

5 EVOLUTION SIMULEE D'UNE NOUVELLE BIBLIOTHEQUE DE CATALYSEURS CHIRAUX

L'évolution dont nous venons de produire les résultats, s'il est vrai qu'elle fonctionne, pêche tout de même par un manque de rigueur et d'efficacité, les deux étant d'ailleurs sans doute liés. Le concept a le mérite d'une certaine originalité ; il importait donc de poursuivre dans cette voie en apportant améliorations et modifications nécessaires.

5.1 Cadre de l'évolution simulée

Comme pour tout projet de recherche, il convenait d'en bien assurer les bases pour que la suite soit fructueuse.

La thématique du laboratoire se focalise sur la catalyse asymétrique, et c'est bien évidemment dans cette voie que nous avons décidé de poursuivre le travail, d'autant plus qu'à notre connaissance, il n'a pas encore été tenté à ce jour d'appliquer la machinerie de l'évolution simulée en catalyse homogène asymétrique. C'est grâce à une collaboration d'une part avec l'équipe du Dr de Bellefon, en ce qui concerne la phase de tests, et d'autre part grâce au Pr Bernadette Govaerts pour le traitement des données, que cette partie put être menée à bien.

5.1.1 Réaction test

A nouveau l'étude se réalisera en testant des systèmes catalytiques par la réduction de l'acétophénone. Toutefois plusieurs possibilités s'offrent à nous dans ce domaine comme le résume brièvement le schéma 112.

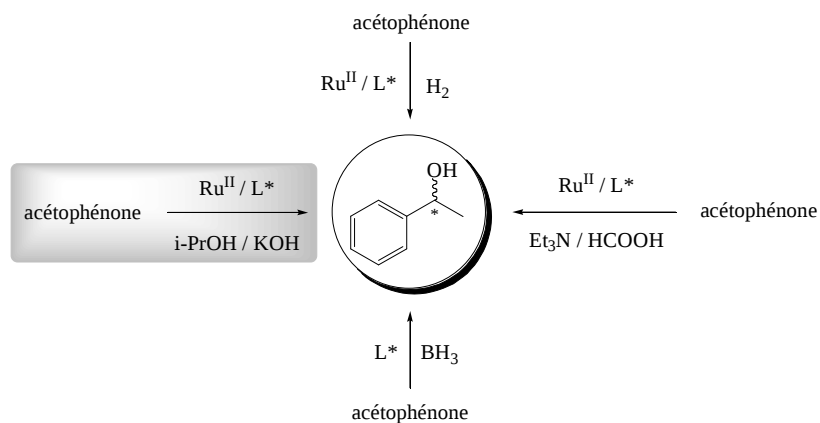


schéma 112

La réduction sous pression d'hydrogène n'a plus à démontrer son efficience ; elle nécessite par contre l'utilisation de matériel spécifique, et d'emblée, l'adaptation difficilement réalisable du matériel de test à Lyon nous invita à ne pas explorer cette manière de procéder. En ce qui concerne la réduction impliquant des borohydrures asymétriques, on se souviendra de la remarque qui fut donnée dans la partie introductive à propos de leur incompatibilité factuelle avec certains groupements réactionnels. Elle ne tient toutefois plus dans le cas de la réduction de l'acétophénone. Cette voie de réduction n'a malgré tout pas été retenue. Le système $\text{HCOOH} / \text{Et}_3\text{N}$ que nous avons aussi mentionné antérieurement offre l'avantage d'une réaction non équilibrée ; par contre son inaptitude à promouvoir la réduction de l'acétophénone par l'intermédiaire d'un complexe résultant de l'association d'un précurseur de Ru^{II} et d'un amino-alcool nous mena à considérer que la réduction classique dans un milieu $i\text{-PrOH} / \text{KOH}$ restait encore la meilleure option. En effet ce système s'avère d'une grande compatibilité, d'une mise en œuvre simple, et peu onéreux. De plus, c'est un système dont nous possédons l'expérience au laboratoire, ce qui nous affranchit normalement de toute mauvaise surprise.

Ensuite, il nous fallut porter notre choix sur un type de ligand dont on pouvait formuler le « code génétique » de manière aussi simple que rigoureuse, et appliquer ce même principe aux précurseurs métalliques.

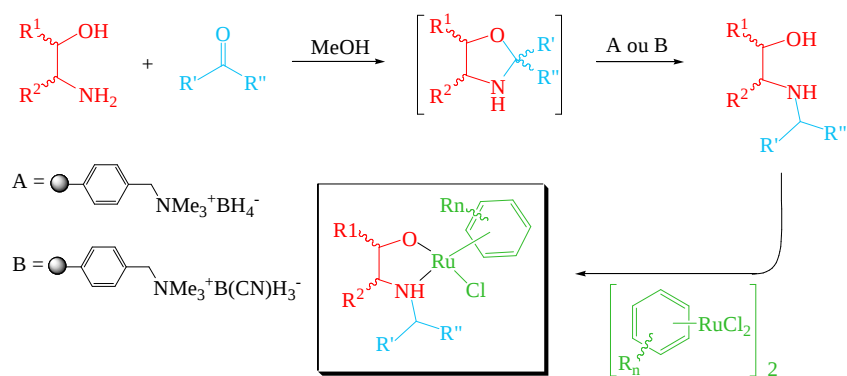
5.1.2 Amino-alcools, carbonyles et précurseurs métalliques

Nous n'avions pas pour ambition de mettre en lumière un nouveau couple métal-ligand particulièrement efficace sur l'acétophénone. La littérature donne de nombreux exemples de systèmes accédant à plus de 95 % ee, que ce soit en hydrogénation ou en transfert d'hydrure. Notre intérêt était de montrer que parmi nos systèmes, nous étions capables de retrouver ceux qui se sont avérés efficaces, sans être obligés de tous les tester. Afin d'éviter le risque d'obtenir une majorité de résultats aux très faibles performances, notre choix s'est porté sur des ligands de type amino-alcools. La pince chélatante « N,O » a par maintes occasions pu prouver son efficacité en transfert d'hydrure. Ces ligands furent construits à partir d'amino-alcools, et nous avons décidé d'ensuite procéder à une réaction d'amination réductrice sur ces derniers. C'est notamment suite aux travaux de Carpentier et van Leeuwen^{69, 71} montrant à la fois l'influence que pouvait avoir le substituant de l'amine, et l'importance de conserver un hydrogène sur cette amine, que nous déterminâmes notre choix. Le code peut alors s'établir simplement en considérant les blocs de construction. L'amino-alcool de départ sera codé $A_1 \rightarrow A_x$; le groupement alkylant $B_1 \rightarrow B_y$ et chaque ligand aura donc un code de type « $A_i B_j$ ».

Les précurseurs métalliques sont des complexes de ruthénium de type $[RuCl_2(\text{arène})]_2$. Eux aussi sont très régulièrement employés en réduction par transfert d'hydrure et ont prouvé leur efficacité. La variation dans ce cas se situera au sein du groupement arène coordonnant le ruthénium en η^6 . Chaque précurseur métallique se verra ainsi attribué un code de type $M_1 \rightarrow M_z$.

Afin de bâtir une banque de composés de taille raisonnable mais suffisante, 12 amino-alcools furent choisis pour servir de base, chacun d'eux serait couplé à 30 groupements carbonyles par le biais de synthèses parallèles. Ces divers couplages réductifs amines-carbonyles, réalisés grâce à des réactifs supportés sur résine, devaient nous permettre de générer une banque de 360 ligands chiraux. Chaque ligand serait alors testé indépendamment face à 6 précurseurs métalliques différents pour permettre la réduction de l'acétophénone. Au total, 2160 systèmes catalytiques sont

donc théoriquement accessibles *in situ* par cette méthode, comme le montre le schéma 113.



"catalyseur type" de code génétique $\text{A}_x\text{B}_y\text{M}_z$

schéma 113

Dans le détail, les 12 amino-alcools choisis pour entrer dans l'élaboration de la banque sont représentés au schéma 114. Certains sont commerciaux, d'autres seront synthétisés par nos soins, comme le détaillera la suite. On notera que les positions en α de l'alcool et de l'amine sont fonctionnalisées de manière à pouvoir évaluer l'impact de l'encombrement stérique et parfois l'influence de la stéréochimie absolue qu'offre cette position. Le carbone en α de l'amine est toujours S. Un des amino-alcools, A_5 , possède une structure tout à fait différente, il devrait être utile si l'on veut imposer au système une variation importante et en vérifier les effets.

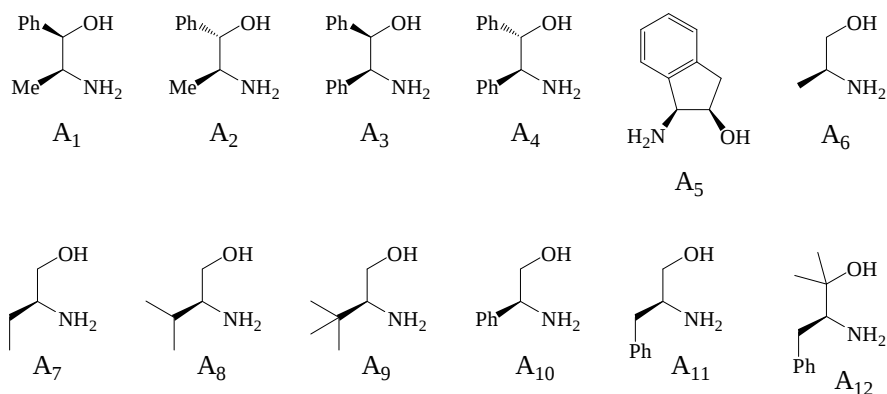


schéma 114

Il est prévu que ces 12 amino-alcools subissent une alkylation par réaction d'amination réductrice. Les candidats carbonyles sont donnés ci-dessous (schéma 115). On peut en dénombrer 4 types dont les caractéristiques apparaissent aisément. D'abord sont présents quelques aldéhydes aliphatiques ; la série se poursuit au moyen de cétones. Les cétones cycliques sont séparées par une unité de taille de cycle uniquement. La part belle est réservée aux suivants : les aldéhydes aromatiques, dont les structures montrent notre souci de varier l'encombrement stérique du cycle. On remarquera la présence de dérivés naphtyles et anthranyles. Pour conclure la série, les 6 derniers aldéhydes possèdent tous un hétéroatome en position β . Cette plus grande différence a été introduite pour favoriser la formation de ligands potentiellement tridentates et induire un mode de complexation différent, ce qui pourrait revêtir un caractère particulièrement attrayant. Tous les groupements carbonyles que nous ne possédions pas au sein de l'unité ont été achetés chez les fournisseurs habituels.

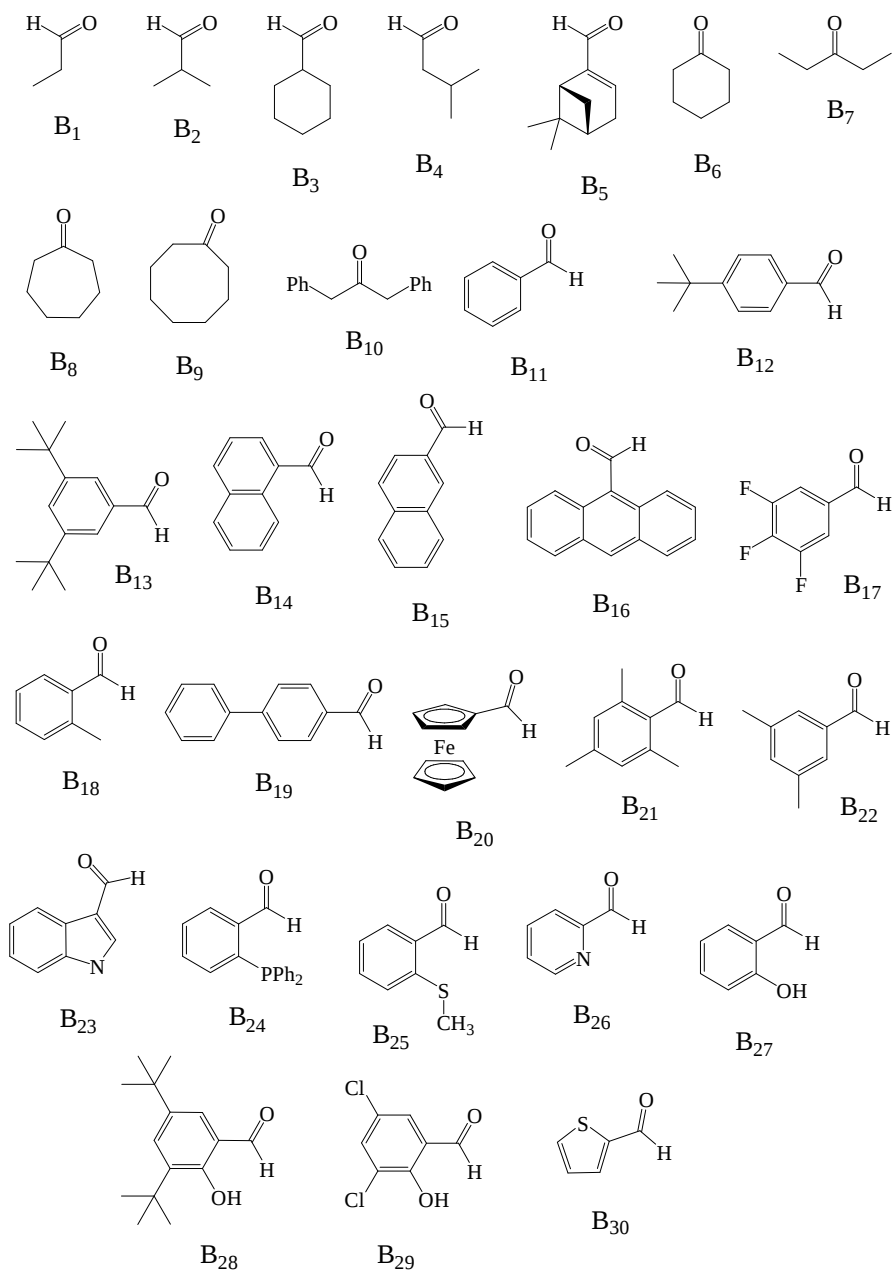


schéma 115

Ainsi que nous l'avons précisé plus haut, il est nécessaire d'introduire une variation dans le chef du précurseur métallique utilisé pour la formation du catalyseur. Le schéma 116 détaille les 6 précurseurs différents qui feront l'objet d'une utilisation. On notera que 5 de ceux-ci sont de la famille de « arène ruthénium ». A l'évidence, nous espérons découvrir des variations dans le comportement de nos catalyseurs en fonction des différents groupements arènes mis en jeu. Le dernier, basé également sur le Ru^{II} est chélaté par 4 molécules de DMSO. Ce mode de coordination original au regard des précédents est ici aussi entrevu sous l'angle d'un changement majeur proposé au système et fournira peut-être des résultats d'intérêt.

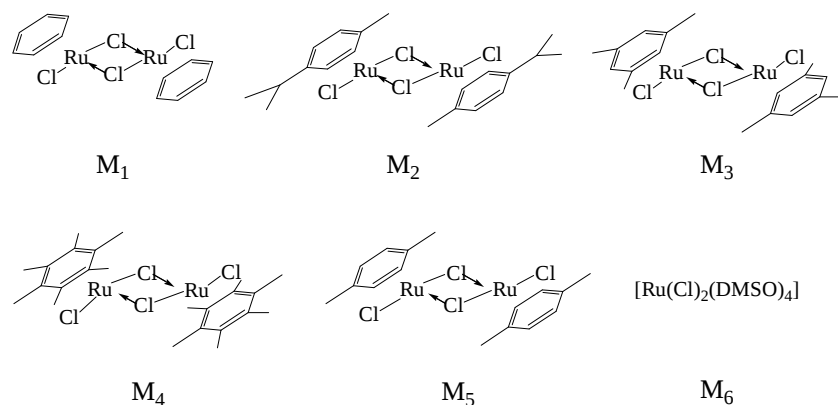


schéma 116

L'ensemble des catalyseurs vont être évalués, et pour ce faire nous supposons arbitrairement que tous empruntent le même cycle catalytique pour réduire l'acétophénone. Ce cycle catalytique est proposé de manière volontairement simplifié au schéma 117. Il est clair pourtant que les catalyseurs formés de ligands tridentates suivront peut-être une voie différente. De même est-il peu probable que l'action du complexe $[\text{RuCl}_2(\text{DMSO})_4]$ soit bien décrite par ce même cycle, tant sa structure est éloignée des 5 autres complexes.

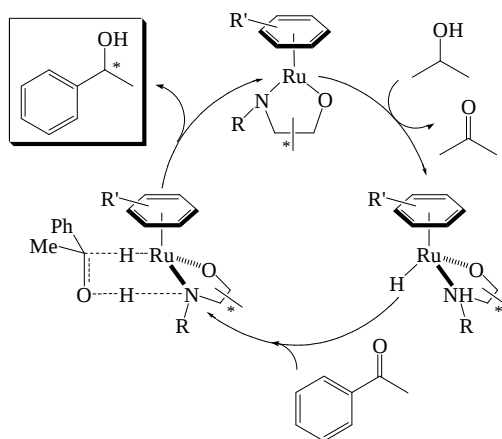


schéma 117

5.1.3 Principe général de l'évolution simulée

D'après ce que nous avons déjà décrit de l'évolution simulée, il est entendu que l'on part d'un ensemble restreint de composés à tester, dans notre cas les catalyseurs. Cet ensemble représente une part congrue de l'ensemble des possibilités que l'on imagine théoriquement pouvoir tester. Pour notre propos, nous chercherons à critiquer et valider la méthode pour la catalyse asymétrique en phase homogène ; ce qui signifie en premier lieu de disposer de tous les résultats afin de permettre toute comparaison, vérification ou investigation utile. Il importe donc, en supposant que tout se déroule comme prévu, que la bibliothèque de catalyseurs soit évaluée dans son entièreté. Une fois ces résultats connus, nous comptons sélectionner des systèmes catalytiques pour en faire notre première génération et la soumettre au processus itératif évolutif. La manière dont cette sélection s'opère fera l'objet d'un point de discussion ultérieur. Cette génération mère subira alors, grâce aux opérateurs génétiques de réplication, de cross-over et de mutation, une métamorphose partielle et deviendra la génération 2, ce que décrit la figure 13. On notera que l'opérateur de régression, utilisé dans le cadre de l'optimisation dirigée, est absent. La raison est double : premièrement, l'algorithme eut été plus compliqué à programmer en tenant compte de cette possibilité supplémentaire, et deuxièmement, la régression si elle intervient sur le métal, entraînera la composition d'un catalyseur *de facto* inactif.

L'évolution sera menée grâce à une méthode informatique, ce qui permettra d'introduire la notion aléatoire qui manquait à l'optimisation dirigée menée au chapitre précédent, de la quantifier et de la contrôler.

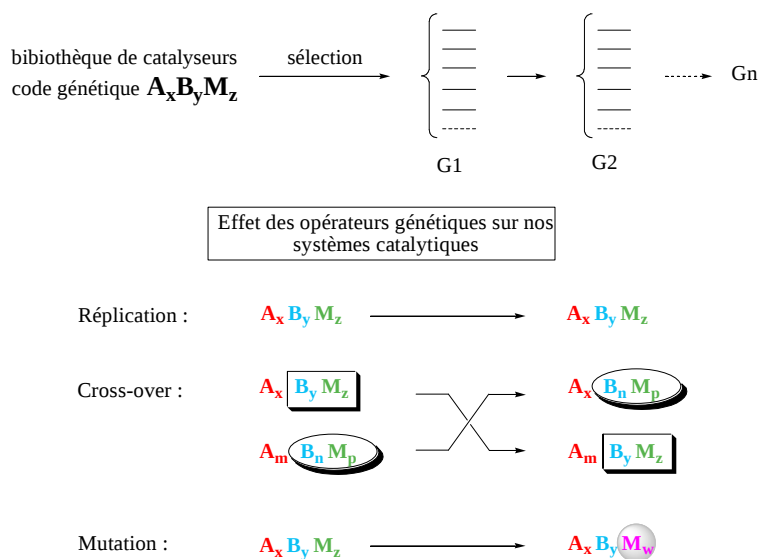


figure 13

Nous procéderons ainsi de suite pour parvenir à certains systèmes hautement performants en quelques générations. Il est entendu que les systèmes que nous qualifions de hautement performants peuvent ne l'être qu'au sein de notre bibliothèque. La figure 14 montre l'évolution typique que nous espérons être à même de réaliser (ici, le nombre de générations n'est pas même indicatif et pourra varier grandement) ; les points de la courbe rendent compte de rendements et d'ee moyens calculés pour chaque génération. Si notre évolution simulée est bien menée, nous devrions à chaque génération constater une amélioration successive de ces deux paramètres d'estimation de la réaction.

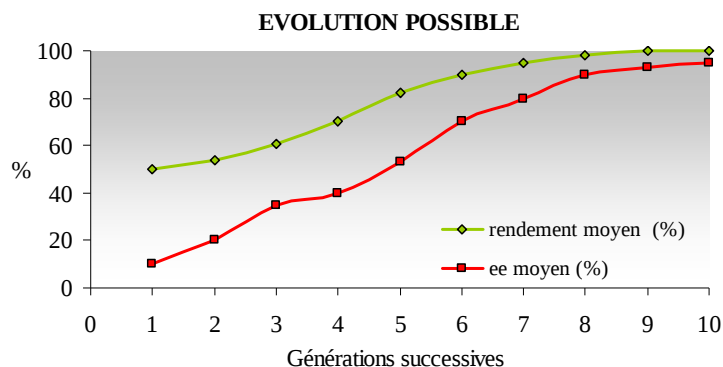


figure 14

5.2 Synthèse des amino-alcools

Parmi les amino-alcools qu'il nous est nécessaire de posséder en vue de la synthèse ultérieure de nos ligands, certains peuvent être obtenus assez aisément par réduction des acides aminés correspondants^{191, 192}.

L'acide aminé est mis en présence de NaBH_4 . Ensuite, et à froid, I_2 est ajouté au milieu, ce qui entraîne un dégagement de H_2 . Le mélange réactionnel est alors porté à reflux pendant 18 h typiquement. On termine la réaction par un work-up basique et le produit de réduction est récupéré. Le schéma 118 donne le détail de cette réaction.

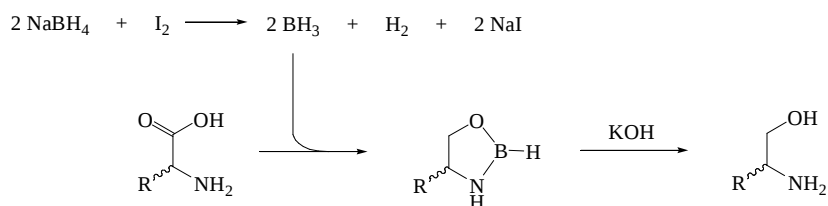


schéma 118

Cette manière de procéder sera utilisée pour produire le *L-tert*-leucinol, le *L*-leucinol et le *L-iso*-leucinol, comme le détaille le schéma 119. On remarquera d'emblée la présence de *L*-leucinol et *L-iso*-leucinol, ce qui n'était à priori pas prévu au sein de la bibliothèque d'амино-alcools envisagés. Toutefois des contraintes liées à la non disponibilité commerciale de certains amino-alcools ou à leurs prix prohibitifs nous ont invités à

poursuivre notre stratégie en modifiant les composés que nous allons utiliser.

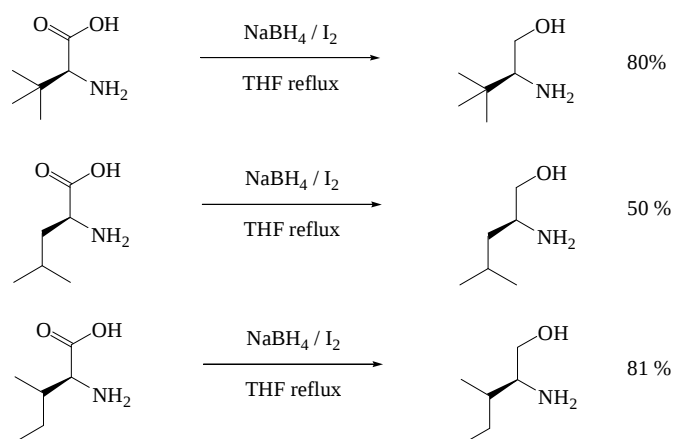


schéma 119

Nous avons également prévu de synthétiser le composé A12 au moyen d'un réactif de Grignard, formé par la présence simultanée d'iodure de méthyle et de magnésium dans l'éther (schéma 120).

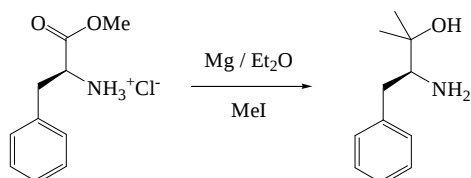


schéma 120

Cette procédure nous confronta, lors de la phase d'extraction au CH_2Cl_2 , à une émulsion particulièrement handicapante puisqu'après avoir utilisé 4 litres de ce solvant, la séparation de phase ne s'opérait toujours pas de manière satisfaisante et rendit l'extraction caduque. Il a donc été décidé d'opérer de manière légèrement différente, en remplaçant le MeI par du MeBr (schéma 121).

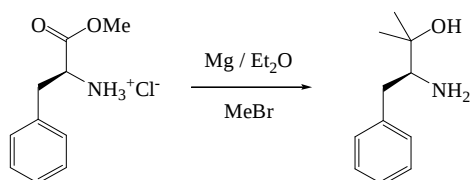


schéma 121

Lors de la réaction, la formation du réactif de Grignard se déroule à froid, s'en suit alors l'addition de l'ester méthylique, puis un chauffage à reflux dans l'éther qui doit assurer la formation de l'alcool di-méthylé. Nous avons dû malheureusement constater l'agglomération du magnésium dès le début de l'étape de chauffe, et malgré une trituration énergique du milieu réactionnel, la prise en masse se réalise à nouveau progressivement. Malgré un reflux d'une nuit entière, jamais le magnésium ne se solubilise. Nous avons jugé inutile de traiter cette réaction, déduisant de la prise en masse nocturne qu'elle n'avait sans doute permis qu'une conversion médiocre, au mieux.

Pour s'affranchir de ce problème, nous avons répété cette réaction en usant de MeBr 0,3 M en solution dans l'éther. Le mélange réactionnel est d'abord visqueux, ce qui nous invite à ajouter la quantité d'éther nécessaire à assurer la solubilité avant l'étape de chauffe, et permet d'éviter la déconvenue précédente. Après une purification au four à boule nous observons un spectre proton qui montre un produit propre. Toutefois, et comme le montre la figure 15, nous voyons apparaître de manière inexplicable deux pics superflus, typiquement dans la région des groupements méthyles.

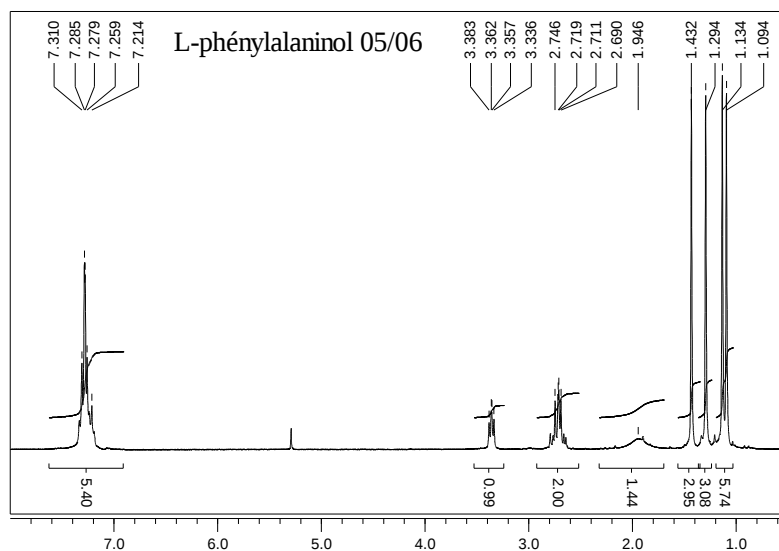


figure 15

Il est difficile d'imaginer qu'une autre position ait pu subir une diméthylation, d'autant plus que les signaux clairs du CH_2 , du proton en α de l'amine, des protons aromatiques et de l'association $\text{NH}_2 + \text{OH}$ sont sans nul doute présents, avec les valeurs d'intégration correspondantes, aux erreurs habituelles près. Notons que le spectre de masse s'il donne un pic correspondant *a priori* à celui de l'ion moléculaire à 220,18 ; fournit aussi un pic de forte intensité à 180,14 correspondant au L-phénylalaninol diméthylé, ainsi qu'un pic à 162,17 résultant de la perte d'un molécule d'eau par ce dernier. Ce composé fut le dernier de la série des amino-alcools que nous essayâmes de synthétiser peu avant de réaliser les dernières aminations réductrices en vue d'obtenir nos ligands. Cet échec apparent nous conduisit à ne pas réaliser la dernière série de couplages.

Le produit de réaction fut stocké tel quel dans une armoire et à l'air. On observa une lente cristallisation. Dernièrement repris pour analyse, il montra une heureuse évolution du spectre proton puisque le produit désiré est finalement obtenu (figure 16), certes trop tard, avec un rendement de 82 %. Les deux pics attribués à deux méthyles superflus ont disparu ; par contre les deux protons du CH_2 sont visibles distinctement à 2,81 et 3,02 ppm, ainsi que le proton en α de l'amine à 2,27 ppm. Les protons du l'amine et l'alcool

associés ne sont plus discernables, ce qui est courant pour les amino-alcools et témoigne de protons libres et donc d'un échange beaucoup plus rapide avec le solvant.

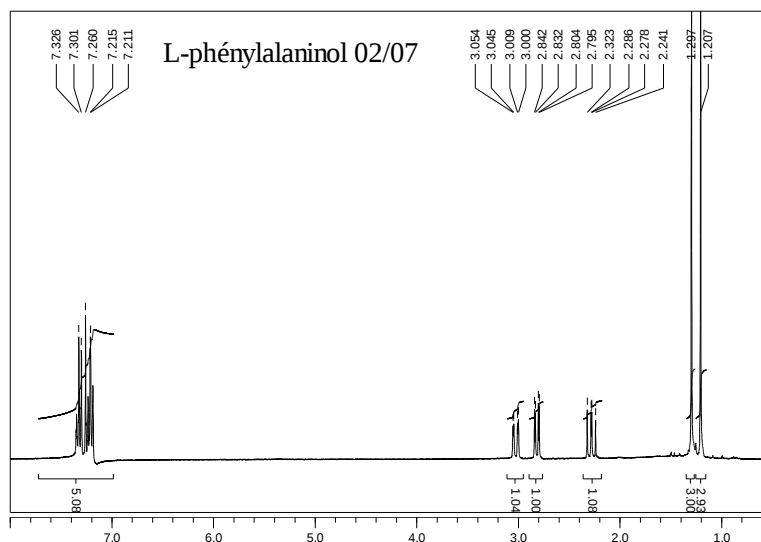


figure 16

La différence entre le premier composé obtenu et le L-phénylalaninol diméthylé pourrait peut-être s'expliquer en admettant l'introduction maladroite d'acétone dans le milieu. Il y aurait alors possibilité de former un cycle à 5 chaînons, pour lequel une lente hydrolyse à l'air est envisageable (schéma 122). Dans ce cas la différence entre les masses molaires des composés est précisément de 40, ce qui correspond au résultat de l'analyse de masse. La présence de deux méthyles supplémentaires visibles en RMN ^1H trouverait aussi une explication.

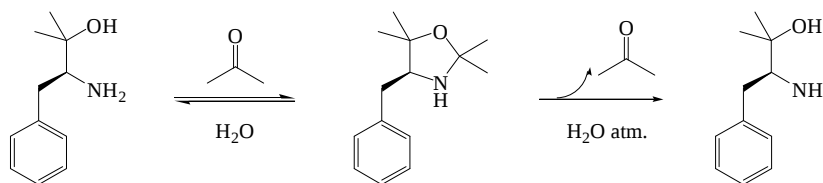


schéma 122

Au final, la série des amino-alcools que nous emploierons effectivement dans les réactions de synthèse des ligands est donnée ci-dessous (schéma 123). La numérotation a été variée, elle rend compte simplement de l'ordre de synthèse de ces familles de ligands, lequel a été dépendant de la disponibilité en amino-alcools.

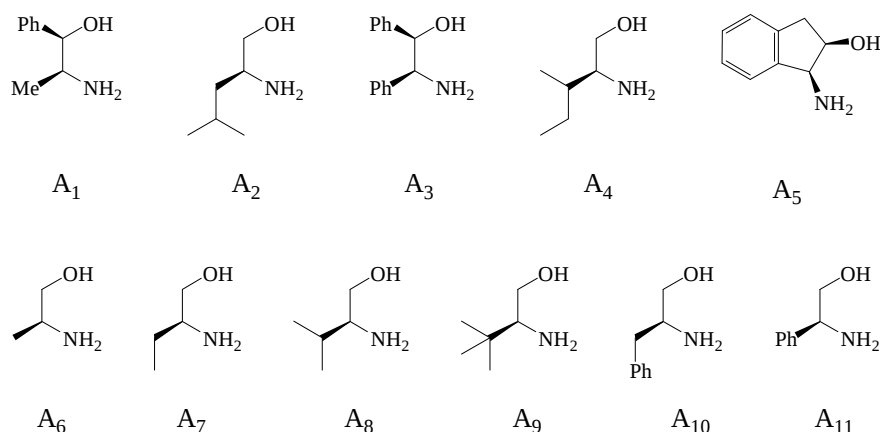


schéma 123

5.3 Synthèse des précurseurs au Ru

La synthèse des six précurseurs métalliques est prévue suivant un mode bien établi¹⁹³ : Le cycle aromatique que l'on veut coordonner au ruthénium sera d'abord réduit en diène par une réaction de Birch ; une oxydoréduction consécutive impliquant du RuCl₃ mènera normalement au complexe désiré. Le schéma 124 détaille cette réaction.

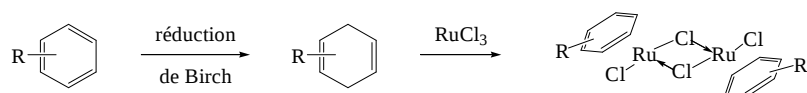


schéma 124

5.3.1 Synthèse des diènes aromatiques

Certains diènes étaient disponibles au laboratoire, il ne nous a donc pas été nécessaire de procéder à la réduction de chacun d'entre eux.

Le 1,3,5-triméthylcyclohexa-1,4-diène nécessite par contre la réduction du mésitylène (schéma 125).

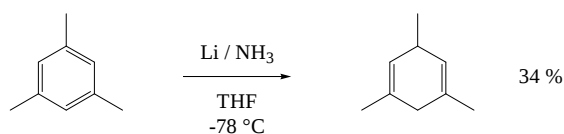


schéma 125

La réduction de Birch se déroule sous atmosphère inerte et débute par la condensation de l'ammoniac à -78 °C, le lithium est ajouté et entraîne l'apparition d'une couleur bleue caractéristique. Le cycle aromatique à réduire est alors ajouté lentement en solution et laissé sous agitation à basse température.

La réaction nécessite évidemment des conditions anhydres strictes, ce qui n'est pas toujours évident à respecter. Toutefois, la solution bleue vire au blanc laiteux dès l'introduction d'une faible quantité d'eau dans le montage. Cette réaction dut d'ailleurs être reproduite plusieurs fois jusqu'à ce que la teneur en diène atteigne 90 % par rapport au mésitylène, après une distillation au four à boules. Dans ce dernier cas, le mésitylène était mis en solution dans le *tert*-butanol, ce qui semble offrir des conditions de réaction plus adéquates. Le rendement final est de 34 %. Notons que la présence de mésitylène ne s'avèrera pas un élément perturbateur lors de la réaction de réoxydation par le RuCl₃.

On utilisera des conditions identiques pour réduire le *para*-xylène en 1,4-diméthylcyclohexa-1,4-diène (schéma 126). Ici aussi, le *tert*-butanol sera préféré au THF, avec un rendement en diène de 36 %.

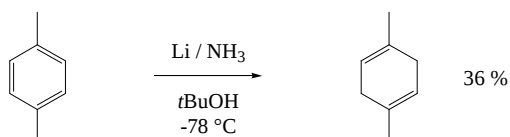


schéma 126

Nous avons également tenté de réduire l'hexaméthylbenzène de cette manière (schéma 127). Aucun produit de réduction n'est observé, ce qui peut

sans doute s'expliquer par le fait que les 6 méthyles que porte l'aromatique, du fait de leur influence de type inductif donneur rendent le potentiel de réduction trop faible, et le couple Li / Li^+ inapte à opérer le passage au diène. Nous devons donc passer par une autre voie afin d'obtenir le complexe du ruthénium correspondant.

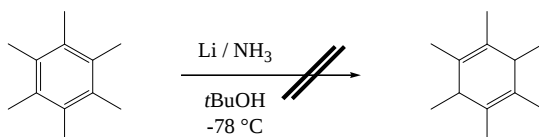


schéma 127

5.3.2 Synthèse des précurseurs métalliques

La synthèse des précurseurs est en général assez simple puisqu'il s'agit de mettre le diène en présence de RuCl_3 et de porter le milieu réactionnel à reflux dans un minimum d'éthanol pendant plusieurs heures pour obtenir le complexe centré sur le métal réduit.

Cette réaction a donc logiquement été appliquée à la synthèse du complexe « M1 » associant le benzène au ruthénium (schéma 128). Le premier essai montra tout de suite l'apparition de la couleur ocre caractéristique de ce genre de complexes, toutefois les déplacements chimiques en RMN ^1H ne correspondaient pas à ceux décrits dans la publication de Bennet¹⁹³. On procéda à un second essai, avec une source différente de ruthénium et une attention toute particulière à ce qu'aucune entrée d'air ne vienne altérer l'atmosphère inerte de rigueur. Le problème se répéta malgré ces précautions supplémentaires. Il nous fut alors possible de comparer nos déplacements chimiques avec ceux que donnaient le complexe d'origine commerciale pour comprendre que les données de la publication étaient fausses et que nous avions obtenu dès le début le di- μ -chloro-bis[(η -benzène)chloroiruthénium II] avec un rendement de 92 %.

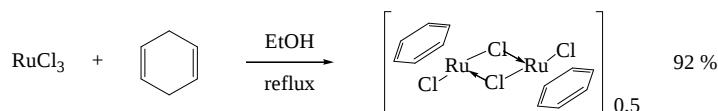


schéma 128

Nous disposions du complexe commercial « M2 », dont le cycle aromatique est le *para*-cymène, en quantité suffisante ; la synthèse nous en fut donc épargnée. Le 1,3,5-triméthylcyclohexa-1,4-diène allait par contre nous servir à obtenir le complexe « M3 » pour lequel le mésitylène coordine l'atome de ruthénium (schéma 129). Si la synthèse de « M1 » nécessitait environ 4 h de reflux, 16 h sont ici nécessaires à l'obtention du produit désiré, avec un rendement de 86 %.

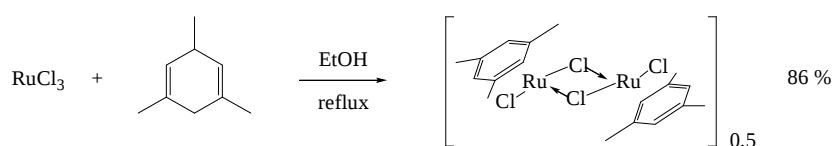


schéma 129

Une dernière fois, cette réaction fut mise en œuvre pour obtenir le précurseur métallique « M5 » (schéma 130). Le reflux fut mené pendant une nuit, car après 5 h de réaction, le changement de couleur caractéristique témoignait d'un état d'avancement incomplet de réaction. Après ce temps réactionnel, le rendement est de 9 %.

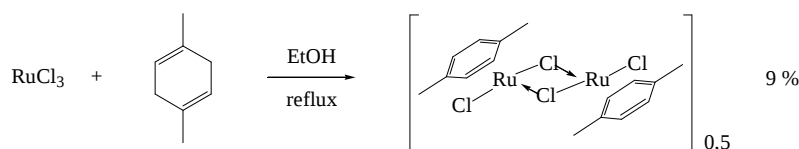


schéma 130

La tentative de réduction de l'hexaméthylbenzène s'était révélée infructueuse, ce qui nous amena à réaliser un échange de ligand pour obtenir « M4 » (schéma 131). Le réactif de base est notre complexe « M2 » ; attendu que le noyau aromatique est coordonné au centre métallique suivant un mode η^6 , on comprendra que la réaction se déroule à reflux directement dans l'hexaméthylbenzène présent à raison de 10 équivalents. Notons encore que la purification se déroule par un lavage à l'hexane. En réalisant ce lavage à chaud, la séparation de l'hexaméthylbenzène est plus efficace encore. Trois lavages successifs dans ces conditions restent néanmoins nécessaires pour parvenir à un mélange composé à 98,2 % de $[\text{RuCl}_2(\text{HMB})]_2$, 1 % de « M2 »

et moins de 1 % d'hexaméthylbenzène, ce que permet de déterminer l'analyse du spectre proton. Le rendement de cette réaction est de 49 %.

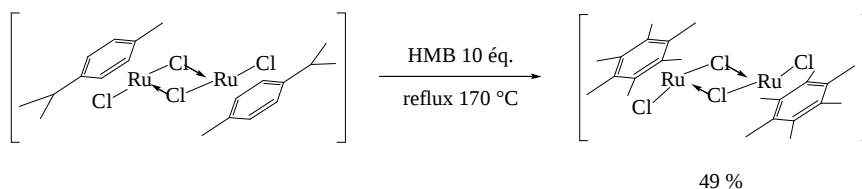


schéma 131

On terminera la partie consacrée à ces précurseurs en synthétisant « M6 »¹⁹⁴. La méthode varie bien entendu puisque le complexe que nous cherchons à produire est de structure tout à fait différente, comme le montre le schéma 132. La réaction se déroule dans le DMSO à reflux et assez rapidement puisque 5 minutes seulement s'avèrent nécessaire. Une précipitation dans l'acétone permet l'obtention de « M6 » avec un rendement de 70 %.

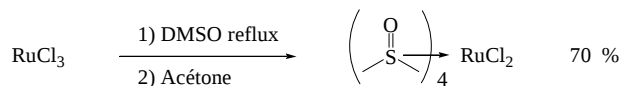


schéma 132

5.4 Synthèse des ligands chiraux

Nous avons donc initialement prévu de nous atteler à disposer de 360 ligands chiraux. Même si ce nombre fut en fin de parcours revu à la baisse du fait de la difficulté à obtenir l'amino-alcool « A12 » dans les temps, le nombre de 330 ligands mérite que l'on s'attarde à bien planifier les opérations. La réaction d'amination réductrice est bien connue ; elle peut être mise en œuvre aisément et permet d'atteindre souvent des rendements de l'ordre de 80 % et plus. Une étape ultérieure de purification par chromatographie sur colonne de silice peut alors être envisagée afin de récolter le produit avec une pureté satisfaisante. Toutefois, il va sans dire que cette méthode de synthèse individuelle n'est plus envisageable d'un point de vue pratique lorsque l'on considère la production de plus de 300 molécules.

Deux variations majeures ont été introduites et qui devaient nous aider à raccourcir les délais de synthèse tout en nous simplifiant la tâche :

- L'utilisation de réactifs supportés sur résine ; de cette manière une simple filtration est requise pour les séparer du milieu réactionnel.
- L'utilisation de Miniblock[®], mis au point par la firme Mettler-Toledo, afin de pouvoir conduire ces synthèses en parallèle.

Toutefois, la synthèse multipuits impose également des contraintes puisque chaque réaction sera menée sous des conditions identiques, notamment de temps et de température, alors que les réactifs engagés déclinent des variations structurales. Ces conditions identiques empêchent que l'on puisse à coup sûr exclure un phénomène partiel de dialkylation (schéma 133) même si les réactions vont normalement être envisagées en plaçant dans le milieu 1 équivalent d'agent alkylant. Cette dialkylation ne serait de toute façon que peu gênante puisque les ligands amino-alcools porteurs d'une amine tertiaire sont connus pour leur inactivité en réduction par transfert d'hydrure au moyen d'un précurseur de ruthénium.

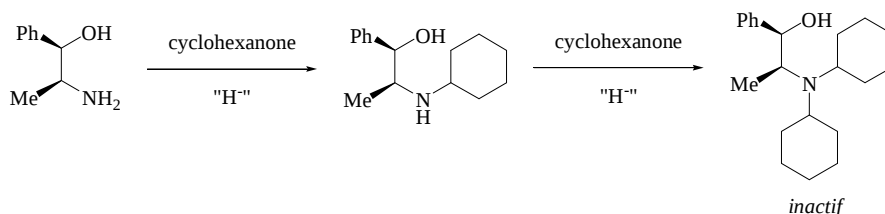


schéma 133

Un haut rendement est préférable également puisque nous savons que l'amino-alcool de départ pourra lui aussi agir comme ligand et s'associer au précurseur métallique pour promouvoir la réduction de l'acétophénone. Nous avons utilement vérifié que les conversions engendrées par l'amino-alcool de départ **139** et le composé alkylé **141** étaient du même ordre de grandeur, ce que présente la figure 17. Après 1 h 30 de réaction, les ee sont eux aussi du même ordre de grandeur (**139** = 63 % ee ; **141** = 72 % ee). De ces observations, il découle donc que si l'amino-alcool est encore présent en part minoritaire après la réaction d'amination réductrice, il ne devrait pas

empêcher que le comportement observé corresponde à celui de l'amino-alcool alkylé. Nous ignorons toutefois si cette relative similitude des activités est applicable pour l'ensemble des ligands que nous envisageons de synthétiser.

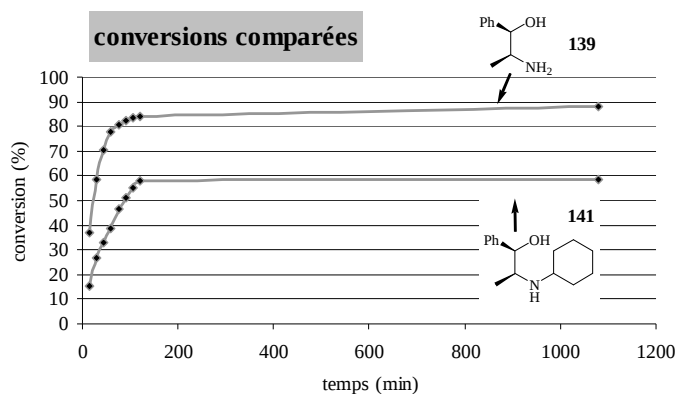


figure 17

5.4.1 Tests de synthèse en réacteur classique et Miniblock®

Le premier pas vers les ligands que nous désirons posséder implique de trouver les conditions générales qui offriront un état d'avancement suffisant du maximum de réactions. A cette fin, nous avons décidé d'engager un amino-alcool simple, achiral et donc bon marché, en couplage réducteur avec l'isobutyraldéhyde (schéma 134) et d'évaluer le temps de réaction nécessaire à l'obtention du composé **138**. Les réactions ci-dessous et celles qui seront décrites ci-après se déroulaient en ballon, le passage au Miniblock® sera l'objet d'une étape ultérieure.

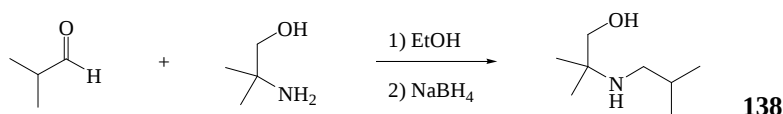


schéma 134

Le milieu réactionnel est d'abord agité pendant un temps défini afin de former l'imine ; le réducteur est ensuite ajouté et l'agitation se poursuit suivant le même intervalle de temps. La réaction fut reproduite plusieurs fois au vu de la disparité des rendements observés, comme l'indique le tableau 6.

tableau 6

Entrée	Temps (min)	Rdt 138 (%)
1	15	/
2	30	30
3	180 (3 h)	/
4	900 (15 h)	30
5	30	60
6	30	82
7	30	26
8	30	60

Il est finalement apparu que le produit sublimait lors de la dernière étape d'évaporation des solvants, le rendement devenant alors majoritairement dépendant de la qualité de la basse pression du rotavapor et de la rampe à vide.

Il a alors été décidé de passer à un composé plus lourd, et de nous mettre en conditions réelles en prenant un amino-alcool dont nous disposions déjà, la (1*S*,2*R*)-(-)- noréphédrine « A1 », et en le couplant à un des carbonyles que nous utiliserions par la suite. Il semblait logique d'utiliser une cétone, en raison de sa moindre réactivité comparée aux aldéhydes à engager. Notre choix se porta sur la cyclohexanone. La réaction est donnée au schéma 135 et sera menée en variant plusieurs paramètres comme le décrira la suite.

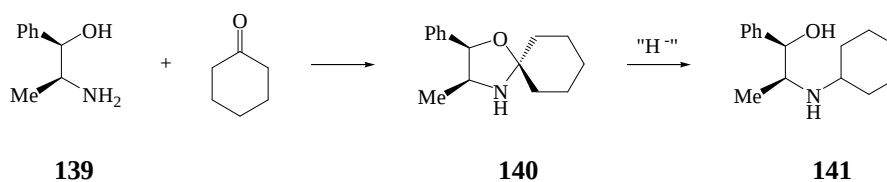


schéma 135

Le produit de réaction **141** sera analysé en RMN ¹H, ce qui permet bien entendu d'identifier la présence résiduelle en réactif **139** ainsi qu'en oxazolidine **140**, comme le décrit le schéma 136. L'oxazolidine peut exister en équilibre avec la forme imino-alcool, que nous ne détectons toutefois pas

dans les premières expériences. Les déplacements chimiques sont donnés pour une analyse réalisée dans le CDCl_3 . Le doublet caractéristique du méthyle ainsi que celui du proton en α de l'alcool serviront à quantifier la teneur du mélange en **139**, **140** et **141**.

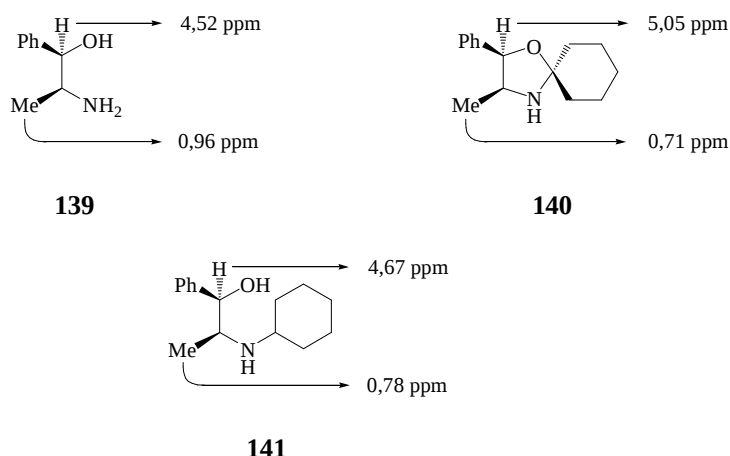


schéma 136

Il apparaît logique de tester en premier lieu le réducteur classique NaBH_4 ainsi que son homologue supporté sur résine que nous nommerons MP- BH_4 dont la firme Sopachem nous avait gracieusement livré un échantillon. Nous travaillons sur 1 mmol de (1S,2R)-(-)- noréphédrine, ce qui correspondra à 1 équivalent. « A1 » (**139**) et la cyclohexanone sont mis en solution dans l'éthanol à 65 °C, le milieu réactionnel est ensuite agité pendant 2 h afin de permettre la formation de **140**. L'agent réducteur est alors ajouté et l'agitation se poursuit pendant 15 h, toujours à 65 °C. Un traitement acido-basique sera réalisé si le réducteur est le NaBH_4 , suivi d'une extraction au CH_2Cl_2 avant évaporation du solvant. Le tableau 7 présente les résultats de ces expériences.

Dans ce tableau, aucun rendement ne sera indiqué, et ce pour deux raisons. Premièrement, la cyclohexanone qui n'aurait éventuellement pas réagi avec l'amine peut aussi avoir été partiellement réduite. Les CH_2 du cyclohexanol se confondent avec ceux de la cyclohexanone, de **140** et de **141**. Deuxièmement parce que dans certains cas, la masse obtenue après

évaporation des solvants était supérieure à ce que l'on aurait pu atteindre théoriquement. Cet excès s'explique soit par la présence de sels résiduels dans les phases organiques, soit par le fait que les résines sont prévues pour être utilisées en combinaison avec un agitateur orbitant et non avec un barreau magnétique. La rotation du barreau entraîne une multitude de contacts et chocs entre les billes de résines et engendre finalement une dégradation de leurs structures ; en témoigne l'aspect légèrement laiteux du mélange réactionnel après quelques heures d'agitation. La filtration ne permet pas d'éliminer ces particules de faible taille.

tableau 7

entrée	éq. C=O	réd	éq. réd	proportions relatives (%)		
				139	140	141
1	1	NaBH ₄	1,5	20	n.d.	80
2	1,5	NaBH ₄	1,5	10	n.d.	90
3	1	MP-BH ₄	1	86	5	9
4	1,5	MP-BH ₄	1	77	7	16

n.d. = non détecté

Les résultats nous enseignent que la méthode classique de réduction au NaBH₄ est efficace, bien qu'une conversion complète en amine alkylée ne soit pas obtenue. Les essais avec résine sont eux beaucoup moins concluants puisque la conversion en **141** reste dans les deux cas proche de 10 %, ce qui invalide la méthode dans de telles conditions ou imposerait éventuellement d'allonger encore les délais de réaction.

Il nous a également semblé important d'évaluer d'autres réducteurs classiques, dans les mêmes conditions de température et de temps, en associant certains d'entre eux à du ZnCl₂ qui jouera le rôle d'acide de Lewis et sera présent à raison de 5 équivalents dès la mise en solution des réactifs. Le réducteur est par contre ici ajouté à raison de 1,5 équivalent, à chaque fois. Les réactions ont lieu dans le MeOH. Le tableau 8 résume les résultats de ces expériences.

tableau 8

entrée	éq. C=O	réd	a.L.	proportions relatives (%)		
				139	140	141
1	1	NaBH(OAc) ₃	/	n.d.	17,8	82,5
2	1,5	NaBH(OAc) ₃	/	n.d.	n.d.	100
3	1	NaBH ₃ CN	/	10	55	35
4	1,5	NaBH ₃ CN	/	5	58	38
5	1	NaBH ₃ CN	ZnCl ₂	15	10	75
6	1,5	NaBH ₃ CN	ZnCl ₂	n.d.	n.d.	100
7	1	NaBH ₄	ZnCl ₂	émulsion		
8	1,5	NaBH ₄	ZnCl ₂	émulsion		
9	1	MP-BH ₄	ZnCl ₂	proportion indéterminable		
10	1,5	MP-BH ₄	ZnCl ₂	proportion indéterminable		

n.d. = non détecté

Deux conditions favorisent l'obtention du **141** avec une pureté maximale : l'utilisation de triacétoxyborohydrure de sodium utilisé seul à raison de 1,5 équivalent (entrée 2), ainsi que le cyanoborohydrure de sodium lorsqu'il est utilisé en combinaison avec ZnCl₂ (entrée 6). Pour cet agent réducteur, on notera le net avantage tiré de l'adjonction dans le milieu d'un acide de Lewis. Toutefois le ZnCl₂ utilisé en combinaison avec NaBH₄ (entrée 7 et 8) mène à la formation d'une émulsion lors du traitement du milieu. Cette émulsion rend l'extraction au CH₂Cl₂ impossible et ne permet pas de déduire les proportions relatives des composés présents au sein du mélange. Concernant les deux dernières expériences qui évaluent la combinaison de MB-BH₄ associé à ZnCl₂ (entrées 9 et 10), les spectres protons n'autorisent pas à déterminer la constitution du mélange ; les protons aromatiques ne sont pas visibles et seule la zone des CH₂ laisse apparaître des pics. Il n'est pas impossible que le produit de réaction soit resté associé en grande partie à la résine.

De ces 10 résultats, on retiendra principalement l'efficacité du NaBH₃CN utilisé en combinaison avec un acide de Lewis, attendu que ce

type d'agent réducteur est disponible supporté sur résine. Notons l'avantage qu'il y a généralement à travailler avec 1,5 équivalent de cyclohexanone

Avant de poursuivre plus avant nos investigations, il a été décidé de contrôler les équilibres en jeu en menant la réaction dans le méthanol deutéré, ce que présente le tableau 9. L'agent réducteur sera le NaBH_4 utilisé à raison de 2 équivalents, et 1 équivalent de cyclohexanone est ajouté.

tableau 9

entrée	temps	réducteur	proportions relatives (%)		
			139	140	141
1	2 h	/	15	85	n.d.
2	4 h	/	15	85	n.d.
3	19 h (15 +4)	NaBH_4 (2 éq.)	17	n.d.	83

n.d. = non détecté

On remarquera que la proportion d'oxazolidine **140** ne varie pas entre 2 et 4 h de réaction, ce qui nous autorise à continuer nos tests sans allonger le temps réservé à générer **141**.

Un des résultats obtenus au tableau 8 (entrée 6) invite clairement à évaluer l'avantage que l'on pourrait avoir à user d'acide de Lewis associé à un réducteur de type cyanoborohydrure pour promouvoir la réaction. Nous avons décidé de ne plus utiliser maintenant que des réactifs supportés sur résine en ce qui concerne l'agent réducteur. Il importe aussi de mener la réaction avec un léger excès (1,1 éq.) de cétone afin de s'affranchir de l'étape de purification. La résine MP-BH_4 fera aussi l'objet de tests concernant notamment la proportion en **141** que l'on peut obtenir en faisant varier la quantité de cette dernière dans le milieu réactionnel. Le tableau 10 décrit ces tests et leurs résultats.

tableau 10

entrée	réducteur	éq. réd.	a. L.	proportions relatives (%)		
				139	140	141
1	MP-BH ₃ CN	2,5	ZnCl ₂ (5 éq.)	34	n.d.	66
2	MP-BH ₃ CN	1,5	AcOH (5 éq.)	n.d.	n.d.	100
3	MP-BH ₃ CN	2,5	AcOH (5 éq.)	n.d.	n.d.	100
4	MP-BH ₄	1,5	Ti(O- <i>i</i> Pr) ₄ (2 éq)	n.d.	n.d.	100
5	MP-BH ₄	1,5	/	25	n.d.	75
6	MP-BH ₄	1,5	/	32	n.d.	68
7	MP-BH ₄	2	/	23	n.d.	77
8	MP-BH ₄	2,5	/	n.d.	n.d.	100

n.d. = non détecté

Si l'entrée 1 nous montre que le mélange contient encore une part importante de réactif amino-alcool. La comparaison avec l'entrée 5 du tableau 8 prouve que la réaction est peut-être sensible à d'autres facteurs peu contrôlables, attendu que cette différence, même faible, reste étonnante.

A plusieurs reprises, nous voyons également que les conditions choisies permettent de mener la réaction jusqu'à une conversion totale (entrées 2-4 et 8). La combinaison MP-BH₃CN et AcOH est particulièrement efficace, ainsi que MP-BH₄ associé au Ti(O-*i*Pr)₄ et le MP-BH₄ seul à raison de 2,5 équivalents. A nos yeux, le résultat correspondant au MP-BH₃CN et AcOH de l'entrée 2 est le plus intéressant. Nous pourrions de cette manière user d'un excès raisonnable de réactif supporté sur résine, ce qui pécuniairement et à plus grande échelle peut s'avérer avantageux. De plus celui-ci fonctionne en combinaison avec un acide à la fois de Brönsted et de Lewis dont l'on peut se séparer après réaction par simple traitement basique

Nous avons donc décidé d'observer l'effet éventuel d'une diminution de la quantité d'AcOH d'une part et d'autre part d'évaluer l'acide D-(+)-10-camphorsulfonique et l'acide *para*-toluènesulfonique, ce que présente le tableau 11 ci-dessous. Pour chacun des essais, 1,1 équivalent de cyclohexanone fut employé ainsi que 1,5 équivalent de MP-BH₃CN.

tableau 11

entrée	acide	éq. acide	proportions relatives (%)		
			139	140	141
1	AcOH	5	5	n.d.	95
2	AcOH	3	12	n.d.	88
3	AcOH	1	n.d.	f.d.	≅100
4	ac. camphor.	5	indéterminable		
5	ac. camphor.	5	f.d.	≅38	≅59
6	ac. camphor.	3	66	10	24
7	ac. camphor.	2	erreur		
8	ac. camphor.	1	140 majoritaire (> 50 %)		
9	APTS	1	51	19	30

n.d. = non détecté ; f.d. = faiblement détecté

Si nous analysons les trois premières entrées du tableau, on constate que 5 équivalents d'AcOH ne permettent plus d'atteindre une conversion complète en amine alkylée **141**, mais 95 % cette fois. Il faut donc admettre que la réaction est sensible à des variations de conditions que nous ne contrôlons pas. Lorsque la quantité d'AcOH passe de 5 à 3 équivalents, on constate une baisse de la proportion en **141** au sein du mélange réactionnel final, alors que celle-ci remonte de manière surprenante lorsque l'on travaille avec 1 seul équivalent d'acide. Ce dernier résultat très intéressant fut répété plusieurs fois, et nous constatons qu'il est possible d'atteindre 100 % de proportion relative en **141** avec 1 équivalent d'acide. Il est donc probable que le résultat de l'entrée 2 eut été amélioré par des tests supplémentaires. Les entrées 4-9 présentent les résultats obtenus avec les acides D-(+)-10-camphorsulfonique et *para*-toluènesulfonique. Dans tous les cas, les proportions relatives mesurées en **141** sont insatisfaisantes, et la (1S,2R)-(-)-noréphédrine reste souvent majoritaire.

A ce stade, et encouragés par les résultats de la combinaison MP-BH₃CN (1,5 éq.) / AcOH (1 éq.), nous avons opté pour que les réactions soient dorénavant conduites en Miniblock[®]. Le système Miniblock[®] offre la possibilité de réaliser les synthèses en parallèle, ce qui sera très avantageux

au regard du nombre de ligands chiraux que nous escomptons tester. La figure 18 montre en exemple le modèle que nous utiliserons.

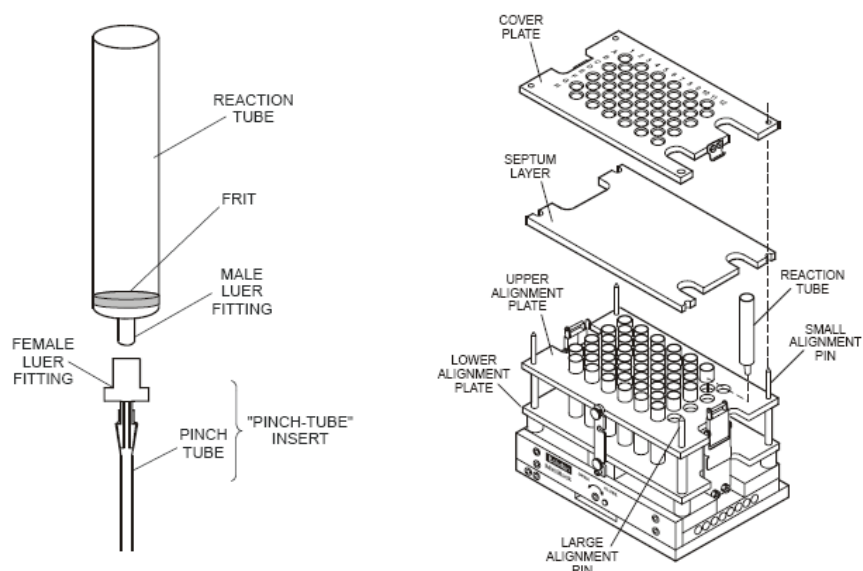


figure 18

Celui-ci se compose de 48 tubes plastiques, le fond des tubes étant doté d'un fritté, enchâssés dans un socle en métal qui assure l'étanchéité de ces derniers. Les tubes sont couverts d'un septum commun sur lequel repose une plaque métallique attachée par clips au socle. Il est possible de chauffer les tubes en plaçant entre le socle et la plaque métallique supérieure un bloc perforé dans lequel circule un fluide caloporteur. Lorsque l'expérimentateur considère que la réaction est à son terme, il est possible de poser le Miniblock[®] sur une pièce métallique qui renfermera en son sein le portoir adéquat de 48 positions accueillant les tubes destinés à récupérer les mixtures de réactions. Ces tubes de récupération peuvent ensuite être placés dans un évaporateur rotatif, ce qui permet de disposer des produits de réaction tels quels.

Dans un premier essai, nous avons placés 48 tubes, et trois d'entre eux abritaient la réaction dans les conditions qui furent déterminées comme assurant une proportion relative de 100 % ou presque d'amine alkylée **141** ;

à savoir la combinaison MP-BH₃CN (1,5 éq.) / AcOH (1 éq.) dans le MeOH à température de reflux très doux (65 °C, bp MeOH = 64,7 °C), et ce pour 0,2 mmol de (1*S*,2*R*)-(-)- noréphédrine et 0,22 mmol de cyclohexanone. Les réactifs sont introduits dès le début, seul l'agent réducteur est introduit après environ 2 h de réaction. Après 15 h de réaction, on procédera à l'addition de carbonate supporté sur résine, MP-CO₃, qui aura pour rôle de neutraliser l'acide présent dans le milieu dont le point d'ébullition se situe à 118 °C. On laissera la neutralisation s'opérer pendant encore 4 h.

Pour ce premier essai, nous avons eu le déplaisir de constater lors du transfert entre tubes de réaction et tubes d'évaporation que le solvant s'était évaporé. Il faut se souvenir que l'ajout du réactif de réduction, une fois la réaction commencée, se déroule en ôtant la plaque métallique et le septum au dessus des tubes de réaction. Cela ne permet que d'entrevoir le contenu des tubes sans deviner le niveau de solvant, sauf si l'entièreté de celui-ci s'est évaporée.

L'essai fut tenté à nouveau dans les mêmes conditions, en portant un soin tout particulier à vérifier que le montage du Miniblock[®] était réalisé strictement dans le respect de la procédure indiquée. Il avait en outre été décidé de contrôler le niveau de solvant lors de l'addition de l'agent réducteur, ce qui peut s'opérer en déplaçant partiellement et précautionneusement le dispositif de chauffage des tubes. Cela nous amena à remarquer que le phénomène d'évaporation du solvant était déjà très avancé après 4 h puisqu'environ 2/3 de ce dernier n'était déjà plus présent dans les tubes. Ajoutons encore le désagrément de constater que cette évaporation n'était pas régulière : les tubes présentaient des variations dans les volumes de solvant qu'ils contenaient encore. Face à cette observation, nous avons alors décidé d'un ajout supplémentaire de solvant au-delà de ce qui était initialement prévu puisque le volume réactionnel fut porté à 4 mL environ au lieu des 2 mL initiaux. Le but avoué était bien sûr que la réaction puisse se dérouler pendant 15 h, de nuit et toujours avec une quantité suffisante de solvant. Une fois encore, lors de l'examen des tubes antérieur au transfert des milieux de réaction, nous avons dû constater qu'ils étaient à sec.

Le problème était inattendu et risquait de compromettre gravement les synthèses de ligands. Les prescriptions concernant les températures de travail indiquaient que les Miniblock[®] n'étaient pas usinés pour mener des réactions sous conditions de reflux. Dans notre cas, la température du fluide caloporteur était de 65 °C dans la cuve de chauffe, celui-ci devait encore passer par deux mètres de tuyaux pour parvenir au Miniblock[®]. Le Miniblock[®] lui-même n'était pas isolé, ce qui permet de douter que la température réelle au sein des tubes de réaction ait jamais été de 65 °C. Une mesure de la température de fluide caloporteur dans sa cuve de chauffe a également été réalisée avec un thermomètre classique à Hg, afin de vérifier que le thermostat fonctionnait parfaitement, ce que la mesure confirma.

Face à cela, nous avons convenu de placer entre le dessus des tubes et la plaque métallique supérieure deux septa, espérant que la pression supplémentaire exercée au dessus des tubes par cette épaisseur additionnelle entraverait l'évaporation constatée. Notons qu'avant de tester cette manière de faire, une discussion avec le représentant Mettler-Toledo avait eu lieu afin de valider la sécurité de cet essai. Aucune contre-indication ne nous a été précisée, et le problème d'évaporation était pareillement considéré comme surprenant chez Mettler. Ils nous proposèrent également d'utiliser le matériel adapté au travail dans des conditions d'atmosphère inerte, selon eux plus étanche. La pose de deux septa n'a toutefois pas permis d'améliorer la situation. En effet, si l'évaporation avait apparemment été stoppée pour la partie haute des tubes, un écoulement se produisait maintenant à la base, provoqué par la pression intérieure à ceux-ci. Comme il nous l'avait été conseillé, un test fut également réalisé en usant du matériel destiné aux réactions sous atmosphère inerte. Ceci imposait, du fait de la communication entre tous les tubes, que chacun soit rempli de solvant. Ici encore le problème d'évaporation persistait dans des proportions plus graves encore.

Les tests ont été poursuivis, en variant notamment les conditions de temps et de température afin de peut-être contourner cet obstacle. Les résultats obtenus sont présentés dans le tableau 12 ci-dessous. On engage à nouveau la combinaison MP-BH₃CN (1,5 éq.) / AcOH (1 éq.) dans le MeOH.

tableau 12

entrée	matériel réactionnel	T (°C)	atm	temps réd. (h)	proportions relatives (%)		
					139	140	141
1	Miniblock [®]	65	air	15	15	ind.	≅80
2	Miniblock [®]	65	air	15	14	9	77
3	ballon	65	Ar	15	n.d.	n.d.	100
4	ballon	65	air	15	n.d.	n.d.	100
5	ballon	55	air	15	3	1,6	95,4
6	ballon	50	Ar	15	8	3	89
7	ballon	50	air	15	ind.	ind.	≅90
8	ballon	50	Ar	48	8	ind.	≅90
9	ballon	50	air	48	ind.	ind.	≅95
10	ballon	40	air	60	4	3	93

n.d. = non détecté ; ind. = indéterminé

Les deux premières réactions décrites par ce tableau (entrées 1 et 2) sont réalisées en Miniblock[®] en prenant soin de vérifier et de rétablir le niveau de solvant toutes les 3 à 4 h. Apparemment, les variations de concentration que cela entraîne perturbent le système. Nous avons également prévu de diminuer la température de réaction afin d'arrêter ou à tout le moins de diminuer le problème d'évaporation. On constatera en s'attachant à l'examen des entrées 5-10 qu'une température plus basse implique une baisse de la proportion relative en amine alkylée **141** au sein du brut réactionnel. Les valeurs restent proches de 90 %. Une augmentation conséquente du temps réactionnel (entrées 8-10) ne permet pas de palier ce désavantage. Notons que le problème d'évaporation subsiste à 50 °C, dans une mesure plus congrue néanmoins (après une nuit de réaction, 2/3 du solvant est encore présent dans les tubes), alors que nous sommes cette fois 15 °C en deçà de la température d'ébullition du MeOH. A 40 °C, on ne constate plus qu'une évaporation très faible, le temps de réaction requis est toutefois quadruplé par rapport aux essais à 65 °C, et la proportion d'amine alkylée **141** est moindre. De plus, l'agent réducteur est insensible à l'atmosphère qui lui est imposée. Des 10 expériences que décrit le tableau 12, il ressort clairement

que nous n'avons pas encore entre nos mains une solution pratique au problème.

Puisque la baisse de température s'avère dommageable lorsque l'on tente de synthétiser l'amine alkylée **141** dans nos conditions, nous avons opté pour un plus grand volume de solvant, en changeant de tubes. Au lieu de travailler avec 48 tubes potentiels, le système à 24 tubes fut monté, avec l'espoir que la quantité de solvant, même si elle diminuait régulièrement, serait suffisante pour assurer la réaction jusqu'à terme. Les trois premiers tubes furent chargés en réactifs et MeOH, selon nos conditions habituelles. Les autres tubes furent remplis de MeOH afin de vérifier le niveau des différents tubes. L'évaporation, à nouveau, fut remarquée. Les deux photographies ci-dessous (figure 19) indiquent clairement l'état des tubes en fin de réaction. La première photographie montre la rangée de tubes au premier plan, ceux de gauche encore remplis de solvant, alors que ceux de droite sont à sec. Les traits noirs sur les tubes indiquent le niveau de solvant. Toute la ligne de droite est à sec. La deuxième photographie permet de distinguer l'irrégularité de l'évaporation, puisque chaque tube a un niveau de solvant différent de son voisin. Les tubes à sec (droite du Miniblock[®]) sont les premiers à recevoir le fluide caloporteur fraîchement chauffé, ceux de gauche sont proches de la sortie du fluide lorsqu'on tient compte du sens de circulation de ce dernier.

**figure 19**

Nos derniers essais ne se sont pas avérés concluants, et nous invitent à envisager une autre façon de contourner le problème de la baisse de température. Nous avons donc entrepris de travailler malgré tout à 65 °C, mais en utilisant des solvants alcooliques à plus hauts points d'ébullition. L'accent a été mis en particulier sur l'éthanol. Le tableau 13 donne les résultats de ces tests.

tableau 13

entrée	matériel réactionnel	solvant	atm	proportions relatives (%)		
				139	140	141
1	ballon	EtOH	Ar	18	4	78
2	ballon	EtOH	air	9	1	90
3	Miniblock [®]	EtOH	air	23	8	69
4	Miniblock [®]	EtOH	air	23	2	75
5	Miniblock [®]	EtOH	air	22	1	77
6	Miniblock [®]	<i>i</i> -PrOH	air	22	51	27
7	Miniblock [®]	<i>n</i> -BuOH	air	ind.	≅30	ind.
8	Miniblock [®]	<i>t</i> -BuOH	air	23	42	35

ind. = indéterminé

L'enseignement des tests se lit assez rapidement. Travailler dans l'éthanol (entrée 1-5) n'offre pas une solution alternative : la proportion relative du brut réactionnel en amine alkylée **141** est légèrement inférieure à 80 % en moyenne, ce qui est nettement insatisfaisant. Les résultats obtenus dans l'*i*-PrOH, le *n*-BuOH ou le *t*-BuOH (entrées 6-8) sont eux franchement mauvais car la part d'amine alkylée voisine avec les 30 %, dans les deux cas où elle put être déterminée. Notons encore que la perte de solvant se produit ici aussi, elle est très nette avec l'éthanol et de 25 % pour l'*i*-PrOH et le *t*-BuOH. Seul le *n*-BuOH conserve un volume inchangé en fin de réaction. Pour rappel, les points d'ébullition sont pour l'EtOH, l'*i*-PrOH, le *n*-BuOH et le *t*-BuOH respectivement de 79, 82, 118 et 82.

Face au problème d'évaporation du solvant, chacune de nos tentatives à lui donner une solution fut vaine. La perte de solvant se produit dans tous les cas, et lorsque les moyens sont mis en œuvre pour s'en accoutumer, voire l'éviter, la proportion relative en amine alkylée **141** dans le brut réactionnel chute en deçà de la limite acceptable. Forcés et contraints, nous nous sommes déterminés à abandonner la synthèse en parallèle à l'aide du système Miniblock[®], pour la mener de manière plus classique à l'aide de ballons surmontés de réfrigérants ; cette méthode ayant été la seule à nous apporter des résultats convaincants.

La synthèse de 3 ligands a donc été mise en œuvre de la sorte, en ballon et sans agitation mécanique. La température de bain d'huile à 70 °C assurait un reflux doux sans ébullition, mais qui devait normalement autoriser une certaine convection au sein du mélange réactionnel. Les ballons seraient de plus agités régulièrement à la main. Les trois carbonyles candidats à la réaction d'amination réductrice furent, ainsi que le présente le schéma 137, B6, B13 et B16. B6 nous servait de jalon ; B13 était engagé en temps qu'aldéhyde standard, et B16, de par le groupement anthranyle, devait témoigner de l'efficacité -ou non- de la réaction lorsqu'un aldéhyde peu électrophile était engagé. Les réactions portent sur 0,2 mmol.

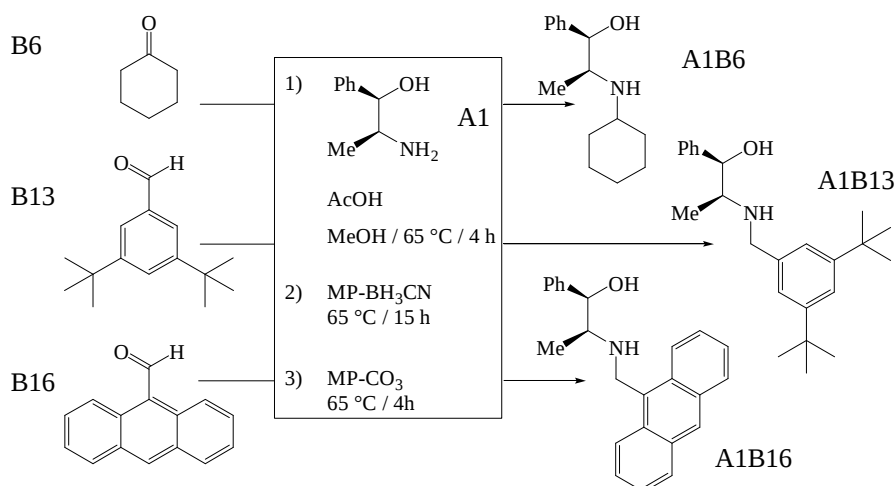


schéma 137

A la fin de la réaction, Il suffit simplement de filtrer les résines ; les solvants sont ensuite évaporés. Les masses des bruts réactionnels et les analyses RMN ¹H correspondantes fournissent les résultats suivants (tableau 14) :

tableau 14

ligand	brut récolté (g)	F.W. (g)	pureté (%)	rdt ligand (%)
A1B6	0,039	233,3	100	83
A1B13	0,065	353,3	76	70
A1B16	0,068	341,5	91	91

Les ligands produits, leurs quantités et leurs puretés sont ici tout à fait acceptables. Les puretés déterminées sont bonnes à excellentes et donnent accès aux rendements réactionnels. La pureté nous importe en premier lieu, vu que les quantités de ligands dont nous aurons besoin ultérieurement pour réaliser les solutions stocks (0.01 mmol) autorisent des rendements de réaction faibles. Il a donc été décidé de poursuivre cette voie, à plus large échelle maintenant.

5.4.2 Synthèses proprement dites des ligands chiraux

Réaliser les synthèses par pas de trois aurait entraîné des retards considérables pour le projet. Nous avons donc monté sous hotte un système assez rudimentaire, mais efficace, permettant de mener en parallèle 10 réactions d'amination réductrice. Les ballons sont attachés à des barres métalliques ; celles-ci restent volontairement libres de vibrer et sont reliées à l'autre extrémité à un petit agitateur oscillant. Cet agitateur, mis en fonctionnement à intervalles réguliers, assurera un mélange au sein des ballons sans que nous ayons à user de barreaux aimantés dont nous connaissons l'effet néfaste sur les billes de résine.

Il y a encore lieu de préciser que le protocole de synthèse en ballon imposait une suite de manipulations assez fastidieuse pour chaque échantillon, ceci dû entre autre à la nécessité de laver la résine au moment de la filtration par un mélange MeOH / CH₂Cl₂ : 50/50. Les résines pouvaient en effet adsorber le(s) produit(s) de réaction et diminuer parfois dramatiquement le rendement de cette dernière. Malheureusement les tubes de récolte des réactions adaptés aux contraintes d'accélération générées par l'évaporateur rotatif étaient moulés dans le polypropylène et se dissolvaient au contact du CH₂Cl₂. L'achat de tubes en verre résistant à plus de 500 G fut envisagé ; leur prix prohibitif eut le mérite de résoudre rapidement la question. Voici les étapes que nécessitait cette manière de procéder :

- Mise en solution des réactifs et réaction
- Filtration des résines, lavage et évaporation du solvant au rotavapor
- Reprise du contenu du ballon et transfert dans un tube de verre, évaporation lente sous hotte.
- Analyse RMN ^1H des candidats choisis, et évaporation sous hotte du solvant après analyse.
- Pesée des tubes à évaporation
- Reprise du contenu des tubes en verres au MeOH chaud, transfert dans les tubes en plastique
- Evaporation des tubes en plastique (évaporateur rotatif Genevac[®])
- Pesée des tubes contenant les ligands chiraux

Nous pouvons maintenant entreprendre la synthèse de notre bibliothèque de ligands chiraux, en combinant successivement les amino-alcools aux composés carbonylés. La première série « A1B i $\rightarrow$ j » sera analysée dans son entièreté par RMN ^1H ; ensuite nous réaliserons les analyses d'un composé sur trois, en décalant à chaque série les candidats sujets à l'analyse, ce qui devrait nous permettre d'accéder à une bonne idée générale de la pureté de notre bibliothèque. Les résultats de ces 330 synthèses sont résumés à la figure 21, qui peut se lire simplement en tenant compte du code de couleur donné juste ci-dessous (figure 20). Chaque case indiquera toutefois en pourcentage la proportion relative de ligand présent dans les tubes.

code couleur							
pureté (%)	>95	85 $\rightarrow$ 94	70 $\rightarrow$ 84	60 $\rightarrow$ 69	50 $\rightarrow$ 59	<50	ind.

ind. = indéterminable ou très approximativement déterminable

figure 20

figure 21

		aldehydes ou cétones																													
		B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	B23	B24	B25	B26	B27	B28	B29	B30
amino-alcools	A1	X	58	92	78	26	100	61	79	69	33	78	76	76	96	75	91	91	88	78	100	100	76	26	59	90	87	89	X	76	47
	A2		66			X			85			71			78			77			79			75			65			X	
	A3			63			100			73			91				94		90			100			62			77			70
	A4	50			76			50			X			83			95			69			83			85			30		
	A5		85			50			100				89			90			100			100		50			82			90	
	A6			100			100			91			72			67			88				100			67			73	0	34
	A7	X			67			85			55			71			100			80			82			100			50		
	A8		X			X			100				76			75		85				92		23			75			X	
	A9			76			74			28			82			85		100				92			X			60			32
	A10	X			77			66			60			50			91			72			58			81			X		
	A11		90			33			96			60			70		X				60			X			X			X	

Il apparaît tout de suite que la majorité de nos ligands offre une pureté supérieure ou égale à 70 %, ce que nous considérons comme tout à fait satisfaisant dans le cadre d'une synthèse parallèle. En outre, dans cette frange, il n'est pas rare de noter des puretés proches, voire supérieures, à 90 %. On remarquera aussi que des puretés de 100 % ou très voisines peuvent être rencontrées à la figure 21. Toutefois, il faut admettre que certaines aminations réductrices sont moins fructueuses comme en témoignent les colonnes correspondant aux composés carbonylés B1, B5, B10, B23 et B24. En ce qui concerne la colonne associée à B28, les analyses doivent être considérées différemment. Les mauvais résultats apparents, ou les indéterminations constatées, relèvent d'une très faible solubilité des ligands synthétisés dans le CDCl_3 . On relèvera encore la très faible solubilité dans le MeOH des ligands obtenus en combinant l'amino-indanol A5 avec la série de composés carbonylés ; au vu des résultats, nous pouvons conclure que cela n'a heureusement pas entravé les réactions.

Le graphe suivant (figure 22) montre comment se répartissent les puretés observées au cours de nos synthèses. On y voit très clairement que peu de candidats analysés exhibent des valeurs inférieures à 50 %. De surcroît, 14,3 % de nos ligands ont une pureté de 100 % environ, ce qui correspond en l'occurrence la plus élevée calculée, et 71,4 % d'entre eux ont une pureté mesurée supérieure ou égale à 70 %. La moyenne calculée de l'ensemble est de 76 %.

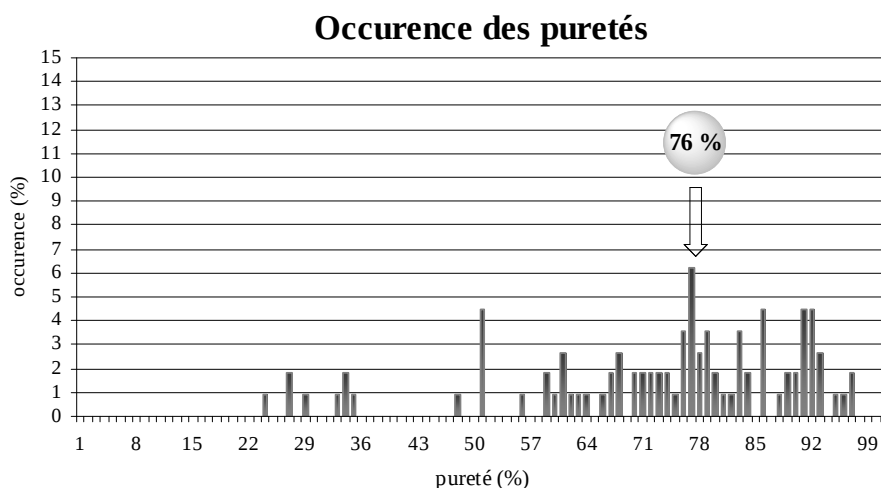


figure 22

5.5 Evaluation de la bibliothèque de catalyseurs par technique HTE

Les complexes de précurseurs métalliques ont été synthétisés auparavant, et les ligands chiraux ont été fraîchement obtenus par réaction d'amination réductrice entre amino-alcools et composés carbonylés divers. Nous pouvions donc maintenant songer à évaluer les 1980 systèmes catalytiques potentiels, générés en associant chacun des 330 ligands aux 6 complexes métalliques de ruthénium. Le schéma 138 rappelle la réaction de réduction énantiosélective qui nous permettra ces évaluations. Un signet, annexé à ce manuscrit, reprenant les structures des ligands et précurseurs métalliques effectivement employés permet une meilleure compréhension des développements suivants.

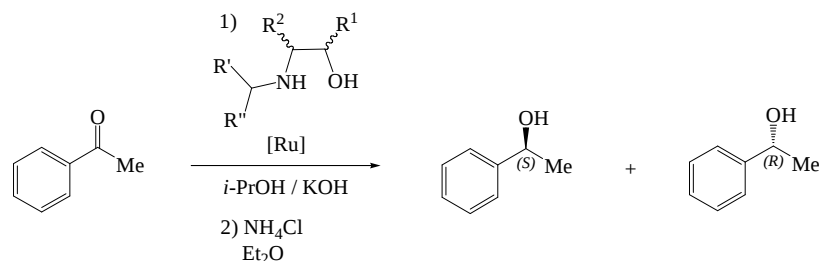


schéma 138

Cette campagne de tests a été réalisée à Lyon, au sein du laboratoire de Génie des Procédés Catalytiques du Dr Claude de Bellefon. Le matériel utilisé pour les tests combinatoires est constitué d'un robot Gilson, piloté par ordinateur, et d'une GC chirale disposant d'un passeur automatique d'échantillon. Le programme de température et de pression permet d'analyser 20 échantillons par heure.

5.5.1 Conditions des tests catalytiques HTE

Avant de commencer les tests proprement dits, il a été nécessaire de se familiariser avec la programmation du robot, pour d'abord configurer les portoirs de réaction que nous allons utiliser, ceux destinés aux solutions stocks, et ensuite les différentes étapes d'addition de solutions et leurs minutages.

Les solutions stocks de précurseurs métalliques ont chaque fois été fraîchement préparées avant emploi. A propos des ligands chiraux, les solutions stocks primaires étaient réalisées à même les tubes d'évaporation qui étaient dégazés et emplis d'*isopropanol*, lui aussi dégazé. Le volume additionné fut calculé pour chaque ligand afin d'obtenir une solution 0,04 M, fonction du nombre de mmol dont nous disposions corrigé, par la pureté moyenne déterminée à 76 % et appliquée à tous nos ligands. Fréquemment, les tubes de solution étaient soumis à une ultrasonication de 15 minutes dans un bain d'eau à 45 °C pour favoriser une bonne solubilisation. La solution stock secondaire était préparée ensuite par mélange de 250 μL de solution stock primaire et de 1250 μL d'*isopropanol*. La solution stock de KOH était préparée et conservée sous argon, la quantité nécessaire était placée avant

distribution dans des piluliers prévus à cet effet. Il est également apparu au cours des tests, et vu les températures estivales courantes voisinant avec 35 °C à Lyon, qu'ultérieurement au quenching par NH_4Cl , l' Et_2O était inégalement distribué. Un phénomène de cavitation fut noté dans le tuyau reliant l'aiguille de distribution au pousse seringue, ce qui imposa de travailler avec de l' Et_2O froid et placé sur le portoir quelque minutes seulement avant la distribution.

L'évaluation des systèmes catalytiques put alors prendre place. Chaque réaction est menée dans un pilulier de 1,5 mL de capacité, surmonté d'un septum que nous scellons par une bague métallique. Celui-ci est dégazé à l'argon avant d'être placé sur le portoir et qu'y soient successivement ajoutées les différentes solutions. Le tableau 15 rend compte des solutions ajoutées en ordre chronologique, de leurs volumes, des quantités de réactifs qu'elles contiennent et de leurs rôles.

tableau 15

solution	molarité (mol.l ⁻¹)	volume (μL)	qté (mol)	rôle
métal	$6,66.10^{-3}$	30	4.10^{-7}	formation du
ligand	$6,66.10^{-3}$	60	4.10^{-7}	pré-catalyseur
acétophénone	0,1428	280	4.10^{-5}	substrat
KOH	0,1333	30	4.10^{-6}	form. catalyseur + début réaction
		$\Sigma_1 = 400$	$\rightarrow 0,1 \text{ mol.l}^{-1}$ en substrat	
NH_4Cl	saturée	100	/	quenching
Et_2O	/	400	/	sép. phase
		$\Sigma_2 = 900$		

On note d'emblée un des avantages majeurs des tests à haut débit : la faible quantité de matériels nécessaires pour mettre en œuvre la réaction. Comme indiqué dans le tableau, nous avons fixé une dilution de $0,1 \text{ mol.l}^{-1}$ en substrat pour conduire nos tests. Le taux catalytique est de 1 %, et la réaction démarre lors de l'addition du KOH à raison d'un ratio $\text{KOH} / \text{Ru} =$

10, qui promeut la formation de l'espèce réductrice sous forme d'hydrure de ruthénium. Le temps de réaction est fixé à 90 minutes, ce qui nous permet de réaliser 120 tests catalytiques en une seule journée, analyses GC comprises.

5.5.2 Campagne de tests catalytiques HTE

5.5.2.1 Reproductibilité

Avant d'ouvrir la session d'évaluation des catalyseurs proprement dite, nous avons à la fois vérifié la bonne programmation des étapes gérées par le robot, et la reproductibilité des tests que nous allons mener. A cette fin, les premières expériences furent 10 évaluations identiques qui associaient le $[\text{RuCl}_2(p\text{-cym})]_2$ (M1) et (1R,2S)-1-amino-2,3-dihydro-1H-inden-2-ol **108**, en raison du comportement bien connu de ce couple. L'une d'entre elles fut cependant invalidée à cause d'un seuil de détection trop bas lors de prélèvement destiné à l'analyse GC. Le schéma 139 présente le détail de la réaction, la conversion moyenne, l'ee moyen ainsi que leurs écart-types respectifs.

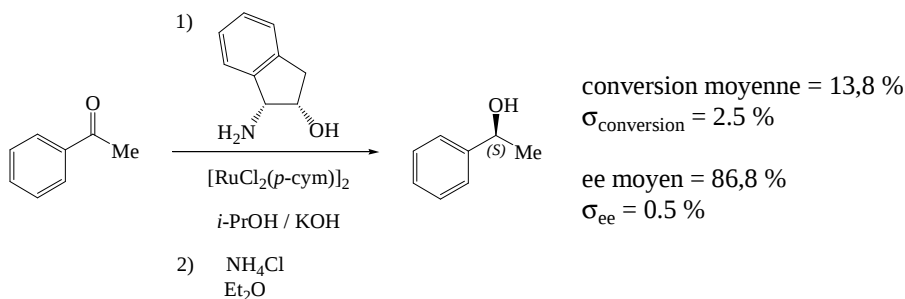


schéma 139

L'écart-type prend une valeur acceptable en ce qui concerne la conversion et particulièrement faible dans le cas de l'ee. Nous considérons donc que les tests offrent une reproductibilité tout à fait satisfaisante.

5.5.2.2 Solubilité des précurseurs métalliques

Lors de notre tentative d'engager la série A1 en évaluation, un problème inattendu survint à la réalisation des solutions stocks des deux premiers précurseurs métalliques. Si M1 $[\text{RuCl}_2(p\text{-cym})]_2$ que nous avons déjà

largement employé dans un autre cadre était parfaitement soluble dans l'isopropanol, M2 $[\text{RuCl}_2(\text{benz})]_2$ ne l'était apparemment pas, du moins à la concentration ici utilisée. Cette différence de solubilité, étonnante au regard de structures relativement similaires, nous obligea à choisir un autre solvant parmi la gamme de ceux qui étaient à notre disposition à Lyon. L'*isopropanol*, CH_2Cl_2 , CHCl_3 , toluène, benzène, éthylbenzène, cyclohexane, éther de pétrole, Et_2O et le THF furent utilisés en ce sens ; on ne constata malheureusement qu'une très faible et insuffisante solubilité dans le CHCl_3 et le benzène. Un mélange de ces solvants à part égale avec l'*isopropanol* n'offrait pas d'autres résultats. Un dernier essai, dans l'eau porta heureusement ses fruits. Le manque de solubilité de M2 nous imposa de vérifier le comportement de tous nos précurseurs métalliques. Le tableau 16 ci-dessous résume les solubilités observées pour nos différents complexes.

solvant complexe												tableau 16
	<i>i</i> -PrOH	CH ₂ Cl ₂	CHCl ₃	Toluène	Benz.	EthylB.	Cy-Hex	E.P.	ET ₂ O	THF	H ₂ O	
M1 [RuCl ₂ (<i>p</i> -cym)] ₂	↑	↑										
M2 [RuCl ₂ (benz)] ₂	–	–	part.	–	part.	–	–	–	–	–	↑	
M3 [RuCl ₂ (mésit)] ₂	–	–	–	–	–	–	–	–	–	–	part.	
M3 [RuCl ₂ (HMB)] ₂	part.	↑	↑	part.	part.	–	–	–	–	part.	part.	
M4 [RuCl ₂ (<i>p</i> -xyl)] ₂	–	↑	↑	part.	part.	–	–	–	–	part.	↑	
M5 [RuCl ₂ (DMSO) ₄]	–	↑	↑	part.	part.	–	–	–	–	part.	↑	
M6 [Ir(COD)acac]	part.	↑	↑	↑	↑	↑	part.	part.	↑	↑	–	

↑ = totalement soluble ; part. = solubilité partielle ; – = très peu soluble, voire insoluble

Il ressort de ces tentatives à mettre en solution nos complexes que les solvants chlorés CH_2Cl_2 et CHCl_3 nous permettront de résoudre le problème, à l'exception de M3 $[\text{RuCl}_2(\text{mésit.})]_2$ pour lequel jamais une solubilité satisfaisante ne fut observée. Dans ce cas précis, d'autres tentatives eurent encore lieu, comme le résume le tableau 17. On constatera qu'aucune d'entre elles ne s'avère satisfaisante.

tableau 17

complexe	mélange de solvants (parts égales)			résultat
M3 $[\text{RuCl}_2(\text{mésit.})]_2$	DMF			-
	DMF	H_2O		-
	DMF	H_2O	<i>i</i> -PrOH	part.
	MeOH			-
	MeOH	H_2O		-
	MeOH	H_2O	<i>i</i> -PrOH	-
	DMSO			part.
	DMSO	<i>i</i> -PrOH		part.
	DMSO	<i>i</i> -PrOH	H_2O	part.

- = insoluble ; part. = solubilité partielle

Manifestement il allait nous être impossible d'exploiter le complexe M3 pour nos tests. Attendu que chaque précurseur métallique devait être associé à notre bibliothèque de ligands, l'éviction d'un seul de ces précurseurs entraînait la perte de 330 résultats. Nous avons donc convenu de le remplacer par un complexe d'iridium commercial dont disposait le laboratoire de Lyon, et qui avait déjà été intégré à une étude de réduction énantiosélective de l'acétophénone en conditions assez similaires¹⁸⁸. Comme le montre le tableau 16, la solubilité de ce dernier est bien meilleure. Ci-dessous (schéma 140) sont présentées les structures des précurseurs métalliques, accompagnées de leur code « M » définitif, qui feront effectivement l'objet d'une association avec notre bibliothèque de ligands.

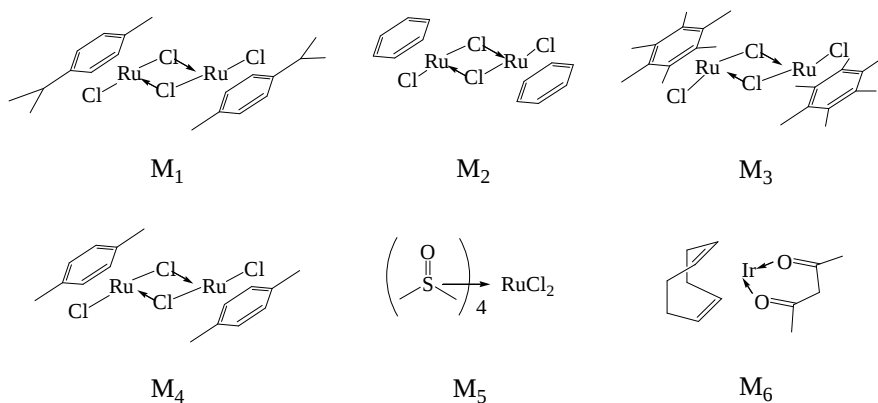


schéma 140

5.5.2.3 Tests catalytiques HTE

A chaque test catalytique correspondront une conversion et un ee. Toutefois, et pour des raisons de traitement mathématique sur lesquelles nous reviendrons dans la partie consacrée à l'évolution simulée proprement dite, il avait été imposé que chaque test soit relié à une valeur unique. Plusieurs manières de procéder auraient sans doute pu être envisagées, et celle que nous allons choisir n'a d'autre prétention que d'être adaptée à ce projet. Lors des premières expériences nous avons constaté que nos conditions de travail menaient à des rendements et des ee au maximum de l'ordre respectivement de 25 % et 85 %. Il était donc nécessaire de prendre un facteur multiplicatif à propos de la conversion, qui permettrait aux cas extrêmes d'être considérés comme attractifs également. Par cas extrêmes, on entend un système faiblement énantiosélectif mais très actif ou l'inverse, un système faiblement actif et très énantiosélectif. De plus nous avons convenu qu'une légère priorité devait être accordée aux qualités d'énantioinduction des systèmes catalytiques. Le tableau 18 permet d'observer ce que donnent les 3 cas proposés confrontés au facteur multiplicatif de 2 que nous avons choisi. Une addition simple de l'ee et de la conversion donnait une priorité presque absolue au premier, ce qui était assez malvenu ; tandis qu'une multiplication par un facteur 3 consacrait la conversion comme paramètre légèrement prioritaire dans les cas extrêmes ce qui allait à l'encontre de notre désir.

tableau 18

conversion (%)	ee (%)	2(conv.) + ee	valeur unique
25	85	→	135
5	90	→	100
30	25	→	85

Nous avons donc pu débiter les tests d'évaluation de la bibliothèque de catalyseurs. Les résultats sont donnés ci-dessous sous forme d'histogrammes. Ceux-ci présentent les combinaisons de la série concernée avec les 6 précurseurs métalliques, et la structure développée du meilleur système catalytique sera arborée par la figure sur son coin supérieur gauche. Nous entamons logiquement les commentaires par la série A1 (figure 23).

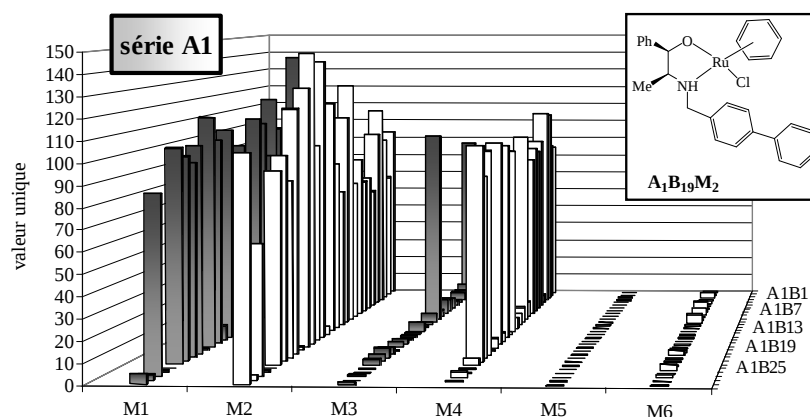


figure 23

Au sein de cette série, on remarquera que les meilleures performances sont obtenues lorsqu'A1 est combinée à M1 et M2, ainsi que M4 dans une moindre mesure. M3 n'exhibe que deux systèmes catalytiques efficaces, quant à M5 et M6, l'activité bien souvent n'est même pas suffisante pour déterminer l'excès énantiomérique. Le meilleur système de cette série est la molécule de code A1B19M2, dont le tableau 19 ci-dessous donne le détail ; il y est également relevé les performances catalytiques du système le plus énantiosélectif ou le plus actif différent du meilleur candidat en terme de

valeur combinée. Le meilleur système sera naturellement « poussé » par un des deux paramètres à sa première place.

tableau 19

code	performance	val. unique	conv. (%)	ee (%)
A1B19M2	combinée	145,6	36,4	72,7
A1B3M4	ee	104,7	7,9	88,9

Comme nous pouvons nous en apercevoir, le meilleur système catalytique en termes de performances combinées doit sa place à la conversion, et l'ee maximum approche agréablement les 90 %.

La figure 24 présente les résultats de la série A2.

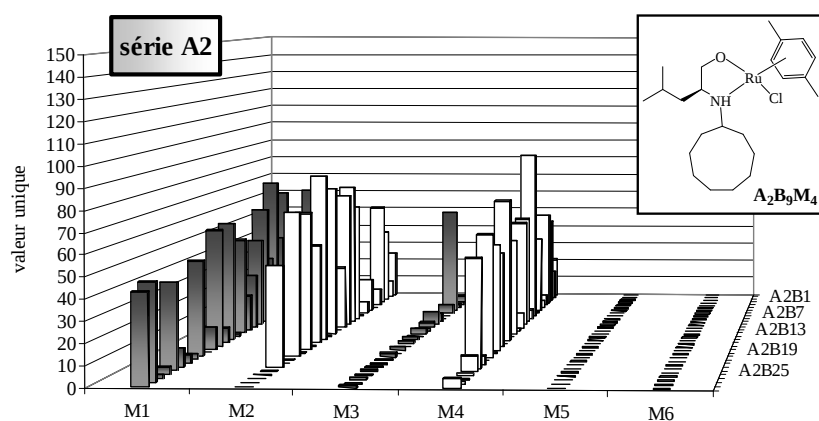


figure 24

On constate une réactivité réduite qu'expriment les valeurs uniques plus basses comparées à celles de la série A1. A nouveau, M3 et surtout M5 et M6 exhibent clairement de faibles performances. De même que précédemment, le tableau 20 rend compte des meilleures performances en termes de conversion ou d'ee et de ces deux paramètres combinés. Pour cette série, on remarquera que la meilleure performance combinée contient le meilleur excès, celui-ci reste toutefois modeste au regard des 89 % obtenus à la série A1

tableau 20

code	performance	val. unique	conv. (%)	ee (%)
A2B9M4	combinée	87,9	11,5	64,9
A2B1M1	conversion	61,1	21,5	18,1

Passons à la figure 25 et aux expériences conduites pour la série A3.

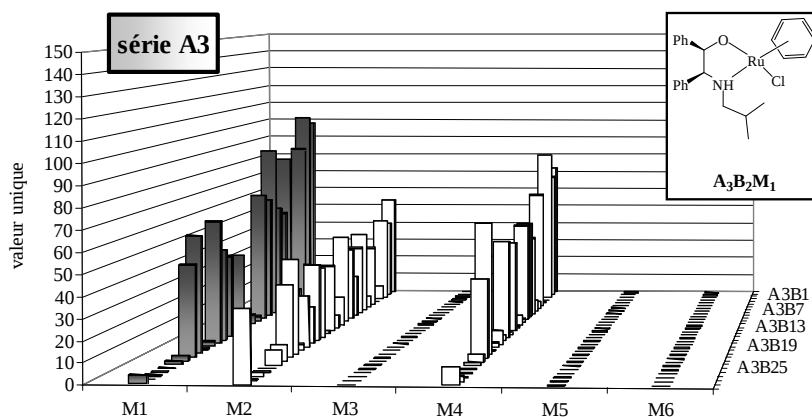


figure 25

Semblablement à ce qui précédemment avait été remarqué, M3, M5 et M6 mènent à des catalyseurs inactifs, tandis que M1, M2 et M4 montrent de meilleurs résultats. Le tableau 21 fait rapport des hautes performances de cette série.

tableau 21

code	performance	val. unique	conv. (%)	ee (%)
A3B2M1	combinée	102,1	27,5	47,1
A3B19M1	ee	61,0	6,1	48,9

Il semble donc se dessiner déjà après 540 tests que les ligands associés à certains précurseurs offrent toujours les meilleurs résultats de la série. En effet, parmi les 6 performances les plus élevées identifiées jusqu'à présent, M1 est présent 3 fois. Les combinaisons avec les groupements carbonyles font ressortir B19 à deux reprises jusqu'à présent.

La série A4 est maintenant décrite par la figure 26 ci-dessous.

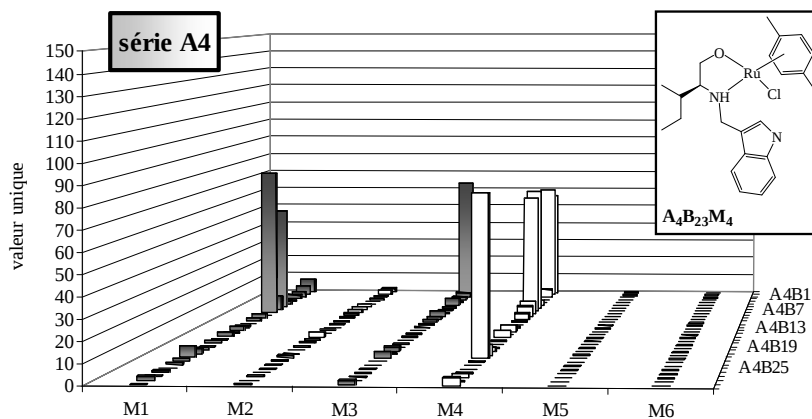


figure 26

On constate une baisse remarquable de l'activité des systèmes par rapport à ce que nous avons pu observer auparavant, puisque seuls quelques catalyseurs se montrent actifs. Ils sont à nouveau confinés dans les essais réalisés avec M1 et M4. Le tableau 22 met en lumière les deux meilleurs systèmes catalytiques de la série. On y voit que la meilleure conversion plafonne à 8 % seulement.

tableau 22

code	performance	val. unique	conv. (%)	ee (%)
A4B23M4	combinée	79,7	6,8	66,1
A4B1M4	conversion	56	8,0	40,0

Attachons-nous maintenant à une brève description de la série A5 par le biais de la figure 27.

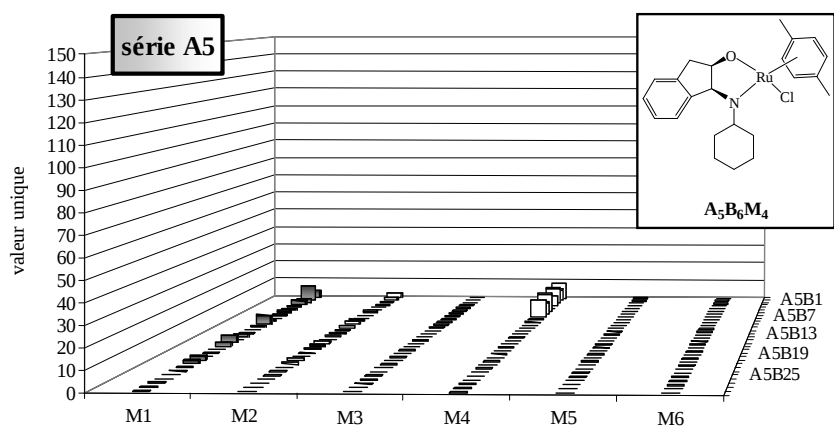


figure 27

Les résultats s'avèrent ici tout à fait décevants puisqu'aucun système ne fonctionne, hormis quelques-uns de manière très faible avec les complexes M1 et M3. Le meilleur système est la structure de code A5B6M4, développée ci-dessus, et dont la valeur unique de 9,6 ne rend compte que d'une conversion de 4,8% ; l'ee n'ayant pu être correctement mesuré ne sera pas pris en compte. Confrontés à des résultats dont le bas niveau nous interpellait, et attendu que cela correspondait à une utilisation de solutions stock fraîches, nous avons supposé un éventuel problème dans le chef de ces dernières. Elles furent refaites avec attention et conduisirent lors de la répétition des tests à ce que les mêmes résultats soient obtenus. Il nous faut donc conclure à la particulière inefficience de ces catalyseurs. Le tableau reprenant les deux meilleurs systèmes en termes de performance ne sera pas donné ici pour d'évidentes raisons.

Nous continuons avec la série A6, décrite ci-dessous par la figure 28.

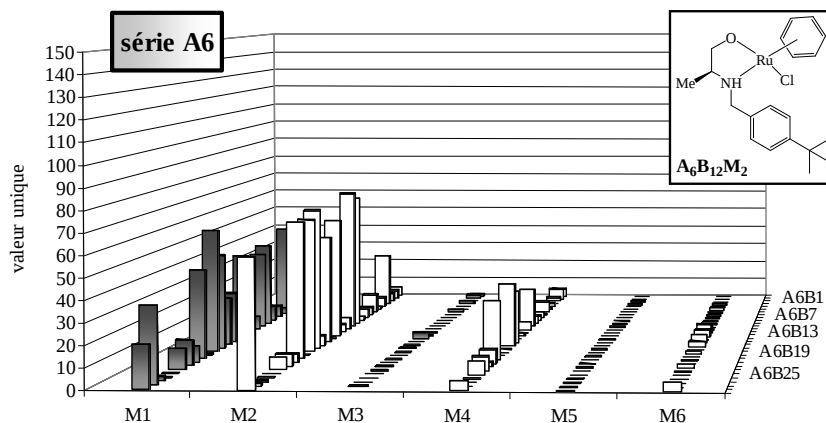


figure 28

L'efficacité des catalyseurs remonte par rapport à la série précédente, tout en restant relativement faible eu égard à ce que nous avons pu observer de mieux jusqu'à présent, soit la série A1. Ici encore la tendance se confirme quant aux performances de M1, M2 et M4. M5 et M6 à nouveau restent confinés à de très faibles niveaux. Le tableau 23 reprend les meilleures performances.

tableau 23

code	performance	val. unique	conv. (%)	ee (%)
A6B12M2	combinée	69,6	8,2	53,2
A6B17M4	conversion	26,0	12,7	0,4

On notera encore que malheureusement le catalyseur remarqué pour sa meilleure conversion promeut la réaction de manière racémique. Il s'agit pourtant d'une structure assez proche, puisque le groupement carbonyle B17 n'est autre que le 3,4,5-trifluorobenzaldéhyde, ce qui indique l'importance que doit sans doute revêtir le caractère électronique du cycle aromatique.

Nous poursuivons avec la série A7 décrite par la figure 29.

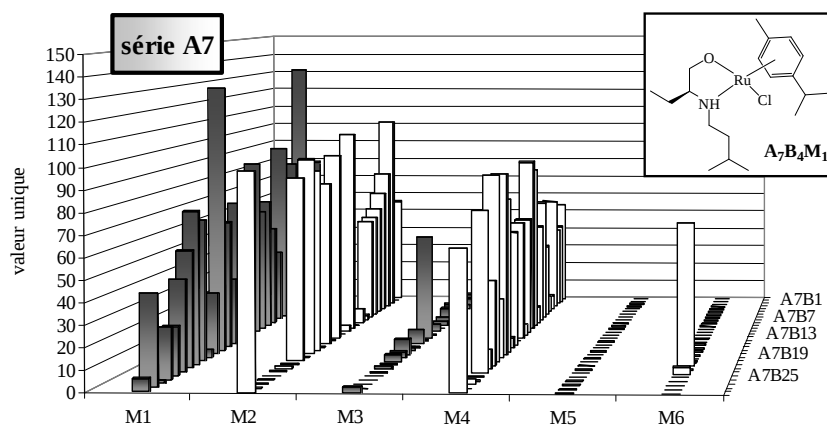


figure 29

La série A7 offre des résultats plutôt bons, comparés à ce qui a pu être observé dernièrement, et qui s'approchent de ceux que la première dévoila. A nouveau M1, M2 et M4 confirment la tendance dont nous sommes coutumiers depuis le début. Le tableau 24 rend compte des meilleures performances de A7.

tableau 24

code	performance	val. unique	conv. (%)	ee (%)
A7B4M1	combinée	131,9	61,5	8,8
A7B15M2	ee	93	5,1	82,9

Il est intéressant de constater ici que la performance combinée dépend grandement de la conversion. A vrai dire, jamais des conversions si élevées n'avaient été remarquées jusqu'ici. Les 7 premiers systèmes les plus efficaces concernant ce paramètre se positionnent tous au dessus de 35 %, et tous à l'exception d'un seul résultat de l'association avec M1. L'énantiosélectivité qui leur est associée reste néanmoins dans chaque cas très basse, pour M1 les meilleurs ee sont de 15 % environ. Le compromis semble donc difficile à trouver pour ces systèmes, l'énantiosélectivité maximale de presque 83 %, ce qui au sein de notre bibliothèque est un très bon résultat, n'est couplée qu'à une conversion faible de 5 %.

On notera encore une surprise provenant de M6 puisque pour la première fois, un catalyseur issu du couplage avec le complexe d'iridium fournit une performance appréciable. Il s'agit de A7B23M6 dont la valeur unique de 68,0 provient d'une conversion de 8,3 % et d'un ee de 51,4 %.

Nous entreprenons maintenant l'analyse des résultats de la série A8 que décrit la figure 30 ci-dessous.

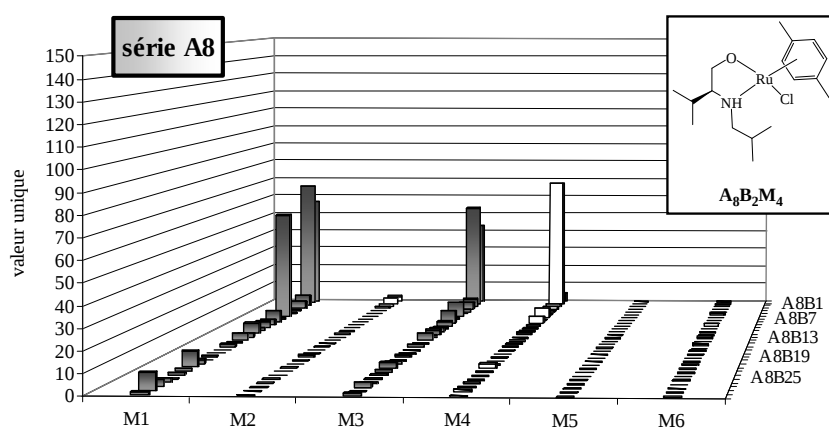


figure 30

A nouveau, très peu de catalyseurs se montrent actifs, et ceux qui le sont restent encore des membres résultant de l'association avec M1, M3 et M4. M2, M5 et M6 ne produisent que des valeurs uniques nulles, ou très légèrement supérieures à 0. Le tableau 25 décrit les meilleures performances, les conversions et les ee restent confinés à des valeurs médiocres.

tableau 25

code	performance	val. unique	conv. (%)	ee (%)
A8B2M4	combinée	68,2	5,5	57,3
A8B2M1	conversion	66,4	11,5	43,5

Analysons maintenant l'efficacité des systèmes catalytiques qu'engage la série A9, ce que décrit la figure 31 ci-après.

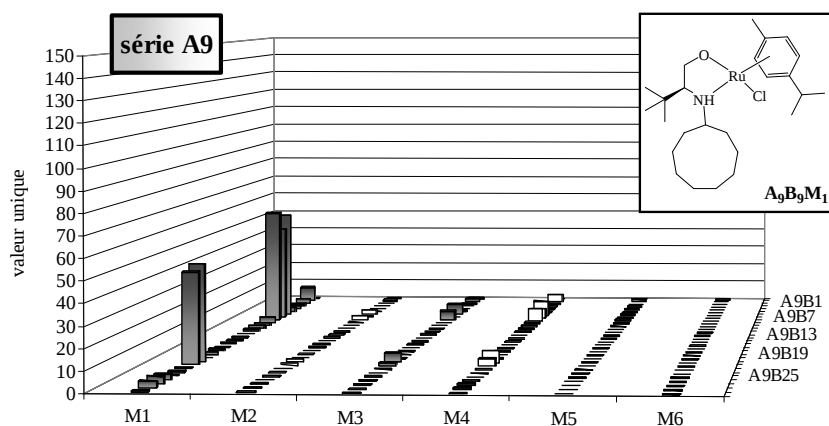


figure 31

Les catalyseurs de cette série n'exhibent que des performances décevantes ici aussi. Le précurseur métallique M1 est encore celui qui fournit les quelques systèmes catalytiques pour lesquels on détecte une efficacité acceptable. Vu le nombre très restreint de réactions qui mènent à des résultats appréciables, on ne s'étonnera pas que A9B9M1 (valeur unique = 57,2) cumule à la fois la supériorité en termes de conversions (9,6 %) et d'ee (38,2 %).

On peut à ce stade s'interroger en constatant l'évolution générale des performances des dernières séries A6 à A9. En effet A6, A8 et A9 donnent des catalyseurs peu efficaces, alors que A7 exhibe des résultats nettement meilleurs. Les séries A2 et A4 doivent aussi entrer dans la comparaison. La structure des 6 aminoalcools impliqués dans la synthèse de ces catalyseurs est rappelée au schéma 141 ; on ne peut que constater leurs faibles différences structurales.

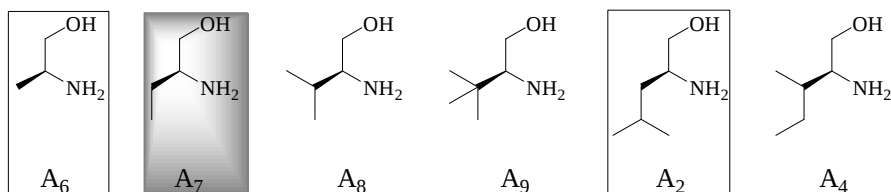


schéma 141

Si A4, A8 et A9 sont des séries quasiment inefficaces pour l'entièreté de leurs membres, A2 et A6 donnent toutefois des performances plus élevées, bien que modestes si on les compare à A7. On notera que le carbone en β de la fonction amine des séries impliquant A4, A8 et A9 sont di- voire trisubstitués, abstraction faite des hydrogènes ; alors ce même carbone est mono- voire non substitué pour A2, A6 et A7, cette dernière série rendant sans doute compte d'un compromis avantageux dans le cadre de ces systèmes de catalyse.

Portons-nous sur les catalyseurs de la série A10 que décrit la figure 32.

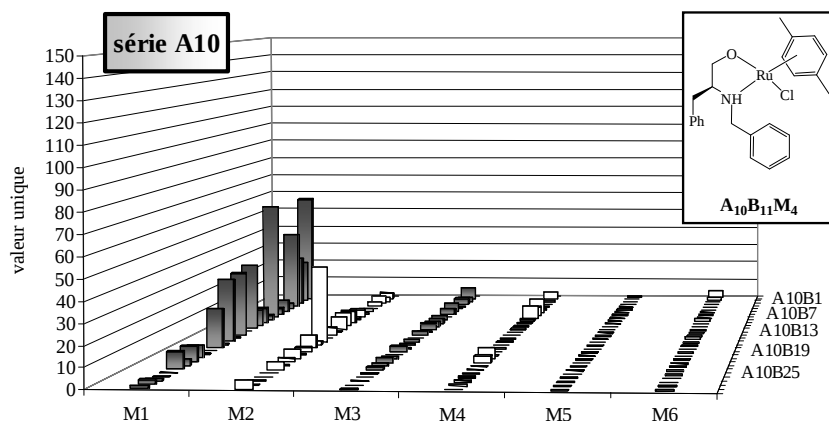


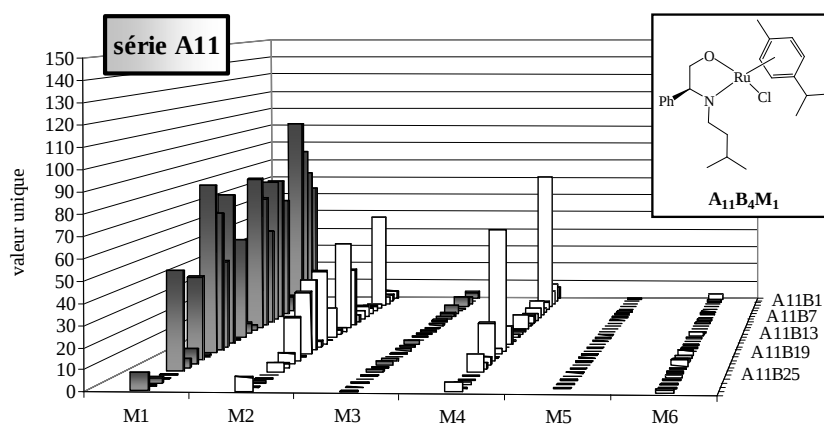
figure 32

Hormis lorsque A10 est couplée à M1, on constate l'inactivité générale des systèmes catalytiques. Un seul résultat au sein de M2 déroge à cette constatation. Le tableau 26 ci-dessous donne les meilleures performances de la série A10

tableau 26

code	performance	val. unique	conv. (%)	ee (%)
A10B11M1	combinée	68,7	14,3	40
A10B4M1	conversion	61,4	19,9	21,6

Nous clôturons cette campagne de tests par la série A11 dont les résultats sont présentés ci-dessous (figure 33).

**figure 33**

Si les combinaisons impliquant M3, M5 et M6 se révèlent inactives, celles avec M2 et M4 présentent une efficacité moyennement basse, tandis que celle associant M1 et nos ligands montre une activité globalement bonne par rapport à l'ensemble des résultats récoltés. De manière peu surprenante, les meilleurs systèmes catalytiques que présente le tableau 27 sont originaires des combinaisons avec M1.

tableau 27

code	performance	val. unique	conv. (%)	ee (%)
A11B4M1	combinée	103,4	20,5	62,4
A11B3M1	ee	86,7	10,1	66,6

5.5.2.4 Systèmes catalytiques remarquables

Après cette suite d'expériences évaluant chacun de nos systèmes catalytiques, il est utile de s'attarder sur les meilleurs candidats. Le tableau

28 ci-dessous décline le « top 20 » en trois colonnes qui naturellement correspondront aux performances combinant conversion et ee, ainsi qu'à ces paramètres pris individuellement. Nous pouvons de plus préciser que tous les ee mesurés au sein de notre bibliothèque de catalyseurs sont en faveur de l'énantiomère S du phényléthanol.

tableau 28

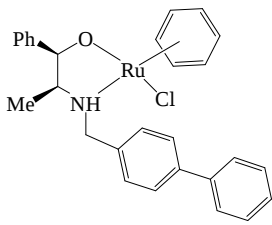
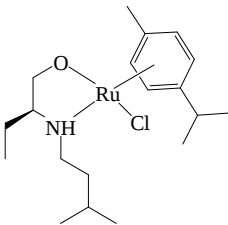
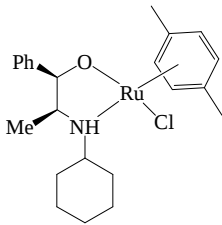
Liste des 20 meilleurs systèmes catalytiques						
ligne	performance					
	combinée		conversion		ee	
1	A1B19M2	145,6	A7B4M1	61,5	A1B3M4	88,9
2	A1B17M2	140,4	A7B19M1	56,0	A1B11M3	88,5
3	A1B4M1	138,0	A7B8M1	43,9	A1B20M1	88,3
4	A7B4M1	131,9	A7B1M1	38,2	A1B11M4	88,2
5	A7B19M1	129,3	A7B6M1	37,9	A1B22M4	88,1
6	A1B20M2	128,3	A1B19M2	36,4	A1B4M4	88,1
7	A1B11M2	125,6	A1B17M2	36,3	A1B17M1	88,0
8	A1B22M2	119,6	A7B13M1	35,7	A1B24M4	87,2
9	A1B15M2	118,3	A7B3M2	35,5	A1B22M1	86,7
10	A1B9M1	116,5	A3B2M1	27,5	A1B11M1	86,6
11	A1B20M1	114,2	A1B20M2	26,2	A1B25M1	86,5
12	A1B12M2	108,9	A1B11M2	26,2	A1B4M1	86,5
13	A1B12M1	107,8	A1B4M1	25,8	A1B12M1	86,2
14	A1B4M2	107,3	A3B1M1	24,9	A1B5M1	86,1
15	A1B4M4	106,2	A1B22M2	24,8	A1B3M1	86,1
16	A1B17M1	105,2	A7B23M1	24,7	A1B17M4	85,7
17	A1B11M1	105,0	A1B15M2	24,1	A1B19M4	85,6
18	A1B30M2	104,8	A7B10M1	22,2	A1B19M1	85,5
19	A1B3M4	104,7	A3B3M4	22,0	A1B24M1	84,9
20	A1B4M1	103,4	A3B1M2	21,9	A1B15M1	83,3

Un regard porté à la colonne des conversions nous informe que les meilleurs systèmes sont globalement situés entre 25 et 35 %, hormis 6 qui sont meilleurs. Si l'on suit les conversions par ordre décroissant, on s'aperçoit que celles-ci tombent déjà en dessous des 20 % si l'on examine le 26^{ème} catalyseur, et sous les 10 % lorsque l'on porte notre attention au-delà du 108^{ème} catalyseur.

Les meilleurs excès énantiomériques observés sont, comme nous le voyons, compris entre 83 et 89 %. Il faut également passer la 26^{ème} place de ce classement pour tomber en deçà des 80 % ee, et la 93^{ème} place pour arriver sous la barre des 50 % ee.

Les trois systèmes catalytiques les plus remarquables au sein de notre bibliothèque en termes de performance combinée, de conversion ou d'ee possèdent les structures que présente le tableau 29 ci-dessous.

tableau 29

	meilleurs catalyseurs		
	combinée	conversion	ee
structure			
code	A1B19M2	A7B4M1	A1B3M4
perf.	145,6	61,5	88,9

Les deux structures qui respectivement permettent d'accéder aux conversion et ee les plus hauts se retrouvent dans le «top 20» établi précédemment, en termes de performances combinées également. Le système A7B4M1 est à la 4^{ème} place de notre classement (ligne 4), et A1B3M4 à la 19^{ème} position (ligne 19). La manière de procéder pour définir

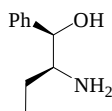
notre valeur unique tablait sur une excellente conversion de 30%, alors que plusieurs catalyseurs furent observés avec une conversion plus élevée que 30 %, ce qui explique la propulsion de A7B4M1 à la 4^{ème} position.

Mettons maintenant en exergue l'occurrence des blocs constitutifs de ces 20 meilleurs catalyseurs. Le tableau 30 ci-dessous la présente pour chaque bloc ; le pourcentage est donné en ne tenant bien sûr compte que de ce que l'on observe dans les colonnes correspondant à la conversion et à l'ee. De plus, il ne sera pas fait mention des groupements carbonylés, dégager une tendance se révélerait hasardeux au regard du grand nombre d'entre eux présent.

tableau 30

occurrence des blocs constitutifs (%)				
amino-alcool	A1	A3	A7	
	67,5	10	22,5	
précurseur métallique	M1	M2	M3	M4
	57,5	20	2,5	20

Comme nous l'observons, l'assise de A1 et M1 est prononcée parmi ces meilleurs résultats. A1 s'avère déterminant comme bloc constitutif lorsque la performance est considérée en termes d'excès énantiomérique puisque les 20 meilleurs systèmes énantiosélectifs en sont constitués. On remarquera par opposition que A7 fournit les 5 meilleures conversions de l'entièreté de la bibliothèque de catalyseurs, elles sont de surcroît largement supérieures aux autres. Cette constatation mise en rapport avec l'évolution que nous avons notée concernant les performances de A2, A4 et A6-A9, nous invite à proposer ci-dessous une structure intermédiaire d'amino-alcool potentiellement intéressante. Elle combine la présence apparemment essentielle du groupement phényle en α de l'oxygène et une préférence pour un substituant éthyle en α de l'azote face au méthyle. On sait toutefois qu'ajouter, au sein d'un même ligand, deux caractéristiques structurales remarquées ne conduit pas nécessairement à des meilleures performances. La stéréochimie est bien évidemment conservée.



Les précurseurs métalliques font eux la part belle à M1 que l'on retrouve dans près de 60 % des meilleurs catalyseurs. Il est associé à A7 pour réaliser les 5 plus hautes conversions observées. Il est également associé à A1 pour atteindre un ee de 88,3 % (tableau 28, ligne 3), qui diffère de 0,6 % du maximum, et sort à peine de l'écart-type mesuré de 0,5 % lors des tests de reproductibilité menés précédemment.

5.6 Evolution simulée

"Une société ne peut durer qu'à condition que chaque génération transmette à la nouvelle génération ce qu'elle a créé et reçu de meilleur."

Cette phrase d'André Comte-Sponville exprime clairement l'impérieuse nécessité pour une population qui veut se perpétuer, à communiquer les améliorations acquises au fur et à mesure et qui lui permettent de s'accommoder au mieux de la pression évolutive qui lui est imposée.

Nous allons procéder par analogie. Au sein de notre bibliothèque seront choisis des catalyseurs, constitutifs de la génération initiale qui se verra examinée en termes d'un indice de performance moyen combinant la conversion et l'excès énantiomérique de chaque catalyseur. La pression évolutive que nous lui appliquerons volontairement est d'améliorer sans cesse cet indice, et si possible, le plus rapidement. Pour ce faire, nous permettrons à la génération d'user des opérateurs génétiques précédemment définis afin de s'améliorer et de se transformer en génération ultérieure.

Conformément à l'état actuel de la science en ce domaine, nous proposons deux voies d'accès à ce type d'évolution, dont une seule sera évaluée pour les raisons que nous décrirons ci-après.

5.6.1 Résultats ordonnés

Avant de procéder à l'évolution proprement dite s'imposait un bref examen des résultats, comme le présente la figure 34. Les divers systèmes catalytiques ont été ordonnés par ordre croissant de valeur unique, qui pour

rappel correspond à la somme de l'excès énantiomérique et du double de la conversion.

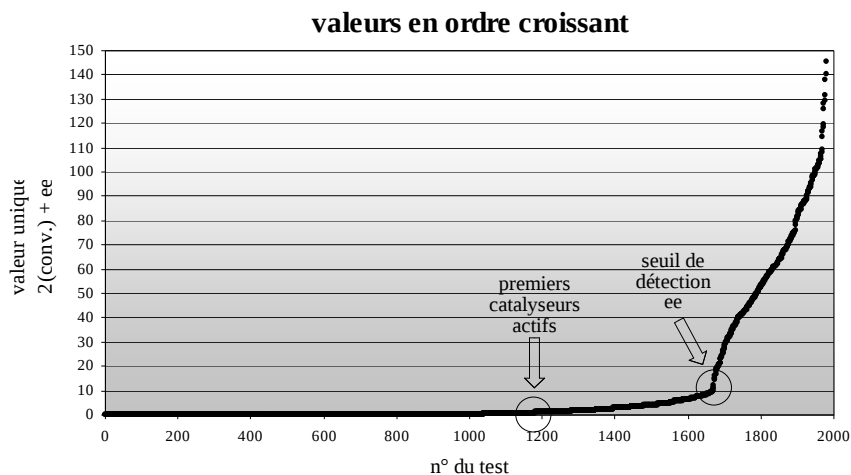


figure 34

Suivant cet ordre, nous avons dû constater à regret et de manière peu attendue que les presque 1200 premiers catalyseurs ne faisaient montre d'aucune activité, soit 60 % de notre bibliothèque. Des conversions très faibles permettent de déterminer les premiers catalyseurs actifs. La courbe d'évolution des performances présente en outre une nette variation peu après les 1400 premiers résultats. Ceci dénote simplement le fait que les systèmes n'atteignant que 4 à 4,5 % de conversion rendaient une détermination non-ambigüe de l'excès énantiomérique difficilement possible. Il a donc été décidé de fixer le taux de détermination à 5 % ; les catalyseurs confinés sous ce seuil voyaient l'ee automatiquement fixé à 0 %. Ajoutons que nombre de chromatogrammes GC associés à des résultats proches de 5 % firent l'objet d'un examen attentif dans la détermination de l'énantiosélectivité. De manière générale beaucoup de résultats supérieurs à 5 % se virent vérifiés, et les cas apparemment litigieux le furent tous.

5.6.2 Réseaux de neurones et construction d'une surface dans un espace à n dimensions – évolution simulée

Nous avons eu la chance de pouvoir discuter de manière approfondie à ce sujet avec le Professeur Michel Verleysen, œuvrant en Faculté des Sciences Appliquées (Machine Learning Group), afin de mettre sur pied un procédé évolutif basé sur la réalisation d'une surface usant de nos résultats.

Les axes utilisés x , y , z , m et n pour bâtir cette surface au sein d'un espace à 5 dimensions correspondraient respectivement aux précurseurs métalliques, amino-alcools et groupements carbonyles d'une part, ainsi qu'à la conversion et à l'énantiosélectivité d'autre part. L'évolution se réaliserait comme suit : les « p » premiers résultats donneraient une approximation très générale de la surface, de laquelle on pourrait dégager une tendance évolutive vers le ou un des maxima éventuels. Une proposition de nouvelles combinaisons en termes d'amino-alcools, de groupements carbonyles et de précurseurs métalliques serait donnée par le biais d'un réseau de neurones. Ces nouvelles combinaisons nous permettraient alors d'accéder à « q » résultats supplémentaires, lesquels seraient associés aux premiers afin de redessiner la surface de manière plus précise et d'en dégager une tendance affinée vers le ou un des maxima éventuels. En procédant par itérations successives, il devrait nous être possible d'accéder *in fine* à quelques unes des structures les plus efficaces de notre bibliothèque, voire même pourquoi pas à la plus performante.

Cette façon de promouvoir l'évolution au sein d'une bibliothèque n'est toutefois pas sans contraintes, comme nous allons maintenant le détailler. Relevons d'abord celle qui est à l'origine de la valeur unique attribuée à chaque système catalytique. En fait, construire une surface de qualité dans un espace à 5 dimensions requiert beaucoup plus de données ($\cong 10^5$) que les 1980 à notre disposition. Il convient donc de réduire cet espace à 4 dimensions en fondant la conversion et l'excès énantiomérique en une seule valeur. Ceci réalisé, d'autres impératifs restent pendants :

- La première exigence requise est de coder les différents blocs de construction en conservant une signification chimique derrière le

code, ce qui implique de passer par des descripteurs moléculaires. S'il est possible de coder des molécules organiques assez simplement, l'exercice s'avère plus compliqué dès lors qu'il s'agit de complexes métalliques.

- Il est ensuite nécessaire que le pas entre les molécules soit constant, ce qui en d'autres termes implique que sur l'axe par exemple destiné aux amino-alcools, l'équidistance entre molécules doit être respectée (plus précisément donc, entre les valeurs accordées aux molécules par le truchement d'un descripteur moléculaire quelconque). En plus de cette nécessaire équidistance, il est impérieux que chaque valeur de pas intermédiaire (entre la première et la dernière de ces valeurs) soit associé à une molécule ; autrement dit la continuité doit exister et être respectée. Il est difficile de croire dans notre cas que ce besoin soit accessible aisément, comme en témoigne la figure 35 ci-dessous. Si l'on conçoit trouver une manière de placer A6-A9 à la suite l'un de l'autre, le rang qu'occuperont A1 et A5 par exemple sera plus problématique. Il sera donc nécessaire de calculer pour chaque série (amino-alcools, groupements carbonyles et précurseurs métalliques) un grand nombre de descripteurs moléculaires jusqu'à en trouver un qui puisse satisfaire aux exigences de pas constant et de continuité.

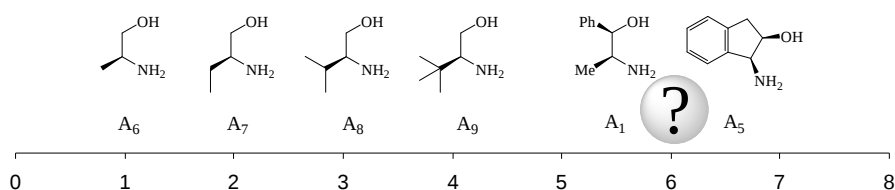


figure 35

- Une nouvelle discussion avec le Pr Verleysen nous apprend la nécessité d'entreprendre dans ce cas le calcul, pour chaque molécule, d'un nombre probable de descripteurs compris entre 1000 et 2000. Avis pris de concert auprès du Dr Robiette et de Marie-Laure Boulanger, la chose est envisageable, mais nécessitera plusieurs mois de calculs et d'interprétation des résultats. En supposant que ce bon

millier de descripteurs ait fourni des données de qualités en termes de pas et de continuité, il faudra absolument pouvoir les corrélérer entre elles puisque nous désirons procéder à l'évolution par le biais d'une surface. Bien qu'il existe des méthodes mathématiques performantes et accessibles en FSA permettant de déterminer les corrélations les plus pertinentes, l'examen de nos données ordonnées comme les présentait la figure 34 nous fait émettre de vives réserves sur l'hypothétique continuité d'une telle surface, et ce principalement à cause des groupements carbonyles qui se distribuent autant parmi les bons que les mauvais résultats.

Il existait donc maintes incertitudes à propos d'abord des descripteurs moléculaires capables de générer entre les molécules un pas constant et une continuité de classement, au regard ensuite du caractère nécessairement corrélable des manières de décrire ces molécules. Ajoutons encore à cela le temps de calcul indispensable à l'obtention de ces données. Ces éléments rassemblés nous ont conduits à ce que cette méthode itérative utilisant les réseaux de neurones ne fasse pas l'objet d'une tentative d'évolution simulée.

5.6.3 Algorithme génétique – évolution simulée

Une autre manière d'envisager l'évolution simulée réside dans l'emploi des algorithmes génétiques, dont nous avons également parlé au cours de l'introduction. Il suffit d'attribuer un code à nos catalyseurs et ceux-ci sont alors utilisables par un algorithme. Nous avons très aimablement pu bénéficier d'une collaboration avec le Pr Bernadette Govaerts, de l'Institut de Statistique de l'UCL, afin de coder l'algorithme génétique suivant nos convenances. Le code canonique est par ailleurs donné en annexe.

5.6.3.1 Fonctionnement de l'algorithme génétique

Il est utile de décrire chaque étape qu'utilise l'algorithme génétique pour promouvoir une évolution au fil des générations. La figure 36 présente en détails la succession d'événements se produisant lors de la transformation d'une génération « G_n » en la génération successive « G_{n+1} ».

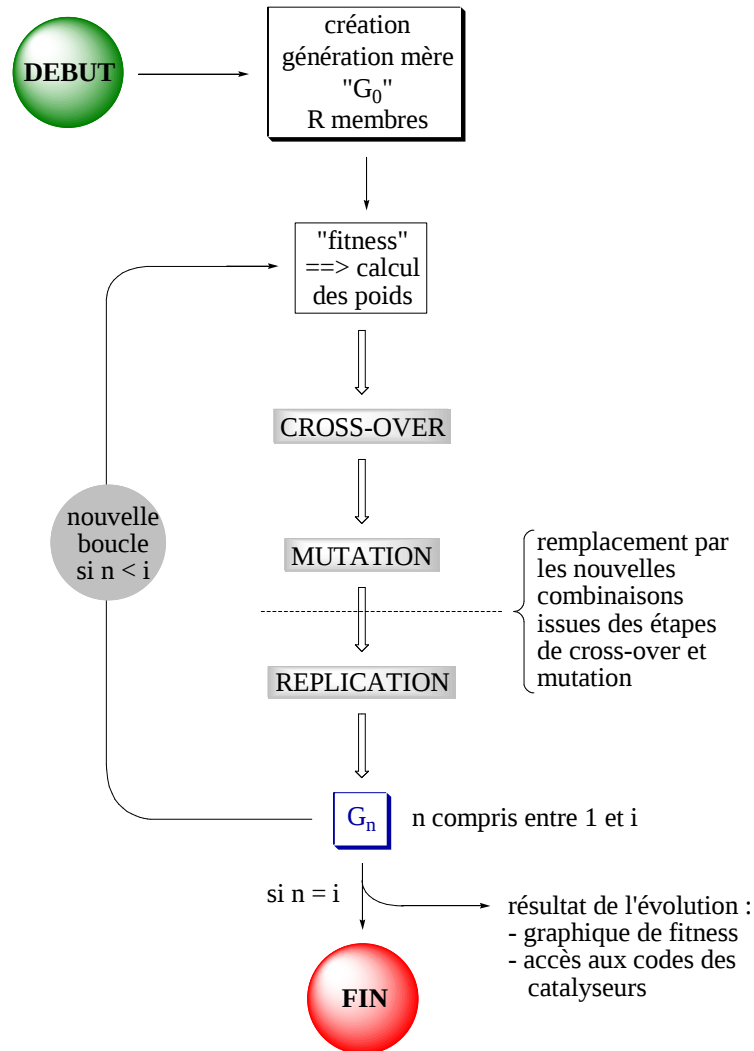


figure 36

Le processus d'évolution simulée débute par la création d'une génération mère. La valeur unique associée à chaque catalyseur est prise en compte pour déterminer une « fitness moyenne ». Pour juger de la qualité relative des systèmes catalytiques au sein de la génération considérée, on divise les valeurs uniques individuelles par la « fitness moyenne », cela servira à établir un ordre prioritaire utile pour les deux étapes suivantes.

Survienent ensuite les cross-over et mutations. On sélectionnera les candidats catalyseurs sujets au cross-over selon une probabilité pondérée par l'ordre établi précédemment. Cet ordre implique que les systèmes les plus efficaces aient une chance plus élevée de participer aux cross-over, attendu qu'ils sont supposés posséder les caractéristiques les plus intéressantes de la génération considérée et que les échanger avec d'autres candidats pourrait s'avérer avantageux. Concernant la mutation, la probabilité est égale pour chacun des membres de la génération traitée.

On constate qu'après ces deux étapes ont lieu les remplacements d'anciennes combinaisons de type $A_xB_yM_z$, et donc de catalyseurs, par les nouvelles combinaisons que les cross-over et mutations viennent de générer. Ces remplacements ne se déroulent pas au hasard, l'ordre établi des molécules permet aussi d'identifier les systèmes les moins efficaces qui seront bien entendu les cibles naturelles désignées à ces remplacements. Après cette étape intervient la réplication qui conserve les meilleurs systèmes catalytiques n'ayant donc pas été l'objet de remplacements. Nous avons à ce stade entre nos mains la nouvelle génération qui pourra être impliquée dans une nouvelle boucle d'évolution.

L'algorithme fut programmé de telle sorte que lorsqu'il recherche les codes génétiques des catalyseurs, il soit capable de détecter si ceux-ci ont déjà été évalués auparavant. Si tel est le cas, il abandonne la solution de modification proposée et en cherche une autre qui sera acceptée à condition qu'elle offre la nouveauté requise. Notons encore qu'il est possible de vérifier si une proposition d'échange déjà antérieurement réalisée a été reproposée par l'algorithme, ainsi que le nombre éventuel de fois où celle-ci le fut. Cette obligation de solution neuve pourrait toutefois présenter le désavantage de confiner l'algorithme dans une boucle de recherche perpétuelle si d'aventure le calcul des poids imposait des échanges univoques, ou si les solutions proposées avaient pour leur grande majorité été évaluées, et ce particulièrement en fin d'évolution. Pour des raisons pratiques sur lesquelles nous reviendrons ultérieurement, il a généralement été décidé de faire tourner l'algorithme jusqu'à ce qu'il génère au minimum 6 nouvelles propositions de catalyseurs jamais encore évaluées.

Lorsque le nombre de boucles algorithmiques est tel que $n = i$, l'évolution simulée se termine. Le programme générera alors deux types de graphiques utiles à un examen visuel rapide et efficace de la qualité de l'évolution qui vient d'être conduite. L'un rendra compte de la fitness moyenne de chaque génération, ainsi que de la fitness moyenne des dix meilleurs candidats de chaque génération également. L'autre montrera si l'on parvient au fur et à mesure des générations successives à retrouver un voire plusieurs des dix meilleurs catalyseurs parmi l'entière des 1980 combinaisons testées. A la fin de l'évolution, les codes génétiques de chaque catalyseur impliqué dans les générations successives, avec le détail de ceux qui auront participé aux cross-over et mutations, seront pleinement accessibles. Cette lecture possible des codes rend le décortilage du cheminement de l'évolution simulée parfaitement lisible à n'importe quelle étape.

5.6.3.2 Application de l'algorithme génétique

Nous allons maintenant avoir le loisir d'examiner les résultats que peut produire l'algorithme. Il est évident que le succès de l'évolution simulée s'exprimera par le nombre et/ou la rapidité avec laquelle l'algorithme est à même de retrouver certains des meilleurs catalyseurs de notre bibliothèque. Nous allons à nouveau utiliser la valeur unique liée à chaque catalyseur pour évaluer sa performance. Notons qu'il serait envisageable de produire une évolution simulée qui utiliserait un « index de désirabilité » pour pondérer de manière plus rigoureuse la conversion et l'ee¹⁹⁵ pour information ; nous avons toutefois jugé que ce dernier point sortait du cadre de notre travail de thèse.

L'efficience de l'algorithme est dépendante des valeurs que l'expérimentateur fixera concernant les paramètres modulables de ce dernier, au nombre de 7 :

- « G_0 », la génération mère, peut être laissée au hasard ou fixée par l'expérimentateur. Le choix de la génération de départ devrait porter à conséquence sur le déroulement de l'évolution, génération après génération.

- « R », soit la taille de la génération mère.
- « RT » le nombre de molécules que l'on décide de conserver pour une réplification obligatoire. La décision de faire passer irrévocablement une molécule d'une génération à l'autre tient au poids relatif qui lui est attribué au sein de la génération. Les meilleures molécules seront sujettes à la réplification obligatoire.
- « RE », le nombre de molécules que l'on veut échanger. Ce nombre rend compte du taux de cross-over que l'on impose à l'algorithme.
- « RM », le nombre de molécules susceptibles de subir l'étape de mutation. Le taux de mutation imposé à l'algorithme déterminera aussi la manière dont se déroule l'évolution
- « RR » détermine l'ensemble de molécules qui pourront faire l'objet d'évincement par l'algorithme, pour les remplacer par des combinaisons nouvelles, résultant des cross-over et mutation. « RR » rassemblera les catalyseurs de faibles poids relatifs.
- « NI », soit le nombre de boucles évolutives, et donc de générations successives, que l'on accorde à l'algorithme avant de le déclarer fini.

Comme nous l'observons, le premier paramètre énoncé est la génération mère, qui peut être aléatoire ou avoir été fixée par le chimiste en fonction de ce qu'il pense connaître comme facteurs importants et liables à la performance considérée. Il nous a semblé intéressant de débiter nos simulations évolutives en vérifiant l'influence éventuelle du choix de « G_0 ».

5.6.3.2.1 Génération mère aléatoire

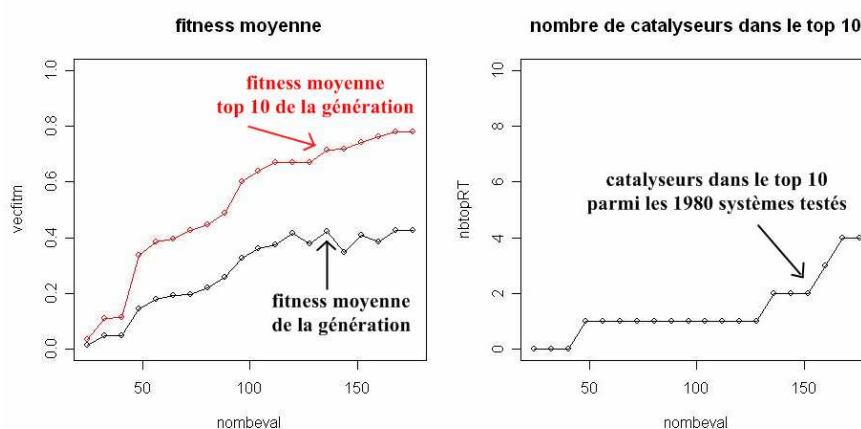
Pour ces premières simulations, nous voulions simplement observer une tendance ; ainsi avons-nous choisi de travailler avec 20 boucles d'évolution, avec 24 membres par génération et des taux de cross-over et mutation respectivement de 3 et 2. Cela imposera que 8 molécules nouvelles soient introduites normalement à chaque génération ; le cross-over, par sa nature d'échange générant 6 nouvelles molécules. Le tableau 31 ci-dessous donne les valeurs de ces paramètres avec clarté.

tableau 31

G_0	R (taille)	RT (top)	RE (ech.)	RM (mut.)	RR	NI
aléatoire	24	10	3	2	16	20

La figure 37 donne les résultats de la première simulation. Au cours de diverses évolutions que nous allons entreprendre, il en sera à chaque fois ainsi. Deux graphes sont générés décrivant pour celui de gauche l'évolution de la « fitness » moyenne de la génération considérée, ainsi que des 10 meilleurs catalyseurs de cette génération. La légende « nombeval » en abscisse précise le nombre de catalyseurs évalués. La « fitness » posée en ordonnée dépeint la tendance de la génération à se rapprocher du meilleur résultat, normé à 1. Nous observerons également dans un second graphe combien de catalyseurs, parmi les 10 meilleurs de notre bibliothèque considérée dans son entièreté, l'algorithme est capable de mettre en lumière.

simulation n°1 (G_0 aléatoire)

**figure 37**

Comme nous aurons l'occasion de le remarquer après, cette première simulation est assez typique d'une évolution moyenne. La fitness moyenne des générations croît jusqu'à la 13^{ème} génération pour ensuite stagner irrégulièrement. La fitness moyenne des 10 meilleurs catalyseurs, qui examine en fait une sous-génération pour chacune des générations, grimpe continuellement, ce qui confirme qu'après la 13^{ème} génération ; les processus de cross-over et de mutation actent à créer de la diversité au sein de la

génération et permettent de trouver d'autre bons candidats en améliorant les caractéristiques des catalyseurs. Ceci est confirmé à l'examen du graphe de droite de la figure 37, puisque le passage de la 14^{ème} à la 15^{ème} génération dénote l'arrivée d'un nouveau catalyseur généré faisant partie du top 10 de notre bibliothèque. Cette croissance continue de manière appréciable puisque l'on note en fin d'évolution la présence de 4 des meilleurs catalyseurs au sein de la 20^{ème} et dernière génération. Cette première simulation marque de plus l'efficacité de la méthode si l'on avait mené le processus directement de manière réelle et pratique.

Examinons d'autres profils que les 20 simulations découvrent comme représentatifs. La simulation n°2 se déroule de manière moins efficiente, comme le présente la figure 38 ci-dessous. On constate que longtemps, l'algorithme peine à trouver une seule molécule faisant partie du top 10.

simulation n°2 (G0 aléatoire)

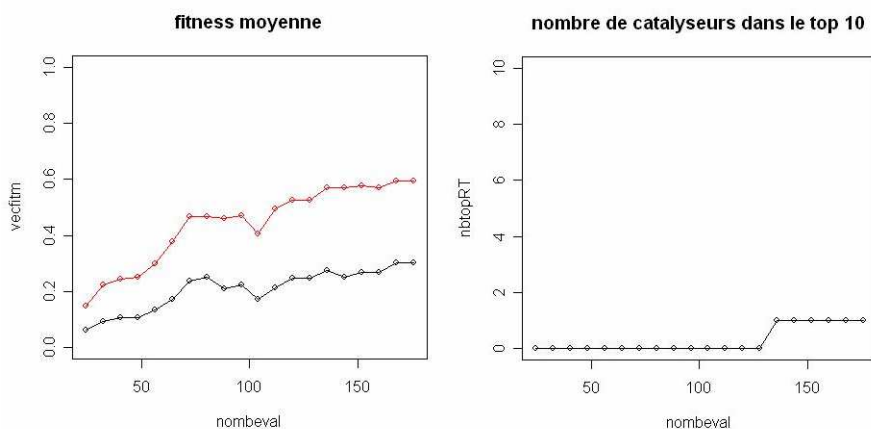


figure 38

La troisième simulation (figure 39) montre que l'algorithme découvre rapidement 3 des meilleurs catalyseurs, et puis stagne ensuite, ce dont témoignent évidemment fitness moyenne des générations et fitness moyenne des sous-génération ne prenant en compte chaque fois que les 10 meilleurs systèmes catalytiques.

simulation n°3 (G0 aléatoire)

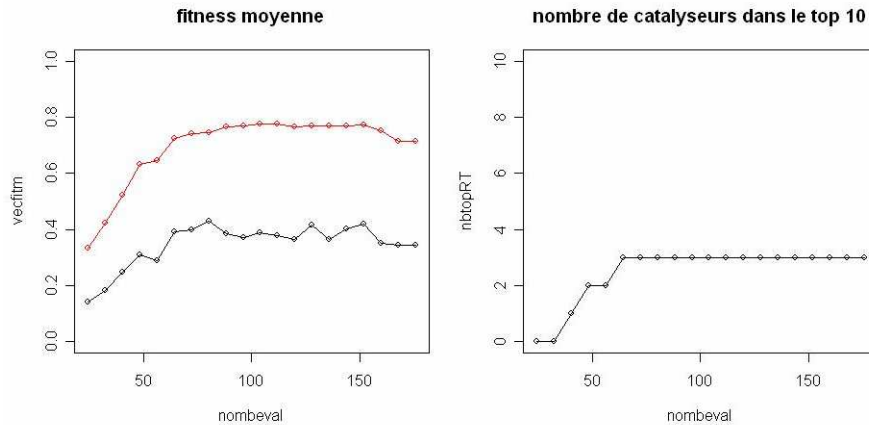


figure 39

Nous observerons aussi des simulations évolutives pour lesquelles la découverte des meilleurs catalyseurs intervient tardivement, ce qui se produit au cours de la 6^{ème} évolution (figure 40).

simulation n°6 (G0 aléatoire)

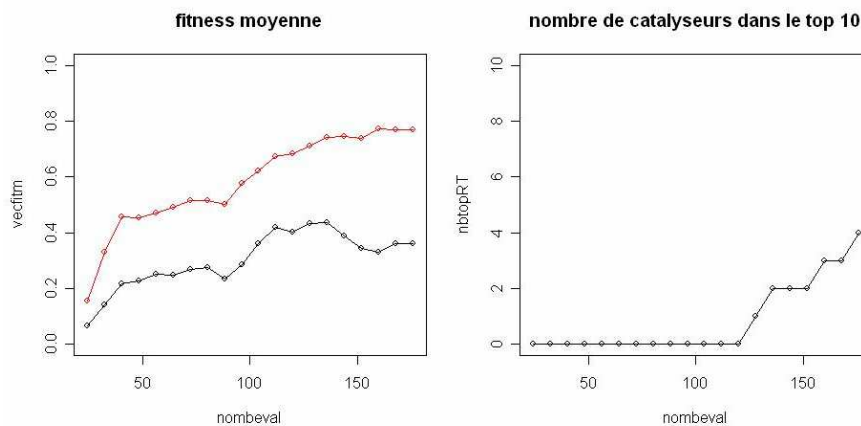


figure 40

Parmi ces expériences simulées, nous aurons le plaisir de constater que certaines d'entre elles se révèlent assez remarquables comme le démontre la 17^{ème} évolution (figure 41). En moins de 150 expériences, il nous aurait été possible dans ce cas d'accéder à 7 des 10 meilleurs catalyseurs ; de plus le

meilleur système catalytique A1B19M2 a été découvert. On remarquera aussi le bond entre la 13^{ème} et la 14^{ème} génération puisque pas moins de 3 nouveaux systèmes parmi les plus performants sont révélés.

simulation n°17 (G0 aléatoire)

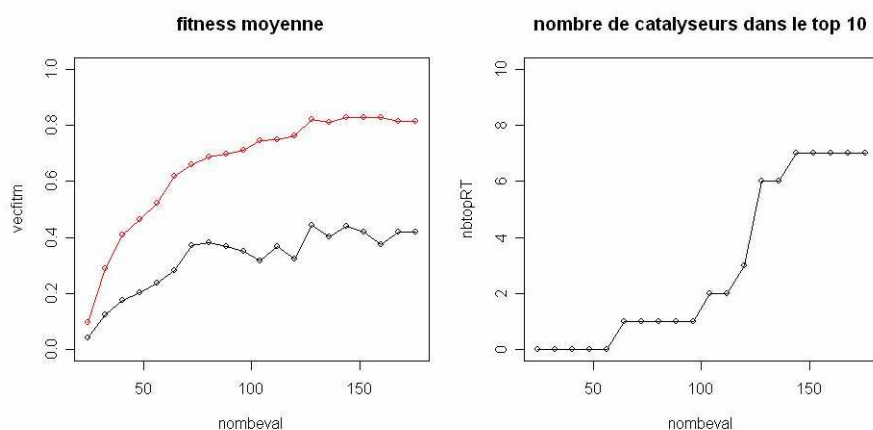


figure 41

Le tableau 32 ci-dessous reprend les résultats obtenus au cours de 20 simulations impliquant une génération de départ aléatoirement choisie. Pour chacune, le nombre de molécules du top 10 que l'algorithme a permis de découvrir est indiqué, ainsi que la présence éventuelle du meilleur catalyseur A1B19M2. Nous retrouvons ce dernier à 7 reprises. Sur ces 20 simulations, on retrouve en moyenne 3,8 candidats du top 10 par simulation.

tableau 32

G0 aléatoire					
n° simul	nbr. top 10	A1B19M2	n° simul	nbr. top 10	A1B19M2
1	4		11	5	X
2	1		12	7	X
3	3	X	13	1	
4	6		14	3	
5	0		15	1	
6	4	X	16	4	X
7	6	X	17	7	X
8	4		18	0	
9	4		19	5	X
10	4		20	7	X
nombre moyen de top 10 = 3,8					

Une autre manière d'évaluer l'efficacité de l'algorithme peut être rendue par sa rapidité à retrouver le plus possible de candidats du top 10 ; il suffit pour ce faire de prendre le rapport pour chaque simulation entre le nombre maximum de candidats découverts et la première génération à partir de laquelle ils apparaissent. La figure 42 présente ces rapports pour nos 20 simulations.

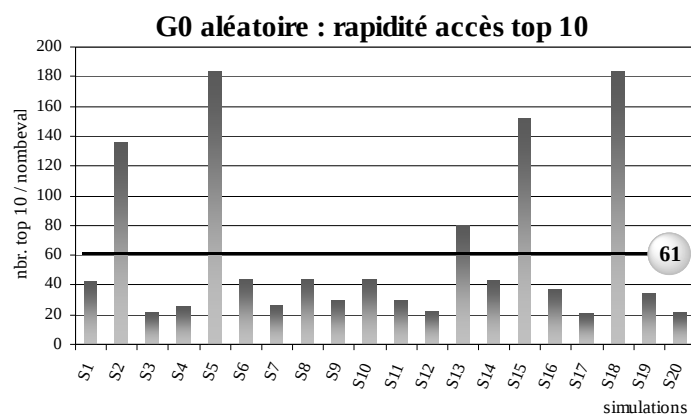


figure 42

Comme le fait remarquer le tableau précédent, deux systèmes ne donnent accès à aucun des catalyseurs du top 10, ce qui aurait conduit rigoureusement à prendre pour eux une valeur infinie dans le graphe ci-dessus et rendu la détermination de la moyenne impossible. Nous avons donc décidé de considérer avec un peu d'optimisme ces processus évolutifs comme s'ils donnaient 1 des meilleurs catalyseurs à la génération suivante. La moyenne à laquelle nous accédons aux systèmes catalytiques est de 62. En d'autres termes, il faut en moyenne tester 62 catalyseurs pour en découvrir un du top 10. La signification de cette moyenne doit être toutefois tempérée, puisque l'écart-type associé prend une valeur élevée de 55,2.

5.6.3.2.2 Génération mère fixe

Nous avons précisé il y a peu que le paramètre de génération mère pouvait être laissé au hasard, ou alors fixé. Face à une évolution que l'on voudrait mener dans un cas réel, le chimiste peut être tenté de définir la génération de départ. Nous avons décidé de construire cette première génération en lui imposant, à notre sens, le maximum de diversité possible et en négligeant volontairement la connaissance que nous avons des résultats que produisent les catalyseurs.

Parmi les 11 amino-alcools, il convient de former trois groupes dont on extraira aléatoirement un des éléments. On rassemblera A1 et A3 en raison de leur double substitution en α de NH_2 et du OH. A5 possède une structure particulière qui nous invite à le retenir d'emblée ; restent ensuite A2, A4 et A6 – A11. Les amino-alcools retenus après tirage au sort par ensembles sont A3, A5 et A7.

Nous définissons 4 groupes pour les composés carbonylés : les aldéhydes aliphatiques de B1 à B5 inclus, même s'il faut considérer une certaine largesse pour que B5 en fasse partie. Les cétones suivent alors, de B6 à B10 inclus. Le troisième groupe se compose des aldéhydes aromatiques, de B11 à B23 inclus. Nous terminons par les aldéhydes qui possèdent un hétéroatome comme point de chélation potentiellement supplémentaire. Le sort détermine pour chaque amino-alcool à qui il sera associé. Si un des candidats a déjà fait l'objet d'une sélection, il est

automatiquement rejeté et un nouveau est tiré au sort. De cette manière, 12 composés carbonylés différents feront partie intégrante de la première génération.

Nous procédons encore de même concernant les précurseurs métalliques, au sein desquels nous définissons deux groupes. Celui des complexes de ruthénium que coordinent un noyau aromatique, et le $\text{RuCl}_2(\text{DMSO})_4$ et $\text{Ir}(\text{COD})\text{acac}$ d'autre part. Le schéma 142 ci-dessous reprend et explicite ces combinaisons.

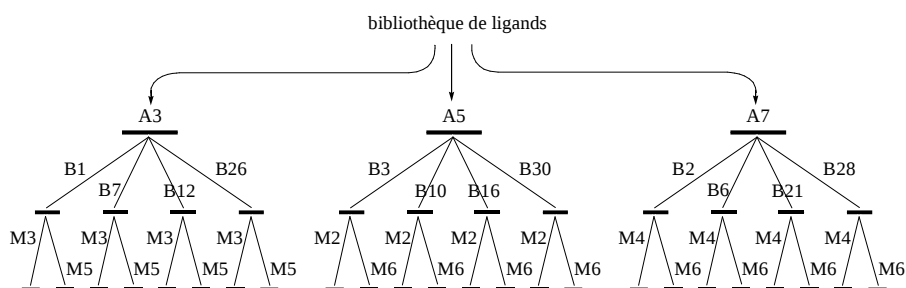


schéma 142

Toutes ces combinaisons ont été introduites dans l'algorithme et définies fixes. Le résultat – *a priori* – surprenant que nous recevons est :

« *trop peu de probabilités strictement positives* »

Différemment exprimé, cela signifie que l'algorithme est incapable de fonctionner. Les codes des catalyseurs sont associés à des valeurs identiques, il est donc impossible à l'algorithme de déterminer les combinaisons à fortes performances susceptibles d'échanger leur caractères, ni les combinaisons à faibles performances sujettes à éviction entre la génération mère et la génération suivante. Un bref coup d'œil aux figures décrivant les résultats associés à chaque série (figure 25, figure 27 et figure 29) permet de comprendre que beaucoup de valeurs sont certainement proches sinon égales à 0. En effet, si A3 et ses combinaisons permet dans beaucoup de cas d'accéder à une valeur combinée utilisable, ce n'est pas le cas lorsque cette série est engagée avec M3 et M6, ce que détermina le hasard. On se remémorera aussi que A5 produisait les résultats les plus décevants de nos

tests puisqu'hormis quelques valeurs sporadiques très légèrement supérieures à 0, l'ensemble des catalyseurs est inapte à réduire l'acétophénone dans les conditions de concentration et de temps que nous employons. En ce qui concerne la série A7, les résultats exhibés étaient plutôt bons, le hasard fit que nous en avons assez peu retrouvé.

En associant les codes des catalyseurs à leurs performances, nous observons que seuls 4 d'entre eux ont des valeurs supérieures à 0 et font donc partie du top 10 de la génération. Cela implique puisque le « RR » de l'algorithme est de 16, soit l'ensemble des molécules de faible poids potentiellement concernées par le remplacement, que celui-ci ne parvient pas à réaliser une discrimination entre les 16 combinaisons potentielles ; les 16 plus mauvaises combinaisons en termes de fitness ont une valeur de 0. Ne sachant quelle voie suivre, l'algorithme se retrouve donc bloqué dès la première génération.

Ce constat d'échec nous oblige à opérer une nouvelle sélection, à agir somme toute en expérimentateur un peu moins avisé et peut-être plus chanceux. Nous décidâmes de laisser une part plus importante au hasard en conservant A3, A5 et A7 mais en optant pour que chacune des 8 combinaisons résiduelles soit sélectionnée aléatoirement. L'algorithme peut cette fois simuler les 20 évolutions nouvelles. La même génération de départ sera utilisée à chaque reprise, et nous conservons identiques tous les paramètres de l'algorithme. Si certaines évolutions fonctionnent également assez bien, on relève un profil typique correspondant à ce que montre la simulation n°27 (figure 43).

simulation n° 27 (G0 fixe)

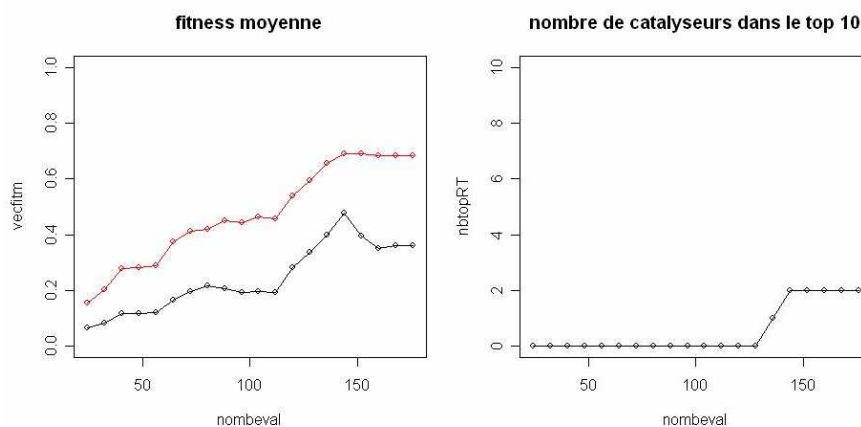


figure 43

Nous rencontrons cette fois-ci à trois reprises une évolution qui ne parvient à découvrir aucun des systèmes catalytiques les plus performants comme le montre la simulation n°24 (figure 44), malgré une fitness moyenne en croissance lente.

simulation n° 24 (G0 fixe)

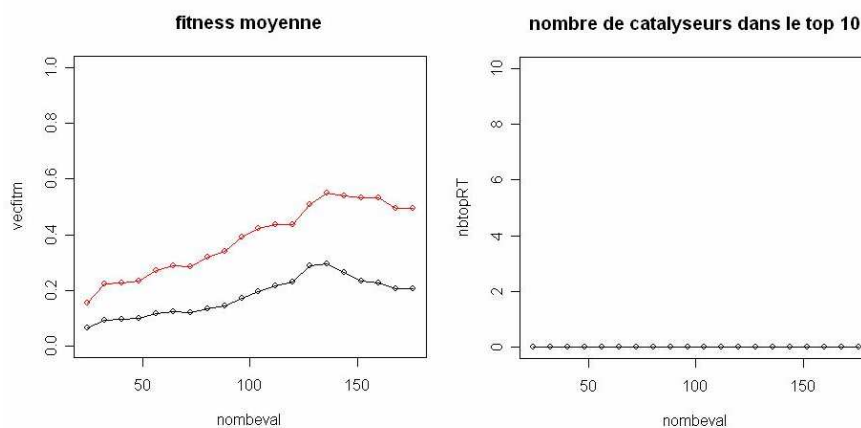


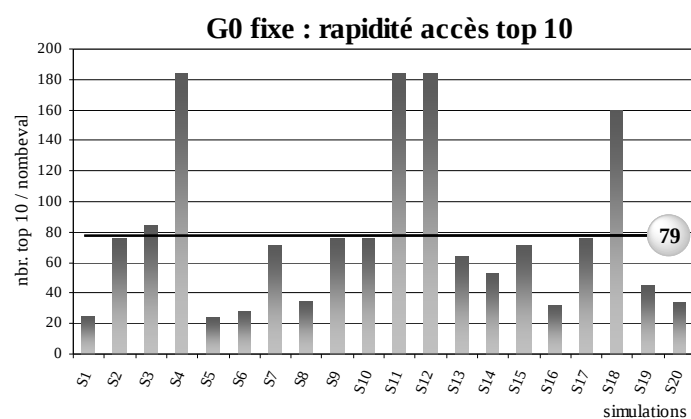
figure 44

Les résultats obtenus par ces 20 nouvelles évolutions simulées sont organisés de la même manière que précédemment, et décrits par le tableau 33 ci-dessous.

tableau 33

G0 fixe					
n° simul	nbr. top 10	A1B19M2	n° simul	nbr. top 10	A1B19M2
21	7	X	31	0	
22	2		32	0	
23	2		33	2	
24	0		34	3	
25	6	X	35	2	
26	5	X	36	4	X
27	2		37	2	
28	5	X	38	1	
29	2		39	6	X
30	2		40	4	
nombre moyen de top 10 = 2,7					

Au cours de ces processus évolutifs, nous retrouvons 6 fois le meilleur catalyseur, par contre le nombre de candidats retrouvés dans le top 10 par simulation tombe à 2,7. En observant, la rapidité moyenne de l'algorithme à retrouver un catalyseur issu du top 10 est de 79, ainsi que le précise la suivante. A nouveau, en raison des grands écarts entre les différentes valeurs répertoriées, on calcule un écart-type très élevé de 98,5

**figure 45**

5.6.3.2.3 Comparaison G0 aléatoire – G0 fixe

Chaque constatation prise séparément sur le comportement des évolutions simulées est insuffisante pour conclure au sens réel de la différence entre un algorithme qui débute soit en usant d'une génération aléatoire, soit d'une génération fixe et déterminée, au moins partiellement, par l'expérimentateur. Les données récoltées et associées en faisceau permettent toutefois de définir une tendance qualitative. Le tableau 34 place côte à côte les variations observées :

tableau 34

	nbr. top moy.	rapidité	A1B19M2	evol. sans top 10
G0 aléatoire	3,8	62	7	2
G0 fixe	2,7	79	6	3

L'algorithme qui dispose d'une génération fournie aléatoirement offre potentiellement 3,8 catalyseurs du top 10 en fin d'évolution, et en moyenne, contre 2,7 s'il commence l'évolution en se basant sur la génération mère fixe. Son efficacité est aussi plus grande, c'est ce que montre l'expression de la vitesse de découverte : 62 tests sont en moyenne nécessaires afin de découvrir un des systèmes catalytiques du top 10 pour une génération initiale aléatoirement déterminée, contre 79 si elle est prédéfinie. A 7 reprises, les simulations menées avec G0 aléatoire conduisent à trouver le meilleur système catalytique, contre 6 avec G0 fixe. Le nombre d'évolutions n'apportant aucun système catalytique du top 10 est également en faveur de G0 aléatoire. A tout cela nous pouvons encore ajouter que les profils d'évolution diffèrent entre ces deux batteries de simulations puisque lorsque la génération mère est fixe, on constate plus souvent une stagnation si on observe le nombre de systèmes catalytiques membres du top 10. Ceci s'explique par le fait que la génération mère fixe est de qualité médiocre, les échanges par cross-over s'avèrent donc moins souvent efficaces, et le salut de l'algorithme à bien mener l'évolution tient alors aux mutations bénéfiques qui seront engendrées. Ce désavantage touche moins souvent un processus d'évolution ayant pour départ une génération aléatoire. Enfin, souvenons-nous encore que la première génération choisie avec ce qui nous a semblé

être du bon sens a carrément empêché l'algorithme de réaliser le moindre processus évolutif. Ce dernier fait est cependant intrinsèquement lié à la nature des résultats obtenus lors des tests catalytiques puisqu'une grande partie des systèmes a dévoilé un manque d'efficacité manifeste, ce que décrivait la figure 34 montrant la valeur unique des tests organisés de façon croissante.

Affiner et quantifier cette tendance requerraient un nombre plus important d'expériences. Toutefois nous pouvons conclure sans gros risque d'erreur qu'une génération de départ laissée au choix du sort constituera un facteur de succès important de nos évolutions.

5.6.3.2.4 Echange par cross-over – mutation faible

Deux des paramètres majeurs de l'algorithme sont les opérations de cross-over et de mutation, et le taux variable avec lequel ils peuvent être indépendamment utilisés pour conduire à diverses fortunes en fin de processus évolutif. Nous allons maintenant examiner les résultats fournis en faisant varier un des ces paramètres, l'autre étant volontairement bloqué à sa valeur minimale. Il aurait été théoriquement possible de travailler uniquement avec les cross-over sans la moindre mutation et inversement. Cette voie ne fut pas envisagée à cause de l'insuccès alors promis ; en effet des échanges internes à la génération sans la moindre introduction de nouveauté conduiraient inévitablement à condamner l'algorithme dans un maximum local. De même, la mutation si elle introduit la nouveauté, ne peut la transmettre sans l'aide des cross-over, et cet effet maximum sera néfaste si la mutation introduit un caractère particulièrement intéressant. La mutation seule, de ce fait, conduirait à une évolution possible, mais forcément très ralentie.

Afin d'évaluer le comportement de l'algorithme soumis à une influence plus grande du cross over, nous avons mené 10 simulations pour chaque variation du taux d'échange, au nombre de quatre ; la mutation reste confinée à un élément de la génération, chaque fois ($RM = 1$). Suivant les enseignements acquis au point précédent, nous débuterons bien sûr avec une génération aléatoirement déterminée de 24 membres, toutefois pour que la

comparaison soit crédible nous conserverons fixe cette génération mère aléatoire. Les évolutions se pratiquent sur 20 générations successives

Pour les 10 premières simulations on détecte un profil habituel que montre la figure 46. Les taux appliqués sont $RE = 1$, et $RM = 1$; 3 nouvelles combinaisons varient donc d'une génération à l'autre. Les simulations sont numérotées en continu, celle-ci est la 4^{ème} de la série.

simulation n° 44 ($RE = 1$; $RM = 1$)

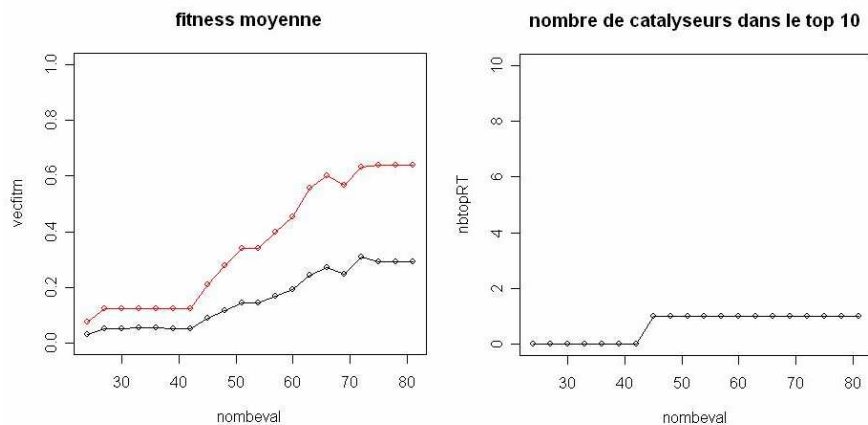


figure 46

Nous ne parvenons par ce biais qu'à découvrir entre un et deux catalyseurs très intéressants par cycle évolutif, ce qui par rapport aux faibles perturbations intergénérationnelles est tout à fait compréhensible. Le meilleur catalyseur n'est découvert qu'à une seule reprise. L'augmentation de la fitness moyenne montre que le système parvient tout de même à s'améliorer malgré les faibles taux de cross-over et de mutation. Un tableau final collationnera les données relatives à cette série, ainsi qu'aux suivantes.

Il est fort difficile de prétendre honnêtement trouver un profil typique lorsque les taux de cross-over sont portés à $RE = 2$ et ensuite à $RE = 3$, les résultats sont plus épars. La dernière série ($RE = 4$) dénote une plus grande stabilité, la figure 47 en montre le profil type. On appréciera l'amélioration qu'engendre une élévation du taux d'échange.

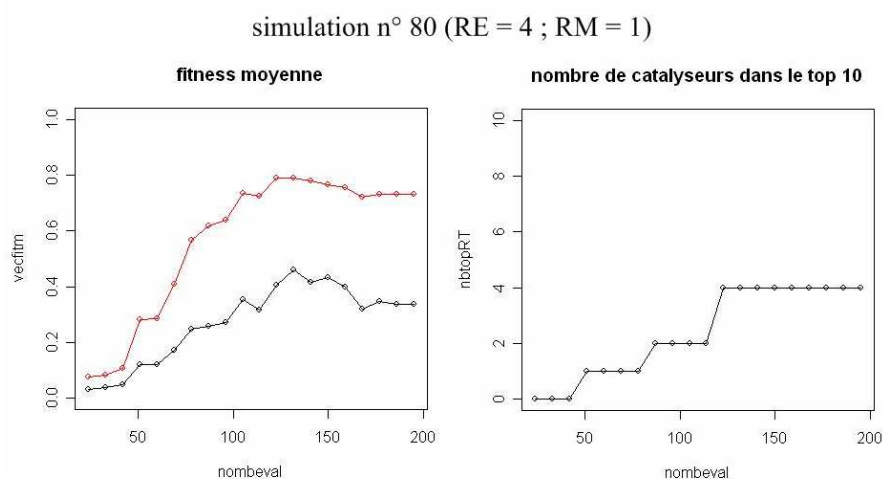


figure 47

Attachons-nous à comparer les divers résultats obtenus. Nous procédons de manière identique en utilisant les points de comparaison établis précédemment que sont : le nombre moyen de candidats du top 10 que l'algorithme parvient à trouver, la rapidité avec laquelle il opère (soit le nombre minimum d'évaluations pour retrouver un des candidats du top 10, en moyenne), la découverte éventuelle du meilleur système catalytique et enfin le nombre d'évolutions n'accédant jamais aux catalyseurs les plus performants. Le tableau 35 présente ces données.

tableau 35

série	RE	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
1	1	1,6	33	6	1	0
2	2	2,3	50	39	4	2
3	3	2,6	48	43	1	1
4	4	4,1	48	24	3	0

Le nombre de catalyseurs du top 10 que l'évolution simulée parvient à mettre à jour en moyenne croît de manière presque graduelle. S'il est risqué d'avancer des conclusions concernant des séries qui se jouxtent, on admettra que la progression entre la série 1 et la série 4 reste valable. L'observation de la colonne relative à la rapidité de découverte, en moyenne, d'un des

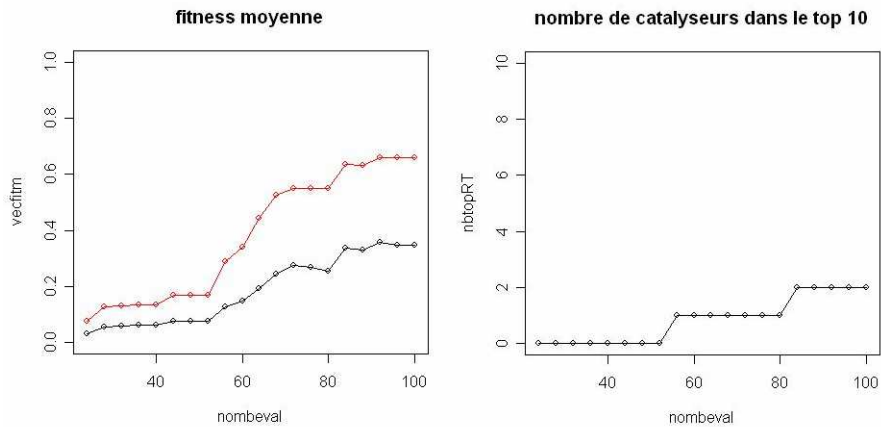
meilleurs catalyseurs doit ici aussi s'examiner en ne comparant que prudemment les séries 1 et 4 ; les écart-types associés démontrent une grande dispersion des valeurs, ce qui avait été constaté par l'examen des profils évolutifs. Tout ceci suggère de manière assez grossière une influence positive de l'augmentation du taux d'échange à laquelle on pouvait s'attendre. On précisera encore que le temps de recherche de l'algorithme augmente fortement entre la 1^{ère} et la 4^{ème} série, ce qui indique sa difficulté à proposer des échanges qui n'ont encore été réalisés.

5.6.3.2.5 Mutation – échange par cross-over faible

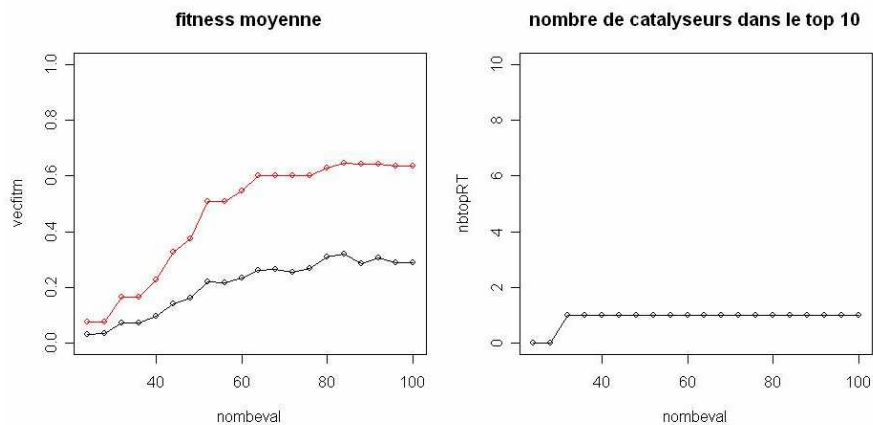
Nous allons pour ce point pratiquer le mécanisme inverse, en reléguant le taux de cross-over à une valeur faible minimale de 1, ce qui impliquera toujours l'échange de deux molécules par génération. Le taux de mutation successivement appliqué au cours des 4 séries suivantes sera de 2, 4, 6 et 8. Le nombre de combinaisons de type $A_xB_yM_z$ échangées à chaque boucle évolutive ne différera que d'une unité par rapport aux simulations menées à propos de l'influence du taux d'échange, et ce afin de rendre la comparaison valable entre les deux. Pour cette raison également, nous partons de la même génération mère fixe, bien que définie aléatoirement.

Lors de la première série de 10 évolutions simulées avec $RM = 2$, nous obtenons des évolutions qui surviennent par à-coups, comme le présente la figure 48. Deux simulations sont présentées ensemble.

simulation n° 84 (RM = 2 ; RE = 1)



simulation n° 90 (RM = 2 ; RE = 1)

**figure 48**

On constate que les fitness moyennes progressent, excepté en fin de processus où la mutation associée à un faible taux d'échange éprouve des difficultés à générer des combinaisons neuves plus performantes que celles déjà présentes au sein de la génération. Au niveau des graphes exhibant le nombre de systèmes découverts du top 10, on observera les profils plutôt monotones, ainsi qu'un accès très restreint à ce top 10.

La série dédiée à $RM = 4$ possède des profils relativement semblables à la précédente. La troisième série ($RM = 6$) se détache en montrant des profils

typiques de l'introduction de plus d'éléments nouveaux au sein de la succession des générations, comme le montre la figure 49. La dernière série (RM = 8) ressemble assez à la troisième.

simulation n° 107 (RM = 6 ; RE = 1)

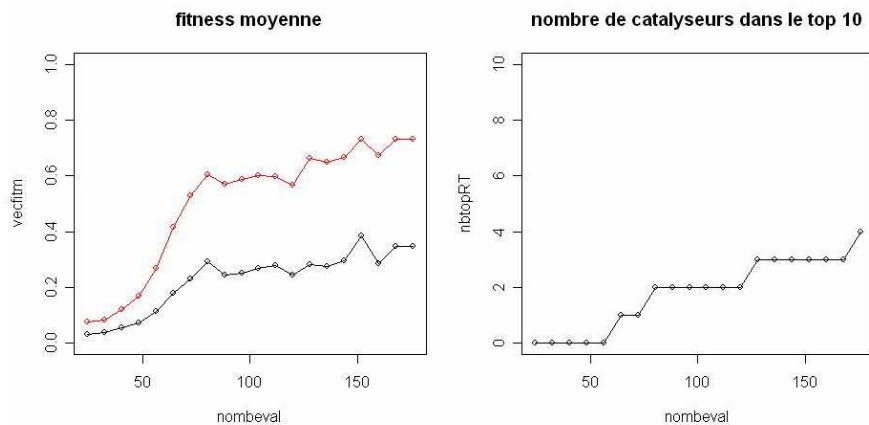


figure 49

De même que précédemment, le tableau 36 rassemble les données collectées pour chaque série.

tableau 36

série	RM	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
1	2	1,4	53	29	2	1
2	4	1,4	94	41	1	2
3	6	2,9	72	59	3	1
4	8	2,4	101	80	2	2

Le changement graduel du taux de mutation s'avère d'une influence plus faible que lorsqu'on examine une variation du taux de cross-over, ce qui était attendu. On se souviendra que si une mutation fait apparaître à l'occasion une caractéristique rendant le catalyseur plus performant, celle-ci peut-être transmise à la génération suivante, mais ne fera jamais que rarement l'objet d'échange. Sa promotion est donc gravement handicapée. La dernière série quadruple le taux de mutation par rapport à la première. Le nombre moyen de catalyseurs du top 10 auquel on accède alors est environ double. La

rapidité moyenne des processus évolutifs semble ralentie. En réalité, les σ_r qui leur sont associés démontrent surtout la variabilité des performances. La dernière série en est un bon exemple puisque deux simulations qui ne découvrent aucun des 10 meilleurs catalyseurs coexistent avec une évolution menant à 7 membres du top 10.

5.6.3.2.6 Evolutions simulées à paramètre préféré

Les variations de taux de cross-over et de mutation ont un effet que nous venons d'apprécier ; les diverses expériences de simulation nous apprennent que les échanges par cross-over entre membres des générations ont une influence de plus de poids que la mutation, même si la quantification précise en est impossible ici. Nous allons conduire les simulations suivantes en prenant premièrement un RE = 4 et RM = 2, suivi de RE = 2 et RM = 4, un paramètre est donc préféré sans que l'autre soit réduit à son minimum. Nous supposons que de tels taux pourraient conduire à peut-être retrouver plus de 8 des meilleurs systèmes catalytiques, ce qui implique de travailler sur une génération de départ (à nouveau déterminée aléatoirement, mais fixe pour les diverses simulations) un peu plus large (30 membres) afin de respecter les conditions qu'impose l'algorithme concernant la taille de l'ensemble des combinaisons potentiellement remplaçables après cross-over et mutation. Le nombre de générations reste déterminé à 20.

Les premières expériences impliquent un taux de cross-over supérieur à la mutation, et comme nous l'examinerons dans le tableau récapitulatif, les résultats sont apparemment excellents puisque l'on parvient à retrouver en moyenne 6 des meilleurs catalyseurs après 20 générations et qu'à chaque fois, le meilleur catalyseur est mis à jour. La figure 50 ci-dessous donne le profil type de cette série.

simulation n° 126 (RE = 4 ; RM = 2)

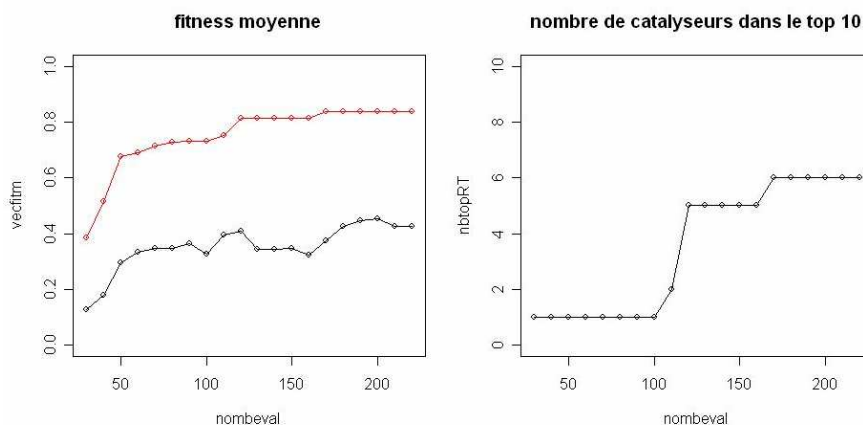


figure 50

Le fait que nous retrouvions invariablement le meilleur catalyseur nous interpella sur la nature de la génération mère. Comme nous le voyons par le graphe de droite de la figure ci-dessus, cette dernière comporte déjà un des systèmes catalytiques les plus performants, nous démarrons avec 1 candidat du top 10. L'examen attentif montrera que G0 dispose en plus du 14^{ème} meilleur catalyseur ainsi que d'une autre combinaison à haute performance. Il est indéniable qu'ainsi avantage, l'algorithme cheminait plus aisément vers « le sommet » de notre bibliothèque. Cela invitait assez naturellement à vérifier d'abord l'influence de ces présences chanceuses au sein de G0 en remplaçant ces trois molécules par d'autres, déterminées aléatoirement mais dont on était certain qu'elles ne faisaient pas partie de l'élite de notre bibliothèque.

Logiquement nous constatons que l'algorithme partant de cette nouvelle génération mère voit son efficacité tempérée. Les données moyennes des deux séries sont rassemblées dans le tableau 37; RE = 4 et RM = 2. Chaque série comprend 10 simulations.

tableau 37

série	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
1	5,8	31	6	10	0
2	2,4	100	73	3	2

La seconde série qui correspond aux simulations basées sur une génération mère exempte de candidats hautement performants, montre une efficacité plus faible. L'algorithme est en moyenne plus lent et l'écart-type rend ici encore compte de la grande variabilité du résultat final.

Observons maintenant le comportement du processus évolutif ayant pour départ la même génération que la série 2 précédente, avec maintenant une préférence en faveur de la mutation. Les taux appliqués sont ici de RE = 2 et RM = 4. Cette troisième série de 10 simulations est présentée par le tableau 38 suivant.

tableau 38

série	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
3	2,6	79	43	3	0

Si ce n'est une rapidité apparemment un peu supérieure, puisque 79 tests sont en moyenne requis pour accéder à un des catalyseurs du top 10 contre 100 précédemment, cette troisième série est fort comparable à la seconde au vu de ses performances.

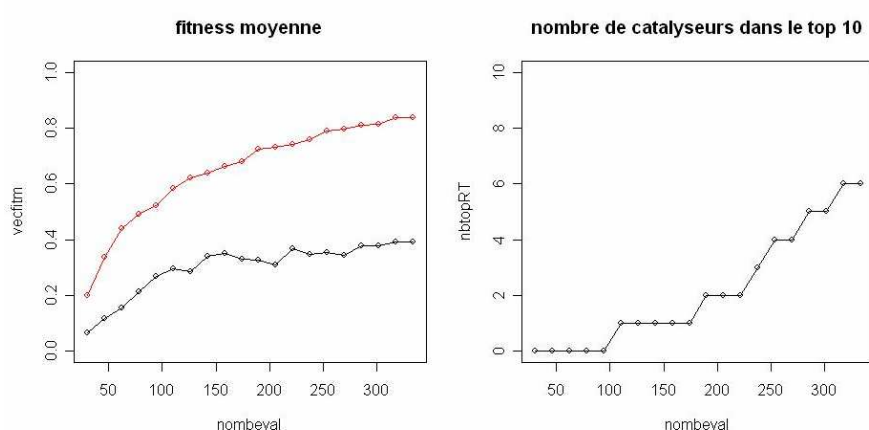
5.6.3.2.7 Valeurs de cross-over et mutation élevées

Est-il possible de réaliser efficacement une évolution simulée si les cross-over et les mutations sont fixés à des valeurs assez élevées de RE = 4 et RM = 8. Le danger est que certaines molécules arrivent par exemple dans la génération par mutation et ne puissent transmettre une de leurs caractéristiques potentiellement intéressantes par échange avant de subir à nouveau une mutation, peut-être à cette position justement. La perturbation beaucoup plus importante imposée à chaque boucle évolutive compensera-t-elle cela ? Il suffit pour s'en convaincre d'observer le profil typique que nous obtenons sous pareilles conditions (figure 51), et de lire les données du tableau 39 relatives à cette quatrième série de 10 nouvelles simulations.

tableau 39

série	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
4	5,7	59	22	7	0

simulation n° 164 (RE = 4 ; RM = 8)

**figure 51**

Nos craintes étaient manifestement infondées puisque les évolutions sont particulièrement performantes au vu du nombre moyen de candidats du top 10 que l'algorithme parvient à découvrir. Notons que la rapidité calculée est agrémentée d'un σ_r témoignant d'une certaine similitude entre les différents processus évolutifs. De surcroît, une évolution assez remarquable est apparue puisque pas moins de 9 des 10 meilleurs systèmes catalytiques sont retrouvés, ainsi que le montre la figure 52. On notera que la fitness moyenne des générations reste faible et particulièrement chaotique en fin de processus, ce qui rend compte de l'introduction importante de diversité génération après génération.

simulation n° 168 (RE = 4 ; RM = 8)

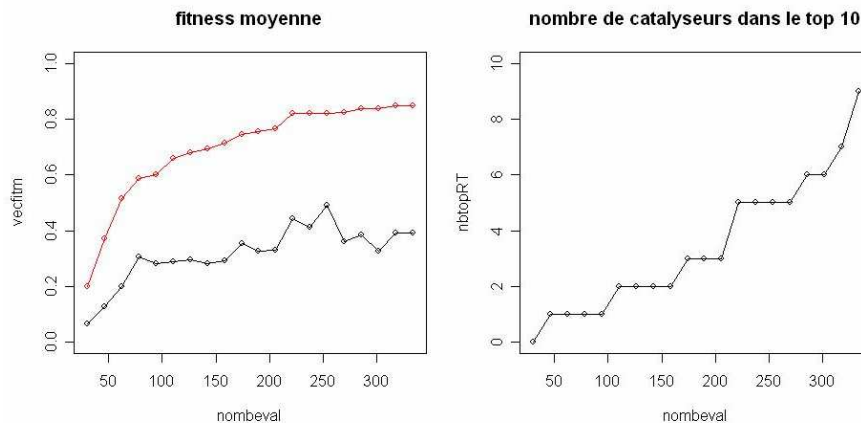


figure 52

5.6.3.2.8 Variation de la taille de l'ensemble de remplacement « RR »

L'ensemble de remplacement « RR » est encore un paramètre qu'il nous est loisible de changer. Il définit le nombre de membres d'une génération que l'on accepte de soumettre à une probabilité d'éviction lors de l'introduction des nouvelles combinaisons générées après les cross-over et mutations. Comme nous l'avons précisé, la probabilité pour qu'un membre soit remplacé par une nouvelle combinaison dépend premièrement de sa présence ou non dans cet ensemble et deuxièmement, est inversement proportionnelle à sa fitness relative au sein de la génération ; cela impose que les moins bons candidats soient soumis aux probabilités de remplacement les plus élevées. Nous démarrons avec une nouvelle G0 de 24 membres tirés au sort, à l'exclusion de candidats en provenance du top 20, afin d'éviter un biais positif originel. RE = 2 et RM = 2 restent fixes pour les 30 simulations ; 6 combinaisons nouvelles sont donc injectées à chaque génération. Les RR seront de 6, 12 et 18 pour les trois séries respectives. En toute logique, RR = 6 devrait conduire aux meilleures performances évolutives moyennes, puisque dans ce cas, les 6 moins bons candidats de la génération font l'objet du remplacement, à chaque boucle. Les résultats moyens sont présentés au sein du tableau 40 ci-dessous.

tableau 40

série	RR	top moy.	rapid.	σ_r	A1B19M2	«évol. « 0 »
1	6	2,8	58	42	4	1
2	12	1,7	73	47	2	2
3	18	2,5	58	40	4	1
2 - « 0 »	6	1,9	65	42	2	1

En comparant les résultats de la première et de la seconde série, nous voyons que les évolutions simulées lorsque $RR = 6$ produisent des résultats en moyenne plus intéressants si chacun des indicateurs est examiné. On voit de plus, et cela est valable pour les autres séries également, que les valeurs de σ_r témoignent d'une assez grande variabilité des performances, mais que cette variabilité est plutôt constante, ce qui autorise une comparaison des moyennes de rapidité. La troisième série est fort semblable à la première. Cela invite à conclure que l'effet que nous attendions, à savoir une diminution des performances moyennes corrélable à l'augmentation de RR , ne se marque pas.

Deux remarques permettent d'expliquer ce phénomène. Tout d'abord, nous voyons dans la dernière colonne que la série 2 comprend 2 évolutions simulées qui ne donnent accès à aucun des catalyseurs du top 10, cela entraîne une chute importante de la performance moyenne de la série. Si nous omettons une des ces simulations inefficaces pour calculer les valeurs moyennes sur 9 simulations, nous obtenons la série 2 – « 0 » dont les valeurs diffèrent déjà nettement moins des séries 1 et 3. En second lieu, il faut se rappeler de la nature même des résultats exhibés par les tests catalytiques menés à Lyon. La figure 34 les présentait de manière ordonnée, et nous avons pu voir que 2/3 des catalyseurs étaient de piètres performances ; 17 à 18 % de notre bibliothèque donne des résultats moyens à bons. Cela implique donc qu'une génération, si elle est en moyenne représentative de la bibliothèque, possède environ 80 % de catalyseurs de qualité faible et peu différents en termes de performances. Or notre espace de remplacement est de 18/24 au maximum, soit 75 %. On comprend qu'ici, la nature des

résultats de la bibliothèque empêche probablement de noter cette diminution attendue si l'on augmente RR.

5.6.3.2.9 Variation de la taille de G0

Si nous avons au départ tenté de déterminer l'influence d'une génération mère aux membres aléatoirement sélectionnés par rapport au choix de départ qu'aurait pu poser un chimiste, il reste encore maintenant à vérifier quelle influence peut avoir la taille de la génération, sachant que naturellement, plus celle-ci est étendue, plus les chances sont grandes d'y inclure de nombreux éléments de diversité et/ou des blocs constitutifs donnant accès aux hautes performances. Ces expériences de simulation seront conduites en prenant des tailles de G0 successivement de 24, 36 et 48 membres, les générations seront dans chaque cas aléatoirement déterminées. En fonction des conclusions du point précédent, nous optons pour un RR maximum qui correspond à la taille de la génération amputée de ses 10 meilleurs membres. On laissera aussi les évolutions se poursuivre sur plus de générations, puisque nous passons à 40 boucles simulées. Les taux de cross-over et mutation sont intermédiaires et resteront invariablement établis à $RE = 2$ et $RM = 3$.

Les 10 premières simulations ($R = 24$) donnent des résultats assez prometteurs, la figure 53 donne le profil type que nous observons dans ce cas.

simulation n° 198 (G0 = 24)

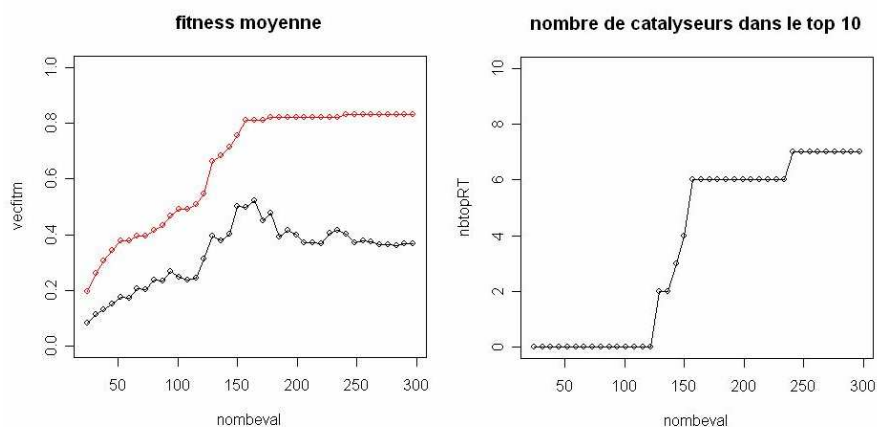


figure 53

Si l'on examine les simulations comme d'habitude en fonction des valeurs moyennes que nous pouvons en dégager, nous constatons le bon fonctionnement de l'algorithme (tableau 41). La vitesse moyenne est très appréciable, par rapport aux valeurs recueillies antérieurement ; de plus nous parvenons presque à chaque coup à accéder au meilleur catalyseur. La colonne reprenant les évolutions n'accédant à aucun des catalyseurs remarquable disparaît.

tableau 41

entrée	Série	top moy.	rapid.	σ_r	A1B19M2
1	1	6,5	40	14	9
2	1 sans top	6,3	42	12	/
3	top de 1	6,7	38	19	/

Au sein de cette série nous avons constaté que certaines simulations démarraient en disposant d'un des catalyseurs du top 10, ce qui pourrait être un biais favorable comme nous l'avons discuté plus haut. Il a donc été décidé de scinder les simulations en deux groupes, dont l'un (entrée 2) rend compte des simulations ne disposant pas dans G0 d'un élément du top 10, et l'autre (entrée 3) rassemble exclusivement les simulations ayant une G0 comprenant un des éléments du top 10. Même si extraire des valeurs

statistiques de respectivement 6 et 4 expériences impose la prudence, on concédera que les valeurs prises par le nombre moyen de molécules du top 10 que l'algorithme met en lumière et sa rapidité moyenne sont fort proches. Ceci tend à montrer que les perturbations imposées à la génération mère sont suffisantes pour s'affranchir du désavantage à ne pas disposer en son sein d'un candidat du top 10.

Concernant les expériences menées avec des valeurs de R de successivement 36 et 48, leur ressemblance avec celle conduites lorsque R = 24 est inattendue. Le tableau 42 présente les valeurs moyennes des trois séries rassemblées.

tableau 42

série	R	top moy.	rapid.	σ_r	A1B19M2
1	24	6,5	40	14	9
2	36	6,5	41	9	8
3	48	6,1	48	15	7

Le nombre de catalyseurs du top 10 que l'on retrouve en moyenne varie peu, ainsi que la rapidité moyenne de l'algorithme. Les écart-types observés témoignent d'une relative homogénéité des processus évolutifs, ce que confirme également le nombre de fois où l'on retrouve le meilleur catalyseur. Dans notre cas donc, il semble acquis qu'une génération de 24 membres soit déjà suffisante pour permettre à terme de découvrir une bonne part des meilleurs systèmes. Augmenter la taille de la génération n'est plus avantageux. Il est probable toutefois que ce facteur soit d'importance si l'on réalise une évolution sur une bibliothèque plus imposante. On notera enfin, et pour achever de convaincre de l'outil performant que peut représenter un algorithme évolutif, qu'une des expériences de simulation menée avec R = 36 permet de retrouver l'entièreté des 10 meilleurs catalyseurs, comme l'illustre la figure 54 ci après. La fitness moyenne, médiocre et chaotique, témoigne de l'introduction constante de diversité au sein des générations successives.

simulation n° 208 (R = 36)

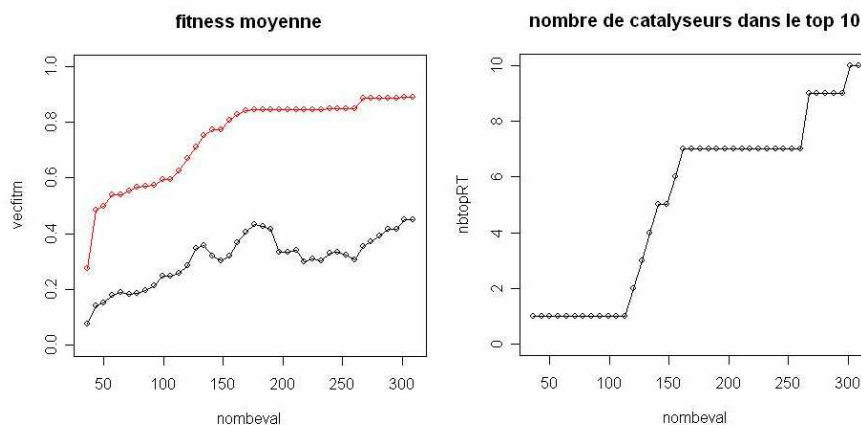


figure 54

5.6.3.2.10 Evolution simulée à double algorithme

Si nous nous plaçons maintenant en situation réelle, que l'on imagine, avoir établi notre bibliothèque virtuellement cette fois, et vouloir démarrer une optimisation grâce à un processus d'évolution soutenu par un algorithme génétique, on prend conscience de l'importance que peut revêtir la génération mère. Nous savons maintenant que la choisir aléatoirement reste un gage de succès, et que la chance peut nous sourire si d'aventure un, voire plusieurs blocs de construction intéressants, sont présents au sein de celle-ci.

Il n'est malheureusement pas impossible non plus que le hasard détermine une très mauvaise génération de base, ce qui pourrait avoir des conséquences fâcheuses sur l'optimisation ainsi menée, alors que nous n'avons droit ici *a priori* qu'à une seule chance et que nous ne sommes sensés n'avoir aucune idée des résultats que pourraient offrir les catalyseurs. On se souviendra dans notre cas que la bibliothèque possède une zone qui la compose à deux tiers et dont les performances catalytiques sont faibles si pas nulles (figure 55, graphe supérieur). De plus, l'algorithme « percevra » notre librairie en fonction de l'agencement des codes, et non des performances ordonnées croissantes, ce qui montre clairement qu'il existe nombre de

maxima locaux dont il devra être capable de s'extraire pour produire une évolution de qualité (figure 55, graphe inférieur).

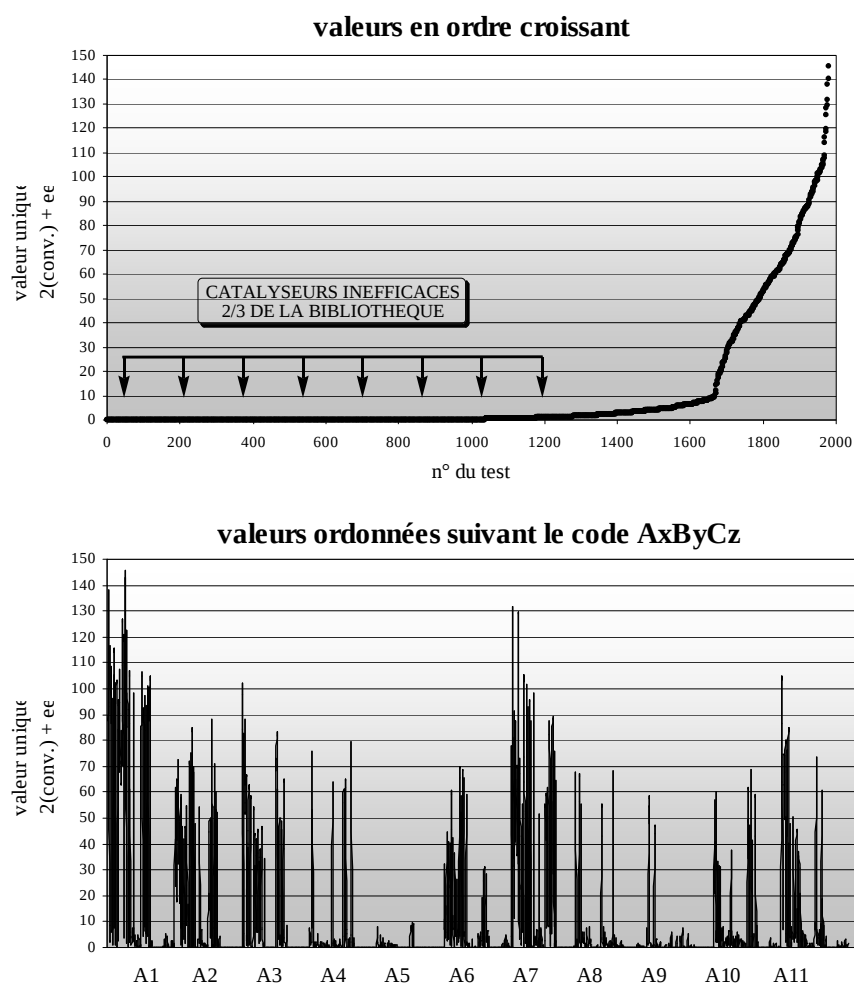


figure 55

Nous savons que maintes séries fournirent des résultats décevants, tant en termes de conversion atteintes que des excès énantiomériques, ce que rassemblent les deux premiers tiers de graphe dédié aux valeurs ordonnées. La probabilité que les 24 membres de la première génération soient intégralement extraits des ces deux premiers tiers s'évalue par l'équation suivante dans laquelle « P » représente cette probabilité, « N » le nombre

total de candidats, « m » le nombre de candidats de la partie de la librairie qui affiche de faibles performances catalytiques, et « n » le nombre de candidats tirés au sort ; ce nombre correspondra donc à G0.

$$P(x = n) = \frac{m!}{(m - n)!} \frac{(N - m)!}{N!}$$

Lorsque cette probabilité s'applique à une librairie assez large, ce dont il s'agit ici, nous pouvons alors approcher cette probabilité par l'équation ci-dessous.

$$P(x = n) = \left(\frac{m}{N} \right)^n$$

Dans notre cas, nous calculons alors une probabilité proche de 0,05 %. Si elle reste faible, nous avons déjà observé dans quelle mesure la négliger pourrait mener à réduire gravement les performances de l'algorithme.

Nous partons donc pour ces dernières simulations avec une taille de génération égale à 24 membres, RE = 2, RM = 3, RR = 14. Nous imposons également que l'algorithme mène au maximum 200 évaluations, ce qui représente à peu de chose près 10 % de notre bibliothèque. Pour corser la donne, nous imposons à l'algorithme une génération mère dont les candidats sont exclusivement choisis aléatoirement au sein des deux premiers tiers de la librairie, dont l'inefficience est connue. Autrement dit, les candidats des séries A1, A7 et A11 sont exclus de cette génération mère, dont la fitness moyenne très faible vaut 0,00566 (fitness comprise entre 0 et 1). En réalisant les simulations telles quelles, nous constatons un profil type pareil à celui-ci (figure 56).

simulation n° 224 (G0 2/3 fit.~0)

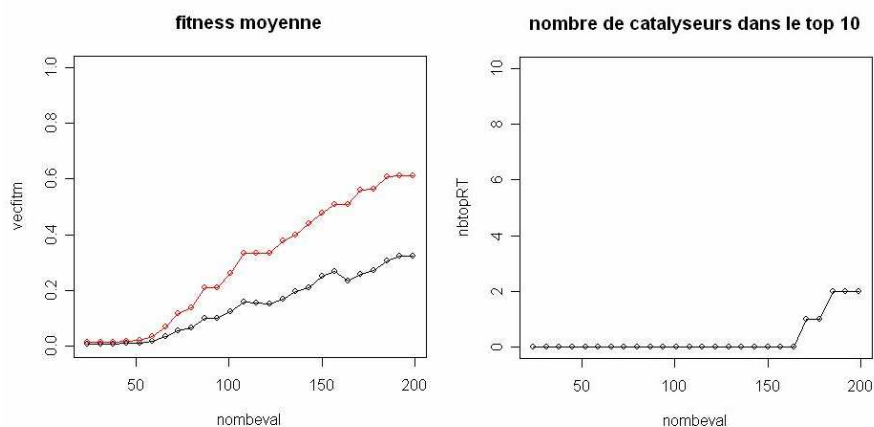


figure 56

De toute évidence, l'algorithme peine à démarrer, comme en témoigne la fitness qui stagne pendant les premières générations. Le retard se marque également dans la découverte de deux catalyseurs du top 10 en toute fin de processus évolutif, alors que 199 systèmes ont été évalués. Les données moyennes relatives à ces simulations seront précisées ci-après lors des comparaisons.

On se doute que le problème ici résulte d'un manque de diversité chimique, et il convient donc de trouver un moyen d'en introduire au plus vite dans le système évolutif. Au rang des solutions possibles, on pensera d'abord à élargir la taille de la génération mère. Passer de 24 à 48 membres initiaux, voire plus, doit conduire statistiquement à rencontrer des blocs de construction d'intérêt. Toutefois, cela va immanquablement augmenter le nombre d'évaluations et nous faire passer outre des 10 % imposés. Il serait également envisageable de pratiquer des taux de cross-over et de mutation élevés ; l'efficacité de la méthode fut discutée auparavant. Ici encore les expériences nécessaires sont d'environ 340 et dépassent le seuil fixé. Nous devrions normalement trouver dans ces conditions en moyenne 3 des candidats du top 10 en 200 expériences, ce qui offre un avantage faible comme nous le verrons bientôt.

Une solution qui permet de satisfaire à la contrainte des 10 % est de pratiquer une évolution par algorithme double. Nous partons de la même génération fixe aléatoirement choisie au sein de la zone de faiblesse. Nous appliquons alors un algorithme à caractère de mutation, uniquement, pendant la première génération. Cela signifie que les 10 candidats les meilleurs de la génération mère seront sauvés, et tous les autres subiront une mutation, ce qui entraînera logiquement l'inclusion de plus de diversité chimique. En terme statistique, la probabilité de retrouver après cette étape à très haut taux de mutation tous les candidats de G1 comme faisant partie intégrante de la zone de très faible efficacité de notre bibliothèque, passe maintenant à approximativement 0,0005 %, ce qui est beaucoup plus acceptable. Une seule étape évolutive à fort taux de mutation permet ainsi de diminuer d'un facteur 100 cette probabilité potentiellement porteuse d'insuccès. On poursuit ensuite l'évolution en réinjectant les 24 membres de G1 dans un algorithme aux taux de cross-over et de mutation identiques à ceux qui furent utilisés juste avant, soient $RE = 2$ et $RM = 3$, et l'on termine ce dernier après 23 générations, ce qui équivaut à 199 évaluations au total des deux algorithmes successifs. Comme nous l'observons directement dans les profils, cette étape est bénéfique à la poursuite de l'évolution sous un jour plus favorable (figure 57).

simulation n° 238 (G0 2/3 fit.~0)

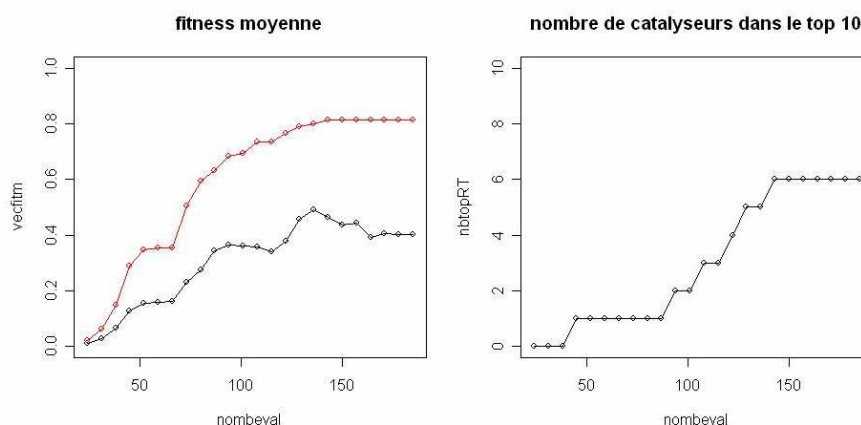


figure 57

L'introduction de nouveauté sous l'impulsion de mutations massives semble suffisante pour assurer le bon déroulement de l'évolution. Si nous comparons maintenant les données moyennes, comme à notre habitude (tableau 43), cela ressort sans équivoque.

tableau 43

série	G0	top moy.	rapid.	σ_r	A1B19M2
1	G0 2/3 fit. ≈ 0	2,1	99	55	1
2	G0 2/3 fit. ≈ 0	5,2	36	9	7
3	G0 aléa	6,5	40	14	9

La première série correspond aux expériences menées avec une mauvaise génération mère et l'algorithme commun. Les valeurs moyennes de la série 2 correspondant à l'évolution menée par l'algorithme double montrent une nette progression dans tous les domaines. De plus l'écart-type associé à la rapidité moyenne se resserre nettement, ce qui prouve la régularité entre les différentes simulations. Enfin, la troisième série correspond aux données obtenues précédemment et relatives au même algorithme partant d'une génération mère aléatoirement déterminée et à chaque fois différente (voir « variation de la taille de G0 »). Les performances moyennes des séries 2 et 3 comparées témoignent de faibles différences, le retard de la série 2 étant assez léger, ce qui démontre l'efficacité de notre solution. De plus on se rappellera que la série 3 bénéficiait dans 4 cas sur 10 d'un des éléments du top 10 dès sa première génération.

5.7 Conclusions

Dans le cadre de ce 5^{ème} chapitre, nous avons proposé et synthétisé une bibliothèque de catalyseurs composée par l'association de ligands amino-alcools et de précurseurs métalliques. Cette bibliothèque fit ensuite l'objet d'une évaluation par le biais de tests à haut débit, et nous avons pris soin de décrire le comportement de chacune des 11 séries de cet ensemble, pour ensuite identifier les 20 meilleurs catalyseurs et leurs performances associées. Notre but était de simuler une évolution sur cette bibliothèque, et

nous avons choisi de travailler avec un algorithme génétique plutôt qu'avec un réseau de neurones pour des raisons de simplicité et de temps que nous avons décrites antérieurement.

La méthode classique d'optimisation d'un catalyseur par exemple consiste à en réaliser le design, pour ensuite le synthétiser et finalement le tester. En fonction du résultat que le chimiste tentera d'exploiter au mieux pour une interprétation, on proposera une nouvelle structure, ou des conditions de travail différentes qui permettront sans doute d'accéder au produit désiré de manière plus efficace, plus rapide, ou encore avec une meilleure chémo-, régio- ou énantiosélectivité. L'optimisation peut aussi se réaliser grâce aux techniques de tests à haut débit ; la différence peut être décrite schématiquement comme suit (figure 58).

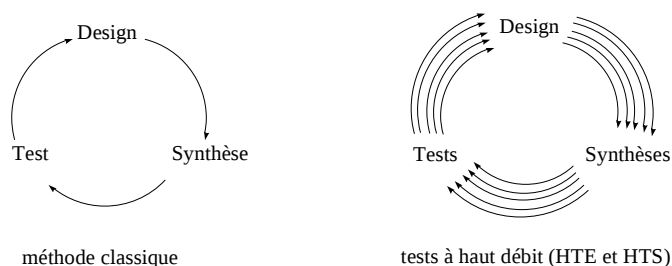


figure 58

Les techniques de tests à haut débit permettent d'exploiter la méthode classique de manière multiple. De ce fait, il est possible par exemple de synthétiser et de tester beaucoup plus rapidement un nombre plus important de catalyseurs. L'avantage immédiat de cette méthode, outre le gain de temps, est de fournir plus de données, et de permettre par leur analyse que l'étape entre les tests et le design des nouveaux catalyseurs s'opère sur des bases plus solides, ce qui doit normalement conduire à une optimisation efficace.

En nous concentrant maintenant sur une évaluation de la méthode d'évolution simulée par algorithme génétique, comparée aux techniques de tests à haut débit, nous noterons plusieurs points :

- L'évolution simulée intervient précisément entre l'étape de test et de design de la nouvelle bibliothèque, fut-ce alors une génération de 24 membres comme dans notre cas. Elle permet à cette étape de se dérouler sur base non plus d'une réflexion, mais d'une pression évolutive qui conserve sa part au caractère aléatoire ; nous avons mesuré l'importance qu'elle peut revêtir.
- Lorsque la méthode classique est opposée à la technique du haut débit, nous avons souligné que le gain de temps pour réaliser un même nombre de tests était évidemment à l'avantage du haut débit. Quant à l'évolution simulée, elle amplifie encore ce gain de temps, puisque le nombre de tests requis pour accéder aux meilleurs systèmes catalytiques se restreint considérablement.
- Il faut aussi considérer le coût de la méthode. Le haut débit permet de travailler sur des petites quantités sans difficultés. Ce bénéfice peut être repris au compte de la méthode évolutive, si l'on utilise les mêmes conditions de travail qu'en HTS, qu'avantage ici aussi le faible nombre de tests requis. On peut pousser la réflexion plus avant, en imaginant que si l'évolution avait été menée dans un cadre réel au laboratoire, il eut été fait usage de carrousel Radley® pour y conduire les réactions. S'il faut admettre que les quantités de réactifs habituellement utilisées sont plus importantes que lorsque l'on travaille en HTS, l'investissement en matériel est par contre drastiquement réduit puisqu'un automate n'est plus requis ; sans compter la part importante parfois que peuvent prendre les consommables au budget d'une campagne de tests à haut débit. Précisons finalement que l'algorithme génétique a été écrit en usant d'un logiciel téléchargeable librement sur internet¹⁹⁶. Nous jugeons comme un des atouts majeurs son caractère financièrement plus léger, surtout dans le cadre d'une application industrielle.
- Nous avons précisé que les techniques de haut débit donnent accès rapidement à un grand nombre d'informations, et ceci est *a priori* un avantage par rapport à l'évolution simulée. Toutefois, la nature même

de l'information récupérée par le biais du haut débit ou de l'évolution simulée est sensiblement différente. En effet, dans le premier cas, le chimiste reçoit un grand nombre de données potentiellement très intéressantes, mais qu'il est nécessaire de trier d'abord, et d'interpréter correctement si l'on veut que le design consécutif soit de qualité. L'évolution simulée donne bien moins d'informations, mais fait d'une certaine manière le tri en se dirigeant toujours vers l'optimum. Ainsi peut-on considérer qu'une information chimique importante, liée à une bonne performance du catalyseur, sera conservée aussi longtemps qu'aucune autre meilleure information ne la remplacera. L'évolution que nous avons réalisée prouve parfaitement que les paramètres significatifs pour ce type de catalyse ont été mis à jour *au sein* de notre bibliothèque. Il est donc possible que soient découverts des indices sur le mécanisme de la réaction par le biais du chemin qu'à suivi l'évolution et la modification des performances qui y est reliée.

- Un des reproches couramment fait aux processus évolutifs tels que celui que nous avons employé est l'incertitude de découvrir le meilleur candidat au sein de la bibliothèque, ce que permet assurément le haut débit. Nous considérons toutefois ce désavantage partiellement justifiable au moins par le fait que l'évolution n'a pas pour but premier de mettre au jour la meilleure molécule ou le meilleur catalyseur d'une bibliothèque, mais d'identifier une structure à haut potentiel dont il serait avisé de tester l'ensemble de la famille. Le but avoué de l'évolution est bien d'accélérer de manière importante la répétition des étapes « tests → design ». Nous avons pu constater cependant au cours des différentes évolutions qu'il était fréquent de retrouver le meilleur catalyseur. De surcroît la méthode à algorithme double simulant le cas où nous partons d'une piètre génération mère permet de retrouver dans 7 cas sur 10 le meilleur des catalyseurs, ainsi que 5 à 6 des candidats du top 10, ce qui peut être considéré comme une bonne performance eu égard aux 10 % seulement de la bibliothèque que nous testons. Cette méthode à double algorithme appliquée en début d'évolution entraîne une diminution drastique de la

probabilité d'un mauvais départ, et offre alors des performances presque'équivalentes à un choix aléatoire de génération mère convenable.

- L'évolution simulée peut aussi être utilisée pour vérifier si une bibliothèque a une chance réelle de conduire à de bonnes performances de ses candidats. En effet, si l'expérimentateur observe que le processus évolutif se révèle en quasi stagnation, même après une ou deux étapes à fort taux de mutation, dont on sait qu'elles réduisent nettement la probabilité de se retrouver avec de mauvais candidats uniquement, il y aura alors fort à parier que la bibliothèque n'est pas ou peu adaptée à offrir de hautes performances par rapport à la réaction par exemple que l'on étudie.
- Soulignons pour finir que l'on aurait pu coder plus que nos structures, et par exemple la température de réaction, le solvant, le taux catalytique, etc... Ceci aurait pour effet d'agrandir l'espace de travail sans beaucoup compliquer la tâche de l'expérimentateur. Il faut par contre admettre qu'une trop grande diversité chimique rend la bibliothèque de plus en plus difficile à coder, et qu'une certaine similitude entre les candidats est requise pour l'évolution, puisque les candidats sont supposés pouvoir échanger leurs paramètres structuraux et réactionnels. Travailler avec des techniques à haut débit permet, il est vrai, de s'affranchir de cette exigence.
- Au terme de ce chapitre, nous observons donc les nombreux avantages de l'évolution simulée pour mener à bien une optimisation. L'expérimentateur a la liberté d'influer sur maints paramètres de l'algorithme dont les plus importants sont sans doute les taux de cross-over et de mutation. Il définira ces paramètres en fonction surtout du nombre de tests qu'il s'autorise à réaliser et donc, du pourcentage de l'ensemble de la bibliothèque virtuelle à investiguer.