé un grand nombre d'hydrazides (700) commercialement disponibles, de plus nous avons estimé à 10.000 le nombre d'acides carboxyliques pouvant être transformés en hydrazides.

Nous avons aussi remarqué que les oxadiazolinones pouvaient servir de squelette de base pour la synthèse de chimiothèques, l'atome d'azote en position 4 pouvant être alkylé, acylé ou intervenir dans des réactions du type Mannich, toutes ces réaction permettant d'introduire un nouveau site de diversité.

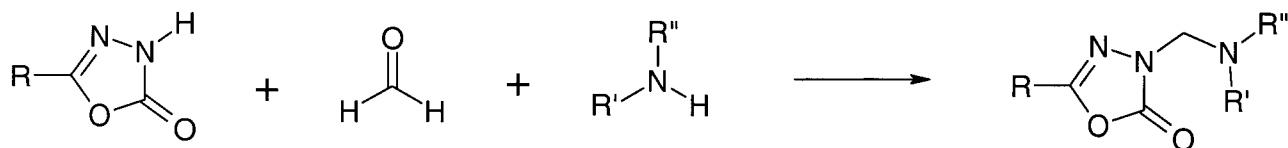
- réactions d'alkylation (Milcent 1991, Gogoi 1990)



- réactions d'acylation (Gogoi 1990)



- réactions de Mannich (Sherman 1961)



Ces constatations nous ont incité à mettre au point un protocole de synthèse robotisée d'oxadiazolinones en prenant en compte les contraintes relatives à toute chimie robotisée et plus particulièrement dans notre cas :

- l'obtention de rendements quantitatifs car la formation de l'hétérocycles sera suivie d'une seconde étape chimique
- l'utilisation d'un solvant de réaction non dangereux et permettant la mise en solution du plus grand nombre d'hydrazides

C-3 Protocole de synthèse robotisée d'oxadiazolinones

A un équivalent d'hydrazide 0.1 M dans le DMF sont ajoutés deux équivalents de 1,1'-Carbonyldiimidazole 0.25 M dans le THF, le milieu réactionnel est placé 12h à température ambiante. Des rendements quantitatifs sont observés dans la plupart des cas.

C-4 Conclusion

La réaction secondaire observée lors des phases exploratoires de synthèse robotisée d'urées nous a permis d'aboutir à un protocole de synthèse simple et robuste pour la synthèse d'oxadiazolinones. L'utilisation du CDI dans ce type de réaction a certes été décrite antérieurement mais le solvant de réaction utilisé était toujours le THF (pouvant limiter le nombre d'hydrazides solubles) et les conditions opératoires, souvent moins douces (utilisation fréquente de systèmes de chauffage) n'aboutissaient pas systématiquement aux excellents rendements obtenus dans nos conditions.

A mild and efficient method for the synthesis of Δ^2 -1,3,4-oxadiazolin-5-ones

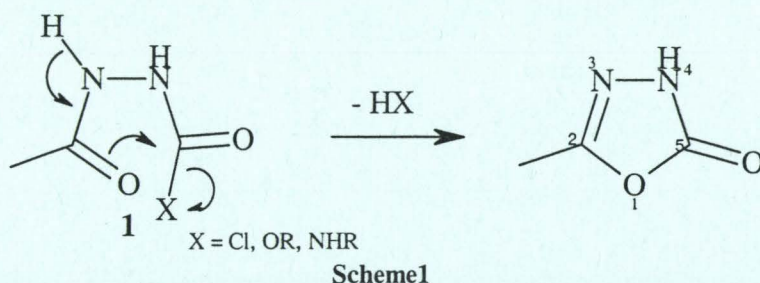
Xavier Williard*, Benoît Déprez and André Tartar

Department of Chemistry, CEREP, 1 rue Calmette, 59000 Lille, France

Abstract : A protocol suitable for the automated synthesis of Δ^2 -1,3,4-oxadiazolin-5-ones is described. These heterocycles are prepared from acylhydrazines and 1,1'-Carbonyldiimidazole. The cyclisation occurs *in situ*, at room temperature, with almost quantitative conversion.

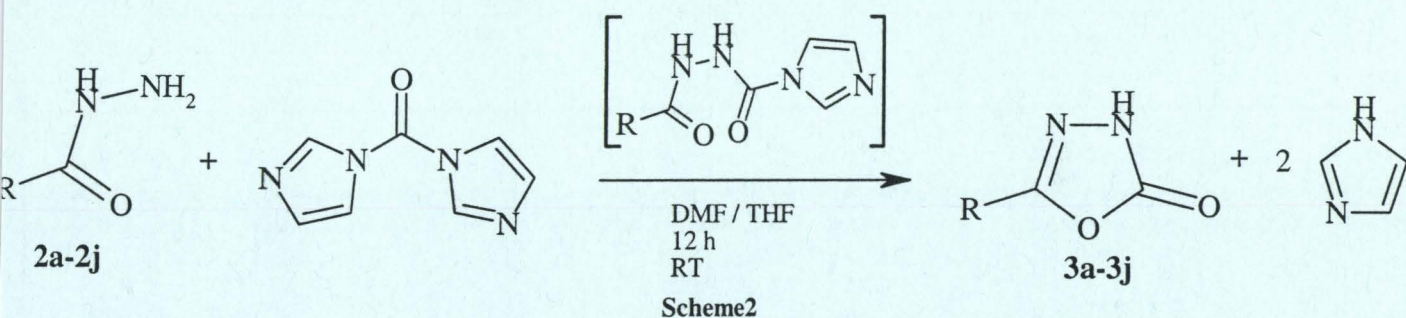
The first biological activity for oxadiazolinones was described in 1955 by Stempel et al.[1]. Tuberculostatic properties were observed for 4-pyridyl derivatives and were attributed to their metabolism to isonicotinoylhydrazide. In 1961, Sherman et al.[2] have prepared antibacterial nitrofurans joined, at the 2-position, directly to an oxadiazolinone ring and more recently insecticides [3], and monoamine oxidase type B inhibitors [4], based on this scaffold, have been reported. Our goal was to develop an easily automated procedure to prepare this heterocycle from a large variety of acylhydrazines. The most important route for the synthesis of 1,3,4-oxadiazoles involves the formation of the O-C(5) bond, usually by nucleophilic attack of the carbonyl oxygen of an N-substituted, N'-acyl hydrazine at the carbon atom which becomes C(5) in the 1,3,4-oxadiazole ring.

Moreover Δ^2 -1,3,4-oxadiazolin-5-ones are conveniently obtained from the thermal cyclisation of acylhydrazinecarboxylic acid derivatives (**scheme 1**).



For example, an acyl chloride (**1**; X = Cl), prepared from acylhydrazines and phosgene [5,6], is usually cyclized *in situ*.

In order to avoid the use of highly toxic phosgene, we performed this reaction in THF with triphosgene [bis(trichloromethyl) carbonate], a solid analogue of phosgene. The reaction of one equivalent of several acylhydrazines 0.1 M in THF with two equivalents of triphosgene 82 mM in THF led at room temperature, to the formation of a single new compound, as shown by HPLC of the reaction medium. These products were identified, by $^1\text{H-NMR}$ and mass spectrometry as the corresponding Δ^2 -1,3,4-oxadiazolin-5-ones. Although high yields were obtained, the scope of the reaction was limited because of the low solubility of most unfunctionalised acylhydrazines in THF. These acylhydrazines were much more soluble in DMF, however, triphosgene is rapidly hydrolysed in this solvent, preventing any automated protocol. This led us to evaluate carbonyldiimidazole as an alternative to triphosgene. Earlier publications [7-12], described the use of this reagent for the synthesis of oxadiazolinones but these reactions were always performed in THF. However when one equivalent of an acylhydrazine 0.1 M in DMF was reacted with two equivalents of 1,1'-Carbonyldiimidazole 0.25 M in THF (RT, 12h) (scheme2), we observed quantitative conversions in most cases.



The compounds were characterised by $^1\text{H-NMR}$, mass spectrometry and HPLC. Table 1 summarises yields obtained from a representative set of 9 compounds.

Compounds	R	% yield ^a 3a-3c	Compounds	R	% yield ^a 3d-3f	Compounds	R	% yield ^a 3g-3i
2a, 3a		95	2d, 3d		50	2g, 3g		81
2b, 3b		89	2e, 3e		83	2h, 3h		92
2c, 3c		86	2f, 3f		76	2i, 3i		90

Table 1 Synthesis of oxadiazolinones (3a-i)

^a: isolated products

For all products, the purity determined by HPLC is greater than 95%

In summary, we have developed a simple and easily automated procedure, using non-toxic reagents and permitting the conversion of acylhydrazines to a wide variety of α^2 -1,3,4-oxadiazolin-5-ones.

General procedure :

To a solution of an acylhydrazine 0.1M (0.5 mmol in 5 mL of DMF) was added 1,1'-Carbonyldiimidazole 0.25 M (1 mmol in 4 mL of THF). After 12h at room temperature, the reaction was quenched with water (1 mL) and evaporated to dryness in vacuo at 40°C. Ethylacetate (10 to 40 mL) was then added to the residues, washed with brine (4 x 10 mL), aqueous layers were extracted with ethylacetate (10 mL) washed with brine (4 x 10 mL). Organic layers were extracted with water (10 mL) and evaporated to dryness in vacuo at 40°C. When soluble in water, products were purified by reverse-phase HPLC (3d) or silica gel chromatography (3c).

Acknowledgements : We thank the Conseil Régional du Nord-Pas de Calais for the research grant of X. Williard, C. Landry for MS analysis and V. Thorel for HPLC analysis.

References and notes :

- 1] Stempel, A. ; Zelauskas, J. ; Aeschlimann, J. A. *J. Org. Chem.* **1955**, *20*, 412-418.
- 2] Sherman, W. R. *J. Org. Chem.* **1961**, *26*, 88-94.
- 3] Rüfenacht, K. *Helv. Chim. Acta* **1972**, *55*, 1174-1178.
- 4] Mazouz, F. ; Lebreton, L. ; Milcent, R. ; Burstein, C. *Eur. J. Med. Chem.* **1990**, *25*, 659-671.
- 5] Dornow, A. ; Bruncken, K. *Ber.*, **1949**, *82*, 121-123.
- 6] Lieser, T. ; Nischk, G. *Ber.*, **1949**, *82*, 527-530.
- 7] Tihanyi, E. ; Gal, M. ; Dvortsak, P. *Heterocycles.*, **1983**, *20*, 4, 571-574.
- 8] Davidson, J. S. *Monatsh. Chem.*, **1984**, *115*, 565-572.
- 9] Mullican, M. D. ; Wilson, M. W. ; Connor, D. T. ; Kostlan, C. R. ; Schrier, D. J., Dyer, R. *J. Med. Chem.*, **1993**, *36*, 8, 1090-1099.
- 10] Boschelli, D. H. ; Connor, D. T. ; Bornemeier, D. A. ; Dyer, R. D. ; Kennedy, J. A. *J. Med. Chem.*, **1993**, *36*, 13, 1802-1810.
- 11] Firoozi, F. ; Javidnia, K. ; Kamali, M. ; Fooladi, A. ; Foroumadi A. ; Shafiee, A. *J. Heterocycl. Chem.*, **1995**, *32*, 1, 123-128.
- 12] Brandstetter, T. W. ; Wormald, M. R. ; Dwek, R. A. ; Butters, T. D. ; Platt, F. M. *Tetrahedron Asymmetry.*, **1996**, *7*, 1, 157-170.
- 13] Spectra data of products **3a-3i**.

Product **3a** : $^1\text{H-NMR}$ (300 MHz, $\text{CDCl}_3/\text{DMSO}(d_6)$ 5/1) : δ 7.46 (3H, m), δ 7.97 (3H, m), δ 9.1 (1H, d, $J=8.44$ Hz), δ 12 (1H, s), EI-MS (m/z) 212(M+, 100%), 168(42%), 153(47%), 28(57%), 127(63%).

Product **3b** : $^1\text{H-NMR}$ (300 MHz, $\text{CDCl}_3/\text{DMSO}(d_6)$ 5/1) : δ 3.65 (1H, s), δ 6.75 (2H, d, $J=8.96$ Hz), δ 7.55 (2H, d, $J=8.96$ Hz), δ 11.75 (1H, s), EI-MS (m/z) 192(M+, 91%), 148(38%), 35(53%), 133(100%), 105(14%).

Product **3c** : $^1\text{H-NMR}$ (600 MHz, $\text{CDCl}_3/\text{DMSO}(d_6)$ 5/1) : δ 7.39 (1H, dd, $J=8, 4.9$ Hz), δ 8.05 (1H, ddd, $J=8, 1.74, 1.74$ Hz), δ 8.65 (1H, dd, $J=4.6, 0.9$ Hz), δ 9 (1H, s), δ 12.3 (1H, s), EI-MS (m/z) 163(M+, 100%), 119(72%), 106(25%), 104(12%).

Product **3d** : $^1\text{H-NMR}$ (300 MHz, CDCl_3) : δ 6.61 (1H, dd $J=3.54, 1.78$ Hz), δ 7.06 (1H, dd $J=3.52, 0.56$ Hz), δ 7.65 (1H, dd, $J=1.75, 0.71$ Hz), δ 10 (1H, s), EI-MS (m/z) 152(M+, 100%), 108(36%).

Product **3e** : $^1\text{H-NMR}$ (300 MHz, $\text{CDCl}_3/\text{DMSO}(d_6)$ 5/1) : δ 6.16 (1H, dd, $J=3.5, 2.6$ Hz), δ 6.65 (1H, dd, $J=3.7, 1.4$ Hz), δ 6.86 (1H, dd, $J=2.6, 1.4$ Hz), δ 11.1 (1H, s), δ 11.7 (1H, s), EI-MS (m/z) 151(M+, 100%), 149(17%), 107(11%).

Product **3f** : $^1\text{H-NMR}$ (300 MHz, $\text{CDCl}_3 + \text{D}_2\text{O}$) : δ 7.18 (1H, dd, $J=5.07, 3.7$ Hz), δ 7.56 (1H, dd, $J=5.07, 1.18$ Hz), δ 7.66 (1H, dd, $J=3.76, 1.15$ Hz), EI-MS (m/z) 168(M+, 100%), 124(26%), 111(55%), 109(12%).

Product **3g** : $^1\text{H-NMR}$ (600 MHz, $\text{CDCl}_3/\text{DMSO}(d_6)$ 5/1) : δ 5.08 (2H, s), δ 6.42 (1H, d, $J=3.5$ Hz), δ 6.79 (1H, d, $J=3.5$ Hz), δ 7.46 (1H, s), EI-MS (m/z) 302(0.8%), 300 (M+, 1.2%), 165(100), 121(11%), 109(33%), 106(12%).

Product **3h** : $^1\text{H-NMR}$ (300 MHz, $\text{DMSO}(d_6)$) : δ 4.83 (2H, s), δ 5.16 (2H, s), δ 6.53 (1H, s), δ 6.68 (1H, d, $J=2.8$ Hz), δ 7.75 (1H, s), δ 8.43 (1H, s), δ 12.75 (1H, s), EI-MS (m/z) 327(M+, 5%), 263(49%), 183(100%), 165(44%), 149(53%), 139(18%), 126(32%), 125(33%), 124(24%), 109(19%).

Product **3i** : $^1\text{H-NMR}$ (300 MHz, $\text{DMSO}(d_6)$) : δ 3.96 (2H, s), δ 7.07 (1H, dd, $J=4.9, 0.9$ Hz), δ 7.41 (1H, m), δ 7.54 (1H, dd, $J=4.9, 3$ Hz), δ 12.24 (1H, s), EI-MS (m/z) 182(M+, 100%), 140(37%), 125(40%), 112(10%), 111(11%), 110(11%).

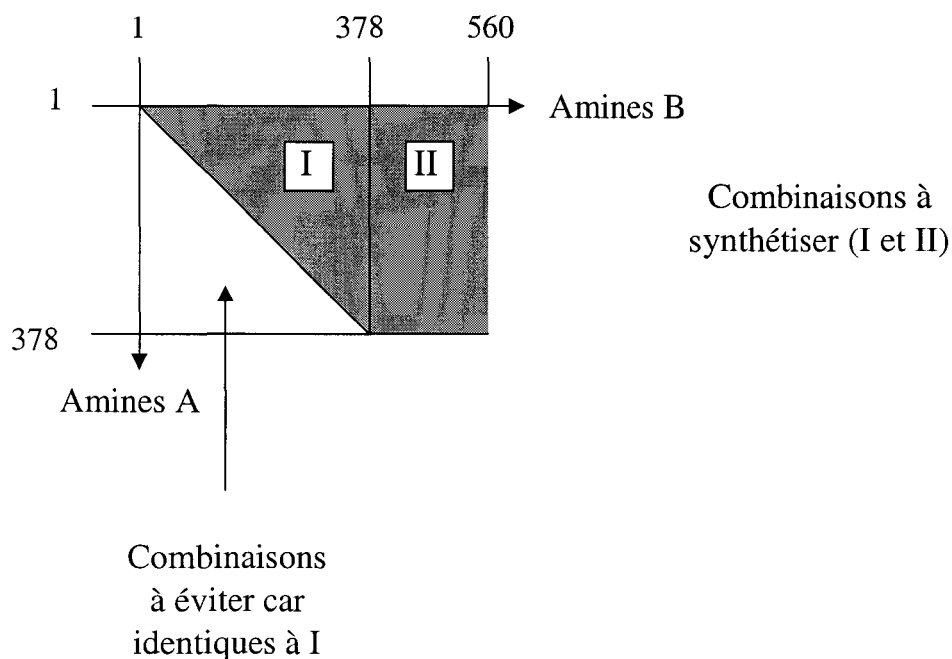
D Application du protocole de synthèse robotisée d'urées à la synthèse de deux chimiothèques

Nous décrivons ici deux chimiothèques d'urées, la première fût produite de manière exclusive pour un partenaire de l'industrie pharmaceutique et la seconde, réalisée en 10 exemplaires, pour divers clients de l'industrie pharmaceutique ou biotechnologique.

D-1 Synthèse d'une chimiothèque de 154720 urées pour un partenaire de l'industrie pharmaceutique

D-1-1 Choix des monomères et architecture de la chimiothèque

Nous disposions alors d'environ 1000 amines pures et solubles (dans les conditions fixées par le protocole). D'après les règles de réactivité et en utilisant les résultats d'une analyse de modélisation moléculaire ayant pour but de déterminer les monomères les plus divers (de masse moléculaire inférieure à 350 g / mole), 378 amines pouvaient être introduites en position A et B et 182 amines supplémentaires, uniquement en position B. L'ensemble de ces combinaisons est schématisé ci-dessous :



Afin de limiter la présence de « doublons » dans la chimiothèque, nous avons donc organisé sa synthèse de telle sorte que seule la demi matrice correspondant aux combinaisons issues des amines indifféremment introduites en position A ou B (378 amines) soit synthétisée, représentée par la surface I du schéma ci-dessus. A ces composés se sont ajoutées les combinaisons issues des amines introduites en position A et celles uniquement introduites en position B (182 amines) représentées par la surface II du schéma ci-dessus.

La synthèse a été réalisée dans des plaques 96 puits (8 x 12 puits) de type « deep-well » possédant des puits d'un volume de 1,2 mL. Seuls 80 puits par plaque ont été utilisés, 2 colonnes de 8 puits étant réservées pour la phase de criblage biologique. Dans chacun des 80 puits utilisés a été distribuée une même amine (l'amine A), chaque puits a ensuite reçu une amine différente (les amines B). Au sein d'une plaque, le partenaire constant était donc systématiquement l'amine A, les partenaires variables, les amines B. Chaque puits contenait théoriquement 6 μ moles du produit attendu.

Remarque : dans un souci de confidentialité, les structures des monomères utilisés ne seront pas décrites (hormis celles extraites du contrôle qualité).

D-1-2 Planification de la production

Nous avons réparti la synthèse sur 25 « jours » de production (les jours de production dépassant en général 24 heures). Chaque « jour » 80 plaques contenant chacune 80 composés différents (sauf pour J23 à J25 lors desquels 58 plaques seront produites), ont été produites cf. schéma ci-dessous :

B	Amines	Amines	Amines	Amines	Amines	Amines	Amines
A	1 à 80	81 à160	161 à 240	241 à 320	321 à 400	401 à 480	481 à 560
Amines 1 à 80	J1	J2	J3	J4	J5	J6	J7
Amines 81 à160		J8	J9	J10	J11	J12	J13
Amines 161 à 240			J14	J15	J16	J17	J18
Amines 241 à 320				J19	J20	J21	J22
Amines 321 à 378					J23	J24	J25

La production s'est concrètement étalée sur 3 mois.

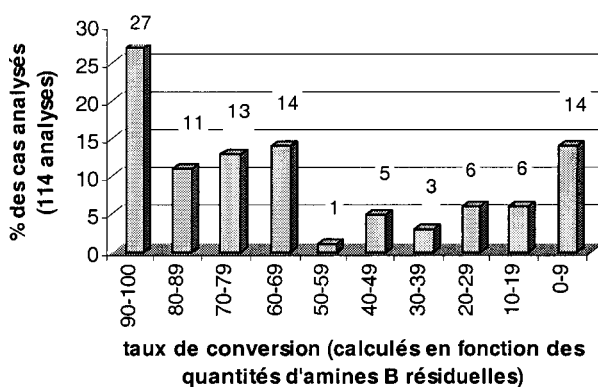
D-1-3 Contrôle qualité de la production

Nous avons alors considéré la phase de contrôle qualité comme une simple validation du système robotique, la composante réactivité chimique ayant été appréhendée précédemment lors de la détermination des classes de N-nucléophiles réactives.

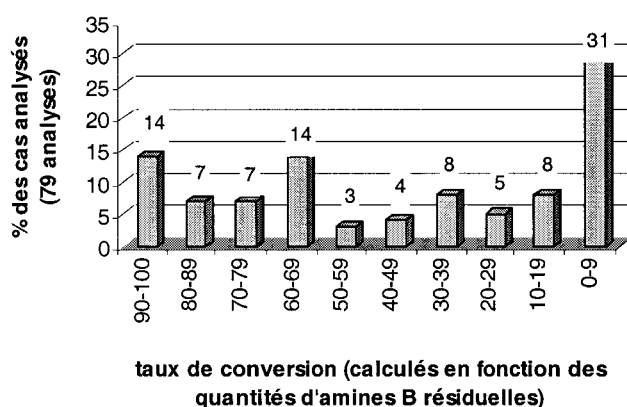
Nous avons donc contrôlé par HPLC un puits toutes les 10 plaques produites, les quantités en monomères résiduels étant déterminées à partir de gammes étalon. Comme nous pouvons sur les chromatogrammes ci-dessous, dans la plupart des cas, les milieux réactionnels étaient conformes à nos attentes : les milieux réactionnels étaient « propres » et les pics principaux identifiés (le plus souvent l'urée asymétrique était majoritaire et les contaminants principaux étaient les produits de départ et l'urée symétrique de l'amine A). Nous avons cependant été particulièrement déçus par les très faibles rendements obtenus lorsque des anilines orthosubstituées par un hétéroatome étaient utilisées. En effet, cette classe de N-nucléophile n'avait pas été validée car nous avons considéré que l'effet mésomère donneur devait permettre de renforcer la nucléophilie de ces anilines. Expérimentalement nous avons pu constater que l'effet inducteur attracteur se manifestait principalement ce qui rendait moins disponible le doublet de l'atome d'azote.

A l'issue de la production, nous avons donc scindé la chimiothèque en deux parties : la première partie impliquait des amines A non anilines. Les conversions obtenues, calculées à partir des quantités d'amines B résiduelles, étaient dans 66 % des cas supérieures ou égales à 40 %. La seconde partie impliquait des anilines en tant qu'amines A. Dans 45 % des cas les conversions étaient supérieures à 40 %, mais dans 30 % des cas de très faibles conversions (de 0 à 10 %) ont été obtenues, celles-ci correspondant aux urées issues d'anilines orthosubstituées par des hétéroatomes.

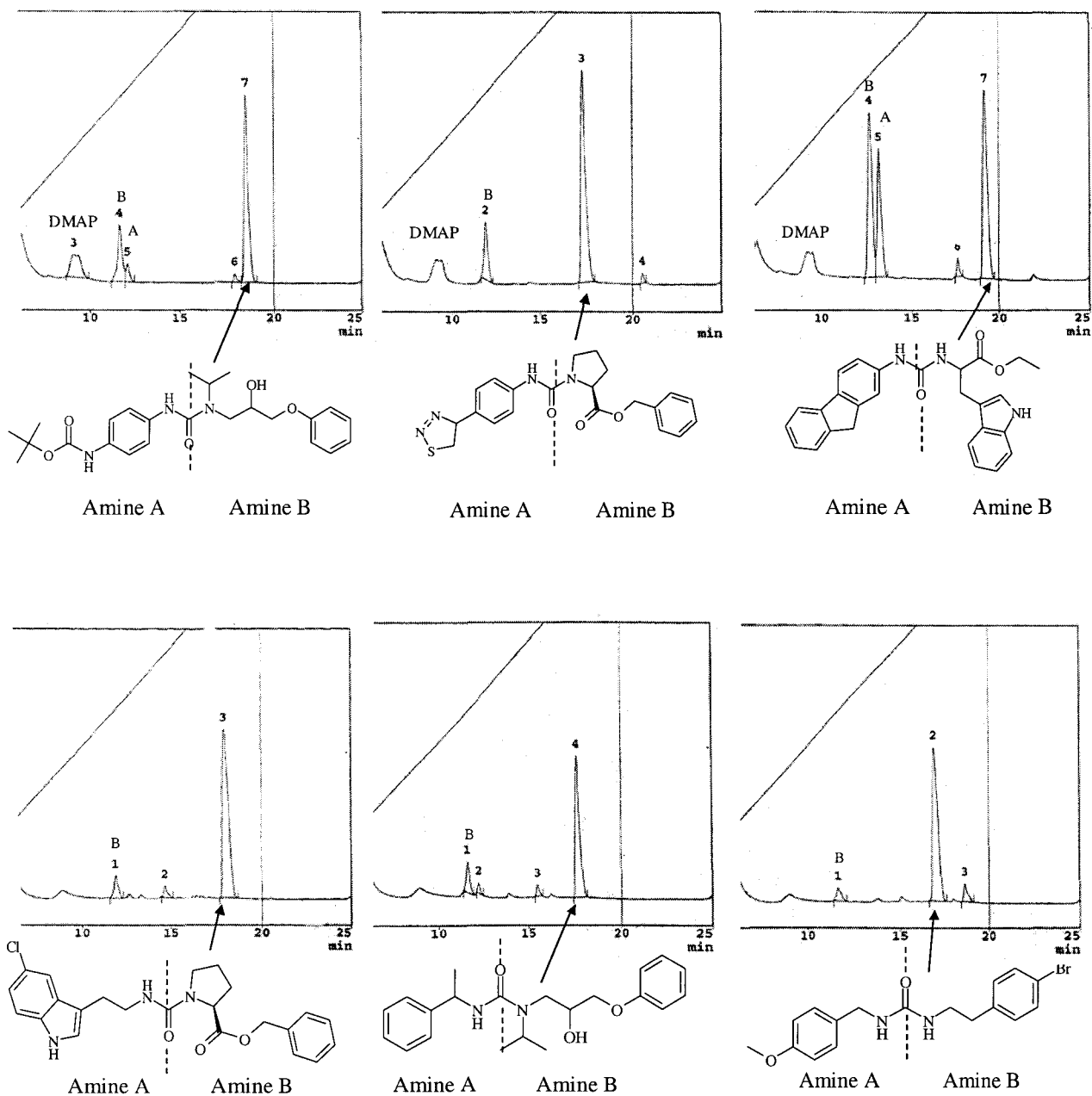
Taux de conversion en urées pour lesquelles l'amine A est une amine aliphatique



Taux de conversion en urées pour lesquelles l'amine A est une aniline



Quelques chromatogrammes issus du contrôle qualité figurent ci-dessous :

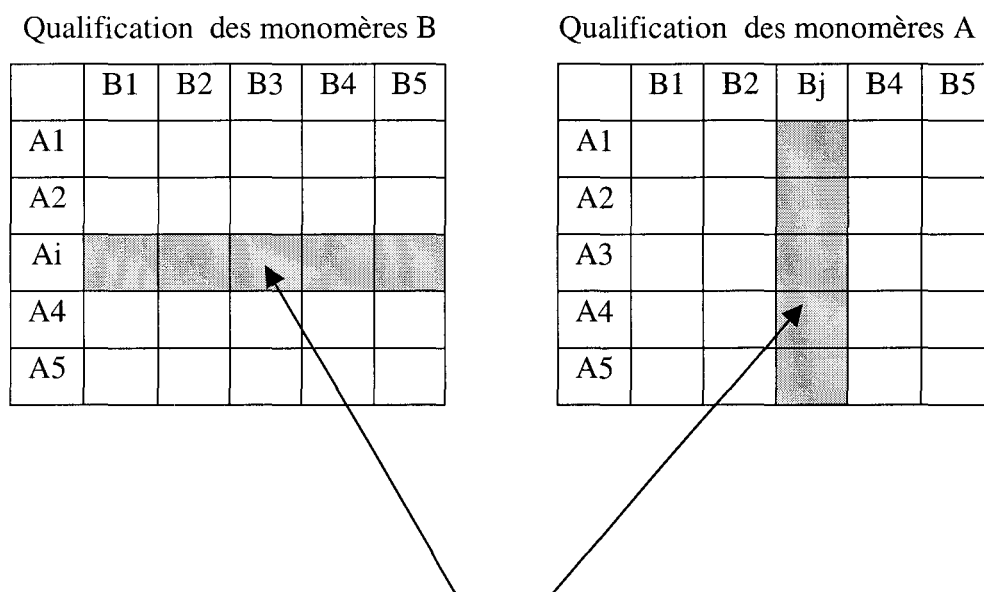


D-1-4 Améliorations à apporter

- La qualification des monomères

Au cours de la phase de contrôle qualité, nous avons pu constater que des N-nucléophiles appartenant à une même classe, telles que nous avons pu les définir, pouvaient avoir des réactivités très différentes. Nous avons donc décidé de tester systématiquement la réactivité de tout monomère pour la (les) réaction(s) intervenant dans la production de toute chimiothèque.

Ainsi, dans le cas d'une chimiothèque conduisant à des dimères A-B, à partir de n monomères A et m monomères B, un monomère A (A_i) choisi comme représentant sera mis en réaction avec chacun des monomères B. Les m couples A_i -B seront synthétisés, et seuls les monomères B donnant des conversions satisfaisantes seront retenus pour la synthèse de la chimiothèque. La même procédure sera appliquée à tous les monomères A afin de synthétiser les n couples A- B_j à partir du monomère B_j , représentant des monomères B. Toute la difficulté sera de choisir les meilleurs représentants A_i et B_j , suffisamment discriminant pour éviter que tous les monomères soient « qualifiés », mais aussi suffisamment réactifs afin d'éviter une sélection trop drastique des monomères.



Combinaisons synthétisées lors de la qualification

- **Le pourcentage de composés contrôlés lors d'une production**

Nous avons désiré accroître le nombre de puits contrôlés lors des productions ultérieures. En effet ce contrôle devait servir non seulement à vérifier le bon fonctionnement du système robotisé, mais aussi à mieux évaluer la qualité de la chimiothèque en terme de pureté. Dans notre première chimiothèque d'urées, nous n'avions contrôlé que 0,12 % de la totalité des composés. Nous nous sommes ensuite fixés comme objectif de contrôler un puits par plaque, soit 1,25 % de la totalité des composés. Notre système analytique était inadapté à cet objectif. En effet, une analyse HPLC (sur colonne de 20 cm) s'effectuait alors en 25 minutes soit une capacité maximale d'analyse d'environ 60 composés par jour. Nous avons alors optimisé notre système en utilisant des colonnes courtes (colonne de 5 cm) contenant des phases de

faible granulométrie (2 μm). Nous avons alors réduit nos durées d'analyse à 6 minutes par composé, soit une capacité de 240 analyses par jour.

- *L'ajout d'un critère d'identité au critère de pureté apporté par l'HPLC*

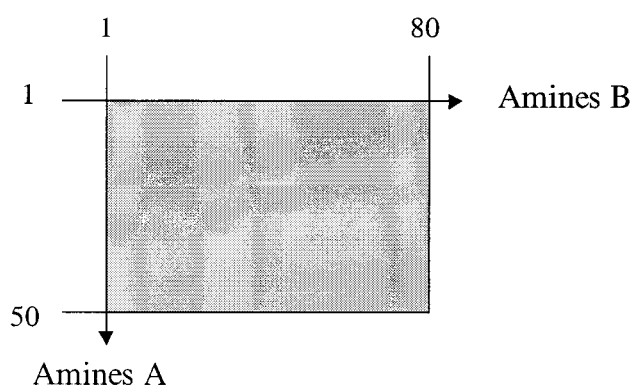
L'acquisition d'un système LC-MS nous a permis, en conservant nos méthodes d'analyse HPLC, d'enrichir la caractérisation des produits contrôlés par l'accès aux spectres de masses de tous les composés séparés au cours du processus chromatographique.

D-2 Synthèse d'une chimiothèque de 4000 urées pour des clients de l'industrie pharmaceutique ou biotechnologique

D-2-1 Choix des monomères et architecture de la chimiothèque

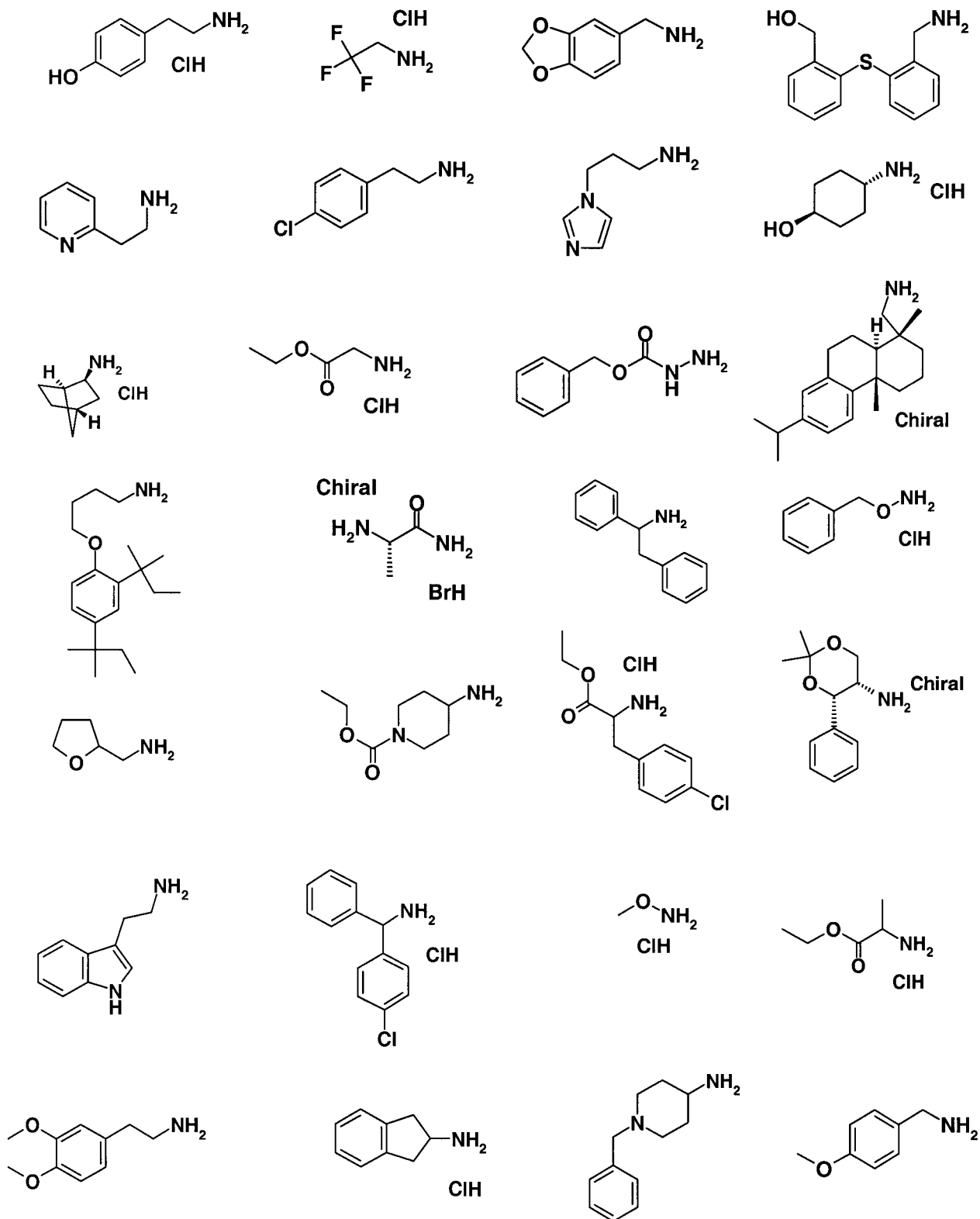
Le but était ici de produire une chimiothèque de petite taille, synthétisée sur la base de monomères divers (diversité évaluée par modélisation moléculaire), purs, solubles, de masse moléculaire inférieure à 350 g / mole et dont la réactivité a été systématiquement contrôlée lors d'une étape de qualification (cf. paragraphe précédent).

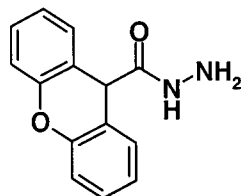
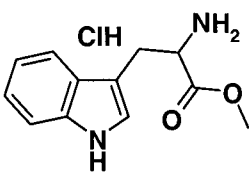
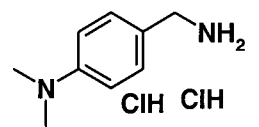
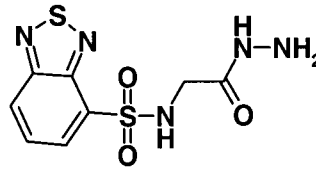
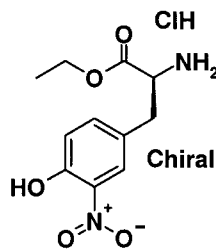
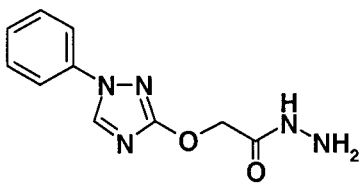
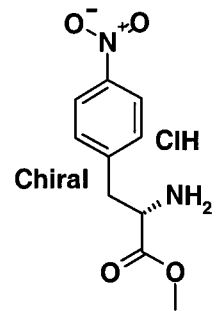
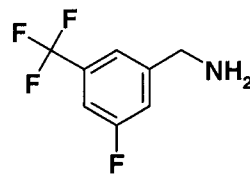
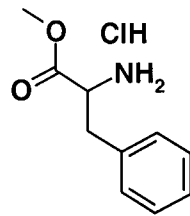
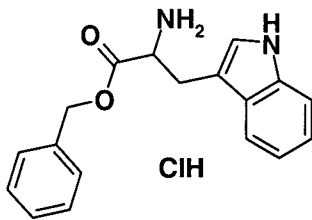
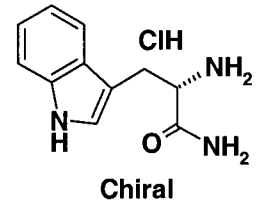
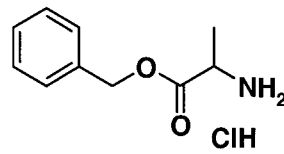
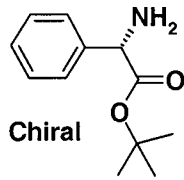
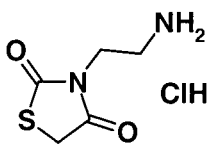
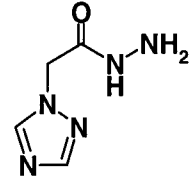
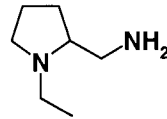
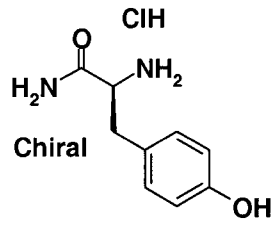
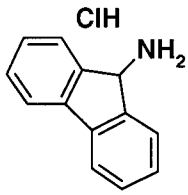
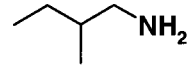
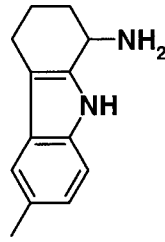
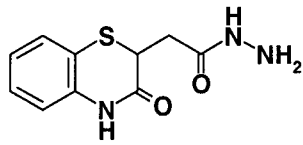
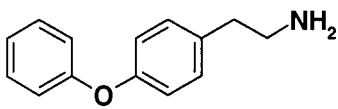
Cinquante amines ont été introduites en position A et quatre-vingt en position B. Aucun monomère ne figurant à la fois en position A et en position B, la matrice complète a pu être réalisée sans conduire à des doublons. La masse moléculaire moyenne des composés obtenus était de 352 g / mole.

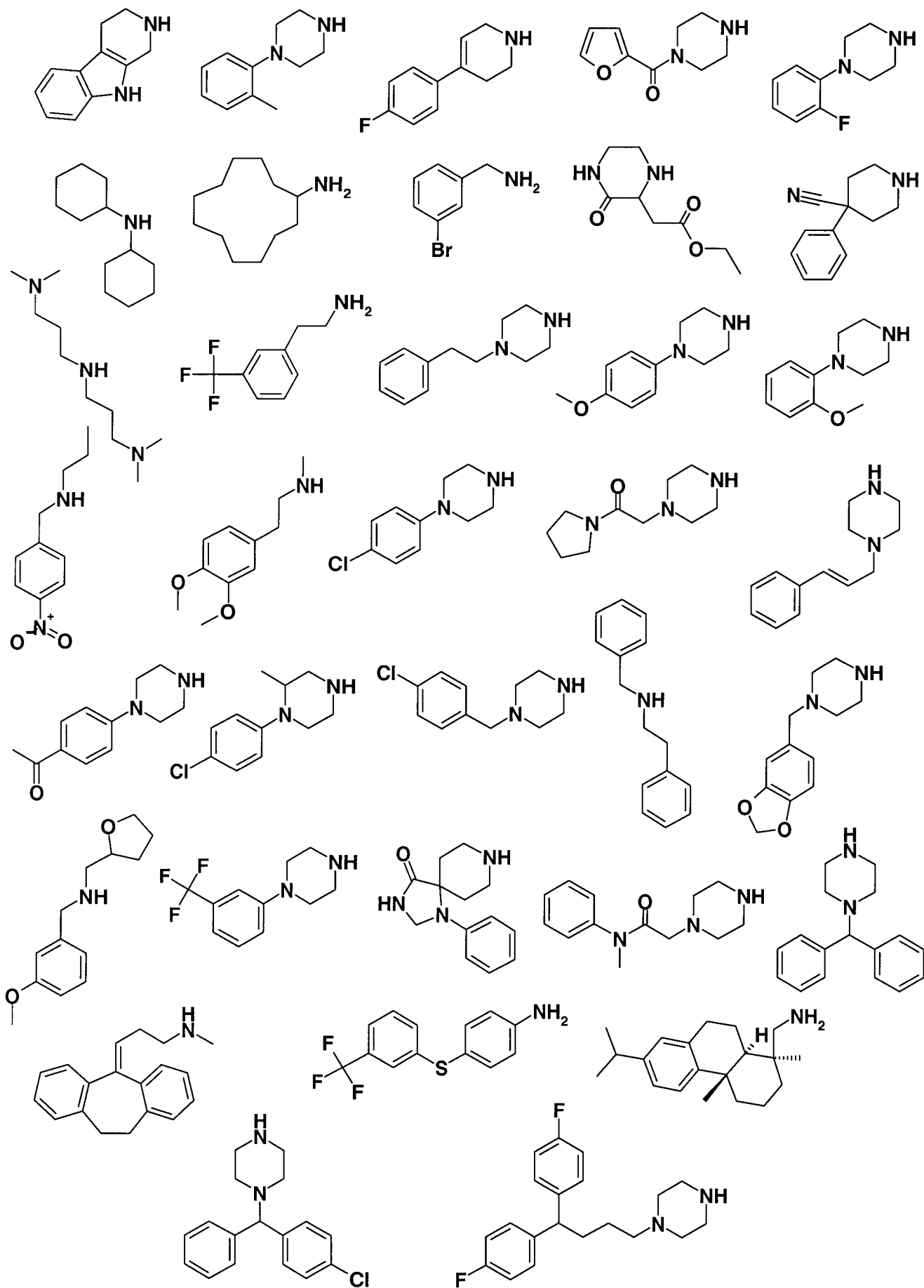


La synthèse a été réalisée dans les mêmes conditions que la chimiothèque précédente (échelle de synthèse, plaques deep-well, protocole de synthèse...). Au sein d'une plaque, le partenaire constant était systématiquement l'amine A, les partenaires variables, les amines B.

Structures des amines A





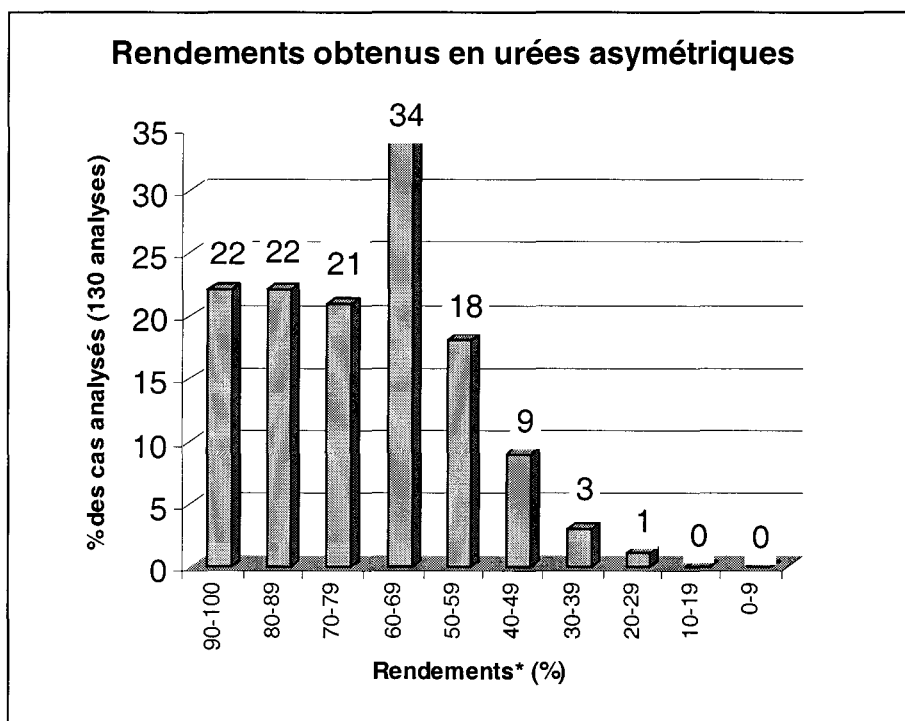


D-2-2 Planification de la production

Nous avons réparti la synthèse sur 6 « jours » de production. Chaque « jour » 50 plaques contenant chacune 80 composés différents ont été produites. Les produits ont été synthétisés en 6 exemplaires, J1 = J2 = J3 ... = J6.

D-2-3 Contrôle qualité de la production

Afin de contrôler, au moins une fois, tout monomère engagé dans la chimiothèque, un puits par plaque a systématiquement été analysé par LCMS permettant ainsi le contrôle de toutes les amines A (amine A = amine commune à tous les puits d'une même plaque, mais différentes d'une plaque à l'autre). En ce qui concerne les amines B (amines constantes d'une plaque à l'autre mais différentes d'un puits à l'autre au sein d'une même plaque), celles-ci ont été contrôlées en analysant par LC-MS les 80 puits d'une plaque prise au hasard.



*calculés en rapportant l'aire du pic HPLC correspondant au produit attendu (déterminé grâce au spectre de masse) sur l'aire de tous les pics

Grâce à l'étape préalable de qualification des monomères, les rendements en urées asymétriques étaient nettement supérieurs à ceux obtenus lors de la synthèse de la première

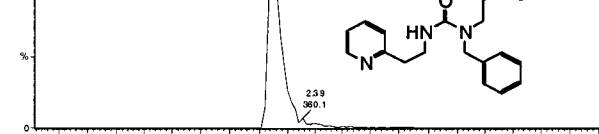
chimiothèque d'urées. En effet dans 96 % des cas analysés, les rendements étaient supérieurs à 40 %. Les analyses ont montré des milieux réactionnels dont tous les pics étaient facilement identifiables, les produits secondaires étant principalement des monomères résiduels ou plus rarement les urées symétrique des amines A. L'histogramme ci-dessus met en évidence une majorité de rendements compris entre 60 et 69 %. Cette proportion est surestimée par les résultats obtenus lors des tests des 80 puits d'une même plaque. En effet, la nature de l'amine A nivelle dans la plupart des cas les rendements obtenus. Pour une amine A réactive, les rendements seront élevés quelque soit la nature des amines B, de même, pour une amine A peu réactive, les rendements obtenus au sein d'une même plaque seront faibles. Ainsi il est préférable de contrôler par LC-MS au minimum un puits par plaque (contrôle de toutes les amines A) car cette méthode permet de balayer l'ensemble de la production, afin d'éliminer, le cas échéant, les plaques ne respectant pas les limites imposées lors du contrôle qualité.

Quelques analyses issues du contrôle qualité figurent ci-dessous, les 6 premières (a) correspondent à l'analyse d'un puits de 6 plaques différentes, les 4 dernières (b) correspondent à 4 puits d'une même plaque. Chaque analyse est composée de deux parties, la partie inférieure correspond au chromatogramme HPLC avec une détection UV à barrette de diodes, la partie supérieure correspond au courant ionique obtenu lors de la recherche de l'ion correspondant à la masse moléculaire + 1 du produit attendu.

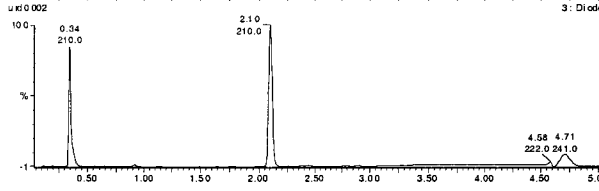
a)

CQ UR DEC %39501

ur00002



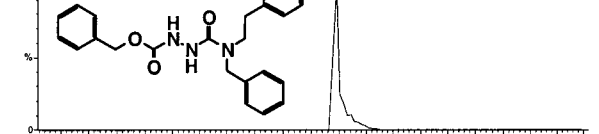
1,1:B,2
1: Scan AP+
360.50
4.18e6



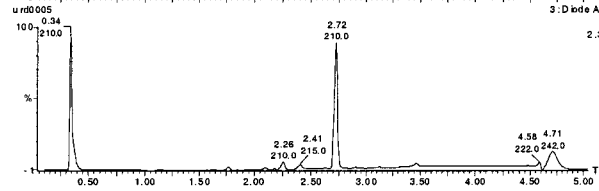
3: Diode Array
TIC
3.27e7

CQ UR DEC %39501

ur00005



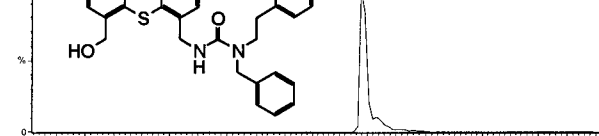
1,1:B,5
1: Scan AP+
404.50
4.18e6



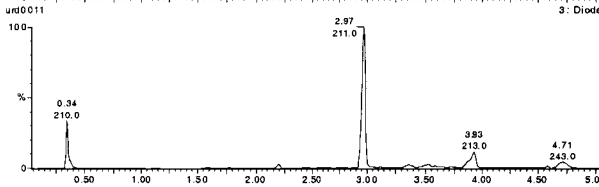
3: Diode Array
TIC
2.30e7

CQ UR DEC %39501

ur00011



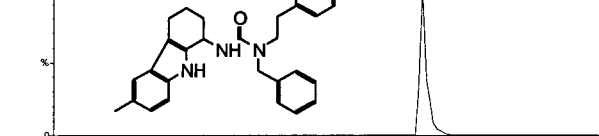
1,1:F,1
1: Scan AP+
483.70
2.06e6



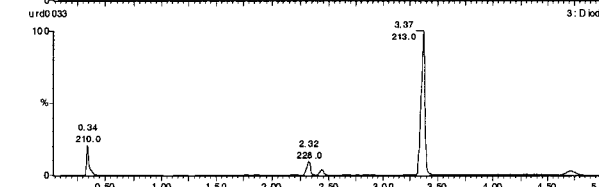
3: Diode Array
TIC
6.53e7

CQ UR DEC %39502

ur00033



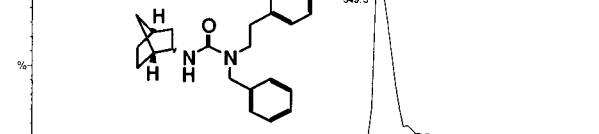
1,2:F,3
1: Scan AP+
438.60
3.21e6



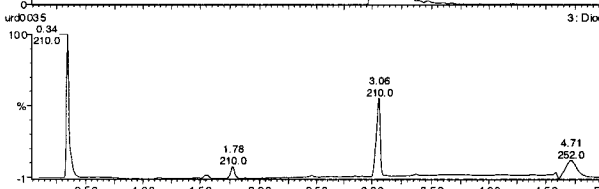
3: Diode Array
TIC
1.08e8

CQ UR DEC %39502

ur00035



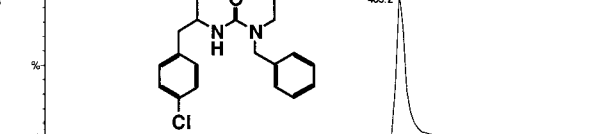
1,2:F,5
1: Scan AP+
349.50
4.18e6



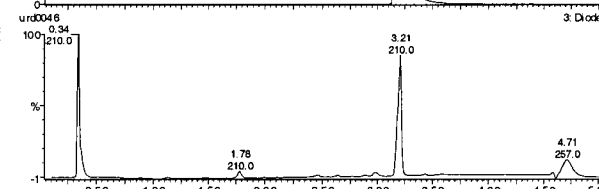
3: Diode Array
TIC
2.83e7

CQ UR DEC %39503

ur00048



2,1:F,1
1: Scan AP+
466.50
3.57e6

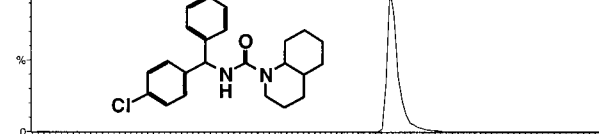


3: Diode Array
TIC
2.61e7

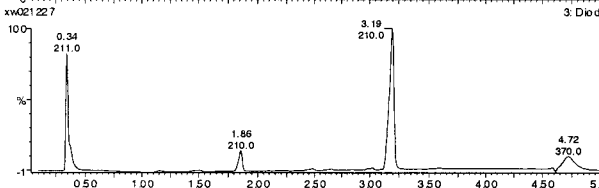
b)

CQ UR DEC %39484

xw021227



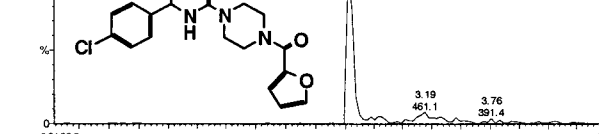
5,1:B,5
1: Scan AP+
364.40
2.87e6



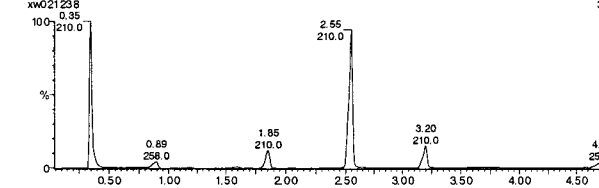
3: Diode Array
TIC
3.15e7

CQ UR DEC %39484

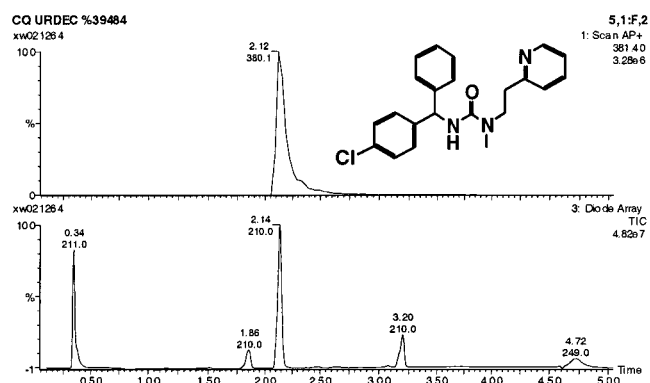
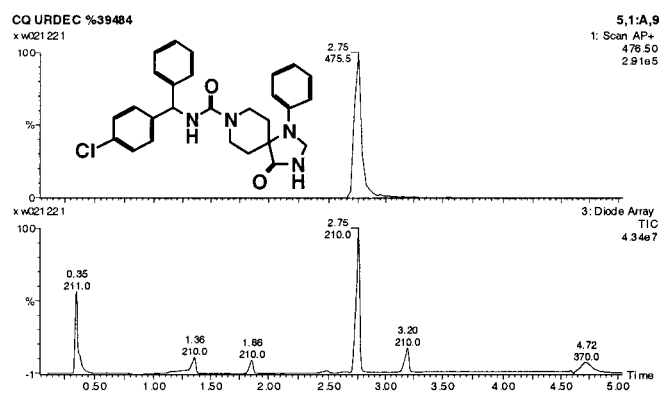
xw021238



5,1:C,6
1: Scan AP+
425.40
3.02e4



3: Diode Array
TIC
7.90e7



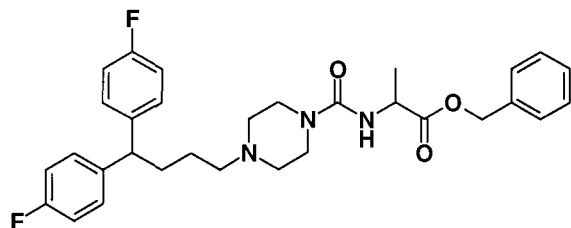
D-2-4 Résultats de tests biologiques

Afin d'évaluer le potentiel de cette chimiothèque d'urées en terme de « taux de succès » (ou hit rate) sur des cibles biologiques, 3120 composés ont été testés sur le récepteur morphinique μ et le récepteur benzodiazépinique BZDc à une concentration de 10^{-5} M (Eau/DMSO-100/0,1). Les taux de succès obtenu, pour des inhibitions supérieures ou égales à 80 % d'inhibition, étaient de 0,16 % sur le récepteur μ et de 0,38 % sur le récepteur BZDc, soit respectivement 5 hits et 12 hits (tous confirmés lors d'un essai de confirmation d'activité dans les mêmes conditions) ce qui constitue une base de travail suffisante pour de futures optimisations. Même si les produits actifs possèdent entre eux des analogies structurales (un même monomère est souvent partagé par plusieurs produits), différentes pistes pourront toutefois être suivies. Dans le cas des hits μ par exemple, la première piste est composée de produits protonnables, contrairement à la seconde.

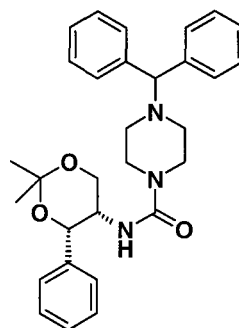
inhibition	Hit rate (%) à 10^{-5} M	
	μ opiacé	BZDc
$\geq 50\%$	3,09	2,88
$\geq 70\%$	0,49	0,83
$\geq 80\%$	0,16	0,38
$\geq 90\%$	0	0,13

Hit rate obtenus lors du criblage de 3120 urées sur les récepteurs μ et BZDc

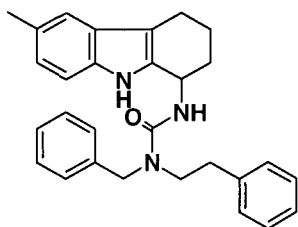
Exemples d'activités sur le récepteur μ , trouvées dans les puits contenant les produits attendus suivants:



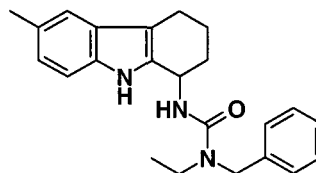
Inhibitions : 72 % et 86 % à $10^{-5}M$



Inhibitions : 88 % et 80 % à $10^{-5}M$

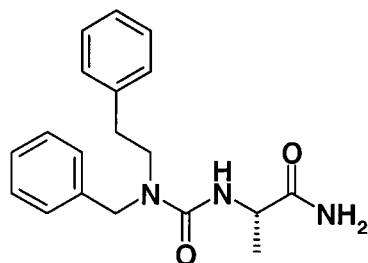


Inhibitions : 80 % et 81 % à $10^{-5}M$

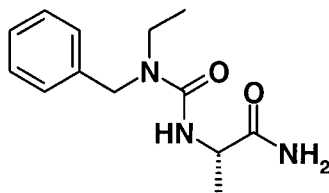


Inhibitions : 84 % et 86 % à $10^{-5}M$

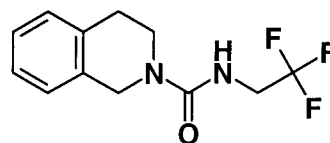
Exemples d'activités sur le récepteur BZDc, trouvées dans les puits contenant les produits attendus suivants:



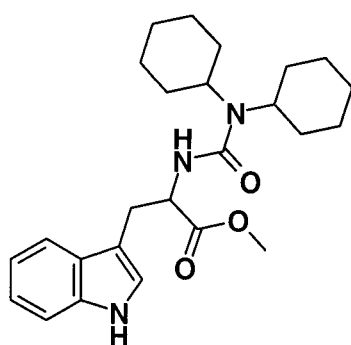
Inhibitions :
87 % et 88 % à $10^{-5}M$
 IC_{50} : 0,8 μM



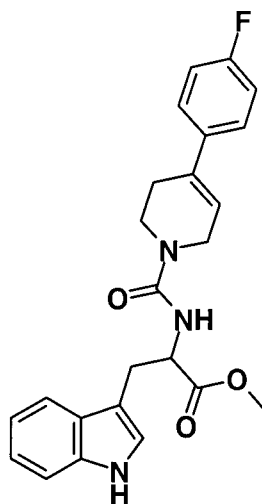
Inhibitions :
85 % et 90 % à $10^{-5}M$
 IC_{50} : 3,7 μM



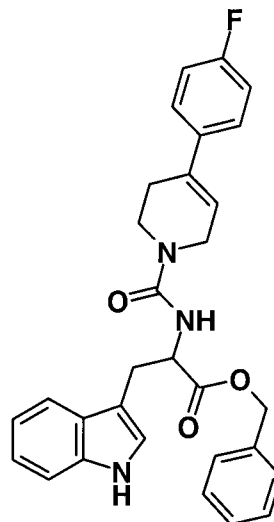
Inhibitions :
82 % et 71 % à $10^{-5}M$



Inhibitions :
84 % et 81 % à $10^{-5}M$



Inhibitions :
89 % et 89 % à $10^{-5}M$



Inhibitions :
84 % et 84 % à $10^{-5}M$

III Conclusion de la seconde partie

Nous avons démontré grâce aux chimiothèques d'urées synthétisées sur nos robots, qu'il était possible de produire rapidement (plusieurs milliers de composés par jour), à partir d'une réaction de dimérisation, un grand nombre de composés. La phase la plus lente du processus a, sans équivoque, été la phase de mise au point de la chimie, qui a nécessité environ une année équivalent temps plein. De nombreuses contraintes imposées par les capacités du système robotisé ont dû être considérées afin d'aboutir à un protocole général de synthèse. La nature du réactif de couplage et des solvants, les temps de réaction, le contrôle des matières premières et des réactifs, la concentration des milieux réactionnels ... devaient être identiques

pour l'ensemble de la production ce qui a constitué la difficulté majeure de ce programme (il est parfois nécessaire de segmenter une production en fonction de conditions opératoires différentes mais l'organisation de la production devient alors complexe et s'adresse généralement à la production de chimiothèques de petites tailles, pour lesquelles des rendements excellents (> 90 %) sont attendus). Sur la base de ce protocole de synthèse, en apportant quelques modifications mineures (temps de réactions, nombre d'équivalents de CDI introduits), nous avons également mis au point les protocoles de synthèse automatisée de carbamates et de carbonates, mais du fait de leur forte analogie avec celui des urées, ceux-ci ne figureront pas dans cette thèse.

Les problèmes associés à la synthèse de composés en mélange, évoqués dans la première partie : le suivi complexe des composés en cours de synthèse, la difficulté d'identifier le(s) composé(s) conférant l'activité à un mélange et le format de synthèse inadapté aux tests de criblage, ont tous pu être contournés. Nous avons d'autre part mis en évidence des activités biologiques avec un taux de succès moyen de 0,27 % (nombre de puits actifs / nombre de puits testés) sur deux cibles biologiques différentes ce qui valide notre approche de découverte d'activité biologique, par le biais de chimiothèques produites par synthèse parallèle à haut débit. Une étude plus approfondie des différentes activités observées sur les cibles biologiques (μ et BZDc) ainsi que l'optimisation de ces activités figurent dans la thèse de Rébecca Poulain soutenue le 30 Septembre 1999 à l'Université des Sciences et Technologies de Lille.

Conclusion générale

CONCLUSION GENERALE

Les travaux réalisés dans cette thèse ont démontré qu'il était possible, en criblant de façon systématique un grand nombre de composés, de trouver des activités biologiques portées par des composés dont la structure ne pouvait en rien présager de telles propriétés. La synthèse peptidique en phase solide a été utilisée pour la synthèse de notre première chimiothèque de 15625 tripeptides (synthétisés en 125 mélanges de 125 composés). La formation du lien amide ainsi que les réactions secondaires pouvant intervenir sur les chaînes latérales de divers acides aminés étant parfaitement documentées, nous avons considéré qu'une sérieuse assurance qualité menée tout au long de la production devait nous permettre d'obtenir une chimiothèque contenant la majorité des composés attendus dans des proportions semblables d'un composé à l'autre. L'identification structurale d'un composé actif, présent dans un des mélanges composant cette chimiothèque, par le biais de chimiothèques orthogonales nous a permis de valider cette nouvelle méthode de déconvolution. Cependant la synthèse de composés en mélange présente des limites difficilement contournables telles que : le suivi analytique de chacun des composés synthétisés, l'apparition de réactions secondaires, l'identification parfois fastidieuse du composé conférant l'activité au mélange, la présence de cumuls d'activités ...

Nous nous sommes donc orientés vers la synthèse parallèle à haut débit. Cette technique demeure étroitement liée à d'autres techniques environnantes, sans lesquelles celle-ci n'aurait pu voir le jour. En effet, la fin du vingtième siècle reste clairement marquée par les progrès de l'automatisme, de l'informatique (pilotage d'automates, stockage de données, traçabilité), de la miniaturisation ainsi que de l'avènement de nouvelles techniques analytiques comme la LC-MS. Les progrès réalisés parallèlement dans tous ces domaines ont largement contribué au succès de notre projet. La phase la plus lente du processus a consisté en la mise au point de conditions opératoires générales, pour la synthèse d'urées asymétriques, transposables à notre système robotisé. De nombreux paramètres intervenant dans cette réaction ont dû être parfaitement maîtrisés avant d'entrer en phase de production. Une fois la production terminée, notre projet ne pouvait être validé sans l'obtention de composés doués d'activités biologiques. Nous avons donc testé une chimiothèque composée de 3120 urées sur deux cibles biologiques (le récepteur morphinique μ et le récepteur benzodiazépinique BZDc). Le pourcentage de

composés actifs au sein de cette chimiothèque était tout à fait satisfaisant : 0,16 % pour μ et 0,38 % pour BZDc, ce qui nous a permis de conclure à l'intérêt certain des chimiothèques combinatoires pour la découverte d'activités biologiques.

Actuellement, l'optimisation de l'efficacité de cette stratégie est menée sur trois fronts.

Le premier consiste à augmenter la pureté des composés testés lors des phases de criblage afin de limiter toute incertitude quant à l'identité de la structure du composé actif. La tendance actuelle se focalise donc sur les techniques de purification à haut-débit, qui ne sont qu'à leurs premiers balbutiements, au regard des capacités de synthèse.

Le second consiste à augmenter le taux de succès en produits actifs au cours des phases de criblage mais aussi à augmenter les chances que ceux-ci possèdent toutes les qualités requises pour constituer de futurs médicaments (composés biodisponibles, dépourvus de toxicité ...). Dans cette optique, le choix des matières premières est donc capital. Celles-ci devront permettre d'accéder facilement à des analogues des composés actifs (afin d'optimiser leur activité et leur sélectivité) et être pourvues des propriétés physicochimiques susceptibles de conférer une bonne biodisponibilité aux composés finaux.

Le troisième consiste à modéliser grâce à des tests *in vitro* les événements se produisant *in vivo* afin de réduire le taux d'échec des composés testés lors des phases cliniques. L'utilisation précoce de données issues de profilage de ces composés sur un grand nombre de cibles biologiques par exemple, permet de choisir les composés les plus prometteurs en terme de spécificité. L'utilisation de modèles de la barrière hémato-encéphalique, de la barrière gastrointestinale, des processus de métabolisation ... constituent autant d'axes de recherches permettant d'anticiper le devenir d'un composé dans un organisme vivant.

Enfin, la chimie combinatoire doit apparaître comme une technique complémentaire à celles de découvertes de nouveaux composés actifs, réputées comme efficaces. Le criblage de composés naturels ou les données recueillies par les ethnobotanistes notamment, permettront d'accéder à des structures totalement différentes de celles issues de la chimie combinatoire et conduiront ainsi à élargir conjointement le champ des maladies vaincues.

Partie expérimentale

PARTIE EXPERIMENTALE

Matériel et Méthodes :

Matériel :

1) Chromatographie Liquide Haute Pression:

Les analyses HPLC ont été effectuées sur des appareils de marque Shimadzu LC10, à l'aide de colonnes coulées au laboratoire (sous une pression de 350 bars, dans un tampon acétate 20 mM à pH 8 en mélange avec 8 parts de méthanol). Le système de solvants utilisé pour les gradients est : Solvant A : 0,05% de TFA dans l'eau et Solvant B : Acétonitrile 80% TFA,0,045%.

Les purifications HPLC ont été effectuées sur le système précédent lorsque les débits utilisés étaient inférieurs à 10 mL / min., pour des débits supérieurs à 10 mL / min le système Shimadzu LC8 a été utilisé.

a) Colonnes analytiques

- Cyano : phase cyano 5-CN Nucléosil (100 Å°, 5 µm), 4,6 x 250 mm
- C18 : phase C18 Nucléosil (100 Å°, 5 µm) 4,6 x 250 mm
phase Tosohaas TSK-gel Super ODS (100 Å°, 2 µm) 4,6 x 50 mm

Gradients :

Run de 55 minutes : à un débit de 1 mL/min Système isocratique à 0% de B pendant 5 min, gradient de 0 à 100% de B en 30 min, isocratique à 100 % B pendant 10 min puis isocratique à 0 % B pendant 10 min. Détection à 215 et 254 nm

Run de 25 minutes : à un débit de 1 mL/min Système isocratique à 0% de B pendant 1 min, gradient de 0 à 100% de B en 15 min, isocratique à 100 % B pendant 4 min puis isocratique à 0 % B pendant 5 min. Détection à 215 et 254 nm

Run de 9 minutes : à un débit de 2,75 mL/min Système isocratique à 0% de B pendant 0,5 min, gradient de 0 à 100% de B en 5 min, isocratique à 100 % B pendant 1 min puis isocratique à 0 % B pendant 2,5 min. Détection à 215 et 254 nm

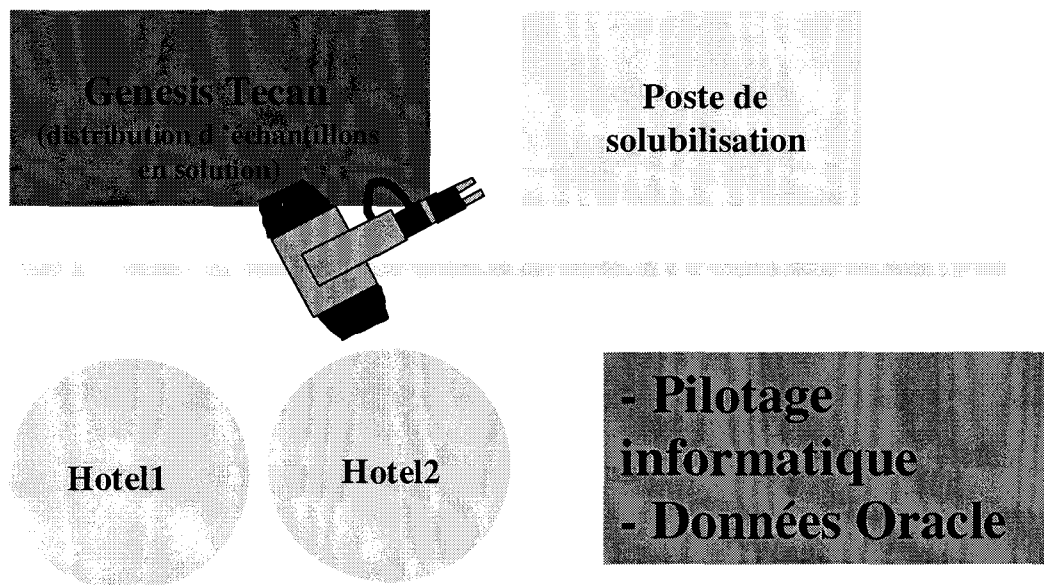
Run de 5 minutes : à un débit de 2,75 mL/min Gradient de 0 à 100% de B en 3 min, isocratique à 100 % B pendant 1 min puis isocratique à 0 % B pendant 1 min. Détection à 215 et 254 nm

b) Colonne préparative

- Colonne : 20 x 500 mm, phase C 18 , Nucléosil (100 Å°, 10 µM)
- Colonne : 25 x 200 mm, phase C 18 , Nucléosil (100 Å°, 10 µM)

Gradients variables

finis), le diluteur Genesis (distribution d'échantillons liquides), le poste de mise en solution (vortex et ultrasons), et les hôtels agitants (stockage des plaques en cours de réaction).



8) Solvants et réactifs :

Les solvants et réactifs commerciaux sont utilisés sans séchage ni purification préliminaire.

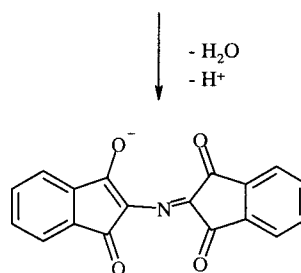
Méthodes :

1) Silanisation de la vaisselle utilisée pour les synthèses en phase solide :

La silanisation diminue l'adhérence des billes de polymère aux parois de verre des réacteurs, permettant ainsi une récupération quasi intégrale de la résine. La vaisselle est silanisée par le dichlorodiméthylsilane (Me_2SiCl_2), la fonction chlorosilylée apparue étant ensuite neutralisée par un lavage au méthanol en excès.

2) Dosage des chlorures par la méthode de Charpentier Vohlard :

Cette méthode a été utilisée pour évaluer le déplacement des chlorures après l'attaque nucléophile d'un acide aminé sous forme carboxylate de césium sur une résine chlorométhylée. A 100 mL de la solution de chlorures à doser sont ajoutés 5 mL d'acide nitrique 6N, 20 mL d'une solution titrée 50 mM d' AgNO_3 et 1 mL d'alun de fer ammoniacal. Ce mélange est titré par une solution 50 mM de KSCN jusqu'à l'apparition d'une faible coloration rougeâtre persistante.



pourpre de Ruhemann

Rq : le pourpre de Ruhemann se forme à partir des acides aminés (sauf la proline) ainsi qu'à partir des amines de type R-CH₂-NH₂.

4) Mise en évidence de fonctions amines secondaires par le test de Christensen :

En plaçant quelques milligrammes de résine dans un tube, en ajoutant 50 µL d'une solution saturée de 2,3,5,6-tétrachloro-1,4-benzoquinone dans le toluène et 200 µL d'acétone et en plaçant cette suspension 5 minutes à température ambiante, il est possible, de façon qualitative, de vérifier si l'acylation des fonctions amines présentes sur la résine a été totale. Dans ce cas aucune coloration n'apparaît dans le tube. En revanche une coloration bleu-noire apparaît sur les billes de résine lorsque des fonctions amines secondaires sont présentes sur la résine.

5) Protocole de déprotection des chaînes latérales des peptides et de clivage du lien peptide-résine, par l'acide fluorhydrique anhydre

Une quantité de tripeptidylrésine correspondant à 150 micromoles (200 mg), est déprotégée par 1x2 et 1x20 minutes de TFA/CH₂Cl₂ 50/50 V/V . Ce lot de résine est traité systématiquement, même en l'absence de tryptophane, par la pipéridine à 10% dans la NMP pendant 2 heures. Après ces deux heures la peptidylrésine est lavée trois fois dans la NMP, trois fois dans le dichlorométhane puis séchée.

Le lot de résine est ensuite mis à gonfler dans un mélange de 125 mg de p-thiocrésol et de 375 mg de p-crésol. Le réacteur étant plongé dans l'azote liquide, on laisse condenser 5 mL de fluorure d'hydrogène anhydre puis on laisse remonter la température à 0°C par immersion des réacteurs dans un bain de glace. Le mélange réactionnel est laissé à 0°C pendant une heure. On évapore ensuite l'acide fluorhydrique sous pression réduite, à 0°C. La fraction peptidique est ensuite précipitée dans l'éther éthylique ajouté goutte à goutte. La suspension est centrifugée à 5000 tr/minute pendant 5 minutes. Le culot de peptide est ensuite recueilli et dissous dans le minimum d'acide trifluoroacétique (< 3 mL). La solution acide de peptide est jetée goutte à goutte dans 70 mL du mélange éther éthylique / pentane, 50/50 placé préalablement à 4°C. Le précipité de peptide ainsi obtenu est recueilli par une nouvelle centrifugation de 5 minutes à 5000 tr/minute. Le peptide, qui est toujours accompagné de la résine et d'un léger surnageant d'éther, est alors repris dans 30 mL d'acide acétique à 5 %. Cette solution limpide recouverte d'une phase étherée contenant la résine, est filtrée sur filtre Millipore Téflon 5 µm, puis, après une évaporation de l'éther résiduel par un flux d'azote, est lyophilisée. Il faut procéder à deux lyophilisations pour obtenir une poudre.

Remarque : un système possédant 6 réacteurs de téflon, permettant de réaliser le clivage HF de 6 résines en parallèle, a été utilisé.



Première partie : -La synthèse combinatoire en mélange : la première étape d'une course à la diversité-

V-B Préparation des chimiothèques orthogonales

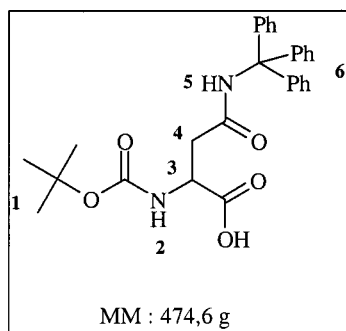
V-B-1 Protections choisies pour les différents acides aminés

Préparation des acides aminé protégés non commerciaux.

- Synthèse de la N α -tBoc(β -trityl) D-Asparagine

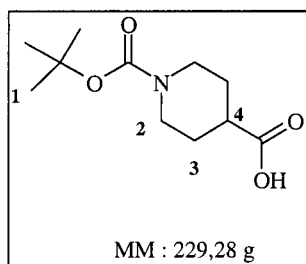
Cette synthèse a été réalisée selon le mode opératoire décrit dans Tetrahedron Lett. Vol32 N°6 p 739sq.

6,1 g (50 millimoles) d'Asparagine, 26 g de triphénylméthanol (100 millimoles), 10 mL (100 millimoles) d'anhydride acétique et 3 mL d'acide sulfurique concentré sont suspendus dans 200 mL d'acide acétique glacial et agité pendant 75 minutes à 60°C. La solution est ensuite lentement ajoutée à 500 mL d'eau à 0°C et le pH de la suspension lentement ajusté à 6 avec une solution de soude 10N. Après une heure à 0°C, les cristaux sont soigneusement lavés par 1,5 L d'eau distillée puis 1,5 L de toluène, puis enfin séchés (10,2 g 37 millimoles, rendement 74%). Les 10 g de H-(trityl) Asparagine-OH ainsi obtenue sont ensuite dissous dans 40 mL d'eau et 40 mL de tertiobutanol avec 1,4 g (35 millimoles) de soude. 40 millimoles de tBoc pyrocarbonate sont ensuite ajoutés. Après 12h de réaction à pH 8, la solution obtenue est diluée par 100 mL d'éther éthylique et 100 mL d'eau. La phase aqueuse est lavée par le toluène et ensuite acidifiée à pH 3 par l'acide sulfurique en présence d'acétate d'éthyle. On extrait la tBoc(β -trityl) Asparagine par 4 fois 100 mL d'acétate d'éthyle. L'évaporation de l'acétate d'éthyle sous pression réduite à $t^{\circ} < 35^{\circ}\text{C}$, conduit à la cristallisation de 11 g de tBoc (β trityl) asparagine contrôlée par CCM $R_f = 0,8$ (CHCl_3 CH_3OH AcOH 80/15/5 V/V/V).



RMN ^1H (300 MHz, $\text{DMSO}-d_6$)

8,67	(s, 1 H : H5)
7,2	(m, 15 H : H6)
6,96	(d, $J = 8,4$ Hz, 1 H : H2)
4,1	(m, 1H : H3)
2,62	(m, 2 H : H4)
1,41	(s, 9 H : H1)



RMN ¹H (300 MHz, DMSO-*d*₆)

3,8	(m, 2 H : H2e)
2,82	(m, 2 H : H2a)
2,41	(dd, $J_{60^\circ} = 3,76$ Hz, $J_{180^\circ} = 14,7$ Hz , 1 H : H4)
4,2	(m, 1H : H2)
1,8	(m, 2 H : H3e)
1,4	(s, 9 H : H1)
1,38	(m, 2 H : H3a)

V-B-5 Déprotection et clivage des tripeptides du support solide

Cf. matériel et méthodes

Seconde partie : -La synthèse parallèle à haut débit : l'outil de choix pour la découverte de nouvelles molécules biologiquement actives-

A Etude comparative de deux réactifs de couplage pour la synthèse d'urées asymétriques : le triphosgène et le carbonyldiimidazole

A-1 Utilisation du triphosgène

A-1-2 Evaluation de l'importance des conditions anhydres

Synthèse de la N-Phényl-N'-(3,4-méthylènedioxybenzyl)urée (3)

Mode opératoire :

Préparer les solutions suivantes :

solution 1 : dissoudre 110 mg (0,37 mmole) de triphosgène (MM : 296,75g, Aldrich) par 2 mL de DCM

solution 2 : placer dans 3,5 mL de DCM, 91 µL (1 mmole) d'aniline (MM : 93,13g, d : 1,022, Aldrich) et 189 µL (1,1 mmole) de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire)

solution 3 : placer dans 2 mL de DCM, 124 µL (1 mmole) de pipéronylamine (MM : 151,17g, d : 1,214, Aldrich) et 189 µL (1,1 mmole) de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire)

Ajouter la solution 2 à la solution 1, placée sous agitation, en une demi heure à raison de 7 ajouts de 500 µL goutte à goutte, à 5 minutes d'intervalle. Laisser agiter 5 minutes puis ajouter en une seule fois la solution 3. Laisser agiter 10 minutes puis évaporer le DCM à l'évaporateur rotatif. Dissoudre le résidu par 25 mL d'acétate d'éthyle. Laver la phase

- CCM (CH₂Cl₂/MeOH-96/4) R_f = 0,71, révélation ninhydrine (-) UV (+) KI-TDM (+) (restes d'aniline et de pipéronylamine détectés)
- HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 20,2 min

Synthèse de la N, N'-Bis(3,4-méthylènedioxybenzyl)urée (5)

Mode opératoire :

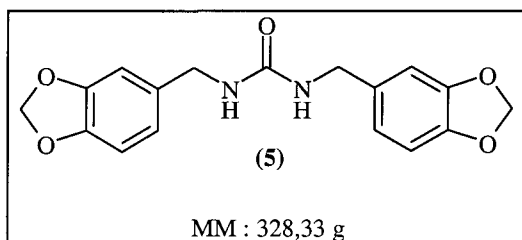
Préparer les solutions suivantes :

solution 1 : dissoudre 110 mg (0,37 mmole) de triphosgène (MM : 296,75g, Aldrich) par 2 mL de DCM

solution 2 : placer dans 3,5 mL de DCM, 91 µL (1 mmole) d'aniline (MM : 93,13g, d : 1,022, Aldrich) et 189 µL (1,1 mmole) de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire)

solution 3 : placer dans 2 mL de DCM, 91 µL (1 mmole) d'aniline (MM : 93,13g, d : 1,022, Aldrich) et 189 µL (1,1 mmole) de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire)

Ajouter la solution 2 à la solution 1, placée sous agitation, en une demi heure à raison de 7 ajouts de 500 µL goutte à goutte, à 5 minutes d'intervalle. Un précipité se forme. Laisser agiter 5 minutes puis ajouter en une seule fois la solution 3. Laisser agiter 10 minutes puis évaporer le DCM à l'évaporateur rotatif. Ajouter 25 mL d'acétate d'éthyle au résidu (un faible dissolution du produit est observé). Laver la phase organique par 3 fois 25 mL d'une solution aqueuse de KHSO₄ à 10 %, par 3 fois 25 mL d'une solution aqueuse de NaHCO₃ à 5 % puis par 25 mL d'une solution aqueuse saturée de NaCl. Le précipité est toujours présent en phase organique. L'acétate d'éthyle est évaporé et le produit sera donc séché une nuit à la pompe à palette.



- CCM (CH₂Cl₂/MeOH 96/4, v/v) R_f = 0,6, révélation ninhydrine (-) UV (+) KI-TDM (+)
- HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 22,4 min

A-2 Utilisation du carbonyldiimidazole

A-2-1 Evaluation de l'importance des conditions anhydres

Synthèse de la N-Phényl-N'-(3,4-méthylènedioxybenzyl)urée (7)

Mode opératoire :

Préparer les solutions suivantes :

solution 1 : dissoudre 81 mg (0,5 mmole) de CDI (MM : 162,15 g, Janssen) par 400 µL de DMF

solution 2 : placer dans 1 mL de DMF, 46 µL (0,5 mmole) d'aniline (MM : 93,13g, d : 1,022, Aldrich)

solution 3 : placer dans 750 µL de DMF, 62 µL (0,5 mmole) de pipéronylamine (MM : 151,17g, d : 1,214, Aldrich)

Ajouter la solution 2 à la solution 1, placée sous agitation, en une demi heure à raison de 7 ajouts de 500 µL goutte à goutte, à 5 minutes d'intervalle. Laisser agiter 5 minutes. (La CCM du milieu réactionnel juste avant l'introduction de la deuxième amine, révèle la présence d'aniline résiduelle ainsi que de l'urée symétrique de l'aniline.) Ajouter en une seule fois la solution 3 et faire des dépôts sur plaque CCM après 10 min., 1h, 3h et 4h de réaction. Après 4 heures de réaction le milieu réactionnel est analysé par HPLC et par spectrométrie de masse

Conditions opératoires de l'étude de l'influence du temps de contact entre l'amine A (aniline) et le CDI avant l'ajout de l'amine B (4-nitrobenzylamine).

La solution 0,2 M d'amine A (aniline), dans le DMF a été ajoutée goutte à goutte, en 6 fois, sur 25 minutes à une solution 0,8 M de CDI, dans le DMF placée sous agitation. Après différents temps de réaction (5 minutes, 15 minutes, 30 minutes, 45 minutes, 1 heure, 1 heure 30) prélever 1/6 de la solution d'amine A + DMF et ajouter 1 équivalent de 4-nitrobenzylamine 0,2 M dans le DMF. Chaque milieu réactionnel a été contrôlé par CCM, 15 minutes et 1 heure après l'introduction de l'amine B et par HPLC, 1 heure après l'introduction de l'amine B.

A-2-2-2 Détermination du nombre d'équivalents de CDI à utiliser

Conditions opératoires :

Un équivalent d'une solution 0,2 M d'amine A (aniline), dans le DMF a été ajouté goutte à goutte, en 6 fois, sur 25 minutes à 1, 1,5 ou 2 équivalents de CDI en solution respectivement 0,8 M, 1,2 M, 1,6 M dans le DMF. Les milieux réactionnels ont ensuite été placés sous agitation pendant 5 minutes. Un équivalent d'une solution 0,2 M d'amine B (4-nitrobenzylamine) a été introduit puis chaque milieu réactionnel a été évaporé 15 minutes après l'introduction de l'amine B et contrôlé par CCM et HPLC

A-2-2-3 Etude de l'influence de la vitesse d'addition de l'amine A au CDI, sur la formation des urées symétriques ou asymétriques

Conditions opératoires :

Un équivalent d'une solution 0,2 M d'amine A (aniline), dans le DMF a été ajouté goutte à goutte, sur des durées allant de quelques secondes à 20 minutes, à 1,5 équivalents de CDI en solution (1,2 M) dans le DMF. Un équivalent d'une solution 0,2 M d'amine B (4-nitrobenzylamine), dans le DMF a été ajouté et chaque milieu réactionnel contrôlé par CCM et HPLC après 90 minutes de réaction.

A-2-3 Evaluation des possibilités d'extension du protocole de synthèse à une amine A aliphatique

Synthèse de la N-(2-indanyl)-N'-phénylurée (10)

Mode opératoire :

Préparer les solutions suivantes :

solution 1 : dissoudre 97 mg (0,6 mmole) de CDI (MM : 162,15 g, Janssen) par 500 μ L de DMF

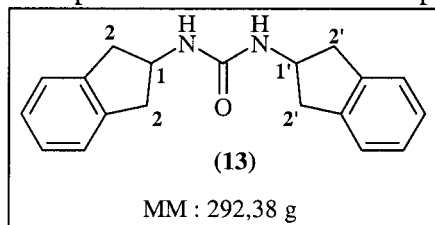
solution 2 : placer dans 2 mL de DMF, 68 mg (0,4 mmole) de chlorhydrate de 2-aminoindane (MM : 169,66 g, Aldrich) et 70 μ L (0,4 mmoles) de DIEA (MM : 129, d : 0,74, distillée au laboratoire)

solution 3 : placer dans 2 mL de DMF, 36 μ L (0,4 mmole) d'aniline (MM : 93,13 g, d : 1,022, Aldrich)

Ajouter la solution 2 sur 10 minutes (en 6 aliquotes distribués goutte à goutte, à 2 minutes d'intervalle) à la solution 1 placée sous agitation. Laisser agiter 5 minutes. Ajouter la solution 3 et analyser le milieu réactionnel par HPLC, après 3 heures et 72 heures de réaction. Les contaminants détectés sont l'imidazolide du 2-aminoindane et l'urée symétrique de l'aniline.

Synthèse de la N, N'-Bis(2-indanyl) urée (13)

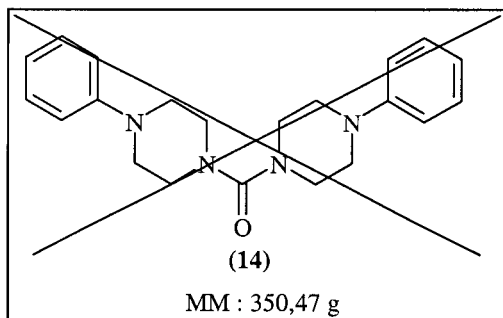
Placer dans 4 mL de DMF, 750 mg (4 mmoles) de chlorhydrate de 2-aminoindane (MM : 169,66 g, Aldrich), 324 mg (2 mmoles) de CDI (MM : 162,15 g, Janssen) et 700 μ L (4 mmoles) de DIEA (MM : 129, d : 0,74, distillée au laboratoire). Placer le milieu réactionnel sous agitation pendant 24 heures. Un précipité est obtenu. Eliminer le DMF à l'évaporateur rotatif. Dissoudre le résidu dans 200 mL d'acétate d'éthyle. Laver la phase organique par 3 fois 100 mL d'une solution aqueuse de KHSO_4 à 10 %, par 3 fois 100 mL d'une solution aqueuse de NaHCO_3 à 5 % puis par 100 mL d'une solution aqueuse saturée de NaCl . L'acétate d'éthyle est évaporé et le produit séché une nuit à la pompe à palette.



- conversion : 85 % (estimée par HPLC)
- CCM ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 96/4, v/v) Rf = 0,5, révélation ninhydrine (-) UV (+) KI-TDM (+)
- HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 24,9 min
- SM (TOF, m/z) : 293,1 g
- RMN ^1H (300 MHz, $\text{DMSO}-d_6$)
 - 7,22-7,11 (m, 8 H : aromatiques)
 - 6,05 (d, $J = 7,3$ Hz, 2 H : NH)
 - 4,33 (m, 2 H : H1, H1')
 - 3,11 (m, 4 H : H2trans, H2'trans)
 - 2,66 (m, 4 H : H2cis, H2'cis)

Essai de synthèse du Bis (4-phényl-1-pipérazine) carboxamide (14)

Placer dans 8 mL de DMF, 790 mg (4 mmoles) de chlorhydrate de 1-phénylpipérazine (MM : 198,7 g, Aldrich), 324 mg (2 mmoles) de CDI (MM : 162,15 g, Janssen) et 700 μ L (4 mmoles) de DIEA (MM : 129, d : 0,74, distillée au laboratoire). Un précipité est immédiatement obtenu. Placer le milieu réactionnel sous agitation pendant 24 heures. Eliminer le DMF à l'évaporateur rotatif. Le produit majoritaire est ensuite purifié par chromatographie basse pression (phase stationnaire : silice normale (100g), phase mobile : $\text{DCM}/\text{CH}_3\text{OH}$ 98/2, v/v).



Le produit attendu n'a pas été obtenu, seul l'imidazolidine de la 1-phénylpipérazine a été isolé.

Le dosage de la benzylamine résiduelle s'effectue en mesurant à l'aide d'un spectromètre UV (à 270 nm), la formation de l'espèce colorée (pourpre de Ruhemann) formée dans les conditions du test de Kaiser (cf. partie expérimentale, généralités).

Mode opératoire :

Dissoudre 57 mg de CDI dans 940 μL de DMF. Prélever 100 μL de cette solution, les placer sur 20 mg de chlorhydrate de benzylamine, ajouter 76 μL de DMF et 24 μL de DIEA : obtention de la solution (1).

Prélever 50 μL de la solution (1), les placer 3 min. à 100°C et ajouter 1,25 mL de DMF : obtention de la solution (2).

Prélever 100 μL de la solution (2) et leur ajouter 100 μL des solutions A, B et C utilisées dans le test de Kaiser. Placer ce milieu 5 min. à 100°C puis le diluer par 5 mL d'éthanol. L'absorbance de cette solution diluée au 1/6^{ème} est ensuite mesurée et la réactivité du CDI donnée par l'équation suivante :

$$\text{Réactivité du CDI (\%)} = \frac{3,46 - (2,31 \times \text{Absorbance})}{1,75} \times 100$$

Mode opératoire :

Préparer les trois solutions suivantes :

Solution A : dissoudre 500 mg de ninhydrine dans 10 ml d'éthanol

Solution B : dissoudre 20 g de phénol dans 5 mL d'éthanol

Solution C : diluer 2 mL d'une solution aqueuse de KCN 0,001 M dans 98 mL de pyridine.

A une masse connue de résine (correspondant à environ 10 μmoles), sont ajoutés 100 μL de chacune des solutions précédentes (A, B et C).. La suspension obtenue est incubée 5 minutes à 100°C et 4,7 mL de méthanol sont ensuite ajoutés. L'absorbance de la solution est lue à 570 nm puis la quantité d'amine présente dans la prise d'essai déterminée grâce au coefficient d'extinction molaire du pourpre de Ruhemann (15000 $\text{M}^{-1} \text{cm}^{-1}$).

Test de réactivité du carbonyldiimidazole, par dosage HPLC

Après avoir synthétisé, purifié et caractérisé (Masse, RMN) 4 grammes de dibenzylurée, nous avons réalisé une gamme étalon des aires du pic HPLC (correspondant à la dibenzylurée) obtenues à 254 nm en fonction de la concentration en urée présente dans chacune des solutions. Les concentrations testées s'échelonnaient de 0 à 2 mM.

Nous avons obtenu la droite de régression suivante (en injectant 20 μL d'une solution de dibenzylurée dans l'acétonitrile, sur l'HPLC Shimadzu LC10 en utilisant le gradient X9) :

$$\text{aire du pic à 254 nm} = [\text{concentration en dibenzylurée}] \text{ mM} \times 129901 \text{ mM}^{-1} + 7777$$

En nous basant sur une réactivité théorique de CDI = 100 %, la préparation de la « **solution test** » suivante devait conduire à une solution 1 mM en dibenzylurée :

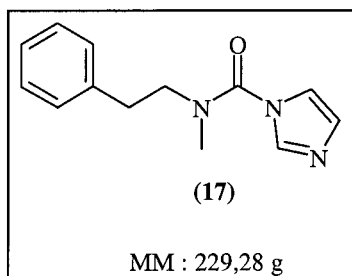
« **solution test** »

Placer dans un tube 70 μL (soit 17 μmoles) d'une solution de CDI 0,247 M puis 15 μL (soit 139 μmoles) de benzylamine. Boucher le tube et laisser réagir 5 minutes à température ambiante puis 1 minute à 100°C. Ajouter 1,62 ml d'acétonitrile. Diluer un aliquote de cette solution au dixième et injecter 20 μL en HPLC.

La réactivité de la solution de CDI sera donc calculée par l'équation suivante :

Synthèse de la N-imidazole-N-méthyl-phénéthylamine (17)

Dans un tube de verre, placer 34 mg (0,25 mmole) de phénéthylamine, (MM : 135,2 g, Aldrich) puis 1 mL de DMF. Ajouter 1 mL d'une solution 0,25 de CDI dans le DMF. Le milieu réactionnel est placé 4h sous agitation, tube ouvert. La LC-MS du milieu réactionnel (5 μ L du milieu réactionnel dilués dans 200 μ L d'éluant B et injection de 20 μ L de cette solution) révèle la formation de l'imidazole attendu avec une conversion de l'amine de départ, supérieure à 95 %.

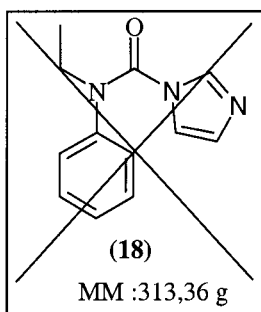


Conversion : > 95% (estimée par HPLC)

LC/MS (phase inverse, C18, barrette de diode, run de 5 min./ APCI⁺, m/z) : t_R = 1,54 min./ 230,7 g (MH⁺)

Essai de synthèse de la N-imidazole-N-éthylaniline (18)

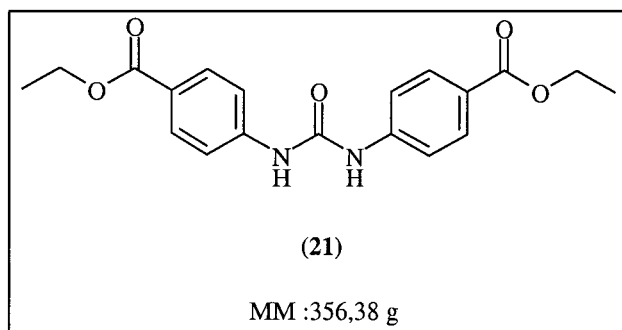
Dans un flacon de verre, 503 μ L (4 mmoles) de N-éthylaniline, (d : 0,963, MM : 121,18 g, Aldrich) et 2 mL de DMF ont été placés. 2 mL d'une solution molaire de CDI dans le DMF ont ensuite été ajoutés. Le milieu réactionnel a été placé 4h sous agitation, flacon ouvert. L'HPLC (20 μ L du milieu réactionnel dilués dans 5 mL d'éluant B et injection de 20 μ L de cette solution) et la CCM du milieu réactionnel n'ont pas révélé l'apparition d'un nouveau produit. La masse du produit attendu n'a pas été détectée par le spectromètre de masse PDMS/TOF.



Synthèse de la N, N'-Bis(*tert*-butyl) urée (19)

Dans un flacon de verre, placer 420 μ L (4 mmoles) de *tert*-butylamine, (d : 0,696, MM : 73,14 g, Aldrich) puis 2 mL de DMF. Ajouter 2 mL d'une solution molaire de CDI dans le DMF. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. L'HPLC du milieu réactionnel (40 μ L du milieu réactionnel dilués dans 3 mL d'éluant B et injection de 20 μ L de cette solution) ainsi que la CCM révèlent la formation de l'urée. 2 mL d'eau distillée sont ensuite ajoutés au milieu réactionnel pour arrêter la réaction puis le DMF éliminé à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette. Le milieu réactionnel est repris par 100 mL d'acétate d'éthyle puis extrait par 2 fois 50 mL d'une solution aqueuse d'HCl à 3 %. L'acétate d'éthyle est ensuite éliminer à l'évaporateur rotatif.

permettant pas d'obtenir l'urée dépourvue de l'amine résiduelle, le milieu réactionnel est évaporé, repris par 100 mL d'acétate d'éthyle puis extrait par 8 fois 50 mL d'une solution aqueuse d'HCl à 3 %.



$\rho = 51 \%$

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,44, révélation ninhydrine (-) UV (+) KI-TDM (+)

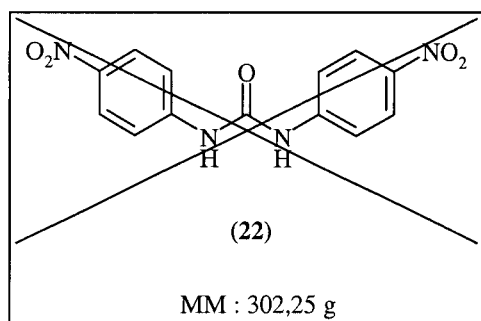
HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 25,5 min.

SM (TOF, m/z) : 356,9 g

RMN ¹H (300 MHz, DMSO-*d*₆)

Essai de synthèse de la N, N'-Bis(4-nitrophényl) urée (22)

Dans un flacon de verre, ont été placés 553 mg (4 mmoles) 4-nitroaniline, (MM : 138,13 g, Aldrich) puis 2 mL de DMF. 2 mL d'une solution molaire de CDI dans le DMF ont été ajoutés. Le milieu réactionnel a été placé 4h sous agitation, flacon ouvert. L'HPLC du milieu réactionnel (40 µL du milieu réactionnel dilués dans 5 mL d'éluant B et injection de 20 µL de cette solution) ainsi que la CCM n'ont pas révélé la formation de l'urée symétrique. La masse du produit attendu n'a pas été détectée par le spectromètre de masse PDMS/TOF.



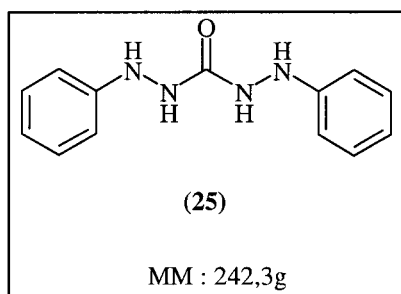
Synthèse de la (4S, 5R)-4-Méthyl-5-phényl-2-oxazolidinone (23)

Dans un flacon de verre, placer 605 mg (4 mmoles) de (1R, 2S)-(-)-Noréphédrine, (MM : 151,21 g, Aldrich) puis 2 mL de DMF. Ajouter 2 mL d'une solution molaire de CDI. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. Le produit attendu ne précipitant pas après l'ajout de 10 mL d'eau, sera purifié par chromatographie basse pression (phase stationnaire : silice normale, phase mobile : DCM/CH₃OH 98/2, v/v).

La CCM et l'HPLC du milieu réactionnel (20 µL du milieu réactionnel dilués dans 6 mL d'éluant B et injection de 20 µL de cette solution) montrent de la (1R, 2S)-(-)-Noréphédrine résiduelle.

Synthèse du 1,5-Diphénylcarbazine (25)

Dans un flacon de verre, placer 400 μL (4 mmoles) de phénylhydrazine (d : 1,099, MM : 108,14 g, Aldrich) et ajouter 2 mL de DMF, une solution jaune est obtenue. Ajouter 2 mL de la solution molaire de CDI préparée précédemment, la solution devient orange après quelques minutes. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. 30 mL d'eau distillée sont ensuite ajoutés au milieu réactionnel, le précipité jaune pâle est isolé par centrifugation, remis en suspension dans 30 mL d'eau, isolé par centrifugation puis placé au dessiccateur. L'HPLC du milieu réactionnel (40 μL du milieu réactionnel dilués dans 3,4 mL d'éluant B et injection de 20 μL de cette solution) montre une conversion > 95 % de l'hydrazine.



Conversion : > 95% (estimée par HPLC)

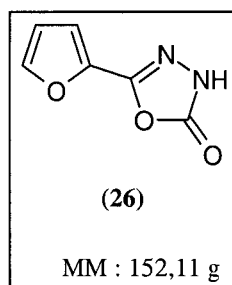
HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 14,5 min

SM (TOF, m/z) : 242,4 g

RMN ^1H (300 MHz, DMSO- d_6)

Synthèse de la 2-(2-furoyl)- Δ^2 -1,3,4-oxadiazolin-5-one (26)

Dans un flacon de verre, placer 504 mg (4 mmoles) de 2-furoichlazide, (MM : 126,12 g, Aldrich) puis 2 mL de DMF. Ajouter 2 mL d'une solution molaire de CDI dans le DMF. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. L'HPLC du milieu réactionnel (40 μL du milieu réactionnel dilués dans 5 mL d'éluant B et injection de 20 μL de cette solution) ainsi que la CCM révèlent la formation de l'oxadiazolinone. 2 mL d'eau distillée sont ensuite ajoutés au milieu réactionnel pour arrêter la réaction puis le DMF éliminé à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette.



Conversion : > 95 % (estimée par HPLC)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,54, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 6,24 min.

SM (TOF, m/z) : 152 g

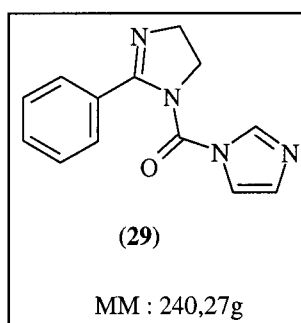
Synthèse de la 1,5-Bis(benzophénone) carbazone (27)

Dans un flacon de verre, placer 785 mg (4 mmoles) d'hydrazone de benzophénone, (MM : 196,25 g, Aldrich) puis 2 mL de DMF. Ajouter 2 mL d'une solution molaire de CDI dans le DMF. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. L'HPLC du milieu

Synthèse de la 1-imidazolide-2-phénylimidazoline (29)

Dans un flacon de verre, placer 585 mg (4 mmoles) de 2-phényl-2-imidazoline, (MM : 146,19 g, Aldrich) puis 2 mL de DMF. Ajouter 2 mL d'une solution molaire de CDI. Le milieu réactionnel est placé 4h sous agitation, flacon ouvert. Le produit attendu ne précipitant pas après l'ajout de 10 mL d'eau, sera purifié par chromatographie basse pression (phase stationnaire : silice normale, phase mobile : DCM/CH₃OH 96/4, v/v).

La CCM et l'HPLC du milieu réactionnel (20 µL du milieu réactionnel dilués dans 6 mL d'éluant B et injection de 20 µL de cette solution) montrent de la 2-phényl-2-imidazoline résiduelle.



$\rho = 48 \%$

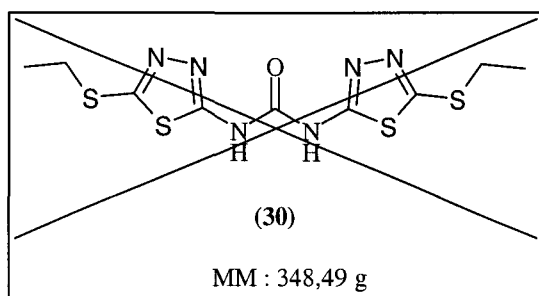
CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,22, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 4,8 et 5,3 min (pic dédoublé)

SM (TOF, m/z) : 240,8 g

Essai de synthèse de la N, N'-Bis(5-éthylthio-1,3,4-thiadiazol-2-yl) urée (30)

Dans un flacon de verre, ont été placés 645 mg (4 mmoles) de 2-amino-5-éthylthio-1,3,4-thiadiazole, (MM : 161,25 g, Aldrich) puis 2 mL de DMF. 2 mL d'une solution molaire de CDI dans le DMF ont été ajoutés. Le milieu réactionnel a été placé 4h sous agitation, flacon ouvert. Après une heure d'agitation un précipité est apparu. L'ajout de 8 mL n'a pas permis de dissoudre le précipité. L'ajout de 700 µL de DIEA a permis de dissoudre le précipité. L'HPLC du milieu réactionnel (120 µL du milieu réactionnel dilués dans 6 mL d'éluant B et injection de 20 µL de cette solution) ainsi que la CCM n'ont pas révélé la formation de l'urée symétrique. La masse du produit attendu n'a pas été détectée par le spectromètre de masse PDMS/TOF, seule la masse du produit de départ a été détectée.



Essai de synthèse de la N, N'-Bis(2-benzothiazolyl) urée (31)

Dans un flacon de verre, 600 mg (4 mmoles) de 2-aminobenzothiazole (MM : 150,2 g, Aldrich) ont été dissous par 2 mL de DMF, une solution jaune a été obtenue. 2 mL d'une solution molaire de CDI dans le DMF ont été ajoutés. En quelques minutes, la solution s'est décolorée et un précipité blanc est apparu. Le milieu réactionnel a été placé 4h sous agitation, flacon ouvert. 30 mL d'eau distillée ont ensuite été ajoutés au milieu réactionnel, le précipité a

Au cours de la purification du composé (32), les produits secondaires suivants ont été isolés et identifiés par RMN :

- l'urée symétrique de la phénéthylamine : 28 mg

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,3, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 20,8 min

- l'imidazolide de la phénéthylamine : 23 mg

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,13, révélation ninhydrine (-) UV (+) KI-TDM (+)

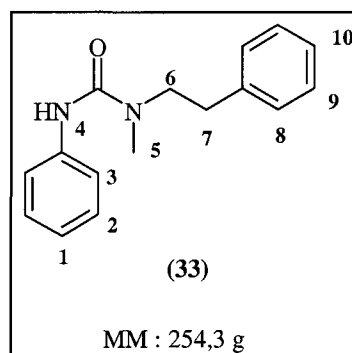
HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 6,6 min

Synthèse de la N-Phényl-N'-méthyl-N'-(2-phényléthyl)urée (33)

Mode opératoire : cf composé (32), en utilisant la solution 3 suivante

solution 3 : placer dans 5 mL de DMF, 73 µL (0,5 mmole) de N-méthylphénéthylamine (MM : 135,21g, d : 0,93, Aldrich)

Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit est purifié par CCE (DCM/CH₃OH 96/4, v/v) : R_f = 0,49



Conversion : 27 % (estimée par HPLC)

ρ = 10 % (produit isolé)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,45, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 21,2 min.

RMN ¹H (300 MHz, CDCl₃)

7,37-7,17 (m, 9 H : 2 H₂, 2 H₃, 2 H₈, 2 H₉, H₁₀)

7,01 (m, 1 H : H₁)

5,94 (s, 1 H : H₄)

3,62 (t, J = 6,9 Hz, 2 H : H₆)

2,99-2,9 (m, 5 H : 2 H₇, 3 H₅)

Au cours de la purification du composé (33), le produit secondaire suivant a été isolé et identifié par RMN :

- l'imidazolide de la N-méthylphénéthylamine : 65 mg

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,29, révélation ninhydrine (-) UV (+) KI-TDM (+)

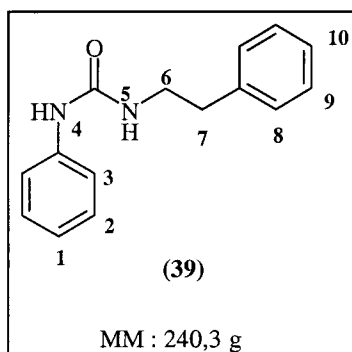
HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 6,8 min

Etude de l'activation du CDI par le LiBr et la DMAP

Mode opératoire :

Préparer les solutions suivantes :

solution 1 : dissoudre 162 mg (1 mmole) de CDI (MM : 162,15 g, Janssen) par 2,8 mL de THF



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 95 %

$\rho = 87 \%$ (produit isolé)

Caractéristiques analytiques : idem au composé 32

Synthèse de la N-Phényl-N'-méthyl-N'-(2-phényléthyl)urée (40)

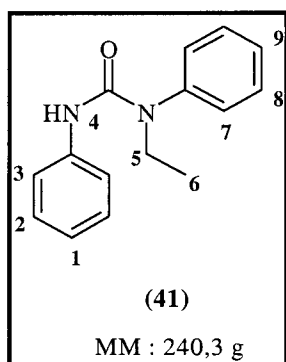
cf. Etude de l'activation du CDI par le LiBr et la DMAP, essai 7

Synthèse de la N-Phényl-N'-éthyl-N'-phénylurée (41)

Mode opératoire : cf composé (39), en utilisant la solution 3 suivante

solution 3 : placer dans 5 mL de DMF, 61 mg (0,5 mmole) de N-éthylaniline (MM : 121,2 g, Aldrich).

Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit est purifié par HPLC préparative (Colonne : 25 x 200 mm, phase C 18 , Nucléosil 10 μ M, 100 A°, gradient : Eluant 1 : 50 % à Eluant 1 : 30 % en 60 minutes puis Eluant 1 : 0 % pendant 60 minutes , 4 mL / min., Eluant 1 : Eau, TFA 0,05%, Eluant 2 : Eau / ACN, 80/20, v/v, TFA 0,0425%).



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 24 %

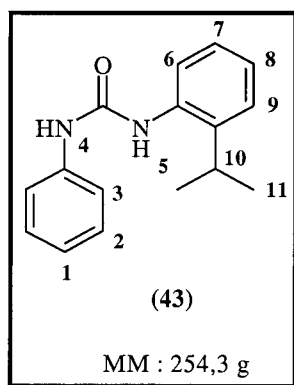
$\rho = 15 \%$ (produit isolé)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,7, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 28,5 min

RMN ¹H (300 MHz, CDCl₃)

7,53	(m, 2 H : H3)
7,45	(m, 1 H : H9)
7,4	(m, 2 H : H7)
7,29-7	(m, 4 H : 2H2, 2H8)
6,11	(s, 1 H : H4)
7,01	(m, 1 H : H1)



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 46 %

$\rho = 26 \%$ (produit isolé)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,5, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 22,3 min

(phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 29,8 min

RMN ¹H (300 MHz, CDCl₃)

7,45-7,19 (m, 8 H : 2 H₂, 2 H₃, H₆, H₇, H₈, H₉)

7,03 (m, 1 H : H₁)

6,93 (s, 1 H : H₄)

3,18 (m, 1 H : H₁₀)

2,8 (s, 1 H : H₅)

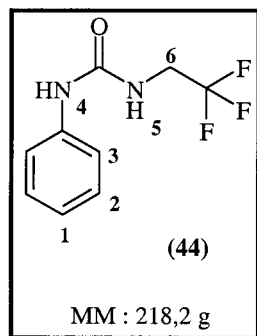
1,19 (d, J = 6,9 Hz, 6 H : H₁₁)

Synthèse de la N-Phényl-N'-(1,1,1-trifluoroéthyl)urée (44)

Mode opératoire : cf composé (39), en utilisant la solution 3 suivante

solution 3 : placer dans 5 mL de DMF, 68 mg (0,5 mmole) de chlorhydrate de 2,2,2-Trifluoroéthylamine (MM : 135,52g, Janssen) et 87 μ L de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire).

Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit purifié par CCE (DCM/CH₃OH 96/4, v/v) : R_f = 0,44



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 46 %

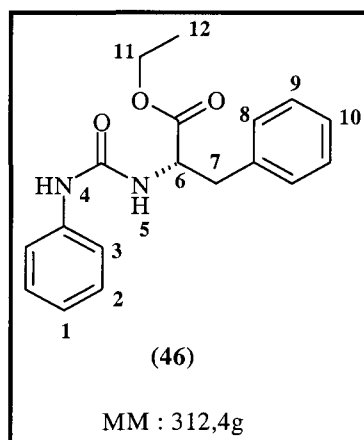
$\rho = 46 \%$ (produit isolé)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,44, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, cyano, 215/254 nm, run de 55 min.) : t_R = 8,6 min

(phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 23,1 min

RMN ¹H (300 MHz, CDCl₃ / MeOD, 95/5, v/v)



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 42 %
 Le rendement en produit isolé, n'a pu être calculé du fait d'une perte mécanique de produit.
 CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,73, révélation ninhydrine (-) UV (+) KI-TDM (+)
 HPLC (phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 28,9 min
 RMN ¹H (300 MHz, CDCl₃)

7,44-7,08	(m, 10 H : H1, 2 H2, 2 H3, 2 H8, 2 H9, H10)
6,9	(s, 1 H : H4)
5,54	(d, J = 7,9 Hz, 1 H : H5)
4,85	(m, 1 H : H6)
4,19	(m, 2 H : H11)
3,1	(m, 2 H : H7)
1,3	(m, 3 H : H12)

(Les signaux sont complexes du fait de la présence de protons diastéréotopiques).

Synthèse 1-Phénylazo-N-phénylformamide (47)

Mode opératoire : cf composé (39), en utilisant la solution 3 suivante

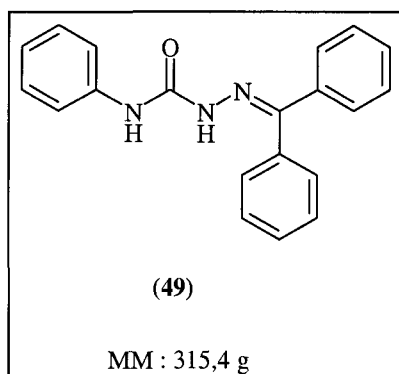
solution 3 : placer dans 5 mL de DMF, 49 µL (0,5 mmole) de phénylhydrazine (MM : 108,14g, d = 1,099 Aldrich). Un contrôle HPLC (phase inverse, C18, run de 55 min.) effectué avant évaporation du DMF montre une consommation totale des deux amines de départ avec apparition d'un nouveau produit (t_R = 25,3 min) correspondant à l'urée. Le contrôle CCM (DCM/CH₃OH 96/4, v/v) met en évidence 4 tâches de R_f : 0,46, 0,54, 0,6 et 0,7. Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) et le milieu réactionnel purifié par CCE (DCM/CH₃OH 96/4, v/v). Le produit majoritaire : R_f = 0,7 est isolé. Les analyses montrent qu'il ne s'agit plus de l'urée mais d'un composé oxydé (dérivé azo) s'étant formé au cours de la purification. Les autres bandes sont constituées de produits instables ou d'intermédiaires réactionnels (chromatogrammes HPLC multi-pics et RMN non interprétables).

8	(s, 1 H : H4)
7,62	(s, 1 H : H9)
7,41	(d, $J = 8$ Hz, 2 H : H3)
7,19	(m, 3 H : H1, 2 H2)
6,86-6,93	(m, 1 H : H7)
6,45	(t, $J = 1,7$ Hz, 1 H : H8)

Synthèse de la 1-Benzophénone-4-phénylsemicarbazone (49)

Mode opératoire : cf composé (39), en utilisant la solution 3 suivante

solution 3 : placer dans 5 mL de DMF, 98 mg (0,5 mmole) de benzophénonehydrazone (MM : 196,25g Aldrich). Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit purifié par CCE (DCM/CH₃OH 96/4, v/v) : R_f = 0,44



Conversion : 29 % (estimée par HPLC)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,69, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 33 min

Le produit se montre instable au cours d'une purification par chromatographie en couche épaisse (phase stationnaire normale, éluant DCM/CH₃OH 96/4, v/v).

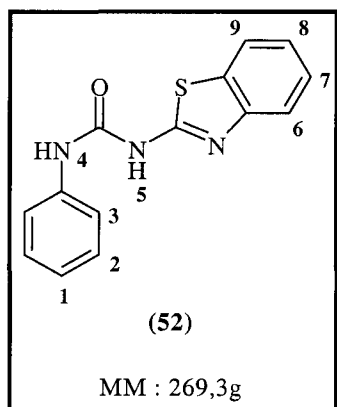
Synthèse de la N-Phényl-N'-méthyl-N'-méthoxyurée (50)

Mode opératoire : cf composé (39), en utilisant la solution 3 suivante

solution 3 : placer dans 5 mL de DMF, 49 mg (0,5 mmole) de chlorhydrate de O,N-diméthylhydroxylamine (MM : 97,6 g, Aldrich) et 87 µL de DIEA (MM : 129,25g, d : 0,75, distillée au laboratoire).

Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit est purifié par CCE (DCM/CH₃OH 96/4, v/v) : R_f = 0,66.

préparative (Colonne : 25 x 200 mm, phase C 18 , Nucléosil 10 μ M, 100 A°, gradient : Eluant 1 : 60 % à Eluant 1 : 25 % en 120 minutes puis Eluant 1 : 25 % pendant 60 minutes, 4 mL / min., Eluant 1 : Eau, TFA 0,05%, Eluant 2 : Eau / ACN, 80/20, v/v, TFA 0,0425%), afin d'éliminer la contamination par l'urée symétrique de l'aniline.



Rendement estimé par HPLC en utilisant une gamme étalon du produit attendu : 5 %

ρ = 4 % (produit isolé)

CCM (DCM/CH₃OH 96/4, v/v) : R_f = 0,58, révélation ninhydrine (-) UV (+) KI-TDM (+)

HPLC (phase inverse, C18, 215/254 nm, run de 55 min.) : t_R = 28,7 min

SM (TOF, m/z) : 269,3 g

RMN ¹H (300 MHz, CDCl₃)

8,75	(s, 1 H : H4)
7,8	(m, 2 H : H6, H9)
7,58	(m, 3 H : H8, 2 H3)
7,38	(m, 3 H : H7, 2 H2)
7,2	(t, J = 7,5 Hz, 1 H : H1)
4,8	(s, 1 H : H5)

Synthèse de la N, N'-Bis(phényléthyl)-N'-méthylurée (53)

Préparer les solutions suivantes :

solution 1 : dissoudre 81 mg (0,5 mmole) de CDI (MM : 162,15 g, Janssen) par 1,4 mL de THF

solution 2 : placer dans 5 mL de DMF, 63 μ L (0,5 mmole) de phénéthylamine (MM : 121,18g, d : 0,965, Janssen)

solution 3 : placer dans 5 mL de DMF, 73 μ L (0,5 mmole) de N-méthyl-phénéthylamine (MM : 135,21g, d : 0,93, Aldrich)

Ajouter la solution 2 à la solution 1. Laisser agiter 15 minutes. Ajouter la solution 3 et laisser réagir 8h. Le milieu réactionnel est alors contrôlé par HPLC et CCM.

Le DMF est évaporé (à l'évaporateur rotatif, à 45°C, sous le vide d'une pompe à palette) puis le produit est purifié par CCE (DCM/CH₃OH 96/4, v/v) : R_f = 0,73

Glossaire

GLOSSAIRE

ACD : available chemical directory (base de données répertoriant les composés chimiques commercialement disponibles)

azoles : hétérocycle aromatique à 5 atomes (fusionné ou non avec un autre cycle aromatique) contenant au moins deux azotes. Exemples : N-acylimidazoles, N-acyltriazoles, N-acylbenzimidazoles...

azolide : amide dont l'azote fait parti d'un hétérocycle aromatique à 5 atomes (fusionné ou non avec un autre cycle aromatique) contenant au moins deux azotes. Exemples : N-acylimidazoles, N-acyltriazoles, N-acylbenzimidazoles...

BOP : benzotriazole-1-yl-oxy-tris-(diméthylamino)-phosphonium hexafluorophosphate

CCE : chromatographie sur couche épaisse

CCM : chromatographie sur couche mince

Chimie d'assemblage : chimie permettant le couplage covalent d'une matière première à une autre

Déconvolution : détermination du (des) composé(s) actif(s) au sein d'un mélange lui-même actif, par fractionnement de ce mélange

DIEA : diisopropyléthylamine

DMA : diméthylacétamide

DMF : diméthylformamide

HBTU : 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetraméthyluronium hexafluorophosphate

Hit-rate : voir taux de succès

HOBT : 1-Hydroxybenzotriazole

HTS : high throughput screening (criblage à haut débit)

Kaiser (Test de) (Kaiser 1970)

Test utilisé dans le cadre de la chimie sur support solide (voir chapitre matériel et méthode) afin de doser les fonctions amines libres, présentes sur le support solide, par la formation du pourpre de Ruhemann.

KI/TDM : iodure de potassium/méthylène-bis-paradiméthylaniline, révélateur pour CCM utilisé pour la mise en évidence de fonctions du type amide primaire NH-CO

MM_{att} : masse moléculaire attendue

MM_{exp} : masse moléculaire expérimentale

Monomère : employé dans cette thèse pour toute matière première pouvant être liée de manière covalente à une autre matière première par le biais d'une chimie d'assemblage

PyBOP : benzotriazole-1-yl-oxy-tris-(pyrrolidino)-phosphonium hexafluorophosphate

R_f : rapport frontal

SM : spectre de masse

T. A. : température ambiante

Taille : la taille d'une chimiothèque ou d'une sous-chimiothèque fait référence au nombre de composés que cette chimiothèque ou sous-chimiothèque contiennent :

- pour une chimiothèque l'expression taille limitée s'emploie pour un nombre de composés inférieur à 10.000, l'expression taille moyenne pour un nombre compris entre 10.000 et 100.000 et l'expression taille importante pour un nombre supérieur à 100.000

- pour une sous-chimiothèque l'expression taille limitée s'emploie pour un nombre de composés inférieur à 100, l'expression taille moyenne pour un nombre compris entre 100 et 1000 et l'expression taille importante pour un nombre supérieur à 1000.

Taux de succès ou « hit rate » : expression utilisée dans le cadre du criblage de composés sur des tests biologiques pour chiffrer le nombre de composés actifs sur le nombre de composés testés

TDM : 4,4'-Tétraméthyl-diamino-diphénylméthane

THF : tétrahydrofurane

TOF : time of flight (technologie utilisée en spectrométrie de masse permettant d'obtenir le spectre de masse d'une molécule par la mesure du temps mis par cette molécule et ses fragments, après leur désorption d'une matrice par bombardement atomique, pour atteindre une cible)

t_R : temps de rétention HPLC

Bibliographie

BIBLIOGRAPHIE

- Abrahart 1936** : E. N. Abrahart, *J. Chem. Soc.*, **1936**, 1273
- Adams 1964** : P. Adams, *U.S. Pat. Dec.* 15, **1964**, 3, 161, 676
- Anderson 1958** : G. W. Anderson and R. Paul, *J. Am. Chem. Soc.*, **1958**, 4423
- Bergmann 1928** : M. Bergmann and L. Zervas, *Hoppe Seyler's Z Physiol. Chem.*, **1928**, 175, 145
- Borchardt 1994** : A. Borchardt and W. C. Still, *J. Am. Chem. Soc.*, **1994**, 116, 373-374
- Brenner 1992** : S. Brenner and R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, 89, 5381-5383
- Buckley 1951** : G. D. Buckley and N. H. Ray *U. S. Pat.*, **1951**, 2, 550, 767
- Bunin 1998** : B. A. Bunin «*The Combinatorial Index* », **1998**, Academic Press Limited
- Carell 1994** : T. Carell, E. A. Winter, A. Bshir-Hashemi and J. Rebek *Angew. Chem.*, **1994**, 33, 2059-2060
- Cook 1967** : L. H. Cook *U. S. Pat.*, Jan. 31, **1967**, 3, 301, 897
- Cortese 1996** : R. Cortese, *Combinatorial Libraries*, **1996**, Walter de Gruyter, Berlin, New York
- Coulson 1947** : C. A. Coulson and H. C. Longuet-Higgins *Trans. Faraday Soc.*, **1947**, 43, 87
- Crain 1984** : D. J. Crain, I. B. Dicker, M. Lauer, C. B. Knobler and K. N. Trueblood, *J. Am. Chem. Soc.* **1984**, 106, 7150
- Curatolo 1998** : W. Curatolo, *Pharm. Sci. Tech. Today*, **1998**, 1, 387-393
- Datta 1975** : A. S. Datta, J. S. Morley *J. Chem. Soc. Perkin Trans. I*, **1975**, 1712
- Déprez 1995** : Benoît Déprez, Xavier Williard, Line Bourel, Hervé Coste, François Hyafil and André Tartar, *J. Am. Chem. Soc.* **1995**, 117, 5405-5406
- Dooley 1993** : C. T. Dooley and R. A. Houghten, *Life Sci.* **1993**, 52, 1509-1517
- Ecker 1995** : D. J. Ecker and S. T. Croke, *Biotechnology*, **1995**, 13, 351-360
- Farbenindustrie 1937** : I. G. Farbenindustrie, *A. G. Brit. Pat.*, **1937**, 462, 182
- Fodor 1991** : S. P. A. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu and D. Solas, *Science*, **1991**, 251, 767-773
- Freier 1995** : S. M. Freier, D. A. M. Konings, J. R. Wyatt and D. J. Ecker, *J. Med. Chem.*, **1995**, 38, 344-352
- French 1926** : H. E. French and A. F. Wirtel, *J. Am. Chem. Soc.* **1926** 48, 1736
- Furka 1988** : A. Furka, F. Sebestyen, M. Asgedom and G. Dibo, *Proceedings of the 14th International Congress, Prague, Czechoslovakia*, July 10-15, **1988**, 5, 47

- Furka 1994** : A Furka, *Drug Development Research*, **1994**, 33, 90-97
- Gante 1965** : J. Gante, *Chem. Ber.*, **1965**, 98, 3334
- Gausepohl 1989** : H. Gausepohl et al., *Int. J. Peptide Protein Res.*, **1989**, 34, 287
- Geysen 1984** : H. M. Geysen, R. H. Meloen, S. J. Barteling *Proc. Natl. Acad. Sci USA.*, **1984**, 81, 3998-4002
- Geysen 1986** : H. Mario Geysen, Stuart J. Rodda and Tom J. Mason, *Molecular Immunology*, **1986**, 23, 7, 709-715
- Gogoi 1990** : Gogoi, P. C. et al., *Indian Journal of Chemistry*, **1990**, Sect B 29, 12, 1159-1160
- Hostettmann 1998** : K. Hostettmann, A. Marston and M. Hostettmann, *Preparative Chromatography Techniques*, 1998, Springer-Verlag Berlin Heidelberg
- Houghten 1991** : R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley and J. H. Cuervo, *Nature*, **1991**, 354, 84-86
- Houghten 1992** : R. A. Houghten, J. R. Appel, S. E. Blondelle, J. H. Cuervo, C. T. Dooley, and C. Pinilla *Biotechniques*, **1992**, 13, 412-421
- Kaiser 1970** : Kaiser, *Analytical Biochemistry*, April **1970**, 34, 2
- Kaldor 1996** : Stephen W. Kaldor, Miles G. Siegel, James E. Fritz, Bruce A. Dressman, and Patric J. Hahn, *Tet. Lett.*, **1996**, 37, 40, 7193-7196
- Kurzer 1963 a** : F. Kurzer, *Org. Syn. Coll.* **1963** Vol. 4, 52
- Kurzer 1963 b** : F. Kurzer, *Org. Syn. Coll.* **1963** Vol. 4, 49
- Lam 1991** : K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, and R. J. Knapp, *Nature*, **1991**, 354, 82-84
- Lastovskii 1946** : R. P. Lastovskii, *J. Appl. Chem. (USSR)*, **1946**, 19, 440
- Le 1995** : H. T. Le, M. Mayer, S. Thoret and R. Michelot, *Int. J. Peptide Protein Res.*, **1995**, 45, 138-144
- Lipinski 1997** : C. A. Lipinski, *Adv. Drug Deliv. Rev.*, **1997**, 23, 3-25
- Lipkowski 1986** : A. W. Lipkowski, S. W. Tam, P. S. Portoghese *J. Med. Chem.*, **1986**, 29, 1222
- Lozoya 1997** : Xavier Lozoya, *Pour la Science, édition spéciale « Les Médicaments »* **1997**, 34-39
- Luetzow 1979** : D. Luetzow, N. Neth, U. Wagner and K. Volkaner *U. S. Pat.*, Dec. 11, **1979**, 4, 178, 309
- Majer 1994** : P. Majer and R. S. Randad, *J. Org. Chem.*, **1994**, 59, 1937-1938
- Mazouz 1990** : F. Mazouz, L. Lebreton, R. Milcent, C. Burstein, *Eur. J. Med. Chem.* **1990**, 25, 659-671
- Merrifield 1963** : Merrifield, *J. Am. Chem. Soc.*, **1963**, 85, 2149
- Milcent 1991** : Milcent, R et al., *Journal of Heterocyclic Chemistry*, **1991**, 28, 6, 1511-1516
- Ming 1988** : Y. Ming and D. W. Boykin, *J. Heterocycl. Chem.*, **1988**, 25, 1729
- Mitchelitsch 1973 a** : W. Mitchelitsch and F. Hofmann *German Pat.*, Aug. 9, **1973**, 1, 768, 256
- Mitchelitsch 1973 b** : W. Mitchelitsch and F. Hofmann *Chem. Abstr.*, **1973**, 79, 146024
- Morton 1988** : R. C. Morton, D. Mangroo and G. E. Gerber, *Can. J. Chem.*, **1988**, 1701
- Norwich 1992** : J. S. Norwich, N. A. Powell, T. M. Ngugen and G.J. Noronha, *J. Org. Chem.*, **1992**, 57, 7364

- Ohta 1982** : S. Ohta, A. Shimabayashi, M. Aono and M. Okamoto, *Synthesis*, 1982, 833
- Ohlmeyer 1993** : M. H. J. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10922-10926
- Ostresh 1994 a** : J. M. Ostresh, G. M. Husar, S. E. Blondelle, B. Dörner, P. A. Weber, R. A. Houghten, *Proc. Natl. Acad. Sci. USA* **1994**, *31*, 11138-11142
- Ostresh 1994 b** : J. M. Ostresh, J. H. Winkle, V. T. Hamashin and R. A. Houghten, *Biopolymers*, **1994**, *34*, 1681-1689
- Otting 1959** : W. Otting and H. A. Staab *Liebig's Ann. Chem.*, **1959**, *622*, 23
- Owens 1991** : R. A. Owens, P. D. Gesellchen, B. J. Houchins and R. D. DiMarchi, *Biochem. Biophys. Res. Com.*, **1991**, *181*, 402-408
- Ozaki 1972** : S. Ozaki, *Chem. Rev* **1972**, *72*, 457
- Pinilla 1994** : C. Pinilla, J. R. Appel, S. E. Blondelle, C. T. Dooley, J. Eichler, J. M. Ostresh and R. A. Houghten, *Drug Development Research*, **1994**, *33*, 133-145
- Pinilla 1995** : C. Pinilla, J. R. Appel, S. E. Blondelle, C. T. Dooley, B. Dörner, J. Eichler, J. M. Ostresh and R. A. Houghten, *Biopolymers*, **1995**, *37*, 221-240
- Pirrung 1995** : M. C. Pirrung and J. Chen, *J. Am. Chem. Soc.* **1995**, *117*, 1240-1245
- Pop 1997** : Iuliana Pop, Benoît P. Déprez and André Tartar, *J. Org. Chem.*, **1997**, *62*, 2594-2603
- Price 1926** : T. W. Price, *J. Chem. Soc.*, **1926**, 3230
- Rüfenacht 1972** : K. Rüfenacht, *Helv. Chim. Acta* **1972**, *55*, 1174-1178
- Salmon 1993** : S. E. Salmon, K. S. Lam, M. Lebl, A. Kandola, P. S. Khattri, S. Wade, M. Patek, P. Kocis, V. Krchnak, D. Thorpe and S. Felder, *Proc. Natl. Acad. Sci. USA*, **1993**, 11708-11712
- Shachat 1963** : N. Shachat and J. J. Bagnell, Jr., *J. Org. Chem.*, **1963**, *28*, 991
- Sherman 1961** : W. R. Sherman, *J. Org. Chem.*, **1961**, *26*, 88-94
- Shingu 1957** : H. Shingu, T. Nishimura and T. Takegami, *Yuki Gosei Kagaku Kyokai Shi*, **1957** *15*, 140
- Shriner 1944** : R. L. Shriner, W. H. Horne and R. F. Cox, *Org. Syn. Coll.*, **1944**, *2*, 453
- Slocombe 1949** : R. J. Slocombe, E. E. Hardy and J. H. Saunders, *U.S. Pat.*, Aug. 23, **1949**, *2*, 480, 089
- Sneader 1996** : W. Sneader, *Drug Prototypes and their exploitation*, **1996** by John Wiley and Sons
- Staab 1956** : H. A. Staab, *Chem. Ber.*, **1956**, *89*, 2088
- Staab 1957 a** : H. A. Staab, *Liebig's Ann. Chem.*, **1957**, *609*, 75
- Staab 1957 b** : H. A. Staab, *Liebig's Ann. Chem.*, **1957**, *609*, 83
- Staab 1959** : H. A. Staab and W. Benz *Liebig's Ann. Chem.*, **1959**, *622*, 23
- Staab 1962** : H. A. Staab, *Ang. Chem.*, **1962**, *1*, 351
- Stankova 1994** : M. Stankova, O. Issakova, N. F. Sepetov, V. Krchnak, K. S. Lam and M. Lebl, *Drug Development Research*, **1994**, *33*, 146-156
- Stempel 1955** : A. Stempel, J. Zelauskas, J. A. Aeschlimann, *J. Org. Chem.* **1955**, *20*, 412-418
- Terrett 1995** : N.K. Terrett, D. Bojanic, D. Brown, P. J. Bungay, M. Gardner, D. W. Gordon, C. J. Mayers and J. Steele, *Bioorg. Med. Chem.*, **1995**, *5*, 9, 917-922
- Thompson 1954** : W. W. Thompson, U.S. Pat. Mar. 30, **1954** 2,673,877

Vollhardt 1990 : Vollhardt, « *Chimie Organique* », **1990**, De Boeck-Wesmael, 225

Walsh 1990 : D. A. Walsh, J. B. Green, S. K. Franzysen, J. C. Nolan and J. M. Yanni, *J. Med. Chem.*, **1990**, 33, 2028

Wermuth 1996 : Camille G. Wermuth, *The practice of medicinal chemistry*, **1996**, Academic Press, 24-28 Oval Road London NW1 7DX

Wieland 1953 : T. Wieland and G. Schneider, *Liebig's Ann. Chem.*, **1953**, 580, 159

Williard 1996 : X. Williard, I. Pop, D. Horvath, R. Baudelle, P. Melnyk, B. Déprez and A. Tartar, *Eur. J. Med. Chem.* **1996**, 31, 87-98.

Zhang 1997 : X. Zhang, J. Rodrigues, L. Evans, B. Hinkle, L. Ballantyne and M. Pena, *J. Org. Chem.* **1997**, 62, 6420-6423

Zhao 1995 : P. L. Zhao, R. Zambias, J. A. Bolognese, D. Boulton, K. Chapman *Proc. Natl. Acad. Sci., USA*, **1995**, 92, 10212-10216,

Zhao 1996 : P. L. Zhao, R. B. Nachbar, J. A. Bolognese, K. Chapman *J. Med. Chem.*, **1996**, 39, 350-352

Zuckermann 1994 : R. N. Zuckermann, E. J. Martin, D. C. Spellmeyer, G. B. Stauber, K. R. Shoemaker, J. M. Kerr, G. M. Figliozzi, D. A. Goff, M. A. Siani, R. J. Simon, S. C. Banville, E. G. Brown, L. Wang, L. S. Richter and W. H. Moos, *J. Med. Chem.*, **1994**, 37, 2678-2685



Résumé :

Les travaux décrits dans cette thèse concernent la mise au point et l'utilisation de techniques de chimie combinatoire et de synthèse parallèle à haut débit, appliquées à la découverte d'activités biologiques. Cette thèse s'articule en deux parties :

→ la première partie concerne d'une part la description, selon un modèle matriciel, de toute chimiothèque produite par chimie combinatoire et d'autre part la comparaison des différentes méthodes de déconvolution (déconvolution itérative, balayage positionnel et chimiothèques orthogonales) permettant de déterminer la structure d'un composé actif au sein d'un mélange. La méthode de déconvolution mise au point au laboratoire, faisant appel aux chimiothèques orthogonales, est illustrée par la description de la synthèse d'une chimiothèque de 15625 tripeptides, issus d'acides aminés non naturels, ayant permis la découverte d'un ligand nanomolaire du récepteur à la vasopressine (V2 porcine).

→ la seconde partie est consacrée à la mise au point d'un protocole de synthèse robotisée d'urées asymétriques à partir d'amines et du carbonyldiimidazole comme réactif de couplage. Les étapes de détermination des solvants, des concentrations en réactifs, des temps de réaction, des catalyseurs utilisés et du choix des réactifs et matières premières y sont détaillées. Ce protocole appliqué à la synthèse robotisée de deux chimiothèques d'urées asymétriques, la première comprenant 154720 composés et la seconde 4000 composés nous a permis de cibler les étapes de nos productions à optimiser, mais aussi de valider l'efficacité de nos chimiothèques pour la découverte d'activités biologiques. Le criblage de la seconde chimiothèque sur deux cibles biologiques : le récepteur morphinique μ et le récepteur benzodiazépinique BZDc, a d'ailleurs révélé un taux de succès de respectivement : 0,16 % et 0,38 % des composés testés.

Development and use of combinatorial and high-throughput parallel synthesis techniques applied to the discovery of biological activities

Summary :

The work described in this thesis concerns the use and adjustment of combinatorial and high-throughput parallel synthesis techniques for the discovery of biological activities. This thesis is organised into two parts :

→ the first part is concerned with (a) the description (using a matrix model) of all the combinatorial libraries and (b) a comparison between the different deconvolution methods (iterative, positional scanning and orthogonal libraries), used to define the structure of the active compound among a mixture. The deconvolution method using orthogonal libraries, developed in our laboratory, is illustrated by the description of the synthesis of a 15625 non-natural tripeptide library which led to the discovery of a nanomolar ligand of the V2 vasopressin receptor (porcine receptor).

→ the second part is dedicated to the development of a chemical procedure for the automated synthesis of asymmetric ureas starting from amines and carbonyldiimidazole. The different steps to determine the solvents, the reagent concentrations, the reaction time, the catalysts, and the reagents and starting materials chosen are described in this part. This procedure applied to the automated synthesis of two ureas libraries, the first one containing 154720 compounds and the second one containing 4000 compounds, allowed us to target the steps to be optimised but also to validate the efficiency of our libraries for the discovery of biological activities. Moreover the screening of the second library on two biological targets : opiate μ and benzodiazepine BZDc revealed a hit rate of respectively : 0,16 % and 0,38 %.

DISCIPLINE : Chimie organique et macromoléculaire

MOTS-CLES : synthèse combinatoire, modèle matriciel, déconvolution, chimiothèques orthogonales, vasopressine, synthèse parallèle, haut-débit, chimie automatisée, criblage, urées, carbonyldiimidazole, oxadiazolinone

ADRESSE ET INTITULE DU LABORATOIRE :

Service Chimie des Biomolécules URA CNRS 1309, Institut Pasteur de Lille
1, rue Calmette
59 Lille