Facts, Figures and Trends in Lead Generation

Rebecca Deprez-Poulain and Benoit Deprez*

UMR CNRS 8525, Institut Pasteur, Université de Lille 2, Faculté des Sciences Pharmaceutiques et Biologiques, 3 rue du Pr Calmette, F-59006 Lille Cedex, France

Abstract: The goal of this paper is to review the variety of approaches adopted to improve lead generation, and make the process easier for the chemist, faster and more likely to succeed in later phases of drug development. Our analysis shows that successful lead generation requires not only an accurate definition of the needs (to define the most relevant assay protocols and readouts), but most of all a good hit as a starting point. It also appears that teams where techniques are combined are more successful in that difficult game.

Key Words: Hit-to-lead (H2L), optimisation, High Throughput Screening (HTS), parallel synthesis, library design, Structure-activity relationships (SAR), Structure-property relationships (SPR), functional assay, druggability, rational drug design.

INTRODUCTION

To illustrate the interests at stake during hit-to-lead, and the difficulties faced by the industry to generate innovative new drugs, one can have a look at recently published figures.

Only 25% of the projects initiated on targets defined as "druggable" actually deliver a lead [1]. According to a recent survey in High Throughput Screening (HTS) groups [2], out of 326 leads issued from HTS-generated hits, only 62 have reached man. In this survey the lead definition was probably less stringent that the one used here and this explains the apparent high attrition rate (lead-to-man attrition: 81%). For every 7,000,000 compounds screened (2000), only one product is marketed, with a lot early dropouts [3]. 40% of the New Chemical Entities (NCE) are rejected due to poor pharmacokinetic properties and 11% for toxicity (2000) [3]. 120,000 "quality" compounds should be screened to find compounds consisting of a real lead series for a therapeutically sound target (1998) [3].

However if high throughput technologies and in particular HTS are often pointed as the major reasons of the low success rate in drug discovery, HTS is undisputedly becoming the major source of leads in industry. For example, at Abbott[1], most of the leads now derive from hits identified in HTS campaigns, as illustrated in (Fig. (**1**)).

LEAD GENERATION: DEFINITION OF NEEDS AND CONSTRAINTS FOR THE MEDICINAL CHEMIST.

Not only driven by a therapeutic need, drug discovery is part of an industrial process, and must be accompanied with considerations on discovery, development and production costs, patient population and route(s) of administration. However, it often begins with a screening step (classical HTS, virtual screening, fragment-based approach…), remotely linked to these considerations. Critical steps in current drug discovery are therefore the experiments and concepts that transform a primary hit into the molecule entitled to enter clinical trials (Fig. (**2**)).

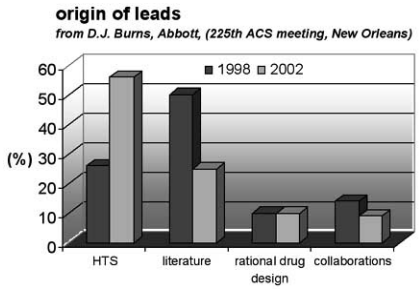

Fig. (1). Origins of leads at Abbott laboratories, in 1998 and 2002.

Hit

The concomitant births of HTS and parallel synthesis, led to defining new terms such as « hit » as the « library component whose activity exceeds a predefined, statistically relevant threshold »[4-5]. More specifically, to distinguish a hit from a simple "positive" and an "active" (Fig. (**2**)), the term "hit" will be used in this review with the acceptance of a compound presenting a confirmed activity (on re-synthesized sample, purity and structure checked) and having reasonable patenting potential. Additionally, thanks to screening of combinatorial libraries, a hit may already be part of a small chemical family displaying rough SAR. Ideally it would show no sign of toxicity (predicted or measured), and a certain degree of selectivity among subtypes of the target, if applicable.

*Address correspondence to this author at Faculte des Sciences Pharmaceutiques et Biologiques, 3, rue du Prof. Laguesse, 59000 Lille,, France; Tel: +33 (0)3 20 96 40 24; Fax: +33 (0)3 20 96 47 09; E-mail: bdeprez@pharma.univ-lille2.fr

[1] D.J. Burns, 225th ACS National Meeting, New Orleans, LA, March 23-27,2003.

1568-0266/04 $45.00+.00

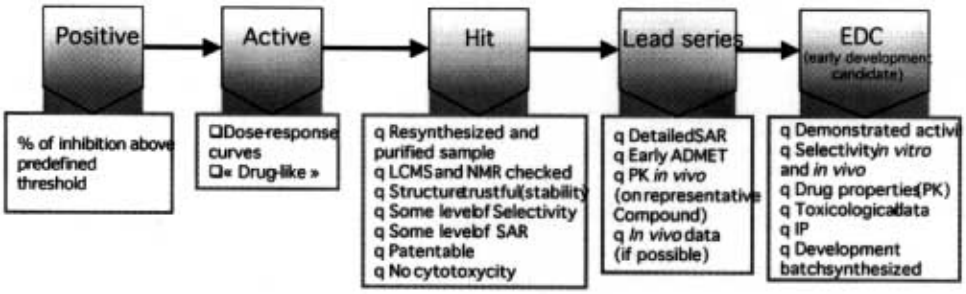

Fig. (2). Drug-Discovery process in large pharmaceutical companies.

To allow the synthesis of analogues displaying varied pharmacophore and scaffold, those hits need to be accessible via a strong retro-synthetic path, applicable to a wide variety of starting materials. Hits devoid of a robust chemical accessibility, such as complex natural compounds are often ranked low during hit selection.

Clinical Candidates and Drugs

A long debate as emerged for the establishment of structural criteria, especially in virtual screening, for the grading of compounds regarding their "ability" to become a drug. A number of groups have tried to identify such criteria among pools of marketed drugs or clinical candidates. Amongst all required compound-related properties for a safe and efficacious drug, a premium has rapidly been put on pharmacokinetics over toxicity issues. Indeed it appears that a number of relatively simple, structure-based rules can guide the chemist into safe grounds in terms of absorption and cell permeation. The rationalization of absorption has been pioneered by Lipinski whose work is condensed into the "Rule of 5" [6]. This rule stipulates that poor absorption and permeation are likely to occur when at least two parameters from the following, lie over an empirically defined threshold: MW (> 500), cLogP (> 5), HBD (> 5), HBA (> 10) [2].

Other criteria taken into consideration by medicinal chemists to prioritize (and sometimes to dismiss) compounds are the number of exocyclic methylene groups, a count for flexibility, and the total number of halogens [7].

A recent paper has linked membrane permeation and oral bioavailability to flexibility and polar surface area (PSA), independently of molecular weight [8]. These simple descriptors, easily calculated, can also be included in the "Drug-likeness" definition.

Besides these PK-related rules, some medicinal chemists consider some groups, such as Michael acceptors and nitro-groups, as bearing unacceptable liabilities in terms of toxicity. They thus dismiss series where such groups play a strong role in the interaction with the receptor.

All these guidelines have emerged thanks to a rapidly growing ability to handle large sets of structures *in silico* and to the electronic storage of biological data over several decades. These rules are becoming even more useful with the generalization of HTS and parallel synthesis since a huge amount of data need to be handle, classified, prioritized and most of all reduced to the most valuable pieces of information.

Leads and Candidates for Lead Optimization

As illustrated by Fig. **2**, a lead displays structural and biological properties that place it somewhere between a hit and a clinical candidate. IUPAC definitions [9] of Lead Generation and Discovery reflect this: *"Lead Discovery is the process of identifying active new chemical entities, which by subsequent modification may be transformed into a clinically useful drug". "Lead Generation is the term applied to strategies developed to identify compounds which possess a desired but non-optimized biological activity"*

It shows that a lead is not a fully optimized structure towards each desired property, but yet bound to undergo successful optimization: the potential to fulfil all stereoelectronic, physicochemical, pharmacokinetic and toxicological properties required for clinical usefulness via structural modifications is the most important trait of the lead compound [9].

Teague *et al*. have long tried to define terms and distinguish between drugs and leads [1,7]. They have helped medicinal chemist to set criteria required for a molecule to become a lead. According to their analysis, a lead: i. is patentable; ii. is chemically tractable; iii. has SAR outlined; iv. has a known mode of action; v. is cell active; vi. has drug-like property (such as a good ADME profile). He observes that very often, optimization of *in vitro* potency and selectivity is achieved by appending additional substituants to the original molecule. Therefore, in order to reserve room for improvement, he recommends that the lead compound should be smaller and less lipophilic than accepted by Lipinski.

Recently however, Proudfoot [10], by analyzing recently launched drugs and their disclosed parent lead, has shown that drugs seem to be more closely related to their parent lead. His conclusion is that a lead can be more drug-like than commonly accepted, i.e. "bigger and fatter" than Teague recommended. Another way of analysing this observation, which is more consistent with Teague's pioneer and widely accepted work, is that the constraints on lead structures (applying between lead and drugs) are higher because lead optimization is multiparametric and thus highly difficult. Indeed, improvements of one property thanks to (even small)

[2] MW : Molecular Weight, cLogP : calculated octanol/water partition coefficient, HBD : Donors of Hydrogen Bonds (NH and OH), HBA : Acceptors of Hydrogen Bonds (N and O).

chemical changes can be deleterious for another property. Great structural steps preserving all previous gains in all properties become more difficult as the project advances (Fig. (**3**)). *De facto*, leads are often structurally similar to drugs. Another important consequence of that situation is that hit-to-lead procedures must be very productive to give the largest choice before the final steps of optimization. Ultimately, hit structures must be very open to analoging (small and chemically tractable), to generate the largest diversity early in the project.

Table **1** summarizes the needs and constraints for the medicinal chemist to define a molecule as a lead, starting from Teague's work. Medicinal chemists now consider a series as a "lead series", if one member has additional *in-vivo* and early ADMET data available[3]. A consequence of the requirement for ADMET results is illustrated by the fact that most PK groups have incorporated discovery teams in large companies, while they were before part of the development teams [11].

HIT-TO-LEAD PROCEDURES

Hit-to-lead procedures used by medicinal chemists has long been rather mono-dimensional, taking into account primarily activity (IC_{50}, K_i, on the target and its close subtypes) to choose the lead series and the best representative molecule. Nowadays hit-to-lead strategies tend to incorporate, not only at the development stages, but also during the early discovery steps, the highest possible diversity of biological and structural parameters. The optimization steps performed by medicinal chemists have shifted from an activity-based procedure to a property-based strategy inside multidisciplinary teams. Thus instead of the classical Structure-Activity Relationships (SAR), the medicinal chemist should provide Structure-Properties Relationships (SPR) for his series.

Table 1. Needs and Constrains for a Lead Structure (Criteria are Ranked According to the Authors' Opinion).

NEEDS	CONSTRAINS
Detailed Structure-Activity	Chemically tractable
Mode of action (agonist, antagonist,…)	Patentable
Cell activity	Medium molecular weight and lipophilicity
In-vivo activity	
eADMET	

As examples of properties of the initial hit structure to optimize (or at least to monitor) in parallel, are i. the functional activity (agonistic or antagonistic activity on receptors such as GPCR or Nuclear Hormone Receptors) and cell activity; ii. the selectivity among subtypes (if appropriate) and specificity versus other targets; iii. the synthetic route for development; iv. the early ADME data. These changes in the medicinal chemist work implies that, taking into account all properties together, some properties are allowed to remain sub-optimal in order to have an overall acceptable property profile for the lead series. The relative importance of these properties for optimization remains to be debated for each new project and at several time points in each project.

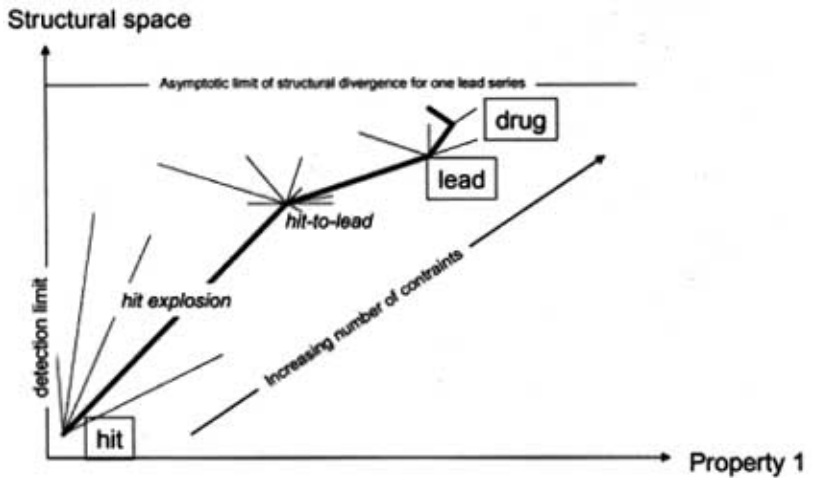

Fig. (3). Walk on the route from hit to lead. Property 1 illustrated here is essential at early stages but become secondary when the compounds is to be selected for clinical trials (this is often the case for activity on the molecular target). The figure also shows that leads are *de facto* very similar to drugs, as Proudfoot observes.

Library Design for Hit Optimization

The use of small, well-designed libraries can be beneficial for identifying lead series.

[3] early ADMET : early Adsorption, Distribution, Metabolism, Excretion and Toxicology data. Usually, these "early results" are obtained in-silico and/or on models (cell-based).

i. Selection of Building-Blocks or Products

Virtual screening can be used efficiently when searching database for analogues of the parent hit or for building-blocks analogues. Several teams have proposed algorithms and multipharmacophore descriptors to optimize hit (or lead), in order to find compounds sharing a similar 3D-arrangements of pharmacophore elements to the hit and/or the building-blocks that compose the hit. A nice example deals with fibrinogen receptor antagonists using an RGD pharmacophore key to enhance the proportion of actives in the library selected for screening [12].

Similarly, Horvath *et al*. published a fast algorithm that retrieve from a virtual library comprising all feasible combinations of available starting materials, all the compound that display a pharmacophore similar to that of the parent hit (irrespectively of the initial compound backbone). Poulain *et al* optimize ligands of the μ-receptor [13] using this method in combination with another method based on the synthesis of compounds having a chemical topology similar to that of the hit compound.

A recent publication also describes, though not disclosing structures, the selection of analogues of hits and the synthesis of libraries to inhibit protein-protein interactions (here Bcl-xL/Bak interaction) [14]. The building-block sets used for the library synthesis should always comprise both historical reagents and specifically designed reagents, taking into account medicinal chemistry experience and project-specific requirements.

ii. Designing the Library: the Use of Parallel Synthesis and the Combinatorial Approach

When designing libraries for hit optimization, it is often useful not to limit the library to the combination of the best fragments but to make *all possible* combinations of analogues of fragments of the hit(s) since many biological properties do not behave linearly in function of chemical structure.

Thus, hit-to-lead can benefit from combinatorial design and parallel synthesis. It is sometimes difficult to synthesized libraries 100% combinatorial. Pickett and co-authors have developed a Monte-Carlo search procedure to select a near-combinatorial subset that fulfils the design criteria, including Lipinski rules and PSA calculations, while enhancing the synthesis efficacy [12]. These combinatorial approaches have the advantage to rapidly provide rather comprehensive Structure-Properties Relationships since all combinations are tested in one campaign. Pickett and co-workers describe the synthesis of TNFα inhibitors with improved absorption properties while retaining a close-to-combinatorial architecture.

Teague *et al*. have also reconsidered the use of fully combinatorial libraries using multi-step synthesis to generate compounds with a high molecular weight [1]. Several groups have focused on the use of building-blocks assembled around small templates in a one or two step procedure [1].

Selectivity and Specificity

i. Selectivity

The need for selectivity is a critical issue, which has long been successfully tackled by medicinal chemists. Numerous groups try to evaluate the structural requirements for molecules to bind a target and not its close subtype, in order to predict the selectivities of compounds prior to testing. Some recent examples discuss the selectivity among enzymes of the same family. Recanatini *et al*. used docking experiments with two inhibitors to identify the parameters associated with affinity for a panel of cytochromes P450 [15]. Gupta *et al*. were interested in selectivity among MMPs (Matrix Metalloproteinase), as well as ChC (*Clostridium histolyticum* collagenase) [16]. They used a QSAR approach to evaluate the structural criteria distinguishing the enzymes, thanks to a set of hydroxamic acid-based inhibitors.

ii. Specificity

Poulain *et al.* have explored the field of Structure-Properties Relationships, by exploring specificity of compounds on a panel of diverse receptors and channels [17]. The study has underlined the fact that specificity is an important criterion to track, while optimizing a hit. However they show that although analogues of non-specific compounds are more likely to be non-specific, minor changes in the chemical structure (as a medicinal chemist would do) can change drastically the specificity of a compound, allowing to reach specificity starting from a "promiscuous" compound. A systematic evaluation of the pharmacological profile of early analogues has proven to provide directions to fine tune the structures and define the best optimization routes.

A recent publication shows that similar molecules have similar biological activity though the similarity is very dependent on the fingerprints (2D) used [18]. The authors point out also that targets can distinguish between two molecules that are similar according to structural fingerprints or to the medicinal chemist eye.

Early ADME

In this domain, the medicinal chemist takes advantage of both experimental and virtual procedures [11]. It is now admitted that activity testing must be linked to physico-chemical properties evaluation (especially when cell-based assays are involved in the qualification procedure). Evaluation of cLogP and PSA (polar surface area) are examples of properties used for the prediction of oral absorption, blood brain barrier crossing and cell availability.

Recently, a group has published a method for predicting the volume of distribution of drugs using experimental logD and plasma protein binding [19]. This method can be automated and can thus provide for numerous compounds this useful information. Using data from experiments of oral bioavailability in rats, Veber *et al*. have pointed out the role of flexibility and the number of rotatable bonds in drugs [20]. They proved that, if one commonly admits that lower molecular weight increases oral bioavailability, one should take into account that a lower number of rotatable bonds, hydrogen bond counts and PSA, benefit to oral bioavailability. This, as they say, allows the discovery of new molecules with a higher molecular weight, but a lower flexibility, that would have been dismissed by "Rule-of-5".

Goodwin *et al*. studied the role of hydrogen-bonding potential and volume in passive membrane permeation [21].

In the meantime, experimental procedures and screening methods for permeation are constantly improved [22]. As for molecular descriptors, a recent publication reports the use of the same descriptors to predict simultaneously the ADME profile and the protein-ligand affinity at the early steps of discovery [23].

Fesik *et al*. have recently published a nice example of bioavailability optimization, prior to potency [24]. In a series of matrix metalloprotease inhibitors, the bioavailability of a hit has been improved by NMR-screening of different chelating fragments in order to replace the alkylhydroxamate moiety, responsible for the activity, but also for the poor PK properties. They identified the 1-naphtylhydroxamate moiety. After having checked its ability to both chelate the Zn^{2+} ion in the active site, and to enhance the bioavailability, they incorporated it as a common template in a focused library, varying the other fragments and linkers (Fig. (**4**)).

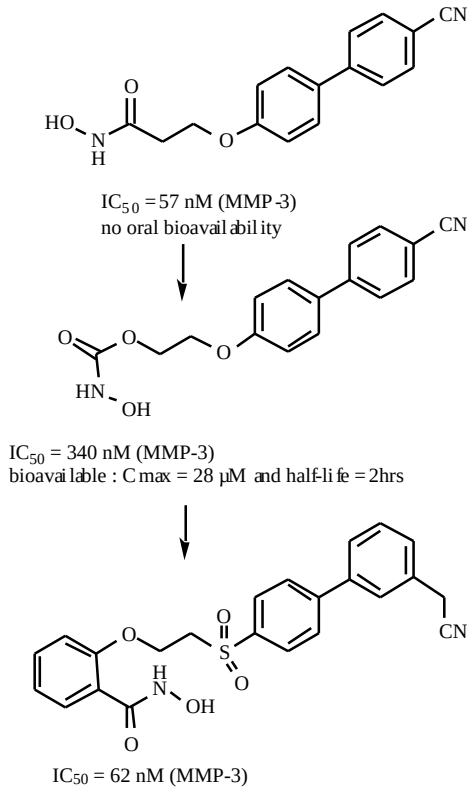

Fig. (4). Optimization of bioavailability prior to potency, by Fesik *et al*.

The Special Case of Peptides

Peptides, either toxins or endogenous ligands of target receptors, fulfil most of the previously cited requirements, with the exception of ADME criteria. Indeed, they are often very potent and selective agonists at their receptor site, the major drawbacks being poor ability to oral dosing and extremely short half-life in biological fluids. However, several, yet irreplaceable, drugs such as insulin and growth hormone, are of peptide nature. The constrains for the administration of peptides, although novel pharmaceutical devices and techniques may overcome this limitations, have hampered the development of peptides. This often leads to the rejection of peptides as hits or leads in the drug discovery process.

However, with the recent combination of several techniques and concepts, peptides might be worth reconsidered as hits and leads. To reflect this move, a full session of the medicinal chemistry meeting of the American Chemical Society in New Orleans was devoted to the use of peptides as starting hits and to techniques developed to turn peptide hits into viable leads.

Freidinger *et al*. at Roche have pioneered the use of peptides as hits by combining innovative chemistry with structural analysis and determination of conformation-activity relationships. By introducing in the peptide molecule aminoacid or dipeptide surrogates that are able to block the peptide in an active conformation they have been able to display the original peptide pharmacophore in a much more compact and rigid molecule. These stepwise modifications, made under careful analysis of conformations by NMR and/or RX structure determination have often improved significantly pharmacokinetics and in some case oral bioavailability [25]. Following the way paved by Freidinger, Lubell and his collaborators at the University of Montreal have initiated an interesting strategy relying on the constitution of a portfolio of conformationnally restricted aminoacids that are dipeptides analogues. These building blocks, once incorporated in a peptide sequence impose locally a precise geometry of both the peptide backbone and the side chains. The collection of 72 fused lactams systems being assembled by Lubell covers almost any naturally occurring dihedral angle in a peptide structure (Fig. (**5**)). They use these building blocks as cassettes in a combinatorial approach to assess conformation-activity relationships of bioactive native peptides. The first step is to identify the conformationnally critical positions in the peptide scanning the sequence with rigid lactams. Once the suitable lactam has been introduced, then they optimize the structure by introducing the desired side chain on the cycle. These reasoned modifications not only result in an improved activity but also in a prolonged half-life. Furthermore, the move to truly non-peptides from these conformationnally frozen compounds is likely to be much easier than from native peptide [26].

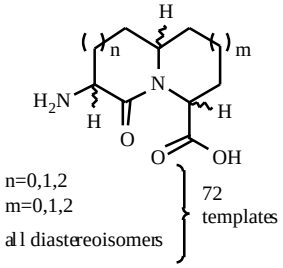

Fig. (5). Library of azabicyclo[x,y,0]alkanes used conformation ind ucers by Lubell *et al.*

In another direction, chemists at Novo Nordisk are developing GLP-1 peptide agonists of long duration of

action by acylating the peptide structure with fatty acids. An example of such lipopeptide currently in phase 2 clinical trials, is liraglutide which has a half-life of 13hours (compared to 1h for the native peptide), retains the potency of GLP-1 and allows for once daily dosing without protracted formulation.

Both types of approach should draw a renewed interest for peptides in medicinal chemistry.

4. STRATEGIES TO GET GOOD HITS OR, BIOLOGICAL-HTS AND ITS ALTERNATIVES

As one can see, hit-to-lead is still laborious since all the critical parameters cannot be optimized at once. Furthermore, experimental and virtual methods to predict or measure such parameters are so numerous and evolving, that it is difficult to be sure to have the right tool at the right time. Therefore in this complex and difficult context, it is of high interest to have the most competent starting points, that is, "good hits".

Until about 5-10 years ago, libraries tested in HTS mode have been rather large. They resulted from the collection of samples obtained from medicinal chemistry programs, combinatorial chemistry, and natural sources. It has often been pointed out that the compounds tested, and thus the hits obtained from synthetic compounds displayed a good chemical diversity but a high logP. Natural extracts, were always difficult to deconvolute and the compounds identified either already known or if new, often difficult to synthesize. All this often resulted in disappointing numbers of exploitable hits.

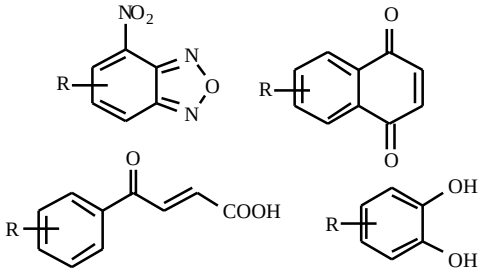

Fig. (6). Frequent hitters as defined by Roche, Gillepsie *et al*.

4.1. Intelligent Screening

The first discovery of an activity is thus very dependent on the nature of the library screened but also on the type of assay used for the primary selection. Also, since HTS can be expensive and a high number of compounds tested is not sufficient to identify exploitable hits, strategies intending to enhance the "value" of hits have emerged. These will be discussed in this part.

HTS groups are often part of the lead-discovery team in which they interact with medicinal chemists. Every key aspects of the initial HTS must result from a concerted choice involving biologists and chemists: the compound file to be tested (in function of the target class) [3], compound pooling options, assay formats, compound dissolution, concentration, number of replicates, instrumentation and finaly assay technology (radioactivity, fluorescence, calorimetry, biochemical or cell-based), in function of both the target class and the compound file chosen.

i. Test Fewer Compounds

It is sometime advantageous to trade off the number (and even diversity) of compounds tested against the relevance and content of information of the assay. Handling a smaller collection of compounds, pharmacologists can afford to design more informative assays, or assays where the target is put in a relevant biological context. For example kinases or proteases can be screened in cells, with their natural protein partners, and therefore correctly folded, activated and processed. The most relevant binding sites and target-state populations are thus presented to the compounds.

ii. Screening at High Concentrations

The choice of the concentration of the compounds for the primary biological test is treated differently in every lead-discovery team. It is a balance between the positive-rate and the relevance of the concentration. Obviously it depends on both the type of screen and the biological target. Putting aside the false positives that can result from "promiscuous compounds", one can say that 10 μM is a usual concentration. Lower concentrations are sometimes used to facilitate hit confirmation (picking, resynthesis, analysis) and prioritization. In that case, 1 to 3 μM seem good concentrations [3].

On the contrary, the choice of testing at high concentrations has been made by the groups doing X-ray or NMR fragment-screening.

David Burns from Abbott demonstrates in his presentation at the 225th ACS meeting that with higher concentrations of compound in the assay, a higher number of chemotypes are identified. Therefore with an increased initial diversity in hand, the room for selection is larger and probability to find a high quality hit is higher.

Underlying the choice of the test concentration is the choice of the threshold for the selection of a positive compound. We have experienced by ourselves in GPCR binding experiment (as it was evidenced by Curatolo) that a percentage of inhibition of 80% is a relevant threshold to distinguish between compounds having a lower or a higher IC_{50} than 2 μM [5].

iii. Different Assay Conditions Give Different Hit Rate and Diversity

Matthew A. Sills *et al* have compared three different screening technologies applied to a tyrosine kinase assay [27]. They have screened 30,000 compounds randomly picked from Novartis corporate library. The three methods are based on the measurement of a residual enzymatic activity. To measure that activity, Scintillation Proximity Assay (SPA), Fluorescence Resonance Energy Transfer (FRET) and Fluorescence Polarisation (FP) have been used to measure the concentration of the reaction product, in a binding experiment. Statistically speaking, the three methods have the same accuracy (equivalent Z' factors). However, they deliver different, partially overlapping sets of hits. FRET and FP technologies give well-correlated results,

while a lower correlation is obtained with SPA. A straightforward conclusion would be to select either FRET or FP for the HTS campaign, because they give more self-consistent results. However SPA gives a higher diversity of hits and for that reason, the authors consider it as the technology of choice for kinase screens.

iv. Screen in a Functional Assay to Find Compounds with the Desired Effect

Rules for agonism and antagonism are often contradictory and shifting a compound from antagonist to agonist is not always possible especially if the primary screen is based on competition against a natural agonist. Therefore, there is a clear benefit to initiate the project with a screen that relates closely to the desired mode of action. Several biotech companies have in the last decade developed tools to assess in HTS mode not only the affinity but also the efficacy of compounds. These technologies rely on reporter-gene or second messenger measurements and identify hits that modify the receptor conformation in the requested way. Hit-to-lead is then eased because the chemist will not have to change the mode of interaction of his compounds with the target, while trying to optimise the other properties.

Kenakin and Onaran, have published a statistical analysis of GPCR-ligand interactions relying on the fact that any given ligand modify the conformation ("micro-states") distribution of the receptor population at the cell surface [28]. According to their model, the particular redistribution of microstates induced by the binding of a given ligand leads to a specific activity. They suggest that old ligands should be studied in various assay conditions to identify new activities such as inverse agonism on constitutively active receptors, or induction / prevention of receptor internalisation.

To exploit further that concept, Arena pharma has developed a technology relying on the constitutive activation of the GPCRs by molecular modification. That technology allows the identification of agonists and antagonists in the same biological setup, even in the absence of an endogeneous agonist [29-30].

v. Screening of a Given Library in Several Assays

At the ACS meeting in New Orleans, David Burns from Abbott also showed the advantage of screening the same compounds in a variety of assays to eliminate promiscuous compounds ("frequent hitter") as soon as possible from the project. He recommends to screen all available targets against all available compounds to identify reliably these nuisance compounds. To reach the throughput required by this strategy, in addition to classical plate-based assays, they devised a technology named µARCS, a micro-array gel-based screening format of very high density that generates very fast with very little material and handling, millions of data points. They also combine the information provided by these bioassays with affinity-based assays employing a mass spectrometric readout that identify promiscuous compounds.

vi. Screening Mixtures

Screening mixtures has always been controversial. If mixtures obtained by combinatorial procedures in the same reactor have been once screened, medicinal chemist now prefer to pool post-synthesis and prior to biological testing, sometimes using orthogonal pooling methods. They use parallel synthesis to avoid all the pitfalls mixture synthesis, and pool from 2 to 10 pure compounds. In these conditions screening is up to 10 times cheaper and deconvolution still easy.

Recently, teams at Agouron and Pfizer have developed a fractionation method to identify the active compound in a mixture, either in a combinatorial or in an post-synthesis made mixture, based on the use of HPLC-MS [31].

4.2. Assemble a Set of Competent Compounds for Screening

One can distinguish three trends in the methods for compounds selection before biological screening.

No selection or random selection is affordable in virtual screening that has the advantage over real HTS to be faster and less expensive.

Rational selection can be made when structural data on the target are available. In that case chemists and modellers will assemble so-called "targeted" or "directed" libraries.

Historical data such as information on side-effects of a known drug or candidate [32] can be very helpful. Recognition of frequently active templates (natural of synthetic) can also give guidelines for the compound selection. All these approaches are discussed in the following paragraphs.

i. Selection Rules

The "Rule of 5" of Lipinski is widely used to select molecules to be tested. These criteria, though valid for a large proportion of drugs, since they were deduced from a set of clinical candidates and drugs, are not always relevant for hits or leads. Indeed, this Rule gathers exclusion criteria, which, if applied too early, may discard compounds that were valuable starting points. The Rule-of-5 can nevertheless be used to prioritize hits and/or lead series [7].

We have pointed out that Michael acceptors and other electrophilic species cannot be considered as leads. Consequently, they are bad hits as well and should be excluded from the screening file. At Sunesis however, chemists use covalent binding between an electrophilic group and a cystein in the vicinity of binding site, to artificially increase the affinity of their compounds above the detection threshold. Once interesting chemotypes are found, the electrophilic group can be removed or replaced by a less harmful substituant. Other compounds that should be discarded before biological or virtual screening are the ones that induce false positive due to their ability to form micelles. They end up being promiscuous compounds active in many screens independently of the target screened [33]. The lack of methods to predict this phenomenon remains nevertheless a problem.

At the ACS symposium in New Orleans, P. Gillespie from Roche presented his conclusions from the evaluation of various *in silico* prediction tools intended to accelerate the selection of compounds after or even before HTS. They

experience in a general manner that tools designed to calculate physico-chemical properties from structure are poorly reliable in the context of drug discovery.

a. Solubility Prediction

Although showed accurate in series of rather non-diverse compounds, these tools turn out to wrongly predict the solubility of hit compounds. Gillespie hypothesized that this would be due to the differences in nature between literature compound sets and the Roche collection.

Simple statistics showed that their compounds have a median molecular weight 200 unit higher than model compounds from literature, log P one unit higher, an average of 4 rotable bonds more and one more aromatic ring, all this signing a much higher complexity of the Roche compounds.

b. Drug-likeness Recognition

In a similar way, "drug-likeness" scoring tools are easily put at fault by sets of complex compounds. Indeed, since drugs are usually complex molecules, these tools often measure drug-likeness more by the presence in the molecule of favorable features than the absence of deleterious ones. The direct consequence of this is the selection by the software of feature-rich but too big molecules.

c. Recognition of Frequent Hitters

Roche *et al*, from Roche have published last year with Gillespie a screening method for the identification of "frequent hitters"[34]. One intrinsic difficulty of this concept lies in the various natures of these "frequent hitters". Indeed they can be really promiscuous compounds as well as molecules that interfere with the assay method, but also genuine "privileged structures", the latter being highly desirable in a screening experiment. In practice, Gillespie confesses that this tool is not systematically used at Roche.

d. "Tox" Alerts

Gillespie (225th ACS Meeting) reports that the software DEREKâ, commonly used to flag potentially toxic compounds automatically attributes more than one alert to more than 80% of their hits. For example, atorvastatin is predicted both carcinogenic and methaemoglobin forming.

In conclusion, none of these prediction tools is used in his groups and preference is given for compound selection, to experimental data and "expert opinion", bearing in minds the strengths and weaknesses of broad scientific consultation. One of his take-home messages was also the poor reliability of HTS. To take that into account, their Hit-to-lead process now includes a critical step after screening by which they retrieve all analogues (not previously screened as well as already screened compounds) having acceptable MW and logP values and retests them in secondary assay. It occurs that 92% of genuine secondary hits are false negatives of the primary hits. From there on, classical experimental data drive the lead selection process : existence of primary SAR, binding affinity, accessibility, solubility (>50µg/mL) and rat PK (F> 50%). Moreover, he points out that a significant proportion of there clinical candidates originate from false negative compounds in the primary HTS campaign!

iii. Privileged Structures

Evans introduced 15 years ago the concept of privileged structures [35-36] with the example of benzodiazepines and biphenyle compounds (Fig. (**7**)). These structures tend to display appropriate ligands for numerous diverse receptors, and further modifications (sometimes minor) of the structure can generate selective and specific ligands for the target of interest.

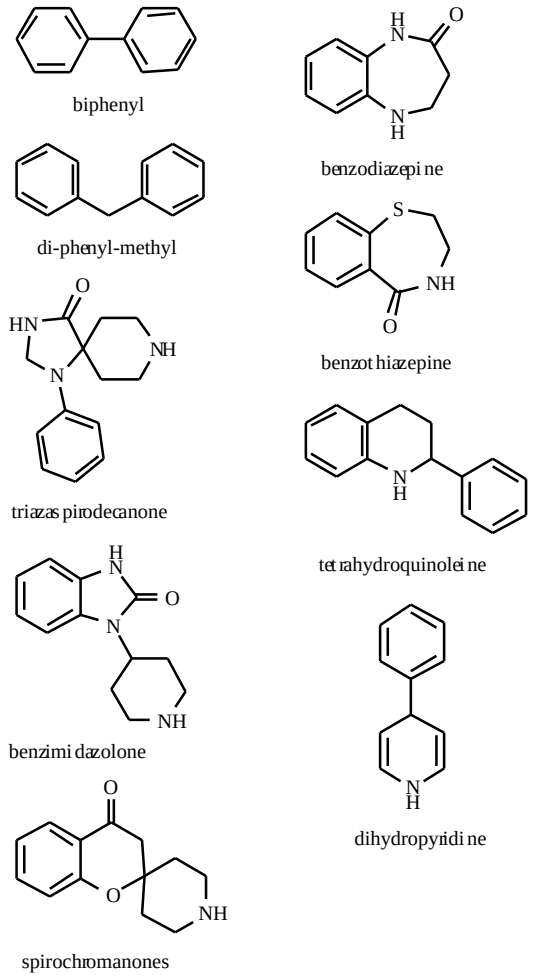

Fig. (7). Some privileged structures identified in drugs and/or in screening.

Medicinal chemist recognise easily some of these privileged structures thanks to their personal experience but during the last 5 years, a number of groups have provided statistical studies to make lists of privileged structures. For exemple, Murcko *et al*. have looked at drugs, using pharmacophore modeling techniques, to enlighten common frameworks (rings and linkers) [37] and functional groups [38]. These studies provide guidelines for molecular design, using the 32 substructures, present in about 50% of all known drugs.

NMR-screening has successfully been applied by Fesik to identify biphenyl and diphenylmethyl moieties as being

able to enhance the binding hit-rates (on a set of 11 target proteins) [39].

Figure (**7**) gathers the most commonly cited privileged structures and some other that we have identified in our work on high throughput profile [17]. Privileged structures can be incorporated in libraries thanks to appropriate building blocks that are joined by a simple, one or two steps, synthetic pathway. This strategy is used by Astra-Zeneca in their "Lead Generation Team" to provide lead-like compounds. Combinatorial chemistry has proved to permit the synthesis of diverse libraries to expand the possibility of finding active molecules. Keeping such privileged structures while incorporating diversity in the rest of the molecule to be tested has thus proven to be a good approach to discover valuable hits, as analyzed by Hann *et al*. [40].

iv. Natural Compounds and Natural Scaffolds

An alternative approach to sometimes disappointing hit-rates when screening synthetic diverse libraries is to tap into natural diversity. Screening collections of natural products has most of the time given interesting activities while identification of the active compound and/or synthesis of analogues remains difficult.

Schneider *et al.* have analyzed which natural structures were not present in trade drugs and have evidenced some new scaffolds that could be used in synthetic chemistry to help the discovery of new biologically active compounds. These natural product based combinatorial libraries could result from the combination of natural-product derived building-blocks and synthetic building-blocks [41]. Interestingly, Schneider *et al*. propose to select natural scaffolds with a lower count in Nitrogen atoms, since synthetic chemistry, and combinatorial chemistry in particular, has mainly focused on nitrogen chemistry.

Many groups are now developing strategies of parallel synthesis of the "natural privileged structures" as a basis for hit generation [42]. This allows not only to access to a natural structure but also to have access to analogues. Numerous examples, by Schreiber's team, make use of stereoselective reactions , C-C bond formation, on solid support or not, to access structures displaying both rigidity and stereochemical complexity to mimic natural structures synthetically [43]. Biaryl rings, related to pterocaryanin C, (Fig. (**8**)) have been tested on protein-binding assays, but also on phenotypic assays (development of zebrafish) in a chemical genomic approach [44].

An alternative approach to building libraries with "enhanced valuable diversity" thanks to privileged structures (inspired from Nature or not), is to take advantage of the information available on the target. Having an experimental 3-D model is the best situation, but knowing the general folding (thanks to sequence alignments within the target family) and guessing the residues forming the active site can also be valuable to drive compound design and selection. Screening targeted libraries and fragment-screening methods using 3D structure of the target are some examples.

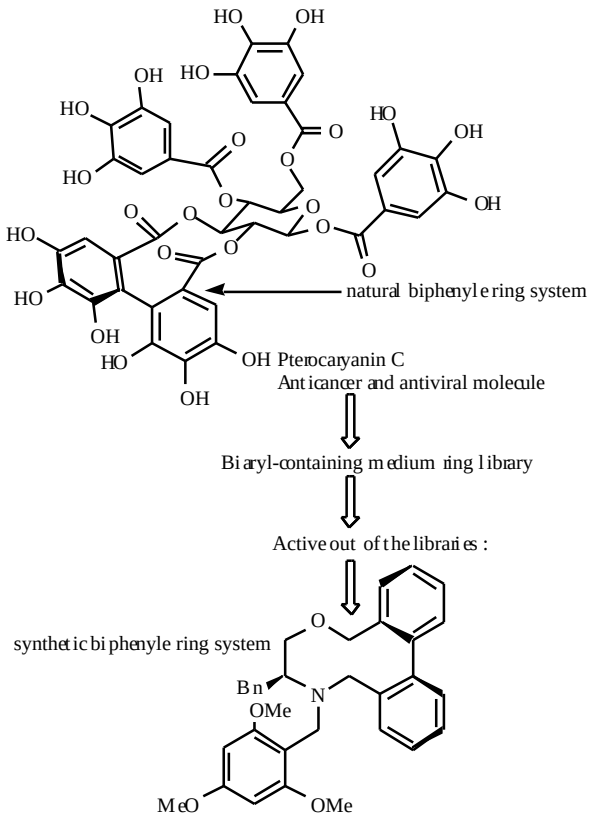

Fig. (8). Some examples of libraries and templates for natural structures synthesis by Schreiber *et al*.

iv. Targeted Libraries

About ten to fifteen years ago, companies were building up their libraries of compounds for HTS, gathering molecules from in-house collections, in-house libraries or commercial sources. The aim was to access to the greater diversity, a kind of "Universal library" as exposed by Combi. Chem. Inc. [42], to enhance the chance to get hits in any biological test.

As pointed out by Martin *et al*., an approach that would discard compounds too similar to a compound already in the library to be tested can be deleterious for the identification of actives [18]. A preferred compromise would be to test between 3 and 10 similar molecules. The concept of testing on any target always the same diverse libraries (sometimes enriched with new mini-libraries) one prefers now to test less "diverse" but well-designed libraries.

In his annual survey on combinatorial library synthesis [43], R. Dolle points out that a majority of publications about combinatorial libraries deal with targeted libraries towards specific families of receptors or enzymes. While keeping chemical positions where diversity can be included, one uses a unique functional group known to interact with the class of targets.

Ellman and co-workers published, for example, the design and synthesis of a library of 170 compounds displaying an hydroxylethyl amino group, known to interact

with aspartyl proteases (Fig. (**9**)). They have tested these compounds in a *Plasmodium falciparum* model [44].

v. Taking Advantage of the Structural Information on the Target

Another example of such strategy is used by Biofocus, that has founded a consortium for GPCR ligands discovery; their strategy is developed in a paper in the current issue. Apart from sequence and 3D-structure (when available), hit itself can also be considered as information on the target!

According to the general accepted fact that hits are a lot more simpler than the final drug, as far as molecular complexity is concerned, Glaxo-Smithkline has developed a strategy for testing very simple molecules to identify new leads [45].

The fragment-based method consists is screening small molecules, usually at a high concentration, to identify fragments of the future lead that bind the target.

This method can be implemented with various assay types. It is particularly useful in NMR and X-Ray based screening. NMR screening has been pioneered by Fesik and co-workers. Other groups have identified hits on FABP4 (h) [45]. This technique is improved regularly. Best methods now give both qualitative and quantitative information (binding constants) on the binding event [47].

Hits can also be identified by X-ray screening[4]. This technique provides moreover information on *how* and *where* the molecule binds to the target. Some companies, after identifying fragments that bind different but close regions of the target, synthetically link them to enhance the binding.

A nice example is given for the use of PRO-SELECT, for the *de-novo* design of potent and selective factor Xa inhibitors [48] (Fig. (**10**)).

vi. Dynamic Combinatorial Libraries

A new strategy has emerged recently, for the making of the compounds by the target itself. Based on a supramolecular approach, combinatorial libraries are designed as a dynamic set of building blocks that interact reversibly together to generate all the library components over time [4]. The target can then trap the preferred ligand (one of the many library components) [49]. Such approach,

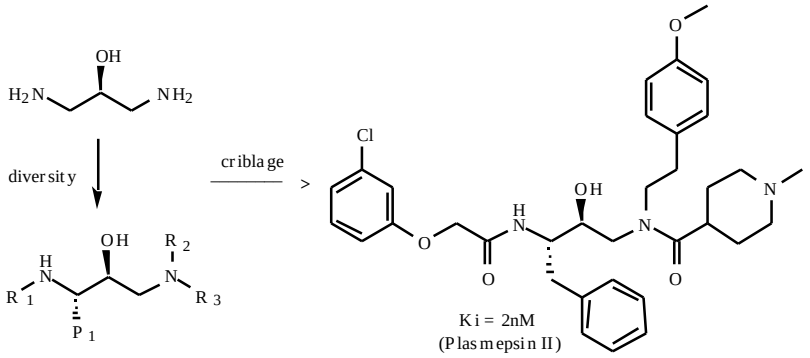

Fig. (9). Structure of the targeted library of *Pf*-aspartyl protease inhibitors by Ellman *et al*.

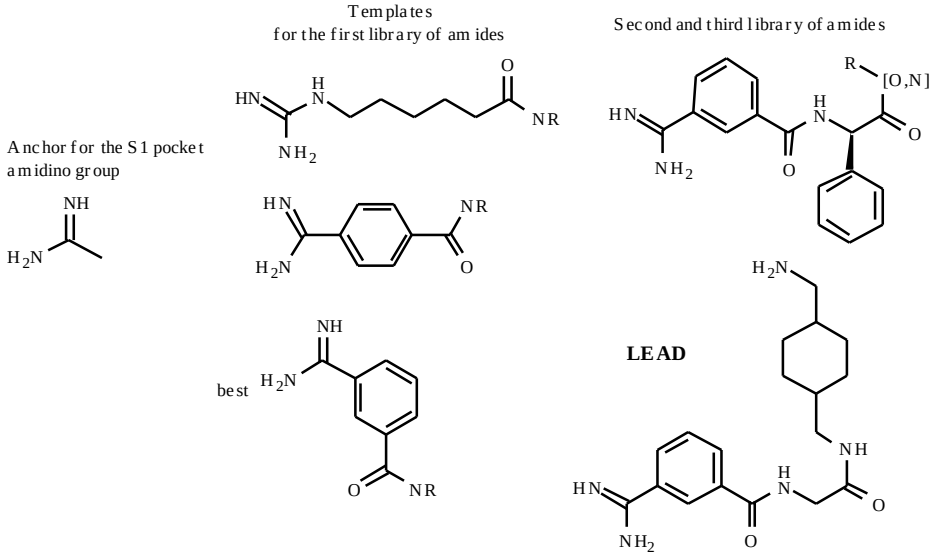

Fig. (10). Inhibitors of factor Xa by Hann *et al*.

though having identified ligands to several targets, needs improvements in the design of the libraries to be compatible with biological conditions such as aqueous media.

5. CONCLUSION

In the recent years, the pressure applied by the market and the shareholders on discovery teams has lead to an explosion of rules for the selection of compounds pre- or post-screening. These rules intend to lower the chances of late failure and to lower the costs of discovery. However it has been much easier to deduce exclusion rules from past failures than to devised positive criteria for inclusion of compounds. One must hope that the concept of privileged structure, combined with the maturation of structure-based design and the better understanding of drug-receptor interaction, will populate pharma libraries with competent series. A growing trend is also to consider the drug discovery process less linearly than previously, and to try to optimize all possible inputs in concerted efforts: HTS library, screening assay, optimization parameters. Thus if well used in concerted ways, the techniques described in this paper will surely provide a majority of valid hits from screening and reduce significantly hit-to-lead and hopefully hit-to-drug attrition.

ABBREVIATIONS

HTS = high throughput screening

PSA = polar surface area

MMP = Matrix Metalloproteinase

REFERENCES

[1] Oprea, T.I.; Davis, A.M.; Teague, S.J.; Leeson, P.D., Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308-1315.

[2] Fox, S. ; Wang, H. ; Sopchak, L. ; Farr-Jones, S. ; High throughput screening 2002, moving toward increased success rates. *J. of Biomol. Screen.* **2002**, *7*, 313-316.

[3] Landro, J.A.; Taylor, I.C.A.; Stirtan, W.G., Osterman, D.G.; Kristie, J.; Hunnicutt, E.J.; Rae, P.M.M., Sweetnam, P.M. HTS in the new millennium The role of pharmacology and flexibility. *J. Pharmacol. Toxicol.* Methods **2001**, *44*, 273-289.

[4] MacLean, D. ; Baldwin, J.J. ; Ivanov, V.T. ; Kato, Y. ; Shaw, A. ; Schneider, P. ; Gordon, E.M. Glossary of terms used in combinatorial chemistry, *J.Comb.Chem.* **2000**, *2*, 562-578.

[5] Curatolo, W. Physical chemical properties of oral drug candidates in the discovery and exploratory development settings, *PSTT* **1998**, *1*, 387-393.

[6] Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.

[7] Oprea, T.I. Virtual Screening in Lead Discovery: A Viewpoint, *Molecules* **2002**, *7*, 51-62.

[8] Veber, D.F. ; Johnson, S.R. ; Cheng, H-Y. ; Smith, B.R. ; Ward, K.W. ; Kopple, K.D. ; Molecular properties that influence the oral bioavailability of drugs candidates, *J. Med. Chem.* **2002**, *45*, 2615-2623.

[9] Wermuth, C.-G. ; Ganellin, C.R. ; Lindberg, P. ; Mitscher, L.A. Glossary of Terms used in Medicinal Chemistry (IUPAC Recommendations 1997). *Ann. Rep. Med. Chem.* **1998**, *33*, 385-395.

[10] Proudfoot, J.R. ; Drugs, Leads, and Drug-likeness : An Analysis of Some Recently launched Drugs. *Bioorg, Med. Chem. Lett.* **2002**, *12*, 1647-1650.

[11] For a recent review on pharmacokinetics : Van de Waterbeemd, H.; Smith, D.A.; Beaumont, K.; Walker, D.K. Property-Based design : Optimization of Drug Absorption and Pharmacokinetics. *J.Med.Chem.* **2001**, *44*, 1313-1333.

[12] Pickett, S.D.; McLay, I.M.; Clarck, D.E. Enhancing the Hit-to-Lead Properties of Lead Optimization Libraries. *J.Chem.Inf. Comput. Sci.* **2000**, *40*, 263-272.

[13] Poulain, R.; Horvath, D.; Eckhoff, C. ; Chapelain, B.; Bodinier, M.-C.; Deprez, B. From Hit to Lead. Combining Two Complementary Methods for Focused Library Design. Application to µ opiate ligands. *J.Med.Chem.* **2001**, *44*, 3378-3390.

[14] Stahura, F.L.; Xue, L.; Godden, J.W. ; Bajorath, J., Methods for compound selection focused on hits and application in drug discovery. *J. Mol. Graph. Model.* **2002**, *20*, 439-446.

[15] Cavalli, A.; Recanatini, M. Looking for Selectivity among Cytochrome P450s Inhibitors. *J.Med.Chem. (Letter)* **2002**, *45*, 251-254.

[16] Kumar, D.; Gupta, S.P. A Quantitative Structure-Activity relationship Study on Matrix Metalloproteinase and Collagenase Inhibitors. *Bioorg. Med. Chem.* **2003**, *11*, 421-426;

[17] Poulain, R.; Horvath, D.; Eckhoff, C. ; Chapelain, B.; Bodinier, M.-C.; Deprez, B. From Hit to Lead. Analyzing Structure-Profile Relationships. Application to µ opiate ligands. *J.Med.Chem.* **2001**, *44*, 3391-3401.

[18] Martin, Y.C.; Kofron, J.L.; Traphagen, L.M. Do Structurally Similar Molecules Have Similar Biological Activity? *J.Med.Chem.* **2002**, *45*, 4350-4358.

[19] Lombardo, F.; Obach, R.S.; Shalaeva, M.Y.; Gao, F. Prediction of Volume of Distribution Values in Humans for Neutral and Basic Drugs Using Physicochemical Measurements and Plasma Protein Binding Data. *J. Med.Chem.* **2002**, *45*, 2867-2876.

[20] Veber, D.F.; Johnson, S.R.; Cheng, H.-Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J.Med.Chem.* **2002**, *45*, 2615-2623.

[21] Goodwin, J.T.; Conradi, R.A.; Ho, N.F.H.; Burton, P.S. Physicochemical Determinants of Passive Membrane Permeability: role of Solute Hydrogen-Bonding Potential and Volume. *J. Med. Chem.* **2001**, *44*, 3721-3729.

[22] Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *J.Biomol.Screen.* **2001**, *6*, 189-196.

[23] Zamora, I.; Oprea, T.; Cruciani, G.; Pastor, M.; Ungell, A.-L. Surface Descriptors for Protein-Ligand Affinity Prediction. *J.Med.Chem.* **2003**, *46*, 25-33.

[24] Hajduk, P.J.; Shuker, S.B.; Nettesheim, D.G.; Craig, R.; Augeri, D.J.; Betebenner, D.; Albert, D.H.; Guo, Y.; Meadows, R.P.; Xu, L.; Michaelides, M.; Davidsen, S.K.; Fesik, S.W. NMR-Based Modification of Matrix Metalloproteinase Inhibitors with Improved Bioavailability. *J.Med.Chem.* **2002**, *45*, 5628-5639.

[25] for an example of cyclic analogue of somatostatine see : Veber, D.F.; Saperstein, R.; Nutt, R.F.; Freidinger, R.M.; Brady, S.F.; Curley, P.; Perlow, D.S.; Paleveda, W.J.; Colton, C.D.; Zacchei, A.G.; Tocco, D.J.; Hoff, D.R.; Vandlen, R.L.; Gerich, J.E.; Hall, L.; Mandarino, L.; Cordes, E.H.; Anderson P.S.; Hirschmann, R. A super active cyclic hexapeptide analog of somatostatin. *Life Sci.* **1984**, *34*, 1371-1378)

[26] Lubell, W.D. From Freidinger lactams toward a universal strategy for peptide mimicry. The importance of structural constraint for studying peptide function. 225th Meeting of the American Chemical Society, New Orleans 22-27 March 2003. For an example of the use of lactams in a series of ORL receptor ligands : Halab, L.; Becker, J.A.J.; Darula, S.; Tourwe, D.; Kieffer, B.L.; Simonin, F.; Lubell, W.D. Probing opioid receptor interactions with azacycloalkane amino acids. Synthesis of a potent and selective ORL1 antagonist. *J. Med. Chem.* **2002**, *45*, 5353-5357.

[27] Sills, M.A. ; Weiss, D. ; Pham, Q. ; Schweitzer, R. ; Wu, X. ; Wu, J.J. Comparison of assay technologies for a tyrosine kinase assay generates different results in high throughput screening. *J.Biol.Screen.* **2002**, *7*, 191-214.

[28] Kenakin, T.; Onaran, O. The ligand paradox between affinity and efficacy: can you be there and not make a difference? *Trends Pharmacol. Sci.* **2002**, *23*, 275-280.

[4] See Robin Carr's paper on Astex's strategy and technology, in this issue.

[29] Chalmers, D.T.; Behan, D.P. The use of constitutively active GPCRs in drug discovery and functional genomics. *Nat. Rev. Drug Discov.* **2002**, *1*, 599-608.

[30] Beeley, N.R. Can peptides be mimicked? *Drug Discov. Today* **2000**, *5*, 354-363.

[31] Phillipson, D.W.; Milgram, K.E.; Yanovsky, A.I.; Rusnak, L.S.; Haggerty, D.A.; Farrell, W.P.; Greig, M.J.; Xiong, X.; Proefke, M.L. High-Throughput Bioassay-Guided Fractionation: A Technique for rapidly Assigning Observed Activity to Individual Components of Combinatorial Libraries, Screened in HTS Bioassays. *J.Comb.Chem.* **2002**, *4*, 591-599.

[32] Teague, S.J; Davis, A.M.; Leeson, P.D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew.Chem.Int.ed.* **1999**, *38*, 3743-3748.

[33] McGovern, S.L.; Caselli, E.; Grigoriell, N.; Shoichet, B.K. A Common Mechanism Underlying Promiscuoous Inhibitors from Virtual and High-Throughput-Screening. *J.Med.Chem.* **2002**, *45*, 1712-1722.

[34] Roche, O.; Schneider, P.; Zuegge, J.; Guba, W.; Kansy, M.; Alanine, A.; Bleicher, K.; Danel, F.; Gutknecht, E.-M.; Rogers-Evans, M.; Neidhart, W.; Stalder, H.; Dillon, M.; Stjögren, E.; Fotouhi, N.; Gillepsie, P.; Goodnow, R.; Harris, W.; Jones, P.; Taniguchi, M.; Tsujii, S.; Von der Saal, W.; Zimmermann, G.; Schneider, G. Development of a Virtual Screening Method for Identification of "Frequent Hitters" in Compound Libraries. *J.Med.Chem.* **2002**, *45*, 137-142.

[35] Evans, B.E. ; Rittle, K.E. ; Bock, M.G. ; DiPardo, R.M. ; Freidinger, R.M. ; Whitter, W.L. ; Lundell, G.F. ; Veber, D.F. ; Anderson, P.S. ; Chang, R.S *et al*. Methods for drug discovery : Development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* **1988**, *31*, 2235-2246.

[36] For reviews on privileged structures see : (a) Patchett, A. ; Nerdung, R. Privileged structures-An update. *Annu.Rep.Med.Chem.* **2000** ,*35*, 289-298. (b) Horton, D.A.; Bourne, G.T.; Smythe, M.L. The combinatorial synthesis of bicyclic privileged structures or privileged substructures. *Chem. Rev.* **2003**, *103*, 893-930.

[37] Bemis, G.W.; Murcko, M.A. The Properties of Known Drugs. 1. Molecular Frameworks. *J.Med.Chem.* **1996**, *39*, 2887-2893.

[38] Bemis, G.W.; Murcko, M.A. Properties of Known Drugs. 2. Side Chains. *J.Med.Chem.* **1999**, *42*, 5095-5099.

[39] Hajduk, P.J.; Bures, M.; Praestgaards, J.; Fesik, S.W. Privileged Molecules for Protein Binding Identified from NMR-Based Screening. *J.Med.Chem.* **2000**, *43*, 3443-3447.

[40] Lewell, X.Q.; Judd, D.B.; Watson, S.P.; Hann, M.M. RECAP-Retrosynthetic Combinatorial Analysis Procedure: A Powerful New Technique for Identifying Privileged Molecular Fragments with Useful Applications in Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511-522.

[41] Lee, M.-L.; Schneider, G. Scaffold Architecture and Pharmacophoric Properties of Natural Products and Trade Drugs: Application in the design of Natural Product-Based Combinatorial Libraries. *J.Comb.Chem.* **2001**, *3*, 284-289.

[42] Wessjohann, L.A. Synthesis of natural-product-based compound libraries. *Curr. Opin. Chem. Biol.* **2000**, *4*, 303-309.

[43] (a) Tan, D.S.; Foley, M.A.; Shair, M.D.; Schreiber, S.L. Stereoselective Synthesis of over Two Milllion Compounds Having Structural Features Both Reminiscent of Natural Products and Compatible with Miniaturized Cell-Based Assays. *J. Am. Chem. Soc.* **1998**, *120*, 8565-8566. (b) Spring, D.R.; Krishnan, S.; Schreiber, S.L. Towards Diversity-Oriented, Stereoselective Syntheses of Biaryl- or Bis(aryl)metal-Containing Medium Rings. *J. Am. Chem. Soc.* **2000**, *122*, 5656-5657.

[44] Spring, D.R.; Krishnan, S.; Blackwell, H.E.; Schreiber, S.L. Diversity-Oriented Synthesis of Biaryl-Containing Medium rings Using a One Bead/One Stock solution Platform. *J. Am. Chem. Soc.* **2002**, *124*, 1354-1363.

[42] Saunders, J. Screening the universal informer Library. CHI Meeting : Exploiting Molecular Diversity, Coronado, 2-4 March 1998.

[43] Dolle, R.E. Comprehensive survey of Combinatorial Library Synthesis : 2000. *J.Comb.Chem.* **2001**, *3*, 1-41.

[44] Haque, T.S. ; Skillman, A.G. ; Lee, C.E. ; Habashita, H. ; Gluzman, I.Y. ; Ewing, T.J.A. ; Goldberg, D.E. ; Kuntz, I.D. ; Ellman, J.A. Potent, low-molecular weight non-peptide inhibitors of malarial aspartylprotease plasmepsin II. *J.Med.Chem.* **1999**, *42*, 4380-4387.

[45] Hann, M.M., Leach, A.R., Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856-864.

[46] Van Dongen, M.J.P.; Uppenberg, J.; Svensson, S.; Lundbäck, T.; Åkerud, T.; Wikström, M.; Schultz, J. Structure-based Screening As Applied to Human FABP4: A Highly Efficient Alternative to HTS for Hit Generation. *J. Am. Chem. Soc.* **2002**, *124*, 11874-11880.

[47] Dalvit, C.; Flocc, M.; Knapp, S.; Mostardini, M.; Perego, R.; Stockman, B.J.; Veronesi, M.; Varasi, M. High-Throughput NMR-Based Screening with Competition Binding Experiments. *J. Am. Chem. Soc.* **2002**, *124*, 7702-7709.

[48] Liebeschuetz, J.W; Jones, S.D.; Morgan, P.J.; Murray, C.W.; Rimmer, A.D.; Roscoe, J.M.E.; Waszkowycz, B.; Welsh, P.M.; Wylie, W.A.; Young, S.C.; Martin, H.; Mahler, J.; Brady, L.; Wilkinson, K. PRO_SELECT: Combining Structure-Based Drug Design and Array-Based Chemistry for Rapid Lead Discovery. 2. The Development of a Series of highly Potent and Selective Factor Xa Inhibitors. *J. Med. Chem.* **2002**, *45*, 1221-1232.

[49] Ramström, O. ; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Disc.* **2002**, *1*, 26-36.