Available online at www.sciencedirect.com
ScienceDirect
Journal of Pharmaceutical and Biomedical Analysis 46 (2008) 839–847
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

A screening study of ChirBase molecular database to explore the expanded chiral pool derived from the application of chiral chromatography

Patrick Piras *, Christian Roussel

UMR "Chirotechnologies: Catalyse et Biocatalyse," Université Paul Cezanne Aix-Marseille III, Marseille, France

Received 5 March 2007; received in revised form 12 July 2007; accepted 28 August 2007
Available online 7 September 2007

Abstract

28,000 chiral compounds separated by chiral chromatography have been selected in ChirBase database. The enantiomers of this library can be considered as potentially being isolable by chromatography at a small-scale level of development. The resulting chiral pool was analysed with the aim to describe compound molecular diversity as well as druglike and leadlike characteristics. In parallel, we have explored the possibility of using this chiral product library as starting materials for the construction of a virtual chiral combinatorial library. The chemical space occupied by the obtained combinatorial collection of new chiral scaffolds revealed that structures of the chromatographic-based chiral library may also be a source of chiral diversity for drug discovery. Another interesting *in silico* approach consisted to release all the protected or derivatized compounds. This procedure allowed us to enrich and expand the existing library with several thousands of original small chiral amines, acids and alcohols.

Finally, sub-similarity searching strategies were found valuable for quickly selecting suitable small chiral precursors that are readily available from chiral chromatography for the small-scale synthesis of some known chiral drugs.

Keywords: Chirbase database; Chiral HPLC; Chiral combinatorial library; Diversity analysis; Drug discovery; Druglike; Leadlike distribution

1. Introduction

In 2005, worldwide sales of single-enantiomer drugs reached $225 billion [1]. A glance at the top 10 best selling prescription drugs shows that eight are chiral and all are small molecules (Table 1). Another sign indicative of the growing importance of chirality in the pharmaceutical industry: the best-selling drug in the world is Lipitor and it is marketed as a single enantiomer, providing annual sales of $12.9 billion.

The growing demand for highly pure enantiomers in the pharmaceutical and fine chemicals industries is changing the way laboratories deal with chiral technologies. In a recent review [2], authors outlined that most of the marketed chiral active pharmaceutical ingredients (API) are derived from the chiral pool and resolution processes. In this review, authors highlighted the intellectual property issues when patenting new chiral ligands and catalysts for use in asymmetric syntheses. Also, as previously noted in 2001 [3] few asymmetric synthesis have found an economically viable process and thus today are used in production. A recent survey of the reactions used for the preparation of 128 drug candidate molecules [4] has shown that 69 (54%) molecules are chiral and among them, 67 are prepared as single enantiomer. This study yet again revealed that much of the generated chiral centers are derived from commercially available chiral starting materials. One reason for this fact is that chiral pool is the cheapest source of chiral APIs. The leading in-house chiral technologies used to generate chiral centers were found based on resolution methods. Only 20% of the chiral centers were generated using asymmetric synthesis. It is obvious that the needs of quick access to optically pure materials at the early stage of the development process have driven this trend. The remarkable success of the resolution methods can also be explained as a consequence of the impulse given by the progress of chiral chromatography [5–8] and principally the preparative applications for large-scale purification of enantiomers with practical applications in production. [9]. The advantages of preparative chiral chromatography in early drug discovery have been well illustrated by Nelson et al. in Merck & Co. [10]. In this paper, authors explained why a rapid and inexpensive racemic synthesis in conjunction with chiral HPLC

* Corresponding author.
E-mail address: Patrick.Piras@univ.u-3mrs.fr (P. Piras).

0731-7085/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2007.08.031

Table 1
Top 10 best selling prescription drugs

Drug name	Market position	Active molecule	Chirality	Chiral centers	Treatment	Revenue (billion)	Company
Lipitor	1	Atorvastatin	Yes	2	High cholesterol	$12.9	Pfizer
Plavix	2	Clopidogrel	Yes	1	Heart disease	$5.9	Bristol-Myers Squibb & Sanofi-Aventis
Nexium	3	Esomeprazole	Yes	1	Heartburn	$5.7	AstraZeneca
Seretide/Advair	4	Salmeterol	Yes	1	Asthma	$5.6	GlaxoSmithKline
Zocor	5	Simvastatin	Yes	7	High cholesterol	$5.3	Merck
Norvasc	6	Amlodipine	Yes	1	High blood pressure	$5.0	Pfizer
Zyprexa	7	Olanzapine	No	0	Schizophrenia	$4.7	Eli Lilly
Risperdal	8	Risperidone	No	0	Schizophrenia	$4.0	Johnson & Johnson
Prevacid	9	Lansoprazole	Yes	1	Heartburn	$4.0	Abbott Labs & Takeda Pharm.
Effexor	10	Venlafaxine	Yes	1	Depression	$3.8	Wyeth

Source: IMS Health, January 2007.

was favoured in Merck & Co. to rapidly produce an enantiopure lactone intermediate for pre-clinical trials. This approach let them time to investigate in parallel the asymmetric synthesis and thus they almost certainly avoided to delay for several months the clinical trials. Another case study at Glaxo SmithKline was reported by McGuire [11] for the synthesis of the Vitronectin receptor antagonist SB 273005. The drug candidate selection process was depending on a racemic synthesis and chiral HPLC. The asymmetric synthesis was later considered for advancing into development and commercial process [12].

In some cases, the chiral chromatography step remains the most practical and economical approach for pilot-scale and commercial scale production plant using for instance simulated moving bed technology (SMB). According to McCormick [13], today seven commercial-scale SMB installations are operating. Among them, Pfizer's antidepressant sertraline is one interesting example actually in favour of chromatographic resolution. A Low cost racemic synthesis provides a cheap racemic tetralone which is submitted to the chiral SMB separation. The unwanted tetralone *R*-isomer is racemised and recycled. No other enantioselective process was found to be superior to chiral separation [14].

As a consequence, since 2001 we have noted a remarkable increase in the number of patents claiming the use of chiral stationary phases for the preparation of chiral materials at preclinical and clinical stages. This observation leads us to develop the idea that chiral HPLC provides an access to a new chiral pool comprising an almost unlimited range of compounds.

In this contribution we are continuing our work in exploring the potential of Chirbase the largest information database on chiral HPLC separations [15,16]. ChirBase contains chiral separation data of thousands of small chiral compounds. Most of them are readily accessible as racemate whether they are readily synthesizable or available from vendors. Once they have been selected as starting material, these precursors may be easily separated by chiral chromatography to provide few milligrams of pure enantiomers.

In this paper, cheminformatic strategies were used in order to evaluate the chemical space covered by this expanded chiral pool. One important aim of this work is to examine if the structural diversity is sufficient for actual drug discovery projects. From this perspective it was found of interest to estimate the proportion of molecules having "druglike" or "leadlike" properties. Further examining ChirBase chiral pool, it becomes apparent that there are a lot of compounds that can be easily cleaved to generate new chiral entities enriching the molecular diversity of the original molecule collection. Additionally, as combinatorial chemistry is today a key technology that aims to enrich chemical libraries, we have constructed one original amide library by virtually reacting some of the chiral precursors available by chromatography. This is to our knowledge the first example of generation of a chiral combinatorial library exploring the potential of chiral chromatography to provide new chiral leads.

2. Methods

2.1. Chiral pool data set

At the time of the study, ChirBase database contained 112,154 chiral separations for 32,451 chiral compounds [17]. The Chiral pool data set was created by selecting in ChirBase all the successful chiral separations at all scales (about 90% are analytical and 10% are semi-preparative or preparative). Data selection was mainly based on chromatographic data (enantioselectivity and resolution values). As such data are not always available, compound selection also relied on chromatograms or when given a qualitative description of the experimental results (e.g. a "baseline separation" is claimed by the authors). As seen in Fig. 1, this chiral pool derived from chromatography is composed of approximately 28,000 chiral molecules corresponding to 56,000 enantiomers which may be seen as a significant source to address the chiral needs in the early stages of discovery when enantiomers are neeeded in small quantities.

In this study we also intended to compare various characteristics of ChirBase chiral pool file with other available database specifically devoted to chiral compounds. However in an internet search we could only find two products:

- A database of over 1500 chiral chemicals available from worldwide suppliers launched in 2004 by Bark Information Services [18].

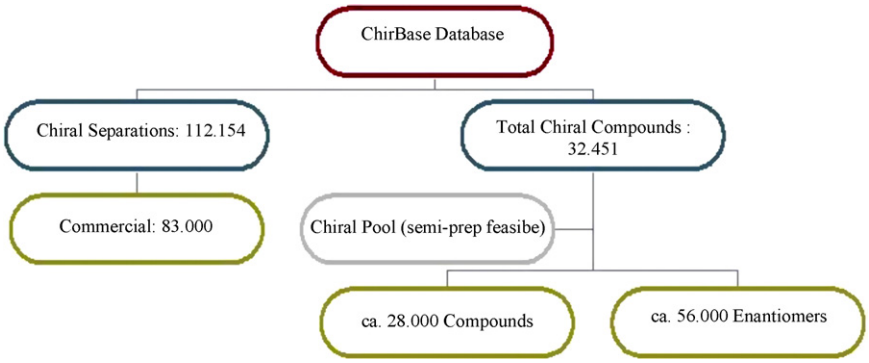


Fig. 1. Summary statistics of ChirBase database.

- A starting material database which contains 2617 chiral precursors (1800 synthetic and 817 commercial in 2005) as part of the Chiron program [19] which is an interactive computer program for synthesis planning.

For a direct comparison with our chiral pool data set, we needed to find a large database with a broad range of small chiral compounds. We finally chose to merge the different SD files available from Sigma–Aldrich Web site [20]. By doing this, we have been able to build two databases that we named:

- *Aldrich precursor database:* 28,159 precursors comprising the chiral as well the achiral precursors (built by merging and cleaning the Aldrich building blocks available as SD-Files).
- *Aldrich chiral database*: 14,102 chiral organic compounds (library of unique structures obtained by merging and removing duplicates of the 40,799 chiral organic compounds supplied as SD-files).

2.2. Computational methods

Chemical databases and combinatorial libraries were developed and managed with the "Screening Assistant" software freely available from the University of Orléans [21]. The Screening Assistant was also used to estimate the diversity of druglike and leadlike compounds.

A number of database pre-processing tasks were needed to be performed to obtain reliable results. For instance, acids and base were neutralized, duplicates and counter ions were removed using the ChemAxon Jchem software [22]. "Reactor" and "Fragmenter" modules of Jchem package were also utilized to generate the different chiral combinatorial libraries.

3. Results and discussion

3.1. Statistics: evaluation of ChirBase chiral pool

One key criterion when evaluating the potential of chemical libraries is the number of compounds sharing the physical and chemical properties that are common to known drugs.

Since its publication in 1997, the Lipinski "rule of five" is the most widely used model for the design and selection of drug candidates. Lipinski et al. [23] have examined and compared the physico-chemical properties of 2245 orally bioavailable drugs which successfully passed the preclinical stages. This analysis revealed that a majority of these compounds respect the following rules: molecular weight $\leq$500, $\log P \leq 5$, number of H-bond donors $\leq$5 (OH and NH groups) and number of H-bond acceptors $\leq$10 (N and O atoms).

These rules are assumed to be good indicators of oral bioavailability and thus may give an idea of the possible drug-like molecules included in a compound library. These rules have been complemented by Oprea [24] who studied the property distribution of drug-related chemical databases. Oprea has shown that some property distributions could characterize the chemical space covered by these commercial databases. He found that $\log P$ and molecular weight produce a Gaussian distribution for all database. However non-drug databases such as ACD (a library of commercially available starting materials for drug synthesis) have lower molecular weight and $\log P$ values than drug database (as MDDR, MACCS-II Drug Data Report).

Property distributions in Chirbase chiral pool are displayed in Fig. 2. Most of the compounds (over 80%) pass the Rule of Five and could be considered as druglike according to Lipinski theory. Also, both molecular weight and $\log P$ produce a Gaussian distribution with a profile quite comparable to that of a small molecule library as described by Oprea [24]:

- 50% of the compounds in the chiral pool have a molecular weight $\leq$287 (ACD database peak is around 300). The regions between the 25 and 75% of the distribution range (called "mid-50%" by Oprea) are between 210 and 337 (ACD database mid-50% is between 224 and 368).
- 50% of the compounds in the chiral pool have a $\log P \leq 3.2$ (ACD peak is at 3.0) and mid-50% is between 2.1 and 4.43 (ACD mid-50% is between 1.64 and 4.45).

However according to Oprea's study, Gaussian distribution trends are not enough significant to distinguish drug and non-drug chemical space. More discriminating are the trends observed from the behaviour of the asymmetrical Gaussian distribution as produced by the number of rings, the number of rotatable bonds or the number of hydrogen donors or acceptors. Database with simple starting material provide distribution with fewer rings, fewer rotatable bonds and also less donors and acceptors. Once again, the distribution pattern of ChirBase chiral

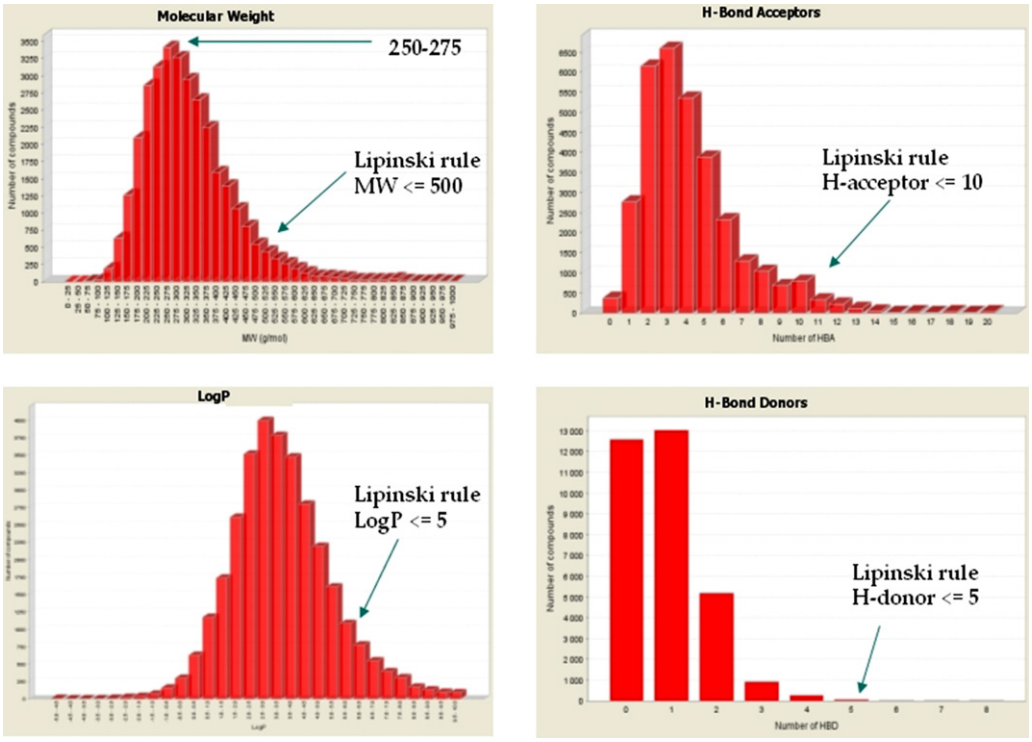


Fig. 2. Molecular property distributions in ChirBase chiral pool (28,000 structures) as generated by the Screening Assistant software [21]. Limits under which the Lipinski "Rule of 5" is fulfilled are indicated.

pool is comparable to a small molecule library. For instance, 62% of the chiral pool compounds have a number of ring ≤ 2 (64% for ACD), 78% have a number of hydrogen donors ≤ 1 (71% for ACD). The distribution of the number of rotatable bonds shows that our chiral pool has an intermediate pattern with a mean peak value of 6 (between ACD peak at 4 and MDDR drug database peak at 7). This is not surprising as chiral chromatography also addresses a lot of pharmaceutical applications.

In further studies, Oprea et al. attempted to define the differences between the lead-like and the drug-like chemical space [25]. Authors could identify a chemical space which revealed that leads are less complex than drugs (lower molecular weight as well number of rings and rotatable bonds) indicating that Lipinski's rules should be reduced.

To evaluate the "leadlikeness" of our chiral pool library, we chose to compute the physicochemical properties recently refined by Verheij [26] who combined Oprea criteria with filters related to in vivo studies [27,28]: molecular weight <450, H-donors <5, H-acceptors <10, $\log P < 4.5$, rotatable bonds <10, log SW (water solubility) > -6.0 and TPSA (polar surface area) < 150.

According to Verheij filters, 61% of our chiral pool library has lead-like properties. This is a significant number if we consider that Verheij found a mean value of 45.9% of lead-like compounds when analysing 45 commercially available libraries (between 19 and 55% of leadlike structures in the biggest supplier database (>150,000 structures)).

Other studies aim to explore the potential of library to supply small chiral building blocks as precursors for the synthesis of new leads or drugs. These approaches are often central at the drug discovery stage which often necessitates screening libraries of low-molecular weight. Leach et al. have recently redefined the approaches that can be applied in the design of a fragment library for lead discovery [29]. Among them, one useful and simple way to select fragments is the "rule of three" suggested by Congreve et al. at Astex [30] who investigate the physicochemical properties of their in-house collection of fragment hits. They determined that most of compounds respect the following criteria: molecular weight < 300, hydrogen bond donors ≤ 3, hydrogen bond acceptors ≤ 3, $\log P \leq 3$.

These rules when applied to our chiral pool library specify that 20% of the compounds may be considered as appropriate potential precursors of chiral lead-like and drug-like molecules. This gives a final set of about 6000 molecules (or 12,000 enantiomers) that are potentially available from chiral chromatography as building blocks for asymmetric synthesis.

In summary computational chemistry offered us a quick and simple approach to explore the property space of chemical structures contained within the chiral pool library. There are clearly no evident spatial limits between small precursor, leads and drugs. As schematically illustrated in Fig. 3, the physical property rules designed to differentiate precursor space, lead-space and drug-space appear to overlap (a precursor can be a lead and a lead can also be a drug).

3.2. Diversity study

The previous statistic evaluation pointed out the drug-like or lead-like character of the chiral pool but it does not reveal if it is characterised or not by a good coverage of chemical

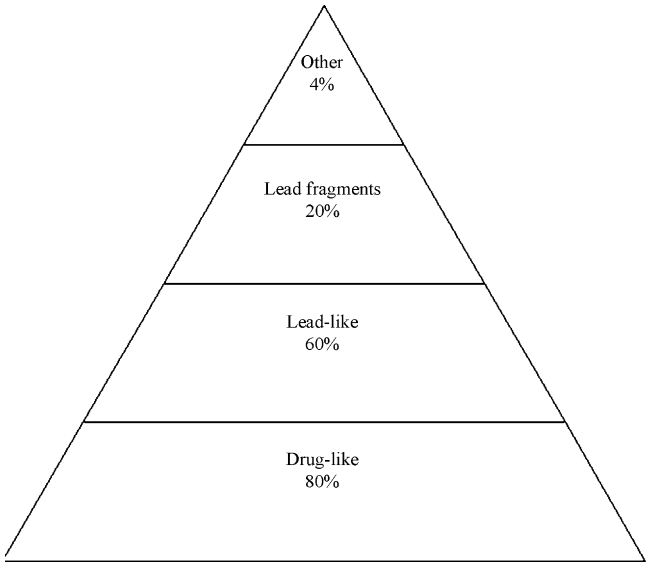


Fig. 3. Percentage of drug-like, lead-like and lead-fragment compounds found in ChirBase chiral pool library. Filters are derived from [30] for low molecular weight compounds (called "lead fragments"), [26] for "lead-like" and Screening Assistant software [21] for "drug-like" and "Other" (with undesirable structural features, or bad property).

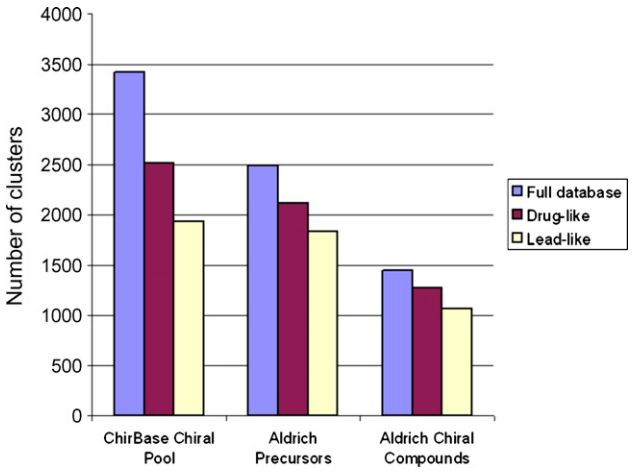


Fig. 4. Cluster distributions aimed at comparing diversity of ChirBase chiral pool (28,000 structures) with two Aldrich catalog files: Aldrich precursor library (28,159 chiral/achiral organic compounds) and Aldrich chiral library (14,102 chiral compounds).

space. Analysis of diversity is a key issue for virtual screening of compound libraries [31].

For this purpose, fingerprint descriptors (such as MACCS SSKeys) that encode the molecular structures into a linear vector are often used. In this study, diversity analysis was performed with the Screening Assistant software which applies a clustering analysis to find similar compounds and form groups. The analysis is based on SSKey-3DS descriptors which are built on the coding of the presence or absence of 32 fragments as well as H bond acceptors, aromatic bonds, or fraction of rotatable bonds [32].

To get some idea about the diversity of our chiral pool library, we chose to compare our compound population with two Aldrich libraries with regard to the number of clusters. For each database, we will assume that a greater degree of diversity correlates with an increasing number of clusters formed. The results of the cluster analysis are displayed in Fig. 4. For comparison, we have added the drug-like and the lead-like cluster distributions as retrieved from the application of the same rules ("Screening Assistant" filters was applied to the three libraries). Chirbase chiral pool which includes almost the same number of structures as there are in the Aldrich precursor database is characterized by the highest number of clusters (3418). However, this figure clearly indicates that this difference of diversity is in part due to clusters occupied by non-leadlike and non-druglike structures. This could be related to the fact that ChirBase is not only the result of our own laboratory activity. Contrary to in-house compound collections, our library is built via a multitude of sources. Interestingly, when we look at the number of leadlike clusters, results are quite comparable: 1929 clusters for ChirBase and 1833 clusters for the Aldrich precursor database. Even if the number of compounds per cluster was not accessible, this finding well complements the above statistic study in suggesting that there is not only a significant number of potential lead-like compounds but also a diverse range of scaffolds. Finally, a comparison with the distribution pattern of the chiral compounds commercially available from Aldrich confirms the promising role of chiral chromatography to substantially enrich the existing chiral pool.

3.3. New chiral library generation

In this section, we will examine how the ChirBase chiral pool can be further expanded if we consider the structures that can be derived or synthesized from the current chiral compounds. High-throughput screening (HTS) plays today a major role in drug development.

The aim of HTS is to discover "hits" which are potential leads. Computational techniques offer to the scientists the ability to perform virtual HTS experiments (such as ligand docking and alignment) often referred as "High-throughput *in silico* screening" [33,34]. The compound collections used for virtual HTS can be a database of commercially available compounds as well as virtual combinatorial libraries.

One way to create a virtual combinatorial library is to perform a chemical reaction on a library of synthesizable, in-house, or commercially available reagents. This strategy when applied to a large number of reactants and investigating all possible products can generate a very large combinatorial library of virtual compounds. Hits found from a virtual screening of the obtained combinatorial library are then expected to be easily prepared being given that the original reagents are all accessible. A benefit of such a generated combinatorial library is that it is intended to further enhance the diversity of the original library and so it is expected to increase the hit rates of virtual screening.

To get some idea of the ability of ChirBase chiral pool to generate original new chiral compounds, we have chosen a simple reaction: synthesis of chiral amides from amine and acid precursors. 1405 chiral amines (R-NH2) and 2935 chiral acids (R-COOH) were extracted from the chiral pool library and saved

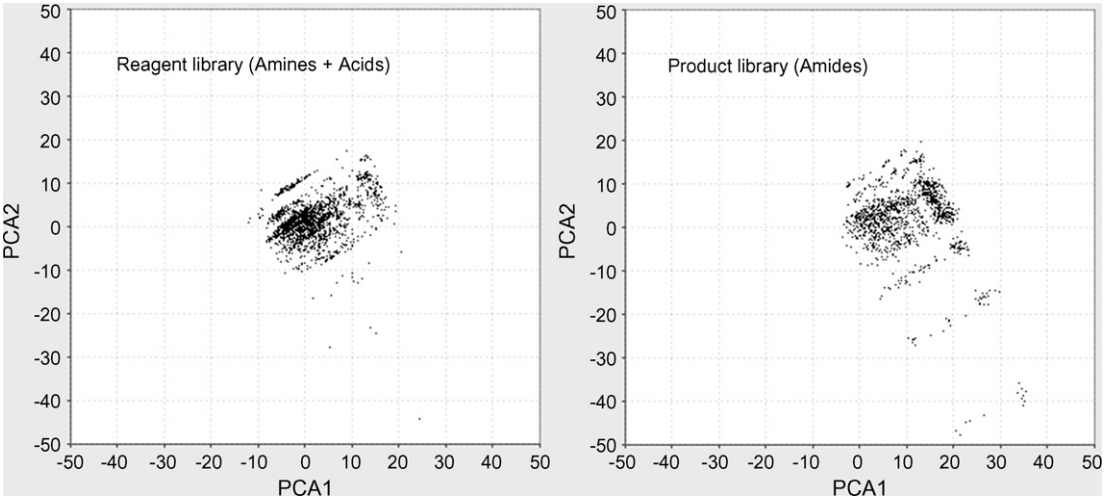


Fig. 5. Graphical comparaison of the diversity within the initial collection of chiral reagents and the generated combinatorial library of amide products. Plots were computed using the same procedure (principal component analysis) and descriptors (SSKey-3DS structural descriptors calculated by the Screening Assistant program).

in two different databases. Virtual synthesis of amides were performed with "Reactor" which is the virtual reaction engine of Chemaxon package (see Section 2). If we consider all possible combinations of these reactants, the "Reactor" engine would provide 4,208,210 potential amide products. In our study, we chose to combine the amines and the acids in a sequential mode (Amine 1 + Acid 1, Amine 2 + Acid 2, etc. . .). We specified to the reactor engine to process the *in silico* reaction only if it is specific (only one possible reaction center in each reactant). The generated virtual combinatorial library consists of 1502 new chiral amides. As at least two asymmetric centers have been introduced in the amide products, this combinatorial library correspond to a potential of ca. 6000 stereoisomers if we considerer all possible pair of enantiomers and diastereomers.

In order to explore the diversity created by this combinatorial approach, we display in Fig. 5 the chemical space occupied by the full initial set of reagents (4360 compounds) and the new amide products. This graphical representation of diversity was produced by subjecting the compounds to a principal component analysis. For both reagent and product set, the "SSKey-3DS" structural descriptors already mentioned above have been used. Obviously, despite a smaller number of compounds (1502 structures), the new chiral amide compounds occupy a larger part of the chemical space as compared to the reagent data set which is more condensed in a particular region of chemical space. Clearly, the new amide compounds occupy regions unexplored in the previous libraries. If we calculate the number of unique molecular frameworks as given by the Screening Assistant software [35], we find:

- 379 distinct molecular frameworks in the reagent library (Acids + Amines = 4340 compounds).
- 868 distinct molecular frameworks in the products library (1502 amides).
- 992 distinct molecular frameworks in the whole molecule data set (4340 reagents + 1502 amides = 5842 compounds merged in one database).

This suggests that the increase of diversity in the combinatorial amide library is related to the creation of 613 new molecular frameworks (992 − 379). Interestingly, it appears that 40% of the virtual chiral compounds included in the amide library are likely to exhibit "drug-like" characteristics and 10% lead-like properties. As seen in Fig. 6, these compounds which are characterized by the presence of multiple chiral centers and the availability of all the stereoisomers via chiral chromatography could be an interesting source of novel enantiopure scaffolds for screening experiments.

Another experiment to further expand the existing ChirBase chiral pool consisted of considering all possible compounds that can be readily cleaved (such as esters, amides, ureas, carbamates. . .) to regenerate the parent alcohols, amines or acids. This is an interesting concept since many compounds in ChirBase bear protecting groups or have been derivatized prior to chiral separation. Evaluation of this approach was performed with Chemaxon "Fragmenter" engine. The "Fragmenter" engine is based on the Retrosynthetic Combinatorial Analysis Procedure (RECAP) method [36]. The program utilizes predefined cleavage rules which correspond to chemical reactions. In our study, we entered the cleavage rules for amides, esters, ureas, carbamates and sulphonamides. The maximum number of molecular fragments generated per molecule was set to three molecules.

When applied to the 28,000 chiral compounds of ChirBase chiral pool, the "Fragmenter" method produced a library of corresponding regenerated amines, acids and alcohols. Removal of duplicate compounds and achiral fragments provided a final compound library consisting of 36,130 chiral molecules that represent an additional chiral pool of potential building blocks that can be prepared by chiral chromatography (Fig. 7). By merging these supplementary chiral materials with the original chiral pool into a unique database, and eliminating the duplicates, this resulted in an increase of the amount of chiral structures potentially available in ChirBase to about 85%. About 20,000 new molecules were added giving a total number of ca. 48,000

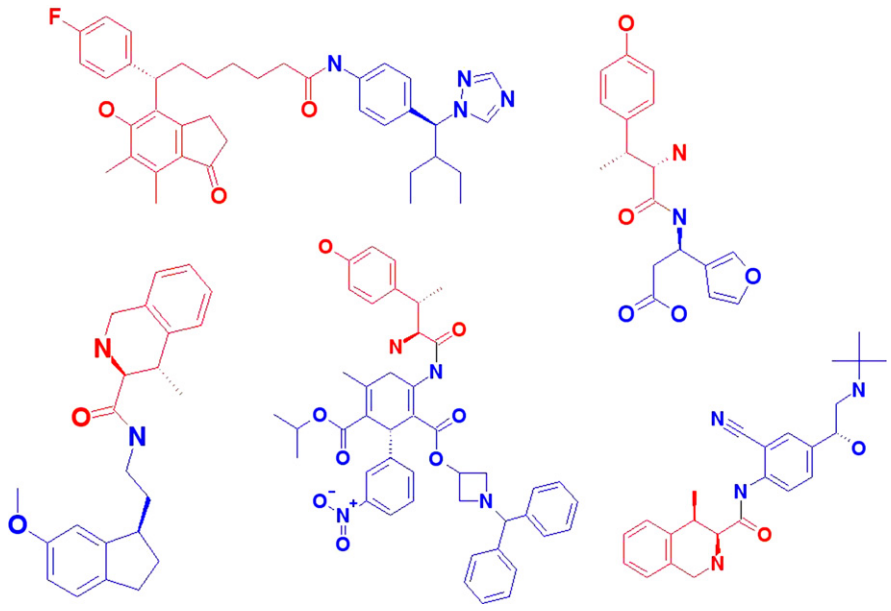


Fig. 6. Chemical structure of five chiral amides randomly chosen from the virtual combinatorial library (1502 amide products).

chiral structures (96,000 enantiomers) which is a better representation of the potentialities of the currently available chiral materials.

3.4. Finding precursors of chiral known drugs

Further extensions of this work focus on addressing the question of whether it is possible to identify in this chiral pool new starting materials corresponding to precursors of known chiral drugs. The idea here is based on the use of the sub-similarity searching capabilities of ISIS/Base database software [37] to detect in ChirBase all the potential synthetic precursors of a given chiral product. Sub-similarity searching is specifically designed to find in a database small compounds similar to structural fragments of a given new or known molecule. A sub-similarity algorithm was recently applied by Tounge and Reynolds to assign a druglike score to a set of reagents [38]. In their approach, authors could attribute a subsimilarity score to small R-group monomers extracted from selected active reagents and utilizes the distribution of these scores to compare databases. In our approach, a subsimilarity score is calculated by comparing the entire structures of each eligible reagent with the chiral drug to be synthesized. Some preliminary results of this searching strategy are shown in Table 2.

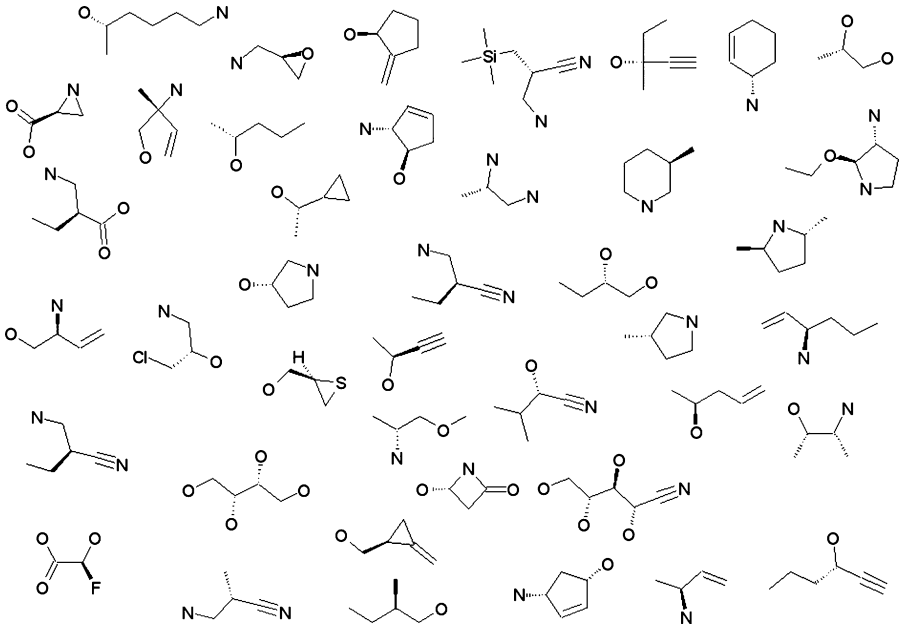
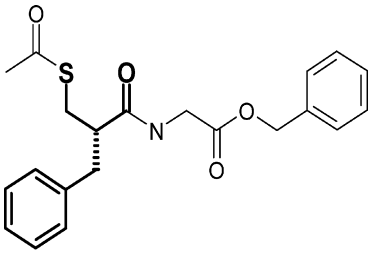
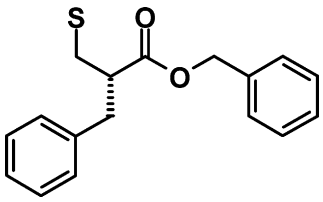
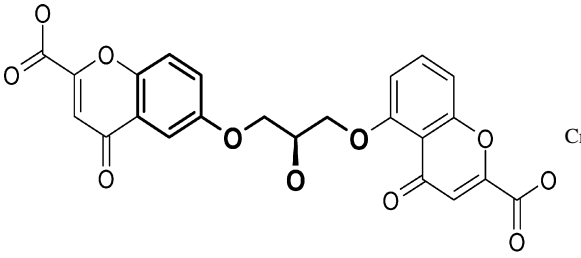
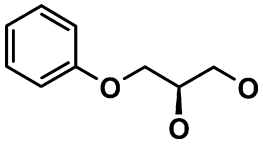
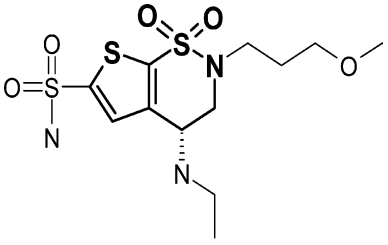
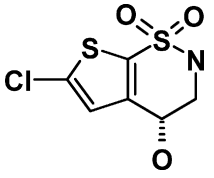
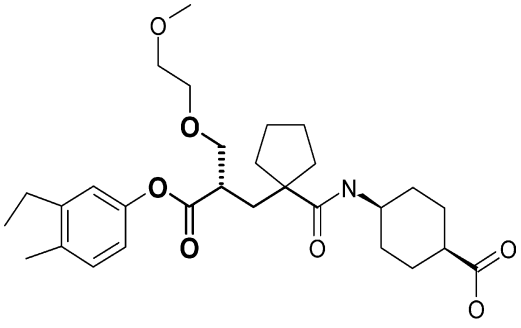
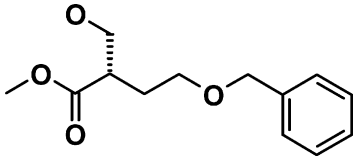


Fig. 7. A panel of the extended chiral pool potentially available in ChirBase. Cleavage of derivatized compounds revealed a rich source of original small starting chiral materials accessible by chiral chromatography.

Table 2
Subsimilarity search examples of known commercial drugs in ChirBase

Commercial chiral drugs	One of the chiral precursors found in ChirBase
 Ecadotril	
 Cromolyn	
 Brinzolamide	
 Candoxatril	

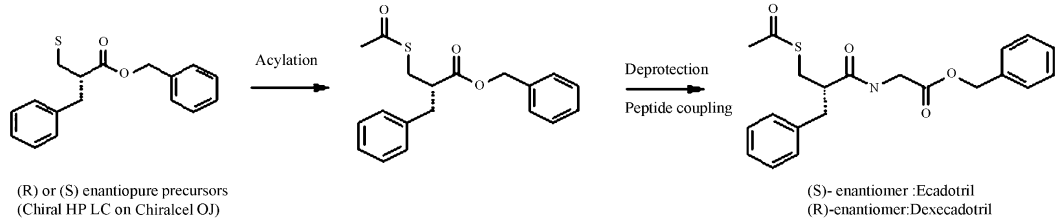


Fig. 8. A new strategic synthesis of ecadotril and dexecadotril enantiomers based on a chiral preparative separation of a precursor found in ChirBase chiral pool.

On the basis of this searching strategy, we could propose a new chiral synthesis of ecadotril and dexecadotril which are both developed as therapeutic drugs. It appears today that the access to both pure enantiomers remains a challenge whatever is the chiral technology employed (chiral pool, chemical or enzymatic resolution as well as asymmetric synthesis) [39]. Starting from the chiral precursor found in ChirBase as displayed in Table 2, we may suggest a new strategy based on the use of chiral HPLC to prepare each single (*R*) and (*S*)-enantiomers (see Fig. 8).

4. Conclusion

In this contribution, we have tried to illustrate some characteristic features of the chiral pool accessible from chi-

ral chromatography. Analyses of the property distributions observed in ChirBase chiral pool revealed a chemical profile of compounds similar to that of collection of commercially available starting materials. In addition, a significant number of druglike (80%) and leadlike (60%) chiral compounds have been identified. As diversity is also a key concern when dealing with compound libraries, we performed a comparative study of the number of single molecular frameworks identified in Chirbase chiral pool and a commercial compound library. As expected, our findings showed that our chiral library has the potential to expand and enrich the scope of the discovery route's toolbox.

To highlight more exhaustively the potential resources of chiral chromatography, we also built a virtual chiral combinatorial library showing that novel chiral lead compounds may be easily generated and screened for drug development. In addition, to complete this study, we have shown that a variety of additional small starting chiral materials could be released from the current chiral pool and supplement the existing chiral library with 20,000 new compounds. As a test of whether the proportion of interesting precursors for chiral chemistry is significant, using subsimilarity strategies we have searched and selected in ChirBase some starting materials that may be suitable for the small-scale synthesis of known chiral drugs.

References

[1] S. Erb, Pharm. Technol. (October 3, 2006).
[2] V. Farina, J.T. Reeves, C.H. Senanayake, J.J. Song, Chem. Rev. 106 (2006) 2734–2793.
[3] H.U. Blaser, F. Spindler, M. Studer, Appl. Catal. A 221 (2001) 119–143.
[4] J.S. Carey, D. Laffan, C. Thomson, M.T. Williams, Org. Biomol. Chem. 4 (2006) 2337–2347.
[5] J. Bojarski, H.Y. Aboul-Enein, A. Ghanem, Curr. Anal. Chem. 1 (2005) 59–77.
[6] C. Yamamoto, Y. Okamoto, Bull. Chem. Soc. Jpn. 77 (2004) 227–257.
[7] G. Gubitz, M.G. Schmid, Biopharm. Drug Dispos. 22 (2001) 291–336.
[8] N.M. Maier, P. Franco, W. Lindner, J. Chromatogr. A 906 (2001) 3–33.
[9] G.B. Cox, Preparative Enantioselective Chromatography, Blackwell Publishing Ltd., Oxford, 2005.
[10] T.D. Nelson, C.J. Welch, J.D. Rosen, J.H. Smitrovich, M.A. Huffman, J.M. Mcnamara, D.J. Mathre, Chirality 16 (2004) 609–613.
[11] M. McGuire, Process Chemistry in the Pharmaceutical Industry, Saddlebrook Resort, Tampa, FL, 2004.
[12] M.D. Wallace, M.A. McGuire, M.S. Yu, L. Goldfinger, L. Liu, W. Dai, S. Shilcrat, Org. Process Res. Dev. 8 (2004) 738–743.
[13] D. McCormick, Pharm. Technol. (May 2, 2006).
[14] G.J. Quallich, Chirality 17 (2005) S120–S126.
[15] A. DelRio, P. Piras, C. Roussel, Chirality 18 (2006) 498–508.
[16] C. Roussel, J. Pierrot-Sanders, I. Heitmann, P. Piras, in: G. Subramanian (Ed.), Chiral Separation Techniques, second ed., Wiley-VCH Verlag GmbH, 2001, pp. 95–125.
[17] Current number of entries is available at http://chirbase.u-3mrs.fr.
[18] BARK Information Services (http://www.chiraldata.com).
[19] The Hanessian group, Université de Montréal, Montréal, Québec (http://osiris.corg.umontreal.ca).
[20] http://www.sigmaaldrich.com.
[21] A. Monge, A. Arrault, C. Marot, L. Morin-Allory, Mol. Diversity 10 (2006) 389–403.
[22] ChemAxon Kft. Máramaros köz 3/a Budapest, 1037 Hungary (http://www.chemaxon.com).
[23] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Adv. Drug Deliv. Rev. 23 (1997) 3–25.
[24] T.I. Oprea, J. Comput. Aided Mol. Des. 14 (2000) 251–264.
[25] T.I. Oprea, A.M. Davis, S.J. Teague, P.D. Leeson, J. Chem. Inf. Comput. Sci. 41 (2001) 1308–1315.
[26] H.J. Verheij, Mol. Diversity 10 (2006) 377–388.
[27] D.F. Veber, S.R. Johnson, H.-Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, J. Med. Chem. 45 (2002) 2615–2623.
[28] C.A. Lipinski, J. Pharmacol. Toxicol. Methods 44 (2000) 235–249.
[29] A.R. Leach, M.M. Hann, J.N. Burrows, E.J. Griffen, Mol. BioSyst. 2 (2006) 429–446.
[30] M. Congreve, R. Carr, C. Murray, H. Jhoti, Drug Discov. Today 8 (2003) 876–877.
[31] T. Lengauer, C. Lemmen, M. Rarey, M. Zimmermann, Drug Discov. Today 9 (2004) 27–34.
[32] L. Xue, J.W. Godden, J. Bajorath, J. Chem. Inf. Comput. Sci. 39 (1999) 881–886.
[33] I. Muegge, S. Oloff, Drug Discov. Today: Technol. 3 (2006) 405–411.
[34] D.C. Fara, T.I. Oprea, E.R. Prossnitz, C.G. Bologa, B.S. Edwards, L.A. Sklar, Drug Discov. Today: Technol. 3 (2006) 377–385.
[35] A. Arrault, A. Monge, C. Marot, L. Morin-Allory, 10th Electronic Computational Chemistry Conference (ECCC10), April 2005 (http://eccc.monmouth.edu).
[36] X.Q. Lewell, D.B. Judd, S.P. Watson, M.M. Hann, J. Chem. Inf. Comput. Sci. 38 (1998) 511–522.
[37] MDL Information Systems, Inc. (http://www.mdli.com).
[38] B.A. Tounge, C.H. Reynolds, J. Chem. Inf. Comput. Sci. 44 (2004) 1810–1815.
[39] T. Monteil, D. Danvy, M. Sihel, R. Leroux, J.-C. Plaquevent, Mini Rev. Med. Chem. 2 (2002) 209–217.