Medicinal chemistry of target family-directed masterkeys

Gerhard Müller

The majority of pharmaceutically relevant drug targets cluster into densely populated target families, thus offering a novel approach that complements the currently favoured screening paradigm in medicinal chemistry. This approach uses a privileged structure concept whereby molecular masterkeys are developed that account for a target family wide structural or functional commonality. Numerous lead compounds, based on multipurpose privileged structures, can be generated that address a variety of targets from a gene family of interest, irrespective of therapeutic area. Several different interpretations of the privileged structure concept will be highlighted, with a strong emphasis on the most stringent application: the optimization of a molecular masterkey for a distinct target family of interest.

Gerhard Müller
Axxima Pharmaceuticals AG
Max-Lebsche-Platz 32
81377 München
Germany
e-mail:
gerhard.mueller@axxima.com

▼ The pharmaceutical industry is undoubtedly one of the world's largest industries, with an associated market that is still growing. However, it faces a steadily growing pressure to release more new chemical entities (NCEs) each year that will evolve ideally into innovative drugs with novel mechanisms of action that target therapeutic areas of unmet medical need. These attributes not only define an ideal set of preconditions for launching a compound with blockbuster potential but also summarize the short term expectations of the upper management team of almost every pharmaceutical company [1–4].

In contrast to these expectations, the field of drug discovery and development has not evolved as efficiently as expected. Despite the steady increase in R&D expenditures within the biopharmaceutical industry, the number of NCEs reaching the market has actually decreased dramatically. Although the number of FDA approvals for 2002 is reported as being 29 (diagnostics excluded) (Table 1), only 14 products deserve the status of a NCE (Fig. 1). In addition, the follow-up times of competitors addressing the same market segments with 'me-too' compounds have decreased, thus limiting profitability [5].

This so-called productivity gap [6–8] represents an alarming signal for an industry that is so focused on pharmaceutical research. To achieve an annual growth rate of approximately 10%, each large pharmaceutical company needs to launch four NCEs each year, with associated annual sales of at least $350 million [9]. However, in reality the average pharmaceutical company launches less than one NCE annually, and not more than 25% of these NCEs achieve sales of more than $350 million [9]. Concomitantly, development times have continued to increase. Since 1964, the total development time has more than doubled from 6.5 years to almost 15 years [10,11].

Given the enormous pressure on pharmaceutical research and the overly optimistic expectations of management teams, numerous paradigmatic changes have been announced over the past ten years aimed at resolving the major bottlenecks in drug discovery and development, which has increased both quantitative, as well as qualitative, output. Indeed, attempts such as the conceptual combination of automated or combinatorial chemistry with HTS or even ultra-HTS have altered dramatically the process of lead finding within medicinal chemistry, resulting in the rapid synthesis and testing of vast numbers of low molecular weight compounds against numerous biological screens. This process is paralleled with the aggressive application of genomics, proteomics and bioinformatics approaches at the initiation of the value-chain of drug development [12–14].

In summary, the past decade of preclinical research has produced some extremely useful technologies that have a huge potential for increasing efficiency. However, the industry still suffers from two persisting bottlenecks:

1359-6446/03/$ – see front matter PII: S1359-6446(03)02781-8

Table 1. FDA-approved drugs from 2002[a,b]

Trade name	Indication	Company
Aranesp	Anaemia	Amgen
Arixtra	Thrombosis	Sanofi/Organon
Avinza	Pain	Elan/Ligand
Avodart	Prostate hyperplasia	GlaxoSmithKline/ Eli Lilly
Bextra	Pain, arthritis	Searle-Pharmacia
Bravelle	Female infertility	Ferring
Doxil	Cancer	Alza
Elidel	Dermatitis	Novartis
Elitek	Hyperuricaemia	Sanofi
Eloxatin	Cancer	Debiopharm/Sanofi
Faslodex	Cancer	AstraZeneca
Focalin	ADHD	Celgene/Novartis
Foradil	COPD	Sepracor/Novartis
Forteo	Osteoporosis	Eli Lilly
Hapsera	HBV	Gilead
Inspra	Hypertension	Pharmacia
Invanz	Antibiotic	Merck
Kytril	Antiemetic	GlaxoSmithKline
Lexapro	Antidepressant	Lundbeck/Forest
Neulasta	Leukopenia	Amgen
Orfadin	Tyrosinaemia	Swedish Orphan
REBIF	Multiple sclerosis	Serono
Remodulin	Hypertension	United Therapeutics
Strattera	ADHD[a]	Eli Lilly
Vfend	Antifungal	Pfizer
Xigris	Sepsis	Eli Lilly
Zelnorm	Bowel Syndrome	Novartis
Zetia	Hypercholesterolaemia	Schering-Plough/ Merck
Zometa	Hypercalcaemia	Novartis

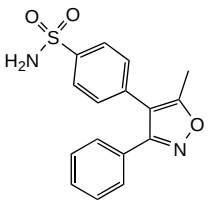
[a]Abbreviations: ADHD, attention deficit hyperactivity disorder; COPD, chronic obstructive pulmonary disease; HBV, hepatitis B virus.
[b](http://www.globalrx.com/new2002.html)
Amgen (http://www.amgen.com); Sanofi (http://www.sanofi-synthelabo.com); Organon (http://www.organon.com); Elan (http://www.elan.com); Ligand (http://www.ligand.com); GlaxoSmithKline (http://www.gsk.com) Eli-Lilly (http://www.lilly.com); Searle-Pharmacia (http://www.pfizer.com); Ferring (http://www.ferring.com); Alza (http://www.alza.com); Novartis (http://www.novartis.com); Debiopharm (http://www.debio.com); AstraZeneca (http://www.astrazenica.com); Celgene (http://www.celgene.com); Sepracor (http://www.sepracor.com); Gilead (http://www.gilead.com); Merck (http://www.merck.com); Lundbeck (http://www.lundbeck.com); Forest (http://www.frx.com); Swedish Orphan (http://www.swedishorphan.com); Serono (http://www.serono.com); United Therapeutics (http://www.unither.com); Schering-Plough (http://www.sch.plough.com).

one closely related to target validation [15], and the other, more relevant for this review, related to the overall quality of compounds in terms of their ability to progress through the value-chain of drug development. Here, the question arises as to whether preclinical researchers, especially in medicinal chemistry, are aware of which approaches are most appropriate for a given project and how to use them in a synchronized fashion. Several years ago, the main emphasis was on generating large numbers of compounds by attempting to scan the multidimensional molecular diversity space as systematically and exhaustively as possible. However, interfacing well-established 'know-how' derived from target classes, for example in the design of moderately sized compound arrays, is now a promising approach to produce tractable lead structures. Consequently, the concept of 'similarity' has evolved as a viable alternative to the concept of 'diversity'.

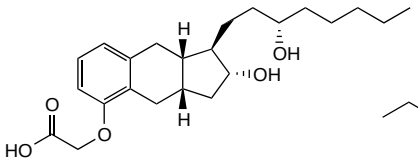
At present, the life science community relies on knowledge based on target structure, mechanism and proven medicinal chemistry approaches towards representatives from densely populated target clusters that a medicinal chemistry strategy could be devised for almost any given new emerging target, even before the corresponding high-throughput campaign is finished. In this context, the concept of target family-directed molecular masterkeys provides an alternative to either blind screening attempts or stringently applied structure based design approaches. The available knowledge on structure and/or function of target families is encoded in low molecular weight substructures that, after modification, deliver high quality lead structures for further expansion towards viable preclinical candidates.

Target families of pharmaceutical interest

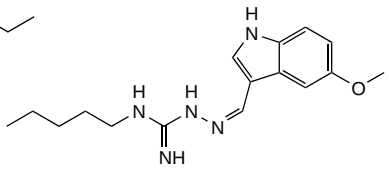
From a medicinal chemist's point of view, it has been interesting to witness the conversion of assessment that has occurred over the past five to ten years when the number of pharmaceutically relevant protein targets within the human genome was discussed. According to a thorough analysis of drugs listed in the 1996 pharmacopoeia, the total number of proteins within humans, for which pharmaceutical research has produced drugs, was less than 500 [6,7]. This analysis from the 'pre-genome' era further speculated on the existence of up to 10,000 potentially relevant proteinogenic drug targets within the human genome, irrespective of their biochemical nature. In this context, it is tempting to believe that a thorough application of genomics and related technologies would generate numerous novel drug targets with significant benefits for drug discovery and development [13]. Only a few years ago, it was even stated that the impact of genomics, if applied rigorously, would halve the cost and time of drug development [13]. The role of medicinal chemistry has been underestimated, because even in the 'post-genome' era, hit generation and subsequent optimization is still essential to produce a viable preclinical candidate. Consequently, a drug target survey 'from a



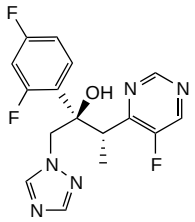
Bextra (SC-65872)



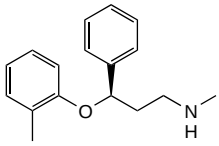
Remodulin (UT-15)



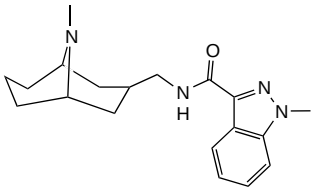
Zelnorm (SDZ-HTF-919)



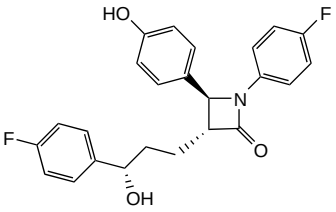
Vfend (UK-109496)



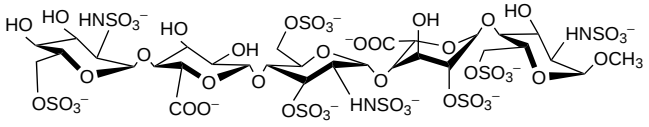
Strattera (LY-139603)



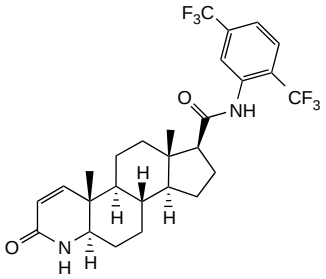
Kytril (BRL-43694)



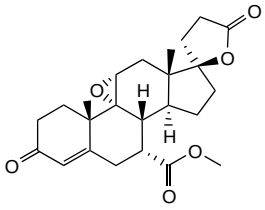
Zetia (Sch-58235)



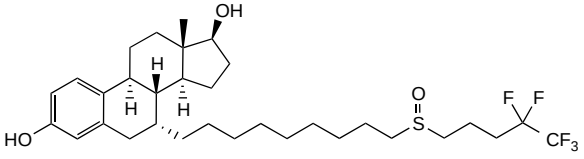
Arixtra (Org 31540)



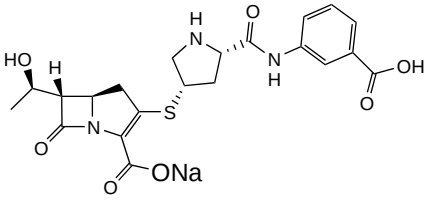
Avodart (GI-198745)



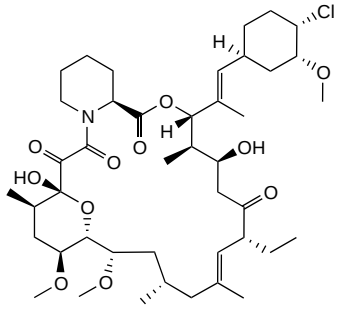
Inspra (SC-66110)



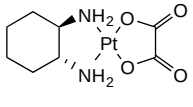
Faslodex (ZD-187780)



Invanz (L-749345)



Elidel (SDZ-ASM-981)



Eloxatin (NSC-266046)

Drug Discovery Today

Figure 1. Approved drugs that are true NCEs from 2002.

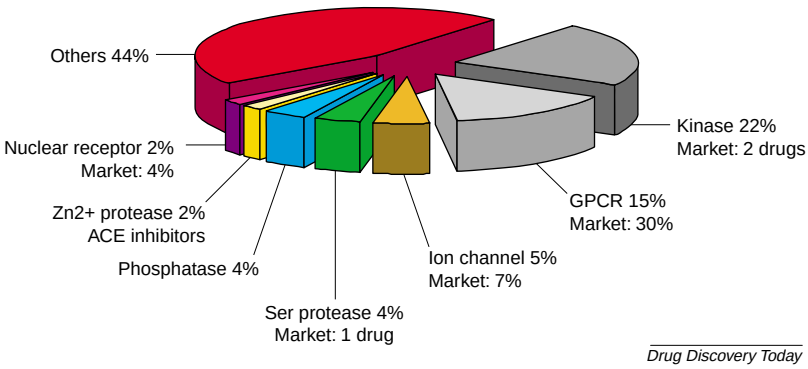


Figure 2. Schematic presentation of the target family distribution within the human 'druggable genome'. If drugs that target members of the given families are on the market, the percentage of their market share, or number of released drugs, is given explicitly.

compound's point of view' seemed to be a more sensible approach towards the often quoted 'druggable genome'. Researchers from Pfizer, (http://www.pfizer.com) mined the human genome systematically for putative drug targets amenable to drug-like high affinity interactions ($K_i < 10\mu M$) with low molecular weight compounds that are compliant with the Lipinski rules [16]. They produced a list of about 400 nonredundant proteins that meet these requirements [17]. Not surprisingly, a large percentage of these targets cluster into target families, such as G-protein-coupled receptors (GPCRs), serine, threonine and tyrosine kinases, serine, cysteine, aspartic acid and metalloproteases, ion channels and nuclear hormone receptors. Based on the assumption that once a member of a target family is amenable by rule-of-five compliant compounds, the entire gene family is druggable, a theoretical number of approximately 3050 putative protein targets was derived by systematic extrapolation within the corresponding gene families (Fig. 2).

Most importantly, the analysis revealed the potentially interesting target classes amenable by low molecular weight compounds, irrespective of any proven disease relation. The fact that these are multimember gene families presents an unprecedented opportunity for systematisation within the discipline of medicinal chemistry in that privileged structures can be tailor-made that account for a family-wide commonality in terms of enzymatic mechanisms and/or molecular recognition elements. The most densely populated target families, notably kinases, GPCRs, ion channels, proteases, nuclear hormone receptors and phosphatases, represent attractive fields of activity for medicinal chemistry. Although approximately 30% of all marketed drugs target GPCRs, approximately 7% address ion channels and approximately 4% bind to nuclear hormone receptors, there are only two drugs that address kinase targets (Gleevec® and Iressa®), one drug that targets a serine protease (Acova®) and a single metalloprotease (angiotensin converting enzyme) is inhibited by a variety of marketed drugs. It is obvious from these results that there is plenty of scope for innovation within the area of hit generation and optimization.

However, it should be emphasized again that the number of principally druggable targets (approximately 3050) by far exceeds the number of drug targets because no aspects of disease relation were considered in this analysis [17]. Within the human genome, the number of pharmaceutically relevant target proteins is estimated to be 600-1500, which is far more realistic when compared with the extremely optimistic expectations several years ago [17].

Privileged structures: room for interpretation

The term 'privileged structure' has appeared more frequently in the literature over the past few years and a refined definition might help to distinguish more clearly between recurring structural elements that elicit false positive responses in biological assays, so-called privileged structures that lack a strict target class relation, and privileged structures that address a protein family wide commonality in terms of the involved molecular recognition phenomenon.

Evans and co-workers first used the term 'privileged structure' in 1988 when they focused on the design of benzodiazepine-based cholecystokinin-1 (CCK_1) antagonists [18]. The definition was later updated by Patchett and Nargund [19]. Evans *et al.* [18] refer to a finding by Chang *et al.* [20] who discovered that the previously described analgesic tifluadom [21], a kappa opioid agonist, also acts as a peripheral CCK receptor antagonist. This documented receptor crosstalk of a single compound to two different target proteins of the same gene family (GPCRs) implies that a single molecular framework is able to provide ligands for diverse receptors. Therefore, judicious modification of such structures could be a viable alternative in the search for new receptor agonists and antagonists [18].

Unwanted privileged structures

From these early studies the most generic definition refers to substructural elements emerging in compounds that showed effects on more than one target protein, irrespective of the corresponding target families that they might belong to. This specific characteristic of compounds is not necessarily a desired profile, for example as hits in numerous different

biological assays covering a broad range of protein targets. The elimination of so-called 'frequent hitters' from compound libraries was described recently by a group from Hoffmann-La Roche (http://www.roche.com) [22] because those compounds were shown to either bind non-specifically to a variety of targets or interfere with the utilized assay read-out methods. These compounds were clearly considered as undesirable starting points for optimization programmes in medicinal chemistry. Obviously, a differentiation between those promiscuous binders and privileged structures is required, thus refining the original definition of Evans and co-workers [18]. This is further supported by a recent study by Shoichet and colleagues [23]. An in-depth study of screening hits that appear non-drug-like with a non-competitive mode of action and contradictory structure–activity relationships has revealed a common mechanism that accounts for the undesired compound profile. The compounds tend to form molecular aggregates, as determined by dynamic light scattering and electron microscopy, with particle sizes of 30–400 nm in diameter. These particles represent the inhibiting entity in numerous enzyme assays. It is noteworthy that this phenomenon is not restricted to compounds classified as non-drug-like by a trained medicinal chemist but also occurs for drug-like molecules such as steroids or kinase inhibitors (Fig. 3) [24].

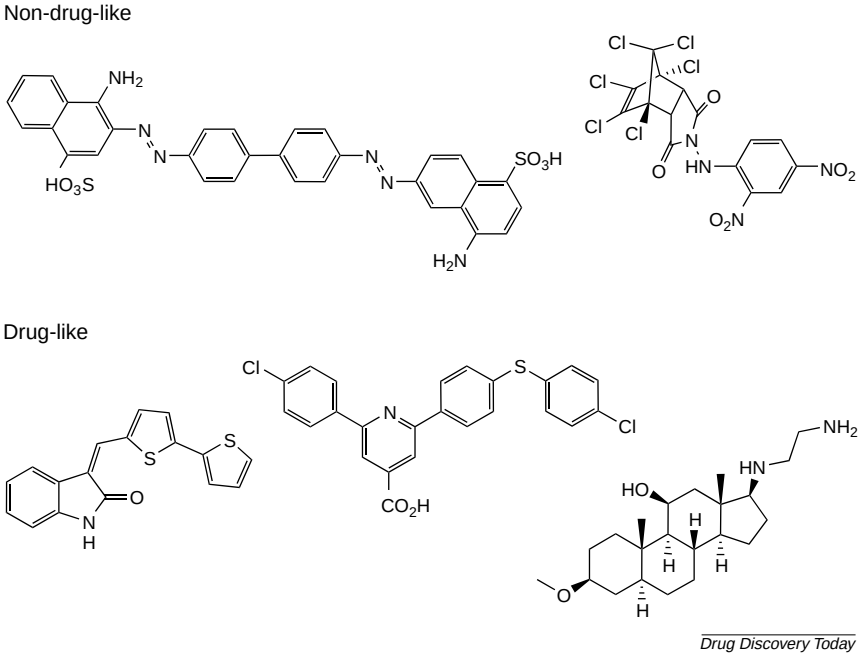


Figure 3. Chemical structures of compounds shown by dynamic light-scattering and electron microscopy to form molecular aggregates, thus interfering in a variety of biochemical assays, producing false positive results.

The findings that 'high-quality' compounds also tend to elicit non-specific biological activities by forming aggregates not only renders the numerous screening hits and associated optimization programmes highly questionable, but also defines a new deselection criterion for hit assessment, which requires biophysical investigations before any significant medicinal chemistry resource assignment.

Privileged structures devoid of target family correlations

Numerous investigations have focused on the identification of desirable privileged structural elements. Two early pioneering studies have provided guidelines for the fragmentation of compounds into core structures and peripheral decoration, and how this approach might drive chemistry programmes based on privileged molecular fragments. (For further details see [25-27]). Both studies reveal frequently recurring substructural elements that can be used for proactively enriching as well as focusing chemistry efforts towards the more productive regions of the often quoted multidimensional molecular diversity space.

Modelled after the Vertex (http://www.vpharma.com) [25,26] and GlaxoSmithKline (GSK; http://www.gsk.com) approaches [27] that relied on *in silico* database mining methodologies, Fesik and co-workers pursued an experimentally based procedure of identifying fragments with a generally high propensity for protein binding [28]. NMR-based screening of more than 10,000 selected fragment-type compounds against 11 target proteins revealed 12 privileged substructures (Fig. 4) that appeared, with statistical significance, in compounds shown to bind to selected targets. The main conclusion from this study refers to the preferential utilization of these substructures in combinatorial libraries to qualitatively enrich an in-house screening compound collection. The quality criterion of these structural elements being privileged relates only to the observation that compounds containing one of the identified fragments might bind to a nonspecified target with higher than average probability. A clear rationale is not apparent on the fragment-target family relation and why these fragments tend to bind.

Once the focus of privileged structures is laid on conformationally constrained core structures, such as tifluadom [18,20,21], that might exhibit prominent orientational characteristics, for example a pharmacophore encompassing

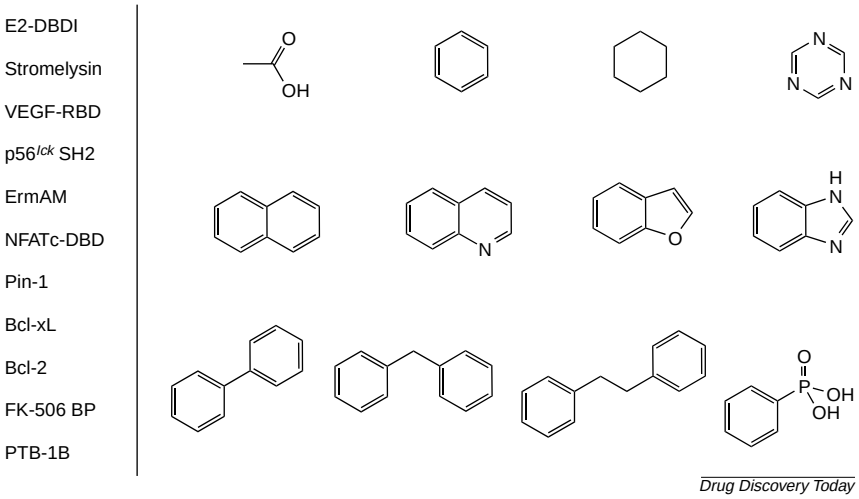


Figure 4. NMR based screening of fragment binding towards a variety of proteins (left) revealed 12 recurring structural elements (right) from Abbott's NMR screening library (http://abbott.com). Protein definitions: Bcl-2, an anti-apoptotic protein; Bcl-xL, an anti-apoptotic protein; ErmAM, a methyltransferase; E2-DBD, DNA-binding domain of human papillomavirus E2 protein; FK-506 BP, FK-506 binding protein; NFATc-DBD, DNA binding domain of NFATc; p56lckSH2, Src homology domain of p56lck; Pin-1, peptidyl-prolyl cis-trans isomerase; stromelysin, a matrix metalloprotease; VEGF-RBD, receptor binding domain of vascular endothelial growth factor.

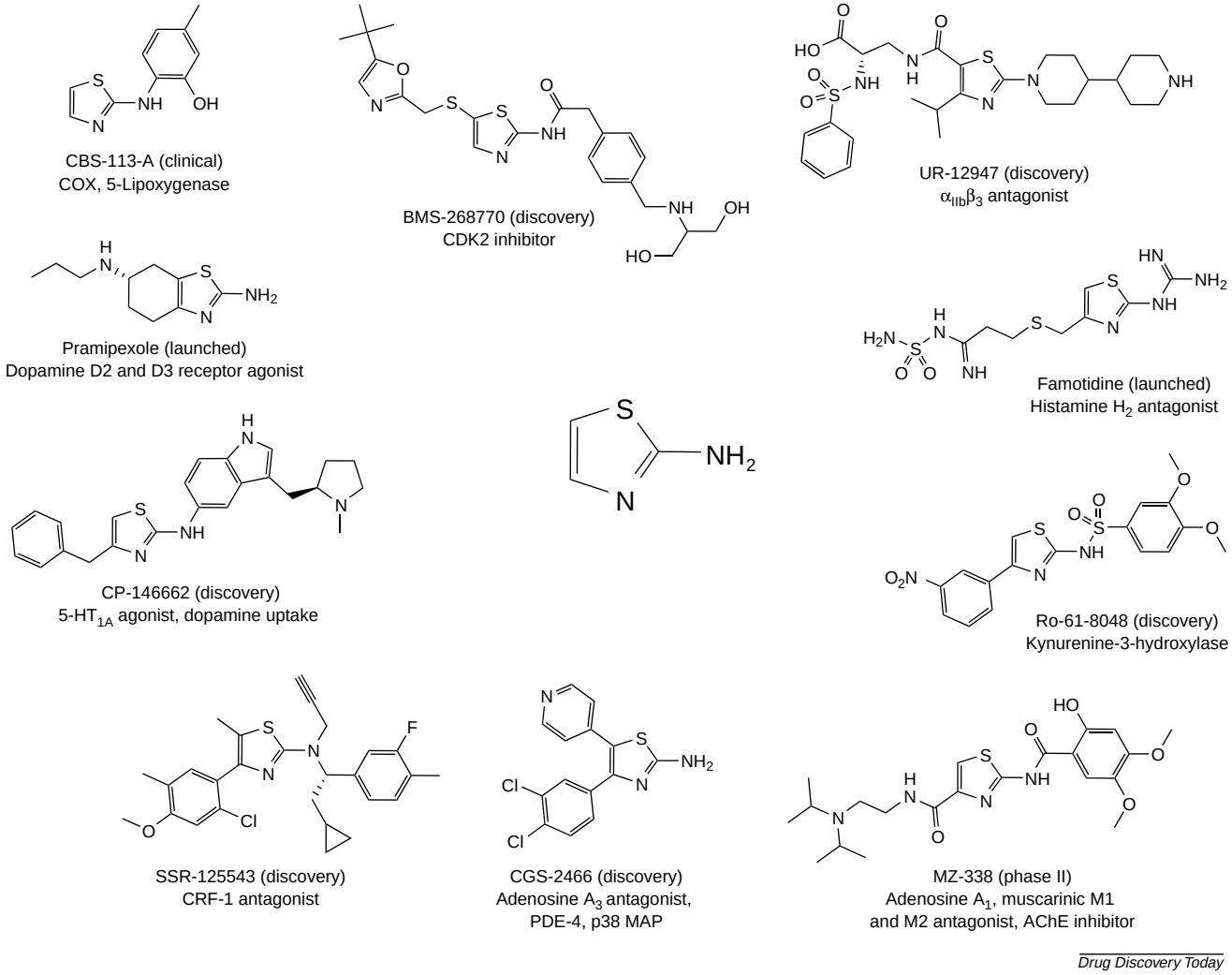


Figure 5. Selection of 2-aminothiazole containing compounds displaying biological activity at diverse target proteins. The research code and the development status (launched, phase III, etc.), together with the protein target, is given for each structure.

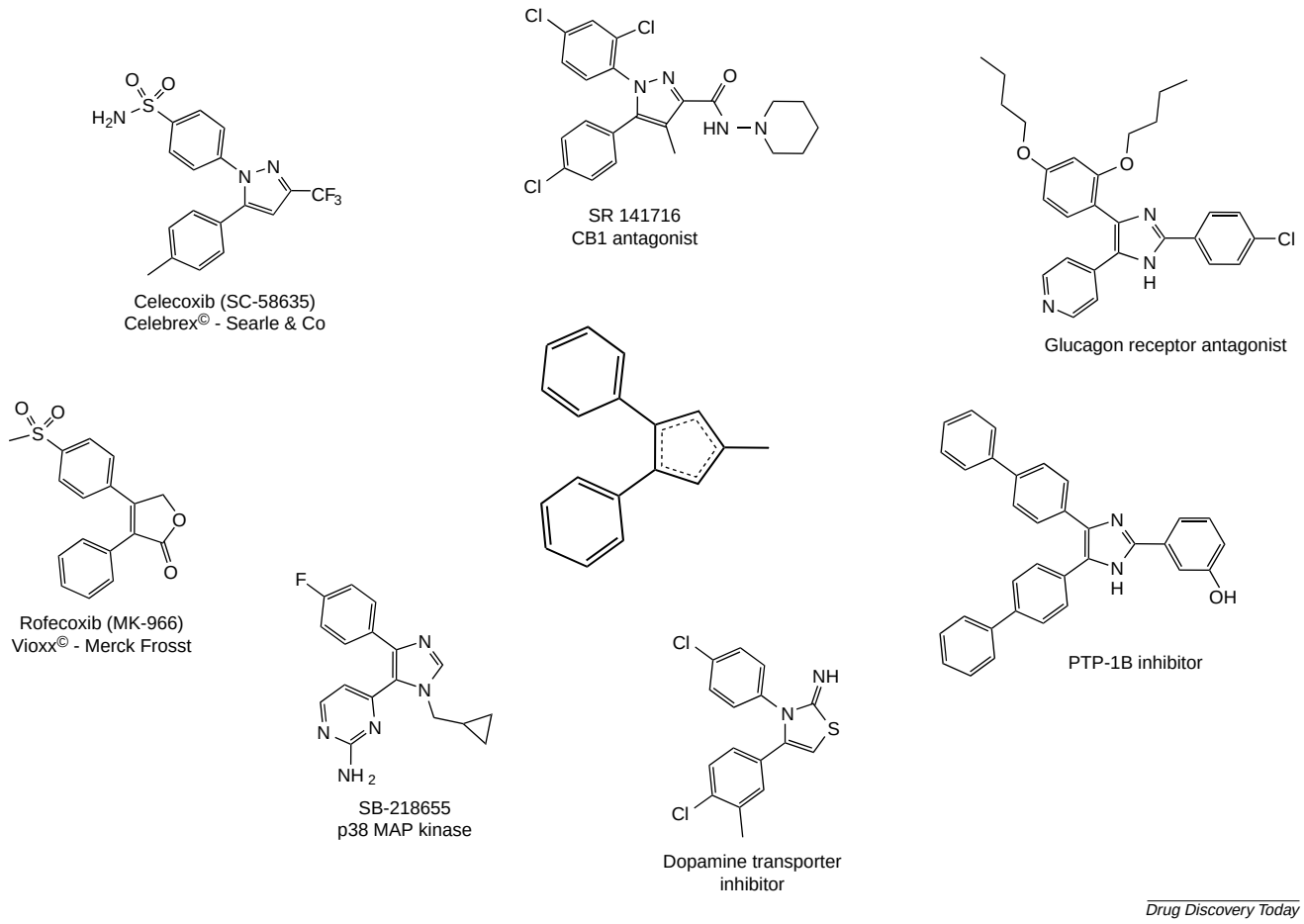


Figure 6. Selection of enzyme inhibitors and receptor antagonists that structurally follow a common underlying template (centre). Celecoxib (Celebrex©) and rofecoxib (Vioxx©) are marketed drugs targeted against cyclooxygenase-2 (COX-2).

molecular periphery, every trained medicinal chemist will easily identify recurring structural motifs. This is exemplified on the 2-aminothiazole core that is found in numerous drugs and clinical and preclinical candidates that address a broad spectrum of targets (Fig. 5).

Again, no target family correlation is evident. Instead, the compounds bind to enzymes such as cyclooxygenases, phosphodiesterases, kinases and acetylcholinesterase through to receptors such as integrins and various members of the GPCR family. The versatile chemistry approaches, delivering decorated 2-aminothiazole-derived compounds, are probably the main reason why this scaffold appears as a recurring structural motif in compounds targeting members from numerous different gene families.

A five-membered heterocycle with a conserved vicinal 1,2-di-phenyl substitution pattern (Fig. 6) serves as a second example for a recurring core structure in cyclooxygenase inhibitors, kinase inhibitors, GPCR antagonists, phosphatase inhibitors and dopamine transporter inhibitors (Fig. 6).

Because the spatial extent of the common underlying 1,2-di-phenyl substituted heterocycle by far exceeds the 2-aminothiazole scaffold, and the nature of the five-membered heterocycle is quite diverse, a versatile chemistry is not the main reason why this structure type frequently occurs in biologically interesting low molecular weight compounds. Although crystallographically derived structures are known for a variety of these compounds in complex with their target proteins, a structural interpretation of the privileged status of the common fragment remains unclear. Again, this indicates that chemical similarity does not necessarily correspond to biological similarity.

Based on these findings the question remains as to how medicinal chemistry could take advantage of those results for future programmes. Because there is no detailed insight in, for example, a highly conserved compound–target interaction mode for the two structural elements described above (Figs. 5 and 6), these findings can only serve to guide combinatorial chemistry initiatives or compound acquisition

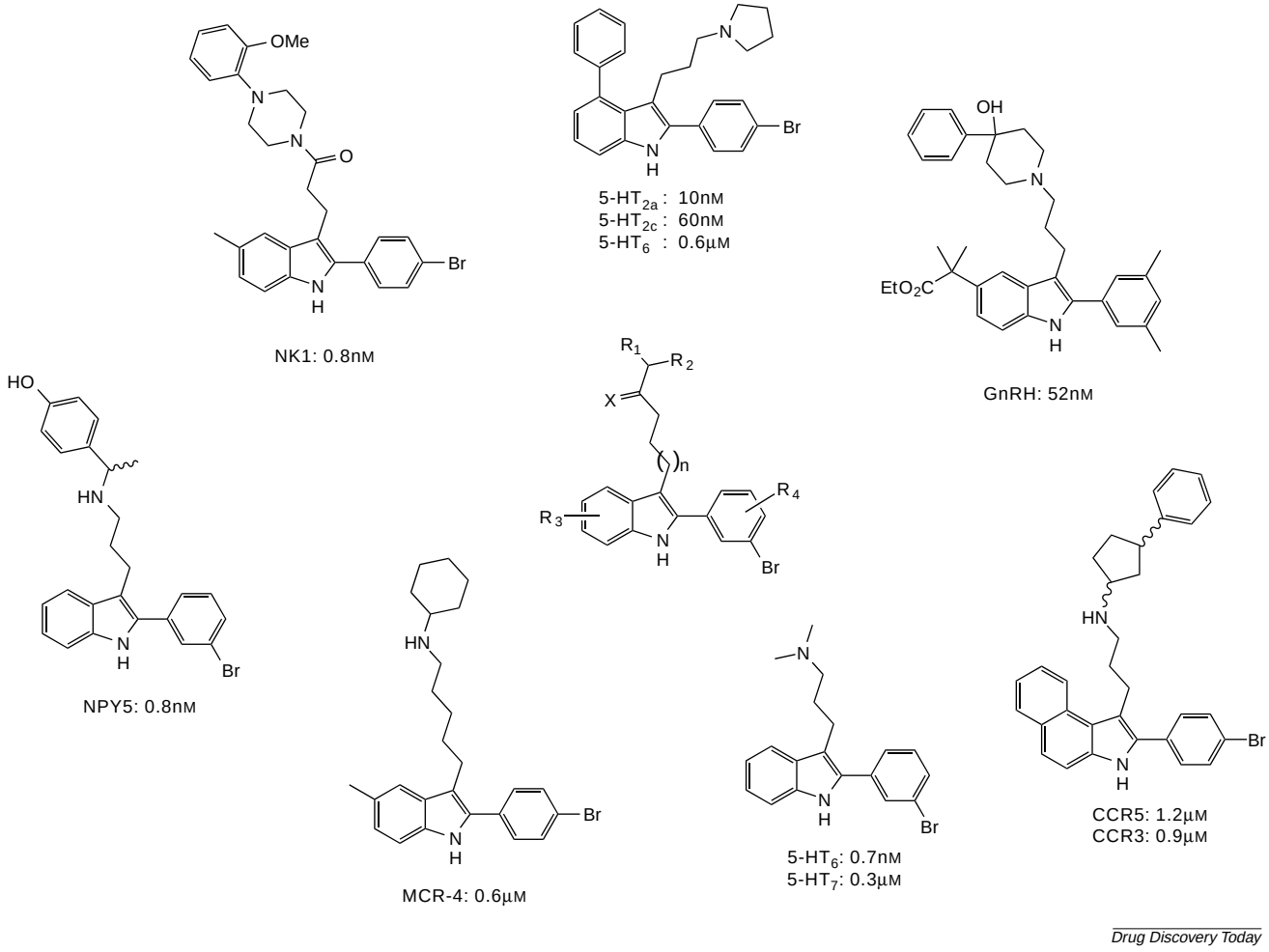


Figure 7. Design principle (centre) of a GPCR targeted library producing highly active and selective compounds targeted towards specific members of that gene family. Abbreviations: CCR, chemokine (C–C) receptor; 5-HT, 5-hydroxytryptamine; GnRH, gonadotropin releasing hormone; MCR, melanocortin receptor; NK, neurokinin; NPY, neuropeptide Y.

so that more emphasis is put on the similarity of those recurring fragments, instead of attempting to scan the molecular diversity universe systematically.

Target family directed approaches

A conceptual combination of synthetic feasibility, on one hand, with a proven compound-target relation, on the other hand, was recently exploited in a combinatorial exploration of 2-arylindole based analogues by medicinal chemists at the Merck Research Laboratories (http://www.merck.com) [29]. This specific scaffold is an ubiquitous element in many biologically active compounds. Additionally, specific experience with indole synthesis resulted in a combinatorial compound library from which highly active and selective GPCR binding indole derivatives emerged (Fig. 7).

In terms of molecular complexity and ease of associated chemistry, each of the depicted compounds (Fig. 7) represents an attractive entry into a corresponding optimization programme. Apart from the above mentioned recurrence of substituted indoles, an understanding for the indole-GPCR relation is not apparent. However, the library design clearly took advantage of principally available information, finally resulting in a significantly increased hit rate towards the target family of interest.

The most stringent interpretation of the privileged structure concept refers to a substructural element with a proven correlation to a target family based on a single or a variety of key structural elements that account for a target family wide commonality in molecular recognition. An illustrative example is the 5,5-*trans*-fused lactam moiety systematically explored as a serine protease directed scaffold by GSK over the past couple of years. Interestingly, the initial finding originated from a screening initiative in which 5,5-*trans*-fused lactone euphane triterpenes, isolated

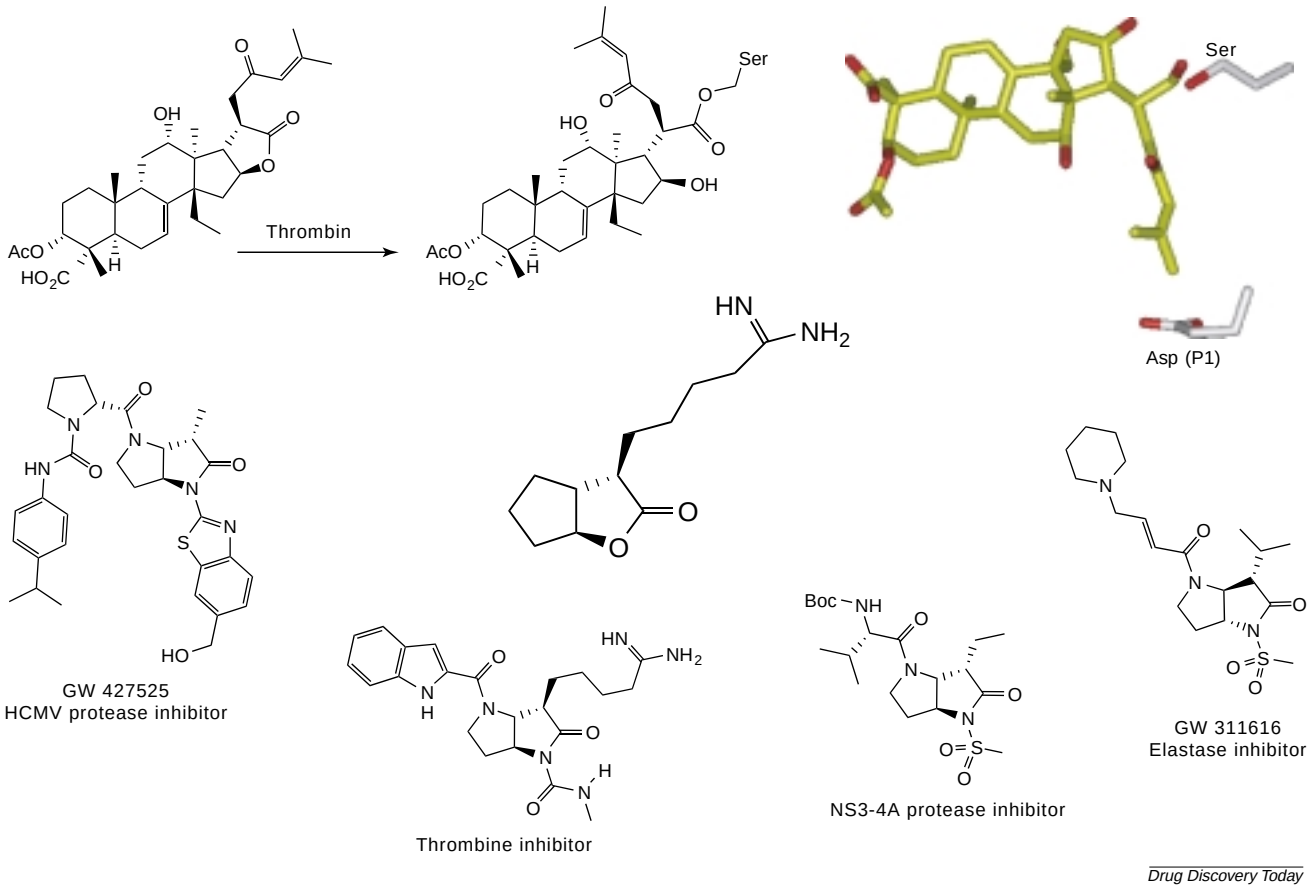


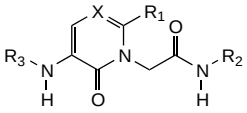
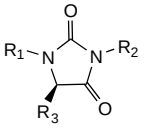
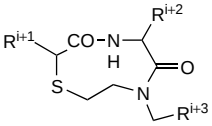
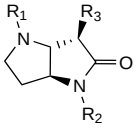
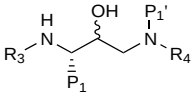
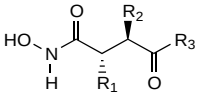
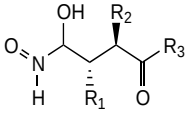
Figure 8. The top of the figure shows a schematic presentation of the thrombin inhibition mechanism of a *trans*-lactone containing natural product (left) together with the crystallographically determined complex structure of the enzyme-bound, ring opened compound (right). The minimal inhibitory fragment that served as starting point for lead optimization is shown in the middle. The resulting 5,5-*trans*-fused lactam based inhibitors for a variety of different serine proteases developed by GlaxoSmithKline (http://www.GSK.com) are shown below.

from methanolic extracts of *Lantana camara* leaves, was found to be a potent inhibitor of the human serine protease thrombin [30,31]. Based on protein crystallography studies the inhibitory mechanism was identified as a lactone ring opening reaction with subsequent acylation of the enzyme's active site serine side chain (Fig. 8) [30].

This mechanism was recognized as a generic serine protease inhibition principle and the medicinal chemistry activities at GSK focused on simplifying the complex natural product to yield the minimum active core structure as a low molecular weight, lead-like [32] building block with considerable inhibitory potency (Fig. 8) [33]. The lactone series was subsequently modified towards a 5,5-*trans*-fused lactam series, yielding not only acylating thrombin inhibitors but also classical reversible, nonacylating compounds with promising lead candidate profiles [34].

This result alone deserves the attention of a broad scientifically interested audience: a complex natural product identified by HTS as a potent thrombin inhibitor (Fig. 8) was converted into a viable lead structure, after a detailed understanding of the mode of action was derived by experimental structure determination of the corresponding protein–ligand complex [30]. At this point, the project can be envisioned as a textbook example of a 'random-goes-rational' approach, converting collected data into broadly explorable knowledge. The derived knowledge, notably that a 5,5-*trans*-fused lactam represents a molecular framework with intrinsic inhibitory potential towards serine proteases and, thus, addresses a family-wide commonality in structure and function of the active site, was further exploited by GSK by targeting additional members of that therapeutically relevant enzyme class. Indeed, in addition to thrombin inhibitors, the scaffold served as a core structure for potent elastase inhibitors such as GW311616 [35,36], human cytomegalovirus (HCMV) protease inhibitors such as GW427525 [37,38] and inhibitors for the hepatitis C virus encoded NS3/4A protease [39,40] (Fig. 8).

Table 2. Schematic overview on selected privileged structures that were proven to produce biologically active compounds for more than one member of a given target family[a]

Privileged core structure	Target family	Distinct members
 β sheet mimic	Serine protease	Elastase inhibitors Thrombin inhibitors
	Cysteine protease	ICE inhibitors Caspase 3 inhibitors
 Reverse turn mimic	Integrin	$\alpha_{IIb}\beta_3$ antagonists $\alpha_v\beta_3$ antagonists $\alpha_4\beta_1$ antagonists $\alpha_v\beta_5$ antagonists
 Reverse turn mimic	GPCR	Somatostatin antagonists Melanocortin antagonists
	Integrin	$\alpha_4\beta_1$ antagonists
 5,5 *trans*-fused lactams	Serine protease	Thrombin inhibitors HCMV PR inhibitors HLE, HNE inhibitors
 Hydroxyethylene dipeptide mimic	Aspartic acid protease	HIV PR inhibitors Cathepsin D inhibitors Plasmepsin inhibitors BACE inhibitors
 Hydroxamates	Metalloprotease	MMP inhibitors TACE inhibitors Aggracanase PDF inhibitors
 Inverse hydroxamates		

[a]Abbreviations: BACE, β-site APP-cleaving enzyme; GPCR, G-protein-coupled receptor; HCMV PR, human cytomegalovirus protease; HIV PR, human immunodeficiency virus protease; HLE, human leukocyte elastase; HNE, human neutrophil elastase; ICE, interleukin 1β converting enzyme; MMP, matrix metalloprotease; PDF, peptidyl deformylase, TACE, TNF-α converting enzyme.

Depending on the specific substitution of the common underlying bicyclic core structure, different serine proteases could be addressed by following the privileged structure approach. It is obvious that it was the detailed structural and mechanistical understanding of the 5,5-*trans*-lactam moiety within the context of serine protease inhibition that allowed a thorough exploration of that target family. This example shows that a privileged structure in its most stringent definition is not necessarily discovered, but might require optimization efforts, as demonstrated with the modification from the natural product encoded lactone towards the lactam structure and the optimization of the derivatization pattern around the core structure.

As soon as a proof-of-principle for a privileged structure hypothesis is delivered, that is a putative privileged structure serves, for example, as a core structure for low molecular weight compounds targeting more than one representative of a protein family, combinatorial chemistry protocols can be established that allow for the generation of target family biased compound libraries with built-in rationales based on experimental validation, rather than on more or less educated guesses. Additional optimized and tailor-made privileged structures, in their most stringent definition, are listed in Table 2.

Conclusion

Current pharmaceutical research needs to pay more attention to the core disciplines that are crucial in time and, more importantly, for the quality of the drug discovery pathway, thereby attempting to widen the persisting bottleneck of generating viable lead structures. Instead of overemphasizing the impact of a steadily growing number of '-omes' and '-omics' on the speed of research, focusing on target selection and medicinal chemistry will aid in closing the productivity gap [6–8]. A shift from biology to chemistry is urgently required. The concept of privileged structures in its most stringent definition can be envisioned as one approach to generate a more systematic view on the druggability of target families, thus finally delivering multipurpose molecular masterkeys for gene families of interest.

Apart from a few pharmaceutical companies [41], the majority of pharmaceutical research operates along the classical approach of working on one target at a time within a clearly defined

disease area. To complement this approach, a target family focus, utilizing molecular masterkeys derived from a privileged structure concept, enables numerous targets to be processed simultaneously across different therapeutic areas. This will require the medicinal chemist to actively look beyond project boundaries. The latter approach might fail in 'big pharma' owing to institutional barriers, but will have more success within smaller and flexible medicinal chemistry oriented biotechnology companies. Systematic exploration of therapeutically relevant target families by tailor-made medicinal chemistry concepts should refine the currently employed definition of chemogenomics, in that the 'chemo-' part is strengthened accordingly.

Acknowledgements

The author wishes to acknowledge Rachel van Rijn (N.V. Organon) for careful reading of the manuscript, and thanks all members of the Lead Discovery Unit of Organon Oss, especially Stan van Boeckel, for numerous fruitful discussions on various topics of drug discovery.

References

1 Drews, J. (2000) Drug discovery: a historical perspective. *Science* 287, 1960–1963
2 Drews, J. (1998) *In Quest of Tomorrow's Medicines*, Springer Verlag
3 Drews, J. (2000) Drug Discovery Today – and Tomorrow. *Drug Discov. Today* 5, 2–4
4 Drews, J. (2003) Strategic trends in the drug industry. *Drug Discov. Today* 8, 411–420
5 Thayer, A.M. (2003) Major drug firms eke out increases. *Chem. Eng. News* 81, 13–14
6 Drews, J. (1996) Genomic sciences and the medicine of tomorrow. *Nat. Biotechnol.* 14, 1516–1518
7 Drews, J. and Ryser, S. (1996) Innovation deficit in the pharmaceutical industry. *Drug Inf. J.* 30, 97–108
8 Drews, J. and Ryser, S. (1997) The role of innovation in drug development. *Nat. Biotechnol.* 15, 1318–1319
9 Bolten, B.M. and DeGregorio, T. (2002) Trends in development cycles. *Nat. Rev. Drug Discov.* 1, 335–336
10 Miller, H.I. (2002) As Biotech turns 20. *Nat. Rev. Drug Discov.* 1, 1007–1008
11 Tufts Center for the Study of Drug Development (2002) *Outlook 2002*, Boston
12 Dolle, R.E. (2002) Comprehensive survey of combinatorial library synthesis. *J. Comb. Chem.* 4, 369–418
13 Tollman, P. *et al.* (2001) A Revolution in R&D: How Genomics and Genetics are transforming the Biopharmaceutical Industry. (http://www.bcg.com/publications/files/eng_genomicsgenetics_rep_11_01.pdf)
14 Collins, F.S. *et al.* (1998) New goals for the U.S. Human Genome Project: 1998-2003. *Science* 282, 682–689
15 Szymkowski, D.E. (2001) Too many targets, not enough target validation. *Drug Discov. Today* 6, 398–399
16 Lipinski, C. *et al.* (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46, 3–26
17 Hopkins, A.L. and Groom, C.R. (2002) The druggable genome. *Nat. Rev. Drug Discov.* 1, 727–730
18 Evans, B.E. *et al.* (1988) Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 31, 2235–2246
19 Patchett, A.A. and Nargund, R.P. (2000) Privileged structures – an update. *Annu. Rep. Med. Chem.* 35, 289–298
20 Chang, R.S. *et al.* (1986) Tifluadom, a kappa-opiate agonist, acts as a peripheral cholecystokinin receptor antagonist. *Neurosci. Lett.* 72, 211–214
21 Roemer, D. *et al.* (1982) An opioid benzodiazepine. *Nature* 298, 759–760
22 Roche, O. *et al.* (2002) Development of a virtual screening method for identification of 'Frequent Hitters' in compound libraries. *J. Med. Chem.* 45, 137–142
23 McGovern, S.L. *et al.* (2002) A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* 45, 1712–1722
24 McGovern, S.L. and Shoichet, B.K. (2003) Kinase inhibitors: not just for kinases anymore. *J. Med. Chem.* 46, 1478–1483
25 Bemis, G.W. and Murcko, M.A. (1996) The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 39, 2887–2893
26 Bemis, G.W. and Murcko, M.A. (1999) Properties of known drugs. 2. Side chains. *J. Med. Chem.* 42, 5095–5099
27 Lewell, X.Q. *et al.* (1998) 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.* 38, 511–522
28 Hajduk, P.J. *et al.* (2000) Privileged molecules for protein binding identified from NMR-based screening. *J. Med. Chem.* 43, 3443–3447
29 Willoughby, C.A. *et al.* (2002) Combinatorial synthesis of 3-(Amidoalkyl) and 3-(Aminoalkyl)-2-arylindole derivatives: discovery of potent ligands for a variety of G protein-coupled receptors. *Bioorg. Med. Chem. Lett.* 12, 93–96
30 Weir, M.P. *et al.* (1998) Novel natural product 5,5-*trans*-lactone inhibitors of human α-thrombin: mechanism of action and structural studies. *Biochemistry* 37, 6645–6657
31 O'Neill, M.J. *et al.* (1998) Isolation of translactone-containing triterpenes with thrombin inhibitory activities from the leaves of lantana camara. *J. Nat. Prod.* 61, 1328–1331
32 Teague, S.J. *et al.* (1999) The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed. Engl.* 38, 3743–3748
33 Finch, H. *et al.* (1998) 5,5-trans lactone-containing inhibitors for serine proteases: identification of a novel, acylating thrombin inhibitor. *Bioorg. Med. Chem. Lett.* 8, 2955–2960
34 Pass, M. *et al.* (1999) Thrombin inhibitors based on 5,5 *trans*-fused indane lactams. *Bioorg. Med. Chem. Lett.* 9, 1657–1662
35 MacDonald, S.J.F. *et al.* (2001) The discovery of a potent, intracellular, orally bioavailable, long duration inhibitor of human neutrophil elastase - GW311616A a development candidate. *Bioorg. Med. Chem. Lett.* 11, 895–898
36 MacDonald, S.J.F. *et al.* (2002) Discovery of further pyrrolidine trans-lactams as inhibitors of human neutrophil elastase (HNE) with potential as development candidates and the crystal structure of HNE complexed with an inhibitor (GW475151). *J. Med. Chem.* 45, 3878–3890
37 Borthwick, A.D. *et al.* (2000) Design and synthesis of pyrrolidine-5,5-trans-lactams (5-Oxo-hexahydro-pyrrolo3,2-bpyrroles) as novel mechanism-based inhibitors of human cytomegalovirus protease. 1. The α-methyl-trans-lactam template. *J. Med. Chem.* 43, 4452–4464
38 Borthwick, A.D. *et al.* (2002) Design and synthesis of pyrrolidine-5,5-trans-lactams (5-Oxo-hexahydro-pyrrolo3,2-bpyrroles) as novel mechanism-based inhibitors of human cytomegalovirus protease. 2. Potency and Chirality. *J. Med. Chem.* 45, 1–18
39 Andrews, D.M. *et al.* (2002) Pyrrolidine-5,5-trans-lactams. 1. Synthesis and incorporation into inhibitors of hepatitis C Virus NS3/4A protease. *Org. Lett.* 4, 4475–4478
40 Andrews, D.M. *et al.* (2002) Pyrrolidine-5,5-trans-lactams. 2. The use of x-ray crystal structure data in the optimization of P3 and P4 substituents. *Org. Lett.* 4, 4479–4482
41 Wess, G. *et al.* (2001) Medicinal chemistry: challenges and opportunities. *Angew. Chem. Int. Ed. Engl.* 40, 3341–3350