

4 OPTIMISATION DIRIGEE D'UNE FAMILLE DE LIGANDS

Le chapitre précédent s'est efforcé de décrire les possibilités de prédiction qui sont actuellement disponibles pour les chimistes. Si certaines peuvent s'avérer précises lorsque les molécules examinées sont structuralement fort proches, on en relève bon nombre encore où les erreurs impliquées dans la prédiction d'ee notamment, voisinent avec les 10 %, quand ce n'est plus. Cette marge d'imprécision nous porte à croire que ces techniques méritent d'une part le soin supplémentaire de développeurs informatiques pour s'affranchir de la difficulté inhérente à la diversité chimique au sein d'une bibliothèque et d'autre part signifie que les méthodes servant à fournir des valeurs, autrement dit les descripteurs moléculaires, nécessitent moins d'approximations, et partant, offrent des données plus fiables encore en termes d'exactitude et de précision. Signalons aussi qu'un laboratoire qui se lance dans l'aventure de techniques de prédiction devra bien entendu prévoir l'achat de programmes *ad hoc* et la formation de son personnel, ainsi que s'accommoder des temps de calculs pour obtenir les données requises, point sur lequel nous reviendrons au chapitre 5, et qui, parfois, peut s'avérer prohibitif.

Pour toutes ces raisons, il nous apparaît qu'envisager la mise en lumière de structures en couplant les techniques HTE et HTS à un principe d'optimisation dirigée reste une manière particulièrement élégante d'arriver à ses fins. Cela a en outre l'avantage que chaque étape d'optimisation se confronte à l'expérience et use des résultats que cette dernière fournit, ce qui logiquement permet d'éviter certaines incohérences. Le choix s'est porté sur un système d'optimisation dirigé, en raison de la simplicité inhérente à sa mise en œuvre, et dont l'équipe du renommé Professeur Reetz continue également à tirer avantage^{149, 150} sous forme d'évolution dirigée, cette fois. Le nôtre croisera utilement l'optimisation dirigée à des opérateurs dits génétiques, et se voudra donc en quelque sorte hybride, ou mixte.

4.1 Cadre de l'optimisation dirigée

Le début de ce travail de thèse rentrait dans un projet déjà en cours que notre laboratoire développait en commun avec l'équipe du Dr Claude de Bellefon¹⁸⁸, directeur de recherche du Laboratoire de Génie des Procédés Catalytiques, de l'Ecole de Chimie Physique Electronique, au sein de l'Université Claude Bernard, à Lyon.

Préalablement, cette optimisation dirigée avait débuté par un criblage par l'équipe lyonnaise d'une bibliothèque de quelques-uns de nos ligands, de structures variées, dont on évaluait les performances par leurs aptitudes à catalyser la réduction asymétrique de l'acétophénone, test bien connu dont rend compte le schéma 94. Cette précaution nous garantissait normalement d'éviter la mauvaise surprise qu'une partie importante des structures révèle une activité inexploitable.

La réduction se déroule sous l'intervention d'un métal associé à un ligand chiral, formant ainsi le catalyseur. Ici, plusieurs métaux seront également évalués face à chaque ligand.

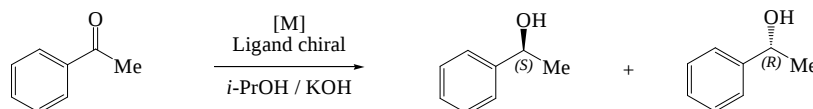


schéma 94

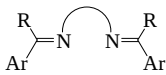
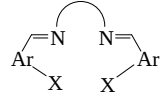
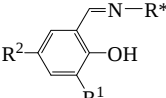
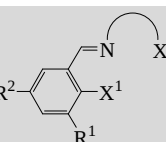
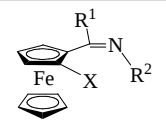
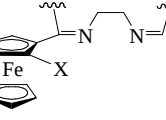
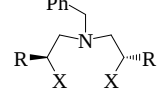
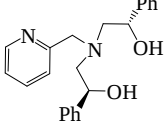
La première génération de ligands fut construite en suivant les principes suivants :

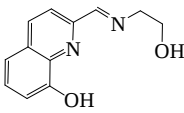
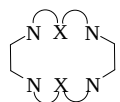
- Les structures des ligands se voulaient représentatives de la plus grande diversité possible.
- Ces structures furent imaginées en tenant compte des paramètres électroniques et d'encombrement stérique, d'une importance capitale en catalyse homogène.
- La disponibilité des sous-structures constituant les ligands et/ou leur synthèse aisée.

- L'incorporation de quelques ligands aux performances antérieurement décrites dans la littérature jalonnait les tests de valeurs qui, eux, nous permettraient de critiquer celles fraîchement obtenues.

La première génération (G1) de 60 ligands fut ainsi dessinée en suivant les 4 principes décrits ci-dessus et en tenant compte des molécules disponibles au laboratoire. Le tableau 2 donne les structures-type de G1, divisée en a-k familles.

tableau 2

Famille	Classification	Structure	Nbre de L*
F _a *	bidentate / C ₂ / structure diimine		8
F _b ‡	tetradentate / C ₂ / structure diimine		6
F _c	bidentate / structure salicylidène		2
F _d #	tridentate / structure salicylidène		19
F _e #	bidentate / structure ferrocène chirale		3
F _f #	tridentate / structure ferrocène chirale		3
F _g #	tridentate / C ₂		3
F _h	tetradentate / C ₂		2

F_i	tetradentate / structure hydroquinoline		2
$F_j^{\#}$	modes de coordination variés / C_2		9
F_k	symétries et structures variées	/	3

* = Groupe aromatique sans hétéroatome coordonnant. ‡ = Groupe aromatique avec hétéroatome coordonnant. $^{\#}$ X = atome ou groupe coordonnant (=N-; NH_2 ; NH-R; NH-Ts; OH; PPh_2 ; S; SO)

Nous expliquerons par la suite le procédé d'optimisation retenu dans cette étude, appliqué successivement sur 3 générations, et qui révéla la famille F_d comme particulièrement intéressante de par les performances exhibées. Nous nous sommes donc attelés à réaliser la synthèse de nouveaux ligands proches de la structure générique de cette famille, jugeant que de nouvelles variations produiraient des analogues potentiellement intéressants et capable d'encore augmenter les ee et/ou les rendements déjà obtenus jusqu'à présent.

4.2 Synthèse de ligands supplémentaires de la famille « F_d », 4^{ième} Génération

Ainsi que le montre la structure générique de la famille F_d (schéma 95), ces ligands possèdent un lien imine, lequel résulte évidemment du couplage entre un aldéhyde et une amine utilisés pour leur synthèse. Ces différentes amines et aldéhydes, ainsi que leurs synthèses éventuelles seront décrits ci-après.

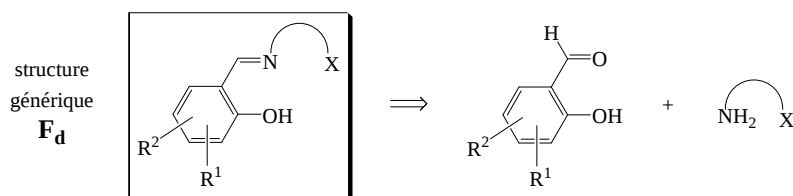


schéma 95

4.2.1 Amines

Deux types d'amines ont été choisis ici comme le montre le schéma 96.

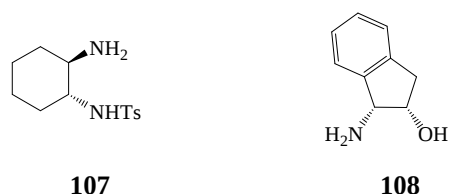


schéma 96

Si l'aminoindanol **108** est disponible commercialement, **107** s'obtient par réaction de tosylation du sel de cyclohexanediammonium correspondant¹⁸⁹, comme l'indique le schéma 97, avec un rendement de 98 %.

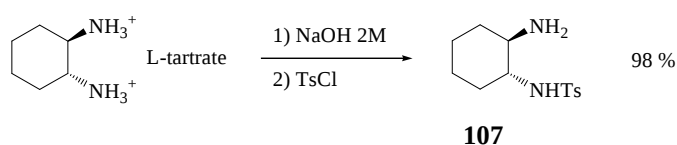


schéma 97

4.2.2 Aldéhydes

Puisque nous avons convenu d'introduire de nouvelles modifications et que seules deux amines seront utilisées pour générer ces nouveaux ligands, il convenait d'avoir en notre possession divers aldéhydes. Certains seront achetés auprès des fournisseurs habituels, d'autres auront été synthétisés par nos soins, comme le présente la suite.

Le salicylaldéhyde **109** et l'orthovaniline **110** firent l'objet d'une nitration (schéma 98). La procédure de nitration aromatique est ici appliquée, nécessitant un mélange AcOH / HNO₃. Un bon contrôle de la température (12-15 °C ici) permet de nitrer préférentiellement la position choisie, et minimise la polynitration. Ceci s'avère particulièrement important pour la nitration du salicylaldéhyde. Les dérivés **111** et **112** sont ensuite purifiés par recristallisation et obtenus avec des rendements respectifs de 16 et 49 %.

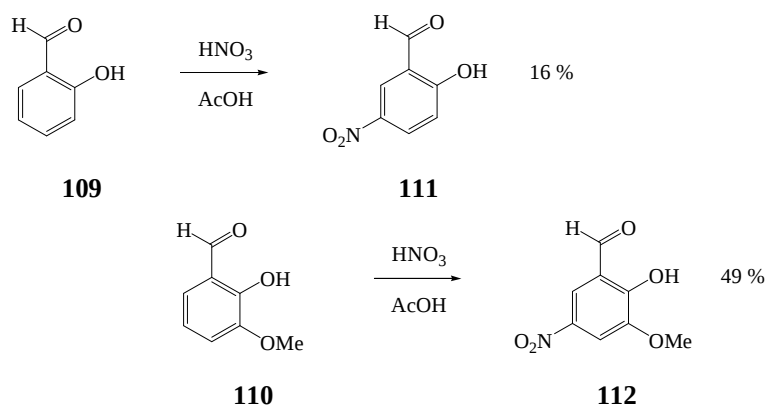


schéma 98

Nous désirions également disposer du 3,5-diiodosalicylaldéhyde **114**, accessible par la réduction d'iodation, présentée au schéma 99, et appliquée au salicylaldéhyde. **114** est obtenu avec un rendement de 9 %.

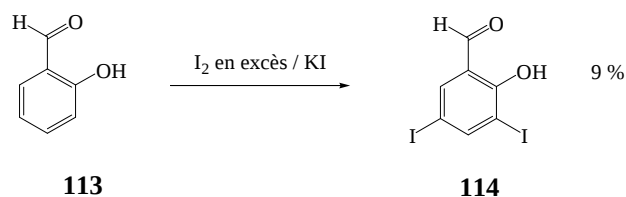


schéma 99

Afin de disposer d'un autre aldéhyde aromatique substitué en position *para* du cycle par un groupement électroattracteur, nous avons tenté une réaction de formylation sur le 4-cyanophénol **115** en vue d'obtenir l'aldéhyde correspondant **116** (schéma 100).

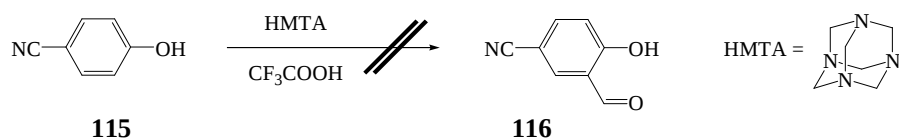


schéma 100

Malheureusement, il ne nous fut pas possible d'obtenir le salicylaldéhyde correspondant, et la faible quantité de **115** dont nous disposions à ce moment ne nous permit pas de réaliser un nouvel essai dans les temps. Notons que

cette méthode eut été normalement efficace si 2 équivalents de HMTA avaient été employés¹⁹⁰ (rendement attendu de 26 %).

Le sésamol **117** s'est aussi avéré être un candidat à la modification pour produire un aldéhyde suivant une voie presque identique, décrite au schéma 101. La réaction se déroule dans l'AcOH à reflux en seulement 30 minutes, et le salicyladéhyde modifié **118** est obtenu pur après recristallisation dans l'EtOH avec un rendement de 69 %.

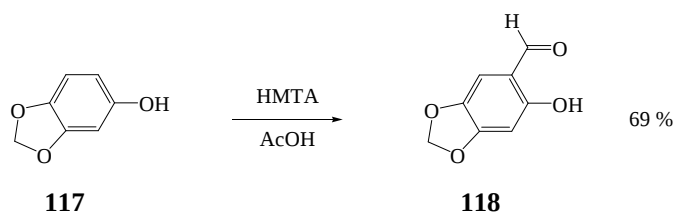


schéma 101

Notons qu'après l'insuccès rencontré lors de notre tentative de formylation du **115**, la quantité de HMTA fut portée de 1 à 10 équivalents. Toutefois, le sésamol dans ce cas s'avère être d'emblée un meilleur candidat à la formylation, puisque la partie dioxolone joue un rôle à la fois inductif capteur et surtout, mésomère donneur, favorable à la stabilisation d'une charge positive, en plus des électrons π du cycle, ce dont ne peut bénéficier **117** comme le décrit le schéma 102 suivant ($E^{\oplus}$ = électrophile).

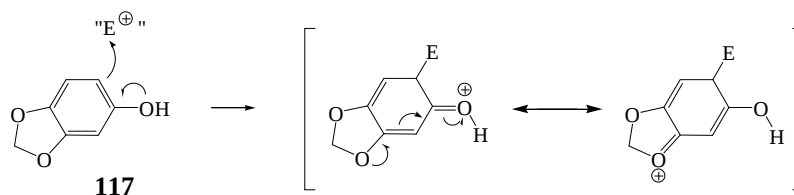


schéma 102

En partant de la nitroaniline **119**, il est possible d'accéder au sel de diazonium correspondant. Ce dernier peut alors prendre part à une substitution électrophile aromatique impliquant la salicyladéhyde, comme l'indique le schéma 103. Le composé diazo **120** s'obtient avec un rendement global de réaction de 72 %.

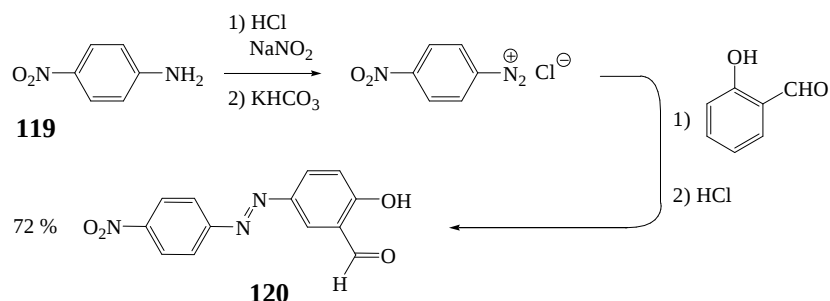


schéma 103

4.2.3 Synthèse des ligands par couplage aldéhyde-amine

Dès lors que les salicylaldéhydes modifiés et les amines étaient en notre possession, le couplage des deux types de composés permettait d'obtenir les ligands de manière aisée puisque ces réactions furent menées simplement dans un solvant à reflux pendant une nuit, avec une quantité catalytique d'AcOH. Les réactions ont pour plus de la moitié été quantitatives et n'ont pas nécessité d'étape de purification. Ces ligands font partie d'une seule famille et furent également testés par nos soins sur l'autre partie du projet de recherche (voir chapitre 6). Ceci nous a amené à les classer utilement en deux sous-familles de ligands, définies comme « Type A » et « Type B », et qui dépendent bien entendu de l'amine engagée lors de la réaction de couplage.

4.2.3.1 Ligands de type A

Ces ligands résultent de l'association entre la cyclohexyldiamine monotosylée et deux des aldéhydes que nous venons de synthétiser en plus de deux composés commerciaux, comme le décrit le schéma 104. Nous obtenons donc quatre ligands disposant d'une pince chélatante tridentate « N,N,O ».

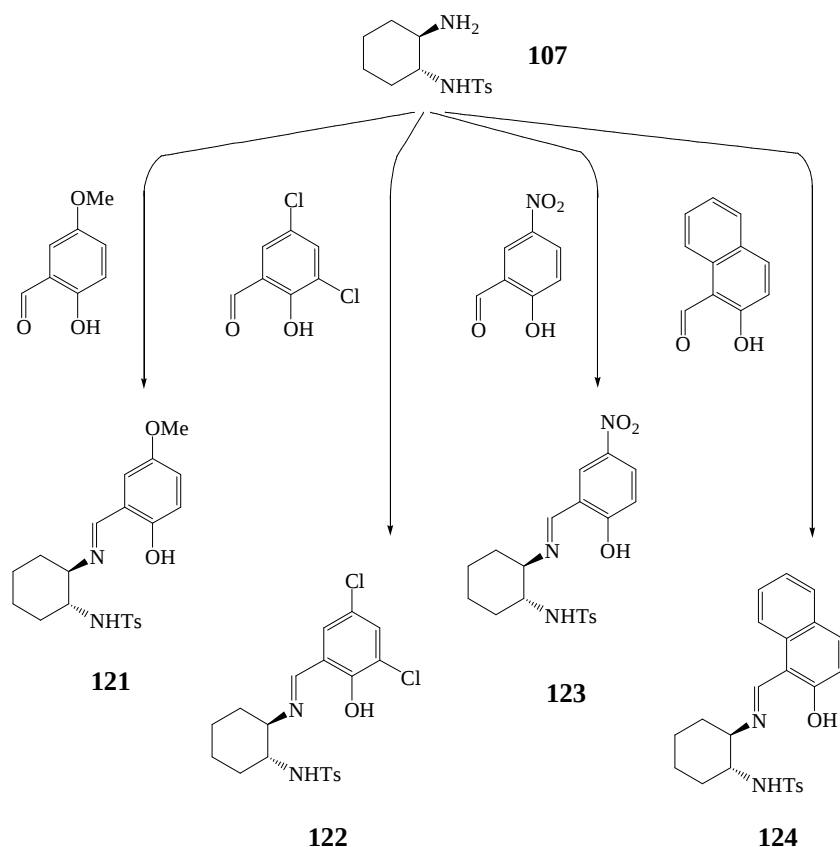


schéma 104

On remarquera particulièrement la présence d'un substituant méthoxy sur **121** et nitro porté par **123**. La nature opposée de ces groupements, électroniquement parlant, pourrait s'avérer intéressante par la suite. **122** porte deux chlores, et se montrera lui aussi plutôt électroattracteur ; toutefois la présence d'un des chlores en position *ortho* introduira un facteur d'encombrement stérique non négligeable.

4.2.3.2 Ligands de type B

Ces ligands sont eux obtenus par la mise en présence du (1*R*, 2*S*)-1-amino-2-indanol commercial et de différents aldéhydes, comme nous allons le décrire ci-dessous. Les ligands de la famille F_d étaient structuralement plus proches encore du type B que du type A ce qui nous a conduits à

produire 13 nouveaux ligands et non pas 4, gageant que les variations décidées apporteraient un bénéfice en termes de performances. Pour plus de clarté, le schéma 105 ne présente que la réaction type, suivie des nouveaux ligands dans le schéma 106.

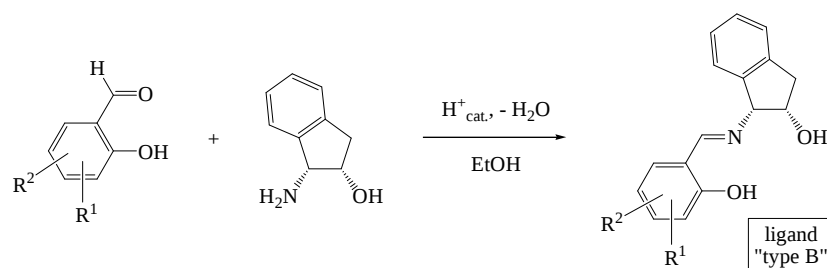


schéma 105

Presque toujours, la mise en présence des deux réactifs au sein du ballon engendrait l'apparition d'une couleur jaune vive prononcée et caractéristique endéans les premières minutes. Dans certains cas, la réaction s'avéra très rapide (5-10 minutes) et il est même arrivé de constater des prises en masse dès adjonction du second réactif, ce qui conduisit à augmenter le volume de solvant afin de permettre le reflux et mener la réaction à son terme. Les structures obtenues sont données ci-dessous (schéma 106).

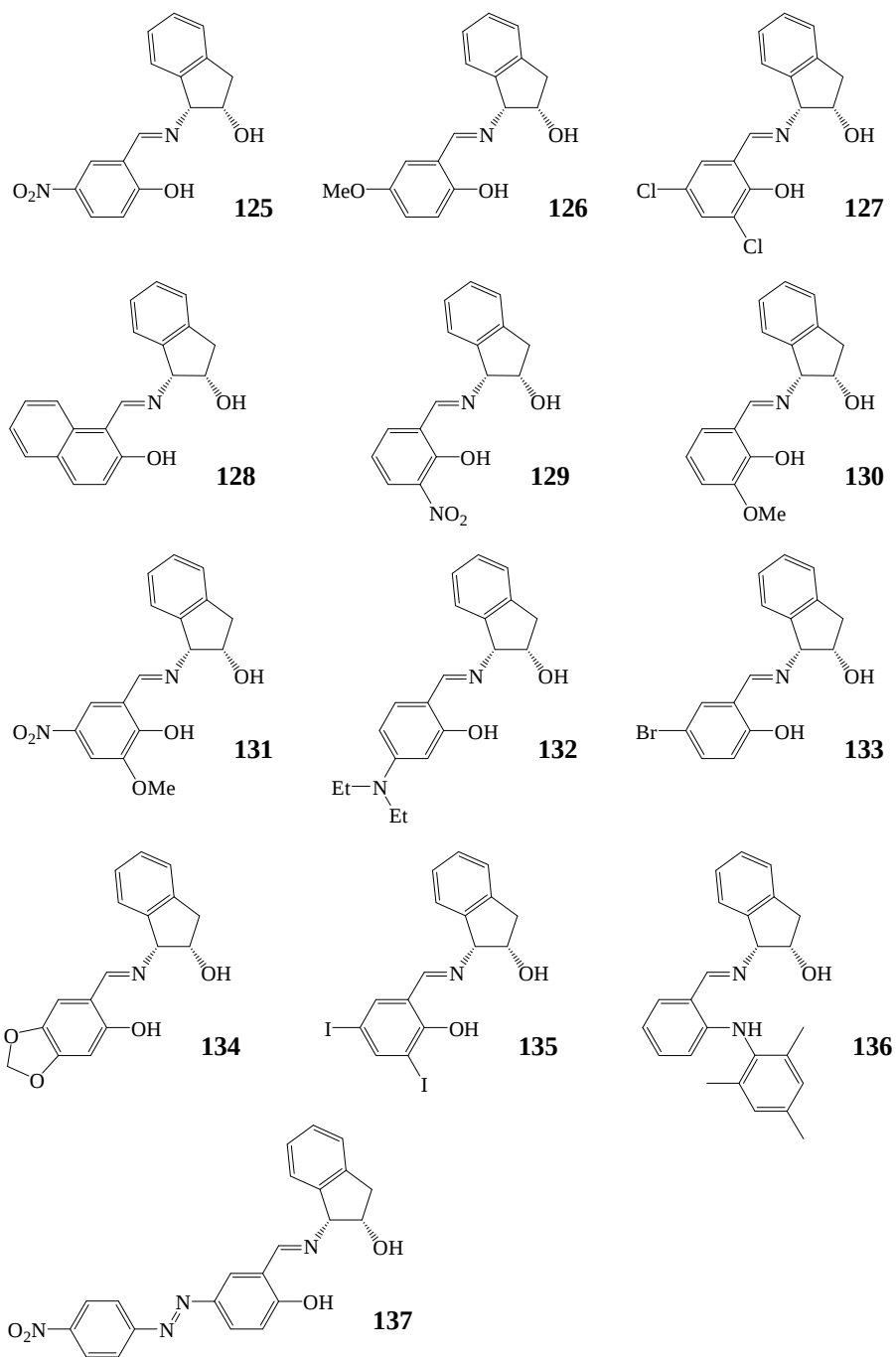


schéma 106

On notera le soin apporté à ce que ces ligands présentent une variation de groupements et de positions. Les substituants nitro et méthoxy en particulier ont été placés en positions *para* et *ortho*, ainsi que rassemblés sur un seul ligand. On notera la présence de substituants diiodé, dichloré, et dioxolone qui peut en première approximation être considéré comme un ligand portant deux méthoxy électrodonneurs. Le 2-hydroxy-1-naphtaldéhyde **128** introduit un groupement plutôt neutre électroniquement. Si nous considérons **132**, la présence d'une amine tertiaire greffée sur le cycle se justifie également dans cette séquence. Des variations plus particulières sont celles observées sur les deux derniers ligands de la série. On constatera que **136** voit son alcool phénolique remplacé par une amine secondaire stériquement encombrée, ce qui doit être compris comme un changement majeur et pourrait avoir une influence marquée lors de l'étape de catalyse, attendu que le mode de coordination a été potentiellement modifié. Le ligand **137** sera normalement marqué par la combinaison de la partie diazo et le deuxième cycle aromatique nitré.

4.3 Déroulement des tests catalytiques

Comme nous l'avons présenté au schéma 94, la réaction test utilisée est la réduction de l'acétophénone en phényléthanol ; les paramètres mesurés sont le rendement et l'ee. Le point suivant détaille la manière dont furent conduites les expériences au sein des 4 générations.

4.3.1 Conditions et paramètres réactionnels généraux

Le grand nombre de tests à réaliser requiert la mise en service de l'appareillage HTS. Celui-ci se compose d'une part d'un automate piloté informatiquement et qui permettra de mener parfois jusqu'à presque 300 évaluations par jour. Il est « seulement » nécessaire alors de préparer les solutions stocks de réactifs et les piluliers où se dérouleront les réactions. Les additions minutées des composés sont laissées à l'automate. L'analyse est réalisée par un appareil GC muni d'un passeur automatique d'échantillons. L'ensemble des expériences fut réalisée par Radwan Abdallah¹⁸⁸, ancien doctorant de l'équipe du Dr de Bellefon.

Préalablement aux tests menés de manière systématique, 4 ligands issus de 3 familles furent évalués avec 4 différentes concentrations en base et il a été décidé de restreindre l'espace de travail à un rapport $[KOH]/[M] = 13$; ce rapport apporte un bon compromis entre la conversion et l'énantiosélectivité de la réaction. Toutefois, l'effort d'évaluation fut mené dans un cadre assez large. Bien sûr chaque ligand fut testé, notamment en le mettant potentiellement en présence de plusieurs précurseurs métalliques : $[RuCl_2(p\text{-cym})]_2$, $[RhCl(cod)]_2$, $[IrCl(cod)]$ et $[Ir(acac)(cod)]$ (commandés chez Strem Chemicals). Seule la dernière famille fait exception à la règle puisque le précurseur au Ru fut uniquement employé, eu égard aux enseignements tirés des 3 générations précédentes. Ajoutons à cela que 3 séries « S » de ratio ligand / métal (L/M) et substrat sur métal (S/M) firent aussi l'objet d'une évaluation, ce que présentent le schéma 107 et le tableau 3 ci-dessous.

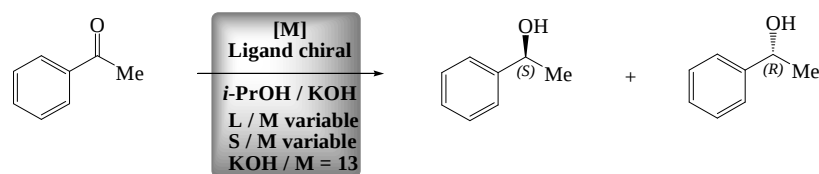


schéma 107

tableau 3

Série	L / M	S / M
S1	1	25
S2	1	100
S3	2	100

Selon les cas, deux ou trois temps de réaction furent aussi examinés : une heure, un jour ou trois jours. Le dernier réactif à être additionné était le KOH et permettait à la réaction de démarrer.

Les tests furent de surcroît réalisés en tenant compte du paramètre de reproductibilité. Cela signifie que chaque ensemble de tests fut au moins exécuté deux fois, et si l'erreur entre les deux ensembles apparaissait comme

trop importante, on resoumettait les systèmes catalytiques à une nouvelle évaluation.

L'entièreté des tests engendrés par cette manière de procéder est résumée par le tableau 4. La dernière génération est décrite par la ligne ombrée, et la dernière ligne du tableau rend compte du nombre total à la fois de ligands utilisés, et de tests effectués.

tableau 4

génération	ligands	métaux	séries	temps	reproductibilité	nombre
G1	60	3	3	2	2	2160
G2	19	4	3	2	3	1368
G3	22	4	3	3	3	2376
G4	16	1	3	1	2	96
Σ	117					6000

L'analyse des échantillons voulait examiner conjointement rendement et ee ; elle fut naturellement entreprise par chromatographie gazeuse (GC). Les paramètres de pression et de température avaient été optimisés et réglés de telle sorte que chaque système catalytique livre ses performances endéans les 3 minutes ; le temps total requis lors de la première génération pour l'examen des 60 ligands est donc de 3 heures. Cela implique que lorsque le premier flacon passe sur colonne chirale après une heure de réaction, le 60^{ème} système ne l'est qu'après 4 h au total. Pour cette raison, seule une analyse qualitative peut être envisagée dans ce cas. Lorsque les temps de réaction sont de 1 ou 3 jours, cette précaution ne s'applique par contre plus, la différence de temps pouvant alors être considérée comme d'influence mineure face aux temps réactionnels.

Comme nous pouvons le constater, 6000 échantillons précisément furent examinés, chacun correspondant à un ee et à un rendement. Attardons-nous maintenant sur les résultats de la quatrième génération de ligands « G4 » produite par nos soins.

4.3.2 Résultats « G4 »

La réaction est ici rappelée en précisant les conditions en usage (schéma 108).

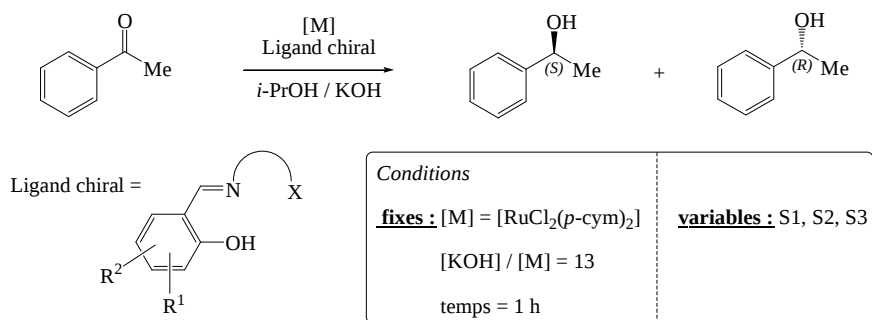


schéma 108

Les résultats obtenus sont présentés à la suite sous forme d'histogrammes rendant compte du rendement et ee produits par chaque système catalytique. Au total, 16 systèmes catalytiques sont testés, chacun dans 3 conditions différentes S1, S2 et S3. Chaque série sera évaluée deux fois en réaction de réduction afin de vérifier la reproductibilité des données, les analyses correspondantes sont nommées « GC 1 » et « GC 2 ». La convention fut prise de considérer l'ee positif lorsque prédominait l'énantiomère « R ».

Les graphes ci-dessous (figure 8) représentent la série S1 : L/M=1, S/M=25.

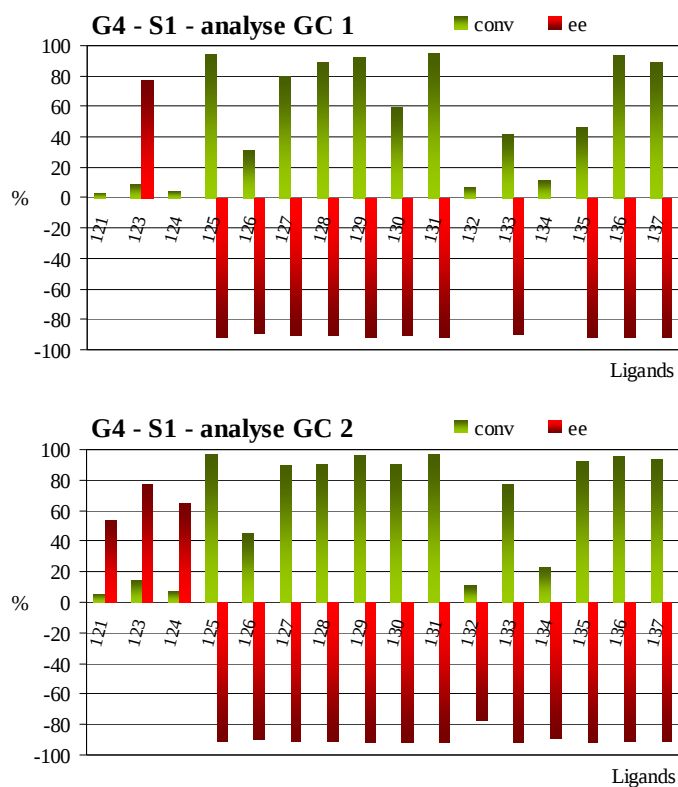


figure 8

On constate, l'absence d' ee pour les ligands **121**, **124**, **132** et **134** au sein de la première analyse. L'explication tient au fait que le rendement faible lors de l'analyse 1 rendait une mesure précise de l'ee impossible, il n'a donc été pris en compte. En ce qui concerne la deuxième analyse, un rendement légèrement supérieur permettait d'afficher une valeur d'ee digne de confiance. Pour les autres résultats, une assez bonne similitude s'observe d'emblée, ainsi qu'en témoignent les valeurs des écarts moyens observés. Ces valeurs sont obtenues en examinant la différence de comportement d'un même système catalytique soumis aux mêmes conditions de tests, et en faisant la moyenne de ces écarts éventuels.

L'écart moyen observé pour la conversion est de 11 %, ce qui peut paraître élevé. Nombre des écarts sont de l'ordre de 5 %, toutefois **130**, **133**

et **135** donnent des écarts compris entre 30 et 45 %. En ce qui concerne l'ee, l'écart moyen est de 0,5% et montre donc une très bonne reproductibilité.

Les graphes ci-dessous (figure 9) représentent maintenant S2 : L/M=1, S/M=100. Par rapport à la figure 8 nous examinons l'effet d'une variation de ratio S/M en passant de 25 à 100.

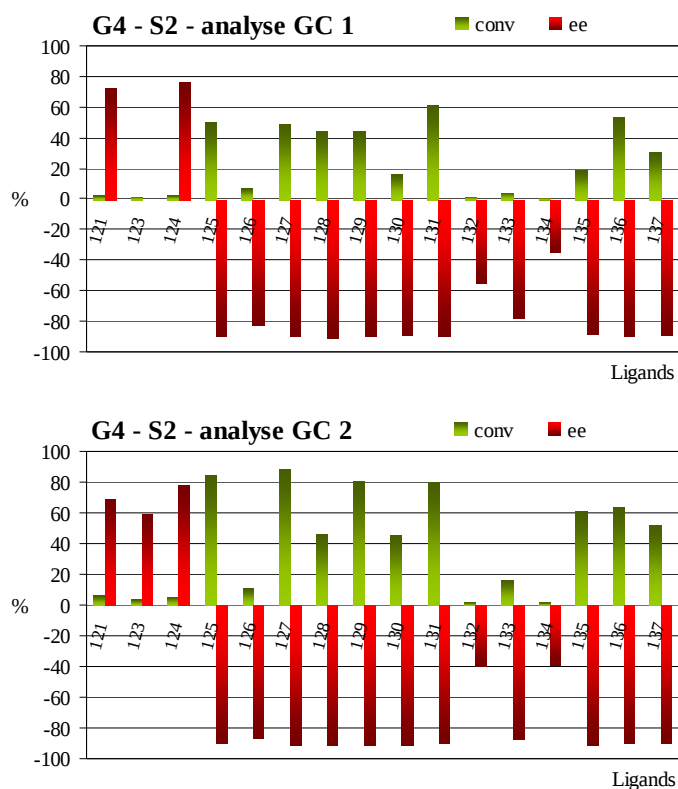


figure 9

L'écart moyen observé pour la conversion est de 16 %, ce qui marque une reproductibilité plus discutable. Parmi ces valeurs, 7 écarts sont en effet compris entre 20 et 40 %. En ce qui concerne l'ee, l'écart moyen est de 3 % et montre donc à nouveau une très bonne reproductibilité. On relèvera toutefois **132** et l'écart exhibé de 15 %.

Les graphes ci-dessous (figure 10) représentent maintenant S3 : L/M=2, S/M=100. Par rapport à la figure 9 nous examinons l'effet d'une variation de ratio L/M en passant de 1 à 2.

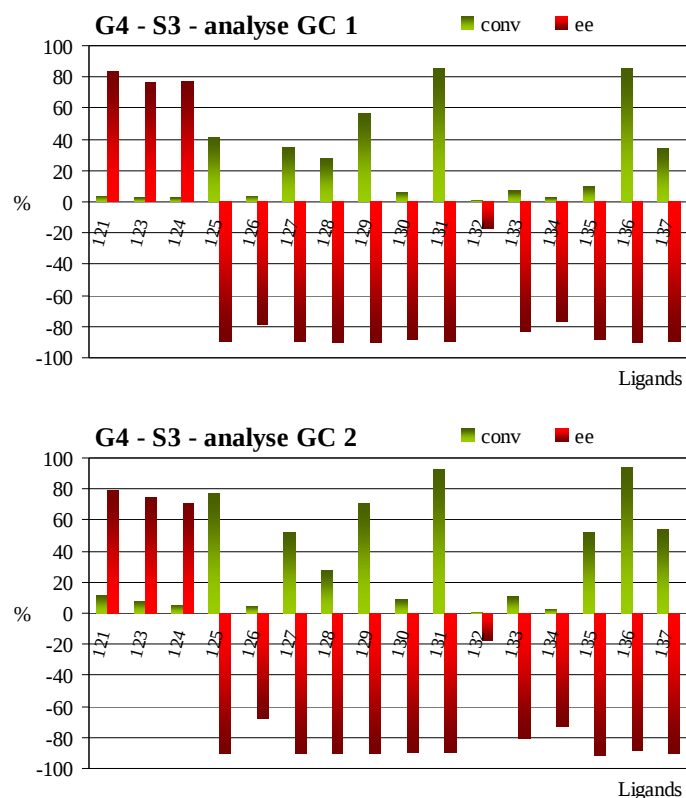


figure 10

L'écart moyen observé pour la conversion est à nouveau de 11 %, et une part plus importante des valeurs présente un écart voisinant avec les 10 %. On relèvera à nouveau trois ligands (125, 135 et 137) associés à des écarts compris entre 20 et 40 %. En ce qui concerne l'ee, l'écart moyen est de 2,3 % et montre cette fois encore une très bonne reproductibilité ; seul 126 présente 10 % d'écart entre les deux analyses.

A l'évidence, les ee sont peu sensibles aux facteurs incontrôlés puisque l'écart à la moyenne est de 3 % maximum, pour l'ensemble des mesures, et que seuls quelques cas se révèlent disparates. Les conversions, bien que

relativement stables pour environ moitié des données, s'avèrent plus sujettes à des variations parfois importantes, puisque des différences allant jusqu'à 40 % furent notées. Nous reviendrons plus tard sur la raison possible de ce comportement, il convient toutefois de se rappeler que les performances des systèmes catalytiques furent examinées après une heure de réaction et que l'analyse complète des 16 échantillons prend alors 48 minutes environ.

Hormis cela, d'autres informations peuvent être déduites de ces expériences :

- Les trois premiers ligands de la série proviennent de la sous-famille de type A, soient ceux ayant été construits à partir d'un aldéhyde aromatique et de la cyclohexyldiamine tosylé comme blocs de construction. La pince complexante est de type « N,N,O ». Nous constatons que les conversions sont résolument faibles toutes séries confondues ; ces ligands sont donc peu actifs. Les ee sont pour la majorité compris entre 70 et 80 % et constituent un score modéré pour cette partie de la dernière génération.
- Les 13 ligands suivants sont tous membres de la sous-famille de type B et ainsi découlent de l'association de l'aminoindanol aux salicyladéhydes modifiés. La pince de complexation « O,N,O » bénéficie aux performances des systèmes catalytiques. Les conversions sont nettement meilleures, particulièrement dans le cadre de la première série. Les ee, meilleurs eux aussi, peuvent être considérés comme bons puisqu'ils sont proches de 90 % et que beaucoup les dépassent. Le maximum de 91,8 % ee est atteint par **131** et **135** au cours de la première série.
- Les ee obtenus par les deux sous familles sont de sens opposés.
- **136**, dont le schéma 109 ci-dessous rappelle les caractéristiques, bénéficie en quelques sorte d'une structure hybride puisqu'il est membre de la sous-famille de type B mais possède néanmoins une pince de complexation « N,N,O ». Comme nous l'avons remarqué aux points précédents, la différence entre type A et B se marque par la

conversion et le sens de l'ee. Les conversions obtenues avec **136** sont souvent parmi les meilleures, et les ee négatifs, ce qui ne désigne pas un comportement hybride mais pleinement de type B.

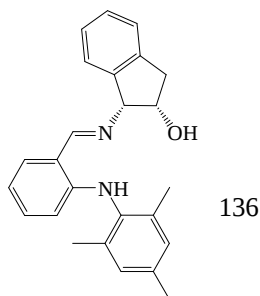


schéma 109

- Au sein de la « sous-famille » de type B, **132** et **135** exhibent les performances les plus basses. Si les autres ligands se comportent de manière irrégulière au regard des conversions, les ee sont eux très similaires, tous ligands de type B et toutes séries confondues. La variation des propriétés électroniques, introduite par le biais des aldéhydes reste d'un effet très modeste sur l'énantiosélectivité. Ce fait associé à la remarque du point 4 pose question sur le mode réel de complexation du ligand au métal en termes de bi- ou tridentate.
- Nous avons examiné 3 séries aux rapports L/M et S/M variables (tableau 3). De ces expériences, nous déduisons que S (L/M=1 et S/M= 25) reste le meilleur compromis pour obtenir les conversions et ee les meilleurs. Les conversions diminuent en passant à des taux catalytiques plus bas (S1 → S2 et S3) sans réellement affecter les ee, **132** et **134** exceptés. Si les valeurs correspondant à ces deux ligands ne sont pas prises en compte, la variation de l'ee reste confinée à 3,3 %. L'augmentation du rapport L/M de 1 à 2 (S2 → S3) ne change quasi rien à la situation.
- L'observation des graphes invite à penser que les expériences répétées ne font peut-être pas partie de la même population, au sens statistique du terme, en témoignent les différences moyennes observées entre les

conversions : S2 = 9 % et S3 = 11 %. Ces expériences s'inscrivent dans le cadre des observations paires, pour lesquelles on formule l'hypothèse nulle « H_0 » suivante : « si les deux méthodes donnent des résultats semblables, alors, en moyenne, la différence entre les résultats des deux méthodes est nulle ». H_0 se vérifie par l'équation ci-dessous :

$$t = \frac{\bar{x}_d - \delta}{\sqrt{\frac{s^2}{n}}} = (\bar{x}_d - \delta) \frac{\sqrt{n}}{s} \quad \text{à } n-1 \text{ degrés de liberté}$$

Dans cette formule, « x_d » représente la moyenne de la différence entre les valeurs se correspondant, « s » l'écart-type de cette différence et « δ » une erreur systématique que nous postulons nulle. Le nombre de degrés de liberté « n » équivaut au nombre de tests effectués, soit 16 dans notre cas. Le nombre « t » obtenu est à comparer à une valeur tabulée de « t » (t de student) et si $|t_{\text{calc}}| < t_{\text{tab}}$, alors l'hypothèse nulle se vérifie. Dans notre cas, nous confronterons $|t_{\text{calc}}|$ à une valeur de t_{tab} pour $n = 15$ et 95 % de probabilité. Le tableau 5 ci-dessous précise les valeurs obtenues pour les trois séries.

tableau 5

série	x_d	s	$ t_{\text{calc}} $	t_{tab}	H_0
S1	11,23	13,80	3,15	2,13	fausse
S2	16,32	15,41	4,10	2,13	fausse
S3	10,84	13,82	3,04	2,13	fausse

Chaque série montre un $|t_{\text{calc}}| > t_{\text{tab}}$ de manière non équivoque, signe clair d'une erreur systématique entre les deux analyses, probable à 95 %. Un graphique de corrélation (figure 11) entre la première et la seconde analyse, toutes deux de la seconde série, permet d'observer clairement un biais positif par l'écart à la diagonale sensée en toute rigueur manifester l'absence d'erreur systématique. La régression linéaire appliquée s'exprime par une équation du premier degré de type « $y = ax + b$ ». L'exploitation éventuelle

des coefficients « a » et « b » imposerait de calculer les écart-types sur a, b et croisé a/b.

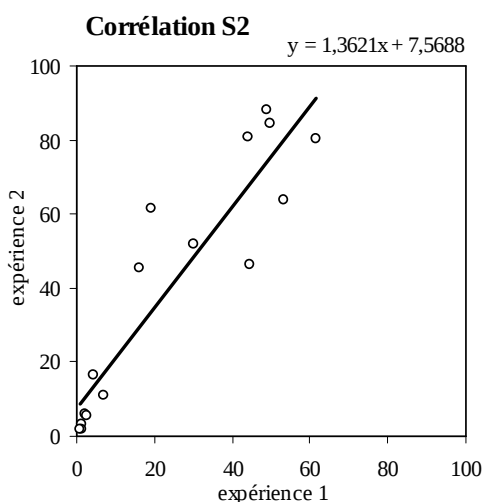


figure 11

Si la présence hautement probable du biais est démontrée, son origine reste par contre inconnue. Il n'est pas interdit de supposer qu'un manque de rigueur dans le minutage des expériences en soit la cause et ait provoqué ces différences de conversions.

4.4 Optimisation dirigée

4.4.1 Principe de la méthode

D'entrée de jeu, nous disposons de plusieurs paramètres dont la variation doit conduire à trouver le ou les systèmes catalytiques performants de notre bibliothèque. Les ligands, leurs blocs de construction et les divers précurseurs métalliques jouissent de cette interchangeabilité ; ainsi que les temps réactionnels. Ainsi est-il possible d'établir un système de codage des ligands, dont un exemple est donné par le schéma 110, qui permet d'envisager ces variations simplement.

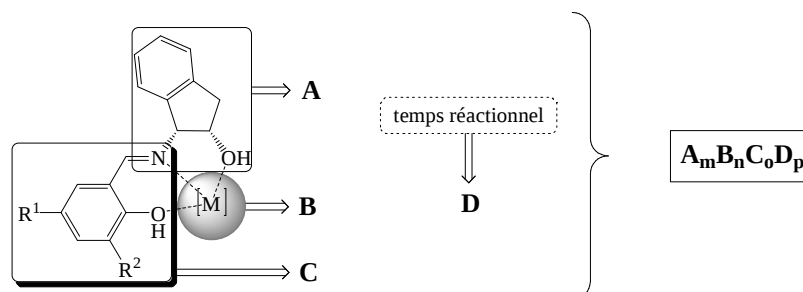


schéma 110

Si elles n'avaient pas été testées complètement à chaque génération, les séries S1, S2 et S3 eussent été sujettes à être, elles aussi, utilisées comme paramètre modifiable. Au cours de cette optimisation, nous nous sommes donc appliqués de génération en génération à évaluer l'influence positive ou non des ces variables, et à leur faire subir ce qui se définit comme des opérateurs génétiques. Ceux-ci sont la réplication, la régression, le cross-over et la mutation. Ci-dessous, le schéma 111 illustre le principe par l'exemple des 4 opérateurs appliqués aux ligands ; il va de soi que cette manière de procéder fut élargie aux autres variables.

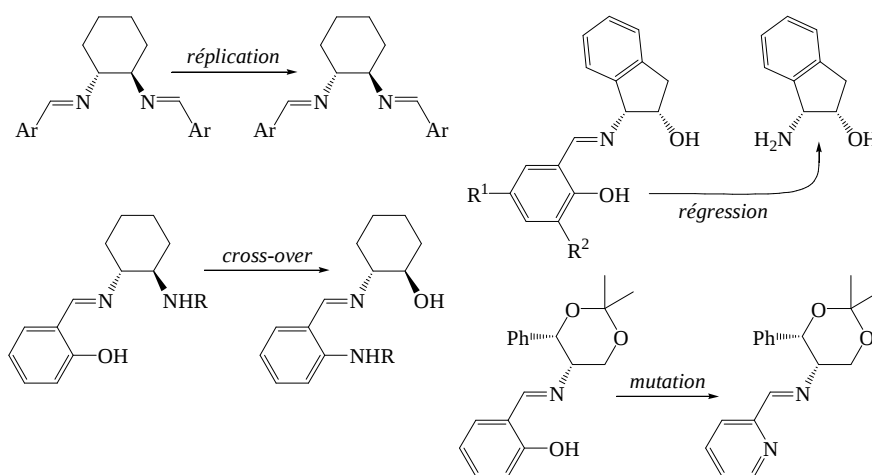


schéma 111

Une génération « parent » conduit logiquement à une nouvelle génération « enfant » grâce aux différences chimiques introduites par ces

opérateurs. Cependant, tous les membres d'une génération n'étaient pas choisis pour transmettre leurs caractéristiques, entièrement ou en partie à la génération suivante. L'évolution se veut d'aller vers une augmentation des performances, et donc délaisse les mauvais candidats ; Le critère de sélection imposé ici était une conversion minimum de 10 % et un excès énantiomérique minimum de 50 %, quel qu'en fut le sens. Tout système catalytique ne réalisant pas ces performances fut éliminé du processus évolutif, les autres pouvaient alors servir de base à la formation de la génération postérieure par le biais des opérateurs génétiques.

4.4.2 Efficacité de la méthode

Attendu que l'excès énantiomérique avait été défini dès le début comme le paramètre le plus important, nous avons décidé d'évaluer l'efficacité de notre méthode évolutive par le biais de l'excès énantiomérique moyen de chaque génération, ou AEEL (Average Enantiomeric Excess of the Library), qui s'exprime par l'équation ci-dessous :

$$AEEL = \frac{\sum_{i=1}^j (ee)_i}{j} \quad (j = \text{nombre de tests par génération})$$

La figure 12 dépeint l'évolution de cet AEEL en fonction du nombre de tests.

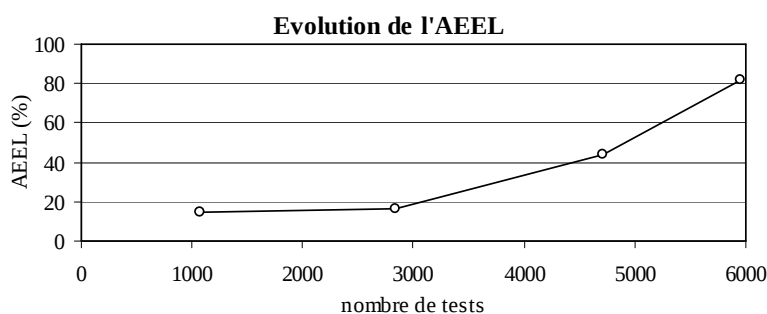


figure 12

Les AEEL calculés sont successivement pour les 4 générations de 15, 16, 44 et 82 %.

Le nombre de combinaisons réalisables au sein de notre « espace réactionnel » était de 4212 ; même si vérifier la reproductibilité pour ceux que nous avons menés nécessita 6000 tests au total. Sur ces 4212 combinaisons possibles, seules 2376 ont été réellement évaluées, autrement dit 56 % de l'espace réactionnel imaginé.

4.4.3 Critique de la méthode d'optimisation

Comme nous venons de le démontrer, ce concept d'optimisation dirigée fonctionne. Ceci étant, il nous faut toutefois convenir de défauts majeurs inhérents à la méthode, que mettent en exergue les points suivants :

- On relèvera d'abord le fait que celle-ci implique de près l'expérimentateur puisque les choix que celui-ci pose conditionnent directement l'efficacité de la méthode. Il est vrai que le chimiste peut, par les connaissances qu'il a du sujet, accélérer le phénomène évolutif et favoriser une voie plutôt qu'une autre ; une erreur d'appréciation de sa part ou un manque de discernement des paramètres influents s'avérerait par contre préjudiciable à la méthode et risquerait de conduire pour un temps au moins l'optimisation dans la mauvaise direction.
- Puisque la méthode est soumise au jugement humain, on doit admettre que la part laissée au facteur aléatoire est réduite. Or on sait que le hasard peut promouvoir l'introduction, dans une génération, de combinaisons nouvelles (métal - ligand - temps réactionnel - série), auxquelles le chimiste n'aurait pas songé directement, voire les aurait exclues pour de bonnes raisons *a priori* ; et que ces nouvelles combinaisons sont parfois utilement porteuses de paramètres enrichissants. Le manque de hasard peut donc ralentir le processus d'optimisation.
- Une part congrue des combinaisons se réalisent en acceptant que les opérateurs précédemment définis jouissent d'un champ d'action un peu plus large dans certains cas. La méthode comporte donc un manque de rigueur occasionnel.

- Un examen minutieux de chaque test effectué permettrait sans doute de mettre en lumière des informations diverses concernant par exemple le mécanisme de la réaction, une éventuelle variation de celui-ci en fonction du ligand, du métal, des conditions, etc... Il est donc honnête de préciser que peu d'informations sont finalement retenues d'une telle quantité d'expériences.

Si nous examinons maintenant la méthode par ses côtés positifs, il convient de préciser qu'elle a le grand avantage de pouvoir être mise en œuvre simplement, c'est-à-dire par un codage aisé des divers systèmes catalytiques et des conditions dans lesquelles on les fait opérer. Cette manière de procéder s'affranchit de tout recours à un outil informatique que ce soit pour décrire catalyseurs et conditions, ou donner une direction à l'optimisation et engendrer successivement les générations ; ce qui peut nettement faciliter la tâche de l'expérimentateur. On notera encore que malgré le grand nombre de tests entrepris, l'optimisation parvint à « trouver son chemin » grâce notamment au regard que porte le chimiste sur les résultats obtenus génération après génération et qui déterminera ensuite quels opérateurs employer, ainsi que l'opportunité du taux avec lequel il faut en user.

La dernière génération, G4, possède un AEEL de 82 % et nombre de ses membres atteignent, voire dépassent les 90 % ee. Pour parvenir à cela, un peu plus de la moitié des combinaisons possibles a été effectivement testée. Par rapport à une méthode n'utilisant que le criblage à haut débit, ceci peut être considéré comme un avantage majeur.

Le but annoncé de cette méthode d'optimisation dirigée à distinguer plus rapidement des systèmes catalytiques performants au sein de l'ensemble des possibilités imaginé est donc atteint.