Beilstein-Institut The Chemical Theatre of Biological Systems, May 24th - 28th, 2004, Bozen, Italy

GPCR HIT DISCOVERY BEYOND HTS

KONRAD H. BLEICHER*,[1], ALEXANDER I. ALANINE[1], MARK ROGER-EVANS[1] AND GISBERT SCHNEIDER[2]

[1]F. Hoffmann-La Roche AG, Pharmaceuticals Devision, CH 4070 Basel, Switzerland
[2]Johann Wolfgang Goethe University Frankfurt, Marie-Curie-Str. 11, D-60439 Frankfurt, Germany

E-Mail: *konrad.bleicher@roche.com

Received: 20th September 2004 / Published: 22nd July 2005

ABSTRACT

High-throughput screening is meanwhile well established in most pharmaceutical companies. Although it is routinely applied for most biological targets, several limitations ask for alternative methodologies. This article will describe two different approaches where highly potent ligands for G-protein coupled receptor targets were identified without the application of random high-throughput screening.

INTRODUCTION

A recent analysis of the targets where drugs have been successfully developed shows that about 45% of these targets belong to receptors of which most are G-protein coupled (GPCRs)[1]. Such proteins are involved in various disease pathways and are recognized as highly valuable targets for most therapeutic indications [2,3]. Besides the obvious commercial aspects, the large number of launched drugs for G-protein coupled receptors also gives confidence in the 'drugability' of such seven transmembrane proteins. Due to their size and high lipophilicity the isolation and crystallization of GPCRs turned out to be extremely difficult, making a structure based drug design approach currently unrealistic. Therefore GPCR programmes are usually initiated after a successful high-throughput screening campaign (HTS) has been accomplished.

Chemists clearly appreciate the luxury of hit lists and compound clusters identified via HTS. Nevertheless tedious downstream work is necessary, such as hit confirmation by re-ordering/ synthesis of compounds, quality assessment as well as structural confirmation to name but a few. It is the task of lead generation chemists to deal with such issues in order to generate as much information as possible from an HTS campaign before the actual hit-to-lead programme is started. Although high-throughput assays can generally be set up independently from the target, such screening exercises are particularly interesting for proteins where no biostructural- or patent data is available. Due to increasing compound inventories, constantly upcoming novel targets and requested selectivity screening, prioritization has to take place as to where and when to initiate a high-throughput screening campaign. It is obvious that there is a constant need for complementary technologies, which allow the initiation of chemistry programmes independent from the high-throughput screening route.

TARGETED LIBRARIES FROM PRIVILEGED STRUCTURES

One such alternative source to generate novel hits is the biased testing of targeted libraries. Such small compound arrays (100-1000 cpds) can be tested manually using e.g., radio-ligand binding assays which are usually available long before an HTS compatible assay is established. Clearly, such biased libraries have to show far higher hit rates to be competitive to a random screening approach where usually an average hit rate of 0.1-1% is observed. Besides the much smaller overhead and the fact that chemistry can be initiated earlier (or even without any HTS) the quality and tractability of the compounds within a targeted library subset delivers a big advantage over historical collections where often undesired compounds such as reactive intermediates, poorly soluble compounds or chemical entities which are synthetically challenging, are discovered. The design of such compound collections is essential in this regard, not only to deliver series with a certain target bias, but also showing favourable DMPK and physico-chemical property profiles. In the area of G-protein coupled receptors, where poor biostructural information calls for alternative approaches, several successful projects have been disclosed [4].

From various strategies the 'privileged structures' approach is probably most often applied. It is still unclear which molecular features are required for a chemotype to result in a 'privileged structure' and how the target proteins recognize such chemotypes.

It has been discussed that common binding sites [5] and first contact motives at the surface of the proteins [6] might be responsible for this promiscuous binding behaviour, but a clear understanding on the binding event of 'privileged structures' to their proteins is still lacking.

The term 'privileged structures' was first coined by Evans *et al.* where benzodiazepines were described to show an inherent affinity to G-protein coupled receptors [7]. This compound class has been exhaustively explored in the GPCR area where modifications such as benzodiazepinones and -diones as well their regioisomers such as the 1,4- and 1,5-benzodiazepines (-ones etc.) and their aza-analogues have been employed.

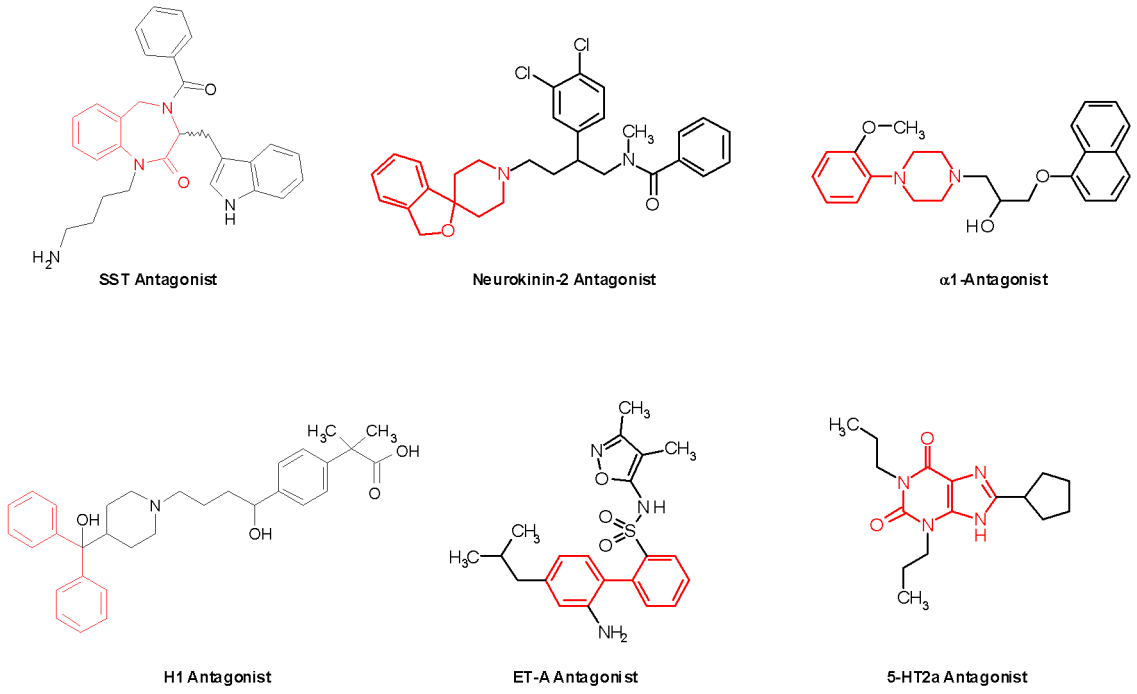


Figure 1. 'Privileged GPCR structures' (highlighted in red) as promiscuous elements of protein ligands (e.g. Benzodiazepinones, Spiropiperidines, Arylpiperazines, Biphenyl methyl-, Biphenyl-Purines).

A number of different 'privileged structures' have meanwhile been discussed in the literature, some of them are depicted in Fig. 1. Exemplified are six different chemotypes representing ligands for six different GPCRs. The spiropiperidines (e.g. represented by the Neurokinin-2 antagonist) have been discovered in various modifications both for GPCR agonists as well as antagonists. Figure 2 exemplifies six different spirpoperidine scaffolds identified for six different G-protein coupled receptors.

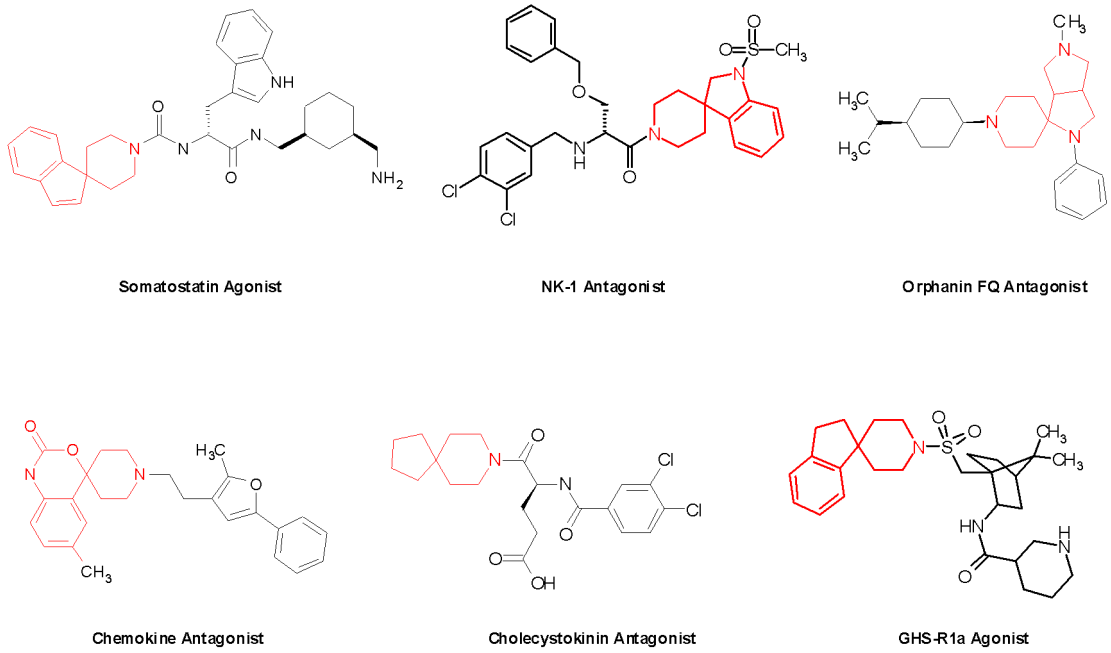


Figure 2. Reported GPCR ligands containing spiropiperdines as recognition motives.

While 'privileged structures' are supposed to show some promiscuity within a protein family, the terminus 'needle' was described as a molecular fragment of ligands binding specifically to particular protein family members. One example for such a needle within the GPCR area is the ortho-substituted biphenyl tetrazole. This motif is well known as a fragment of AT-1 ligands and can therefore be regarded as an 'AT-1 needle' since it appears in most angiotensin-1 antagonists reported in the literature [8].

For the identification of novel small molecule ligands targeting the neurokinin-1 receptor we initiated the generation of focused libraries based on the design strategy to combine the two concepts of 'privileged structures' and 'needles' [9,10]. All three known neurokinin receptors (NK-1, NK-2 & NK-3) belong to the target family of 7-transmembrane G-protein coupled receptors (family 1b). The NK-2 receptor is mainly expressed in the periphery whereas the NK-1 and NK-3 receptors are mainly expressed in the central nervous system indicating their potential therapeutic utility ranging from CNS indications to respiratory and gastric diseases.

The endogenous ligands for the neurokinin receptors are the tachykinins, a group of peptides that share a common C-terminal amino acid sequence Phe-X-Gly-Leu-Met-NH_2 where X is either Phe or Val. The most prominent member of this peptide family is the undecapeptide 'Substance P' (X = Phe) which shows highest affinity for the NK-1 receptor, whereas NKA and NKB (X = Val) are both decapeptides that bind preferentially to the NK-2 and NK-3 receptor, respectively.

Although GPCRs with such large peptide ligands as natural substrates are supposed to be rather difficult to be modulated by small molecules, several drug like NK-1 receptor modulators have been reported in the literature, some representatives of which are depicted in Fig. 3.

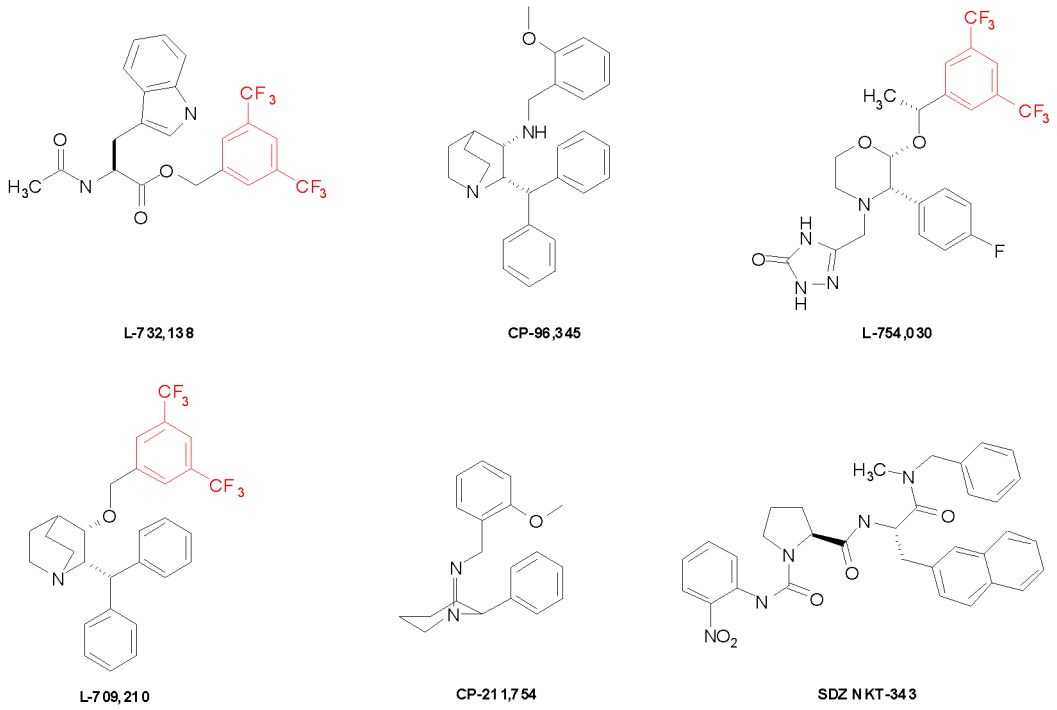
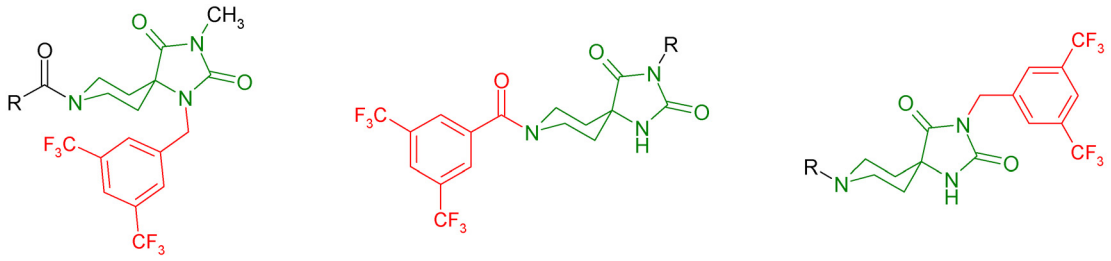


Figure 3. Reported small molecule NK-1 receptor ligands.

As indicated in Fig. 3, many NK-1 ligands have been reported that do possess the 3,5-bis-(trifluoromethyl)phenyl group (highlighted in red) which is meanwhile well recognized as an 'NK-1 needle'. The combination of promiscuous 'privileged structures' and the fairly target specific 'NK-1 needle' gave rise to a targeted library design where the remaining exit vectors in the scaffold (spiropiperidines) were decorated randomly (Figs 4 and 5).

Instead of extensively exploiting the chemotypes by analogizing each exit vector in the templates simultaneously we decided to either concentrate on one additional modification or employ only relatively small building blocks to keep the molecular weight and the resulting lipophilicity in an appropriate range.



1a **1b** **1c**

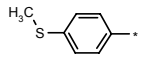
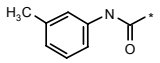
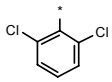
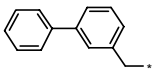
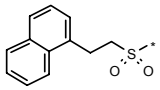
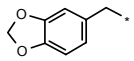
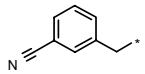
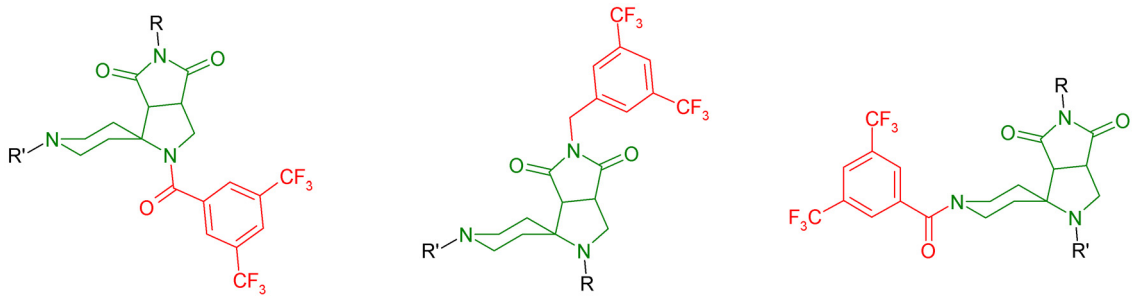
Cpd	R	pKi (NK-1)	Cpd	R	pKi (NK-1)	Cpd	R	pKi(NK-1)
2a		6.34	**2d**		6.88	**2g**		7.43
2b		6.19	**2e**		6.86	**2h**		7.13
2c		5.83	**2f**		6.61	**2i**		6.97

Figure 4. Spirohydantoins as novel NK-1 receptor ligands.

The first compound arrays were based on the spirohydantoin template. These compounds can be generated rapidly from the corresponding amino acid (4-Amino-piperidine-4-carboxylic acid). The 'NK-1 needle' was either introduced *via* the 3,5-bis(trifluoromethyl)benzyl chloride at the α-amino acid nitrogen **1a**, 3,5-bis(trifluoromethyl)benzoic acid at the piperidine nitrogen **1b**, or the 3,5-bis(trifluoromethyl)phenyl- or 3,5-bis(trifluoromethyl)benzyl isocyanate at the imide nitrogen **1c**, respectively.

For all three sublibraries, ligands were identified that showed decent affinities ($pK_i > 5$) in a radio-ligand displacement assay. The three most active ligands for each sublibrary are disclosed in Fig. 4 (**2a-i**). From a set of 136 compounds, 97 molecules showed a binding affinity of pK_i(hNK-1) > 5 which corresponds to a hit rate of 71%.



3a **3b** **3c**

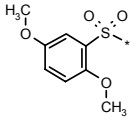
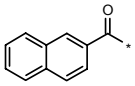
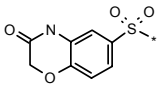
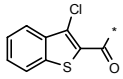
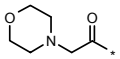
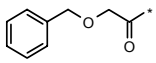
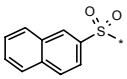
Cpd	R	R'	pKi	Cpd	R	R'	pKi	Cpd	R	R'	pKi
4a	Phe		7.52	**4d**	Mes		6.58	**4g**	Phe		8.29
4b	Phe		7.51	**4e**	H		6.30	**4h**	Phe		8.08
4c	Phe		7.12	**4f**	H		5.64	**4i**	Me		7.41

Figure 5. Spiropyrrolo-pyrrols as novel NK-1 receptor ligands.

As a second example we started from the correspondingly orthogonal protected or resin-bound spiropyrrolo-pyrrole. This template was identified in the orphanin FQ (OFQ) area, a target protein that also belongs to the super-family of peptide class 1 GPCRs. Again the fairly specific 'NK-1 needle' was introduced into the scaffold at all three possible positions (**3a-c**) and the remaining vectors decorated using solution- and/or solid-phase chemistry.

As already observed for the spirohydantoins, independently from the position of the 'NK-1 needle' within the spiropyrrolo-pyrrole template, various low nanomolar binding molecules were identified. From an array of 132 compounds submitted for testing, 91 ligands were identified with a binding affinity of pK_i(hNK-1) > 5, resulting in a hit rate of 69%. Nine representative hits are depicted in Fig. 5 (**4a-i**).

The two examples discussed above exemplify the application of 'privileged structure' based library design for the generation of small molecule NK-1 ligands. As powerful as this approach can be it is obvious that the search for novel 'privileged structures' and/or 'needles' is essential. *De novo* design tools such as Skelgen [11] or TOPAS [12] have been described that allow so-called 'scaffold hopping', a method of moving into novel and therefore patent free chemotypes.

TARGETED LIBRARIES FROM VIRTUAL SCREENING

In contrast to the physical high-throughput screening approach compound filtering and clustering on a computational basis allows the elimination of undesired structures in terms of chemically reactive functionalities, predicted liabilities (e.g. frequent hitter, hERG, CYP450), drug-like properties etc. or group them based on certain similarity/diversity criteria [13]. This is most often applied in the context of compound purchasing to ensure the quality of molecules to be brought in and also to prevent duplication of topologically similar structures already represented in the compound inventory.

A further step downstream the drug discovery chain is the application of such virtual screening tools to pre-select compounds from the corporate collection for directed or biased screening efforts. This is usually applied when a high-throughput screening assay is not available or not appropriate. Several retrospective screening analyses have shown the validity of such an approach where compounds were predicted to be active and increased hit rates were finally observed. To a lesser extent prospective experiments are described showing that biased screening efforts can lead to novel hits where an HTS campaign did not deliver reasonable structures. The integration of virtual screening and HTS seems to be a logical consequence of the maturation of both disciplines but is often still regarded as being competitive rather then complementary and therefore not as intensively applied in many drug discovery companies as one might expect [14].

For the *de novo* generation of targeted libraries basically the same type of algorithms can be applied by computationally screening not physically available compound collections, but rather virtual libraries. Since the number of available structures from such libraries is essentially infinite, the tools and strategies to be applied for computational screening often differ from those used for physically available compound collections. Usually a whole cascade of virtual screening algorithms ranging from 1D to 3D tools are applied to narrow down the huge number of theoretically available compounds to some 100-1000 predicted actives [15].

We applied TOPAS (a 2D topological screening algorithm) as a *de novo design* tool in our search for novel cannabinoid receptor ligands [16]. These receptors also belong to the class 1 GPCR family where two subtypes are currently known. The centrally expressed protein (CB-1) is supposed to modulate appetite by the binding of Anandamide, a lipid GPCR ligand, making it a promising novel drug target for obesity and related diseases.

Starting from a pool of ~ 1500 GPCR fragments (generated from a set of known GPCR ligands) *de novo* designs were generated by linking such fragments randomly and comparing the resulting virtual molecules with known CB-1 ligands based on their topological similarity (CATS) [17]. Two of the most promising proposals are depicted in Fig. 6. Besides the drug-like 'look' of such molecules their chemical tractability is of utmost importance to allow the very rapid validation of the proposed structures. A parallel synthesis approach must be feasible to compensate the fuzziness of the design proposals by generating and testing compound arrays rather than single molecules.

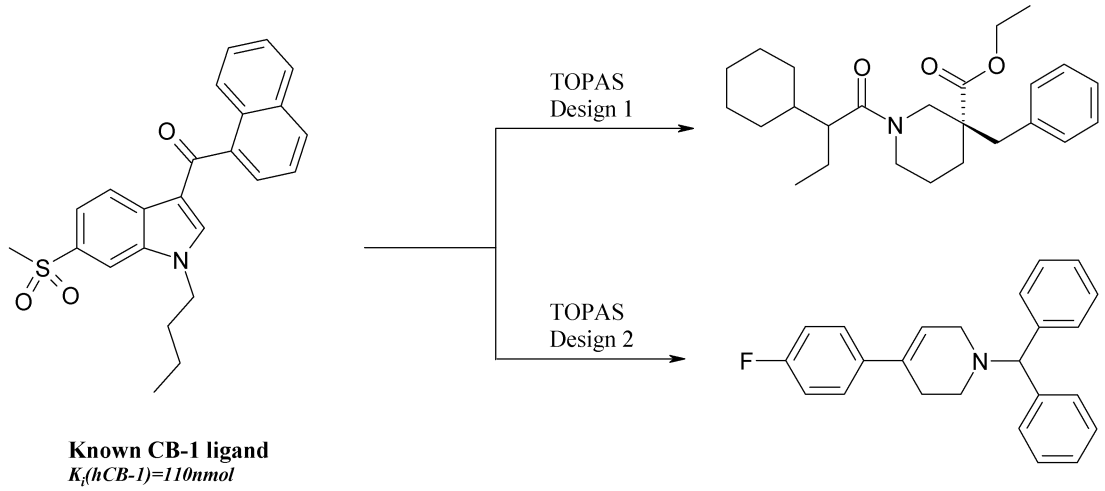
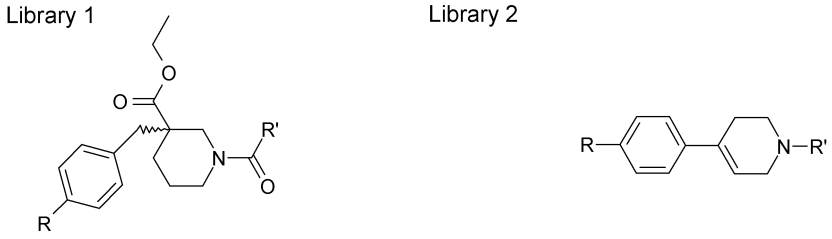


Figure 6. 'Scaffold hopping' proposals generated de novo by TOPAS.

For Design 1 a small array of 83 analogues was prepared and tested in a radio-ligand binding assay. Five hits were identified corresponding to a hit rate of 6%. Fifty analogues of Design 2 were generated, again delivering five hits corresponding to 10% hit rate. Based on the chemical tractability for both design proposals follow-up libraries were generated rapidly to give compound series with low nanomolar binding affinity, the desired functionality (inverse agonists) and preferable physico-chemical properties. In addition, the compounds were further investigated concerning their microsomal stability and cytochrome P_{450} interactions, indicating no issues in this regard. To exclude polypharmacology, a representative lead structure was further tested against CB-2 binding as well as 80 unrelated protein targets before moving into the lead optimization phase.

This 'scaffold hopping' concept can significantly speed up the hit and lead generation process by efficiently combining the advantages of virtual screening and rapid combinatorial chemistry.

The whole project consisted of four consecutive steps: i) fragment-based *de novo* design of virtual analogues by TOPAS; ii) selection of preferred building blocks for synthesis; iii) parallel synthesis, purification and characterization of two small compound libraries and subsequently, iv) biological testing for activity in a human CB-1 receptor binding assay. Three representatives of each chemotype are depicted in Fig. 7.



1a-c **2a-c**

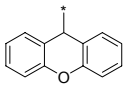
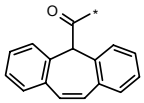
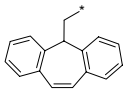
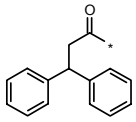
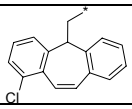
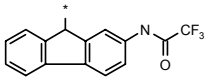
Cpd	R	R'	Ki (umol)	Cpd	R	R'	Ki (umol)
1a	Cl		2.0	**2a**	2,4-CF_3		0.3
1b	tBu		2.0	**2b**	4-Cl		1.5
1c	tBu		4.5	**2c**	H		3.0

Figure 7. Representative CB-1 ligands derived from design proposals 1 & 2 (Figure 6).

Essentially the same strategy was applied to a virtual library where molecular structures are not generated on a randomly combined fragment-basis but rather on templates and building blocks defined by the chemist (RADDAR: Roche Adaptive Drug Design And Refinement). Such libraries were computationally screened to generate predicted hits where, in contrast to the *de novo* approach, the chemistry is already established and the building blocks readily available. Again small arrays were generated based on a the same topological similarity algorithm (CATS) using known CB-1 ligands as seed structures to result in compound libraries that showed much higher hit rates (> 10%) compared to random screening approaches (structures not shown).

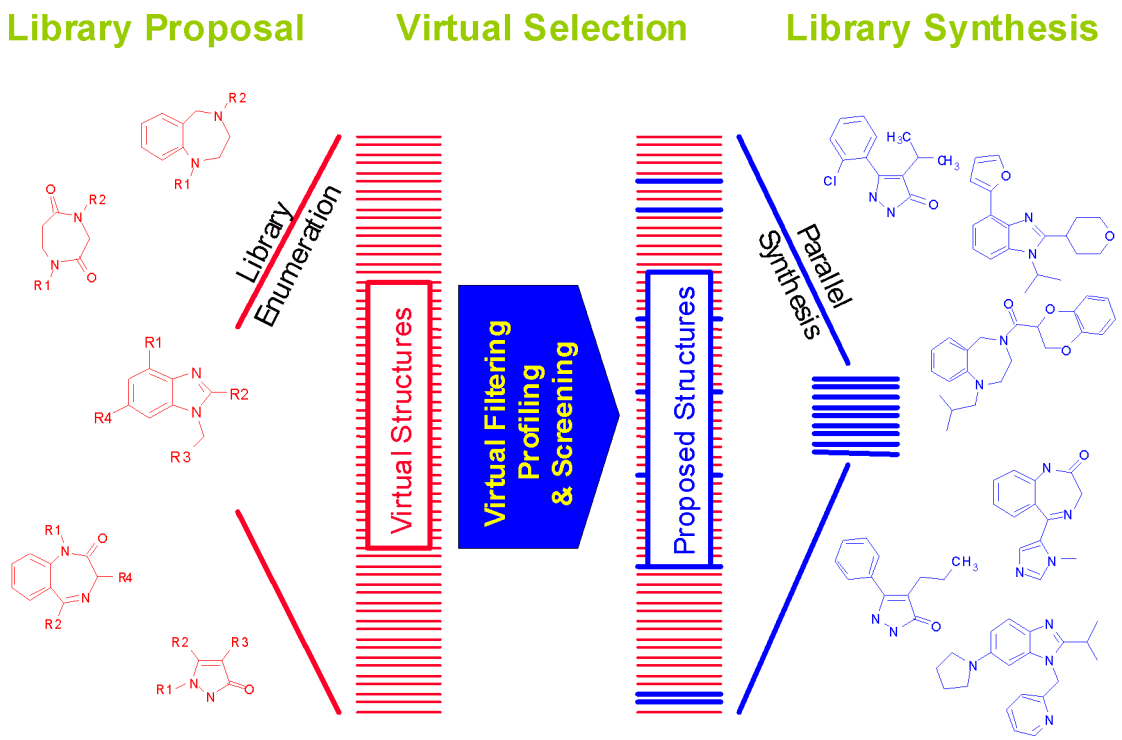


Figure 8. RADDAR: a virtually infinite number of chemical structures can be generated by computationally enumerating combinatorial library proposals. 'Cherry picking' of interesting candidates (proposed structures) based on defined computational algorithms allows the chemist to synthesize these subsets of theoretically accessible molecules that are predicted to be of relevance.

CONCLUSION

Since the generation of lead compounds encompasses much more than affinity for particular targets, it is essential that all the early phase drug discovery issues are addressed. Multidimensional Optimization (MDO), the simultaneous investigation of affinity, selectivity, function, molecular properties and DMPK (*in vitro* and *in vivo*) has been set up in many pharmaceutical research organizations to ensure that compound series with potential liabilities are identified early so as to concentrate on candidates with an appropriate profile, instead of only optimizing affinity in the hope of subsequently finding a remedy for other issues (such as solubility or drug-drug interactions etc.).

It is here that novel methodologies like computational prediction algorithms can have a major impact since the generation of a physico-chemical properties and ADMET profile for whole compound libraries is not only time consuming but also very costly. Although the level of predictive accuracy for ADMET profiles is still suffering from the lack of datasets associated with chemical structures, huge efforts are underway to gain more confidence in those prediction tools which will allow a more informed decision making process when lead series have to be evaluated for further optimization studies and clinical candidate selection [18].

REFERENCES

[1] Drews, J. (2000) Drug discovery: a historical perspective. *Science* **287**:1960-1963.

[2] Watson, S. Arkinstall, S. (1994) *The G-Protein Linked Receptors Facts Book*. Academic Press, New York.

[3] Flower, D.R. (1999) Modeling G-protein-coupled receptors for drug design. *Biochem. Biophys. Acta* **1422**:207-234.

[4] Bleicher, K., Green, L., Martin, R., Rogers-Evans, M. (2004) Ligand identification for G-protein coupled receptors: a lead generation perspective. *Curr. Opin. Chem. Biol.* **8**:287-296.

[5] Bondensgaard, K., Ankernsen, M., Thorgensen, H., Hansen, B., Wulff, B., Bywater, R. (2004) Recognition of privileged structures by G-protein coupled receptors. *J. Med Chem.* **47**:888-899.

[6] Perlman, J., Colson, A.-O., Jain, R., Czyzewski, B., Cohen, C., Osman, R., Gershengorn, M. (1997) Role of the extracellular loops of the thyrotropin-releasing hormone receptor: evidence for an initial interaction with thyrotropin-releasing hormone. *Biochemistry* **36**:15670-15676.

[7] Evans, B.E. *et al.* (1988) Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* **31**:2235-2246.

[8] Burnier, M., Brunner, H. (2000) Angiotensin II receptor antagonists. *Lancet* **355**:637-645.

[9] Bleicher, K., Wüthrich, Y., Deboni, M., Kolczewski, S., Hoffmann, T., Sleight, A. (2002) Parallel solution- and solid-phase synthesis of spirohydantoin derivatives as neurokinin-1 receptor ligands. *Bioorg. Med. Chem. Lett.* **12**:2519-2522.

[10] Bleicher, K., Wüthrich,Y., Adam, G., Hoffmann, T., Sleight, A. (2002) Parallel solution- and solid-phase synthesis of spiropyrrolo-pyrroles as novel neurokinin receptor ligands. *Bioorg. Med. Chem. Lett.* **12**:3073-3076.

[11] Lloyd, D., Bünemann, C., Todorov, N., Manalack, D., Dean, P. (2004) Scaffold hopping in *de novo* design. Ligand generation in the absence of receptor information. *J. Med.Chem.* **47**:493-496.

[12] Schneider, G., Clément-Chomienne, O., Hilfiger, O., Schneider, P., Kirsch, S., Böhm, H.-J. Neidhart, W. (2000) Virtual screening for bioactive molecules by evolutionary de novo design. *Angew. Chem. Int. Ed. Engl.* **39**:4130-4133.

[13] Böhm, H.J., Schneider, G. (Eds) (2000) *Virtual Screening for Bioactive Molecules*. Wiley- VCH, Weinheim.

[14] Bajorath, J. (2002) Integration of virtual and high-throughput screening. *Nature Rev. Drug Discov.* **1**:882-894.

[15] Walters, W. P., Stahl, M. T., Murcko, M. A. (1998) Virtual screening - an overview. *Drug Discov. Today* **3**:160-178.

[16] Rogers-Evans, M., Alanine, A. I., Bleicher, K.H., Kube, D., Schneider, G. (2004) Identification novel cannabinoid receptor ligands via evolutionary de novo design and rapid parallel synthesis. *QSAR Comb. Sci.* **23**:426-430.

[17] Schneider, G., Neidhart, W., Giller, T., Schmid, G. (1999) "Scaffold-hopping" by topological pharmacaphore search: a contribution to virtual screening. *Angew. Chem. Int. Ed. Engl.* **38**:2894-2896.

[18] Davis, A., Riley, R.J. (2004) Predictive ADMET studies, the challenges and the opportunities. *Curr. Opin. Chem. Biol.* **8**:386-386.