16

SUCCESS STORIES OF COMPUTER-AIDED DESIGN

HUGO KUBINYI

Contents

Computer Applications in Pharmaceutical Research and Development, edited by Sean Ekins.
ISBN 0 471737798

16.1 INTRODUCTION

Rational approaches have been applied in drug discovery for at least a century. A striking example, with a surprising outcome, was the design of acetylsalicylic acid (ASS). In 1897, Felix Hoffmann synthesized this compound as a more tolerable "prodrug" of salicylic acid. Seventy years later it turned out that ASS has a unique mechanism of action through irreversibly inhibiting the enzyme cyclooxygenase. Many other drugs were developed from natural products and endogenous transmitters, by rational design. Nowadays, the term "rational design" is most often—incorrectly—applied as a synonym for structure-based and computer-aided design, which developed in the early 1970s. With the progress in protein crystallography, Peter Goodford was the first to use protein 3D structures to design ligands that fit a protein binding site [1, 2]. Two successful applications were published by his group.

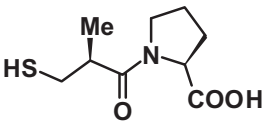
(1) The 3D structure of the 2,3-diphosphoglycerate (DPG) complex of hemoglobin (Hb) served to derive simple aromatic dialdehydes that mimic the function of DPG as an allosteric modulator of the oxygen affinity of Hb. Some of the resulting compounds were as active and even more active than DPG, the natural ligand [1–3].

(2) Trimethoprim analogs were designed as dihydrofolate reductase (DHFR) inhibitors, starting from the observation that a certain distance from one methoxy group of trimethoprim there is the guanidinium group of an arginine, which can favorably interact with a newly introduced acidic group of the ligand. Analogs with significantly enhanced affinities to bacterial DHFR resulted from this approach [1, 2, 4].

However, in the very end both projects failed with respect to "drug design": The Hb ligands do not permeate the erythrocyte membrane, and the trimethoprim analogs lost the high selectivity for bacterial DHFRs.

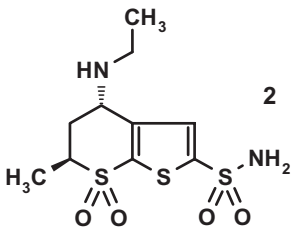
The design of the angiotensin-converting enzyme (ACE) inhibitor captopril [5, 6] may be considered as the first real success of structure-based drug design. Long-lasting attempts to derive bioavailable small molecule inhibitors from snake venom peptides were without much success. A breakthrough resulted from the 3D structure of carboxypeptidase A, another zinc protease, in complex with its inhibitor L-2-benzylsuccinic acid. A model of the ACE binding site guided the way to the weakly active ACE inhibitor lead structure *N*-succinoyl-L-proline (IC_{50} = 330 μM). The antihypertensive drug **captopril 1** (IC_{50} = 23 nM; Fig. 16.1) resulted after minor modifications, namely, the introduction of a methyl group (mimicking an alanine side chain) and an exchange of the carboxylate group with a sulfhydryl group [5, 6].

The topically active antiglaucoma agent **dorzolamide 2** (K_i = 0.37 nM; Fig. 16.1), a carbonic anhydrase inhibitor, may be considered as the first drug in the market that originated from the experimentally determined X-ray structure of its target protein. In the very last steps of its design, a favorable conformation of the six-membered ring was stabilized by the shift of a methyl



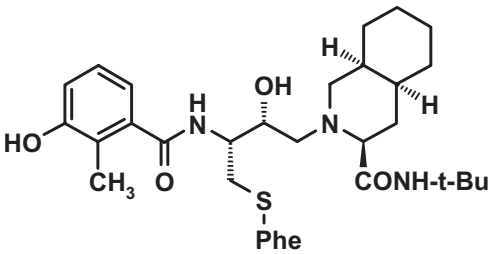
1

Captopril, ACE inhibitor,
IC_{50} = 23 nM



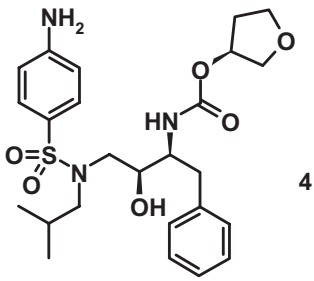
2

Dorzolamide, carbonic anhydrase
inhibitor, K_i = 0.37 nM



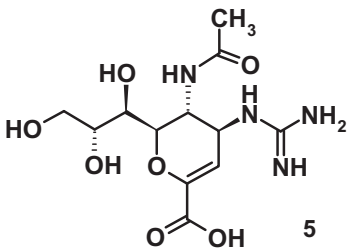
3

Nelfinavir,
HIV protease
inhibitor,
K_i = 2.0 nM



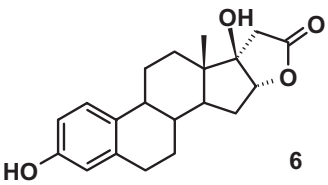
4

Amprenavir, HIV protease inhibitor,
K_i = 0.6 nM



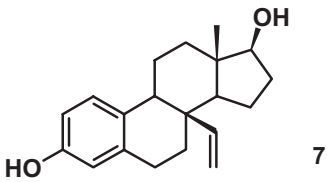
5

Zanamivir, neuraminidase inhibitor,
K_i = 0.1-0.2 nM



6

ERα-specific agonist,
40% E2 activity, 300-fold selectivity



7

ERβ-specific agonist,
50% E2 activity, 190-fold selectivity

Figure 16.1 Some success stories of structure-based design.

group of an *N*-alkyl substituent to this ring, in this manner enhancing the affinity of the molecule by a factor of two [7].

The very first HIV protease inhibitors in human therapy, saquinavir, indinavir, and ritonavir [8, 9], are often considered as typical examples of structure-based design. However, in reality they resulted from classic medicinal chemistry strategies (as did so many drugs in the decades before), starting from the peptide sequence of the cleavage site of the substrate. The 3D structure of HIV protease, which was available only relatively late, may have helped in understanding the details of the structure-activity relationships, but it did not contribute too much to the design. In the 1980s, several companies started to apply 3D structure-based ligand design as a strategic concept in drug discovery. The two most prominent companies, Agouron Pharmaceuticals and Vertex Pharmaceuticals, were both successful in designing the HIV protease inhibitors **nelfinavir 3** (K_i = 2.0 nM; Fig. 16.1) and **amprenavir 4** (K_i = 0.6 nM; Fig. 16.1), respectively. These drugs resulted from structure-based design, but both also contain some structural elements that were discovered in the design of the first HIV protease inhibitors [8, 9].

An example of a straightforward 3D structure-based design was published by von Itzstein and his group [10]. The enzyme neuraminidase (also called sialidase) is an essential coat protein of the influenza virus. It enables the virus to penetrate into the cell and to leave the cell after reproduction, by cleaving sialic acid from carbohydrate side chains at the cell surface. Correspondingly, the 3D structure of neuraminidase constituted a promising starting point for a structure-based design of anti-influenza drugs. Inspection of the surface of neuraminidase with the computer program GRID [11] indicated a pocket that could accommodate a relatively large positively charged group. Exchange of the –OH group of the weak transition state inhibitor Neu-5Ac-2en (K_i = 1 μM) with an ammonium group produced an inhibitor with K_i = 50 nM. If the larger guanidinium group was introduced instead, the strong inhibitor **zanamivir 5** resulted (K_i = 0.1–0.2 nM; Fig. 16.1) [10].

The design of estrogen receptor subtype-selective (ERα and ERβ) ligands is an exciting success story of homology modeling and structure-based design [12–14]. Hillisch et al. investigated the known 3D structure of the human ERα ligand-binding domain (LBD) to derive a homology model of the human ERβ LBD. There are minor but distinct differences in the estradiol binding cavity of the subtypes. Whereas the β side, "above" the steroid ring system, is relatively narrow in ERα, because of a leucine side chain in position 384, there is more space in ERβ, because of a flexible methionine side chain. On the other hand, a methionine in position 421 of ERα is replaced by an isoleucine in ERβ, making the α side of ERβ, "below" the steroid, narrower. Estradiol (E2) analogs were designed to use these structural differences for subtype selectivity, producing the ERα- and ERβ-selective **ligands 6** (40% E2 activity, 300-fold selectivity) and **7** (50% E2 activity, 190-fold selectivity) (Fig. 16.1) [12–14].

There are now many success stories of structure-based design of potent and selective ligands. As these examples have been extensively discussed in books [15–17] and reviews [e.g., 18–24], they are not be repeated here. When combinatorial chemistry and high-throughput screening developed as new approaches to synthesizing and screening thousands, tens of thousands, or even hundreds of thousands of new compounds, it was anticipated that this would generate an unprecedented number of new drugs, marking a milestone in drug discovery. However, the opposite was the case [25, 26]. Most often, screening hits could not be validated or optimized to leads and preclinical candidates. Many such compounds were too large and too lipophilic, too greasy, and they only showed up in the biological tests because of nonspecific binding. It was the merit of Chris Lipinski to take a closer look at the physicochemical properties of biologically active molecules. By an inspection of 2245 drugs and clinical candidates from the World Drug Index he formulated
1 his now famous "rule of five" (Lipinski rule, Pfizer Ro5): To achieve permeability (which is a precondition for oral bioavailability), a molecule should not violate more than one of the following rules: the molecular weight must not be larger than 500, lipophilicity should not be larger than log P = 5, and the molecule should not contain more than 5 hydrogen bond acceptors and not more than 10 N + O atoms (as a rough measure of hydrogen bond acceptors) [27]. However, the rule of five defines only druglike properties, not necessarily druglike character, as expressed by structural features that are typical for drug molecules. This differentiation can be achieved by neural nets that have been trained with drugs (e.g., the World Drug Index or the MDDR) and nondrugs (e.g., the Available Chemicals Directory) [28, 29]. Such neural nets do not allow a discrimination between active and inactive compounds, but they separate druglike structures from mere chemicals, that is, from compounds that contain atypical chemical features, providing about 80% correct assignments to each group.

Molecular modeling plays an important role in all steps of lead discovery and lead optimization. Several computer-aided techniques for automated database searches and docking into protein 3D structures have developed over time. If only ligand structures are available but no 3D structures of the biological target, as until recently was the case for all membrane-embedded proteins, pharmacophore generation and 2D or 3D searches in structural databases are the method of choice [e.g., 30–33]. Starting with the programs DOCK [34] and LUDI [35], the docking of ligands into the binding sites of various proteins, for which 3D structures are available, is now a well-established technique [e.g., 36–45]. A certain problem is the poor reliability of the scoring functions that rank the docking results [e.g., 46–49]. Extensive comparisons of different docking programs and scoring functions [e.g., 50–53], to rediscover known ligands within 3D databases provide evidence that there is no unique solution to the problem. Certain docking and scoring combinations are appropriate for one target, whereas they fail with another target. Consensus scoring, that is, the simultaneous use of several different scoring

functions, has been proposed to solve this problem [54, 55]. However, for the most common programs the quality of the obtained results seems to depend more on the experience and skill of the modeler than on the options used. Scoring functions also tend to overestimate the affinity of large molecules [56]. In this context, a posteriori inspection of all docking results is of utmost importance.

By combination of several techniques, from simple filters and pharmacophore searches to docking and scoring, virtual screening developed as a new paradigm in computer-aided ligand design. In contrast to real, "wet" biological screening, virtual screening opens new dimensions: It offers a number of different approaches for the selection of compounds or sublibraries out of huge in-house inventories, compound libraries of commercial suppliers, or virtual libraries, that is, structures that exist only in the computer, not in reality. Such techniques are rule-based or quantitative filters, neural nets, QSAR, 2D and 3D pharmacophore-derived models, and docking and scoring.

Drug research has often been compared with the search for a needle in a haystack. Indeed, this comparison is valid, for two reasons. First, huge numbers of candidates must be investigated in drug research to discover a lead that can be further optimized to a drug candidate. Second, special technologies should be applied to find a needle in a haystack, for example, a magnet; in the very same manner, virtual screening solves the haystack problem of drug discovery by searching for compounds with favorable properties, be it drug-like character, bioavailability, the fit to a pharmacophore, or the complementarity to a binding site. Despite the fact that virtual screening is a relatively young discipline, it has already been reviewed in books [57–59] and in many dedicated publications [60–77].

Retrospective virtual screening studies, in which only known actives are retrieved, are not included in this review, as well as mere enrichment studies and virtual screening, from which some predictions but no experimental confirmation have resulted. Only a few pharmacophore studies without additional filters or docking and scoring are included. To keep this chapter to a reasonable size, no details or references are provided for the individual biological targets and test systems, lead structures, databases, compound collections and libraries, and computer programs that were used in the virtual screening; for all these details the reader is referred to other chapters of this book and to the references of the individual case studies (some references for the most popular computer programs are given in Section 16.6).

Because most often several different techniques of virtual screening are applied in certain combinations, the discussed examples are not ordered by the applied approach but according to the biological targets. Nevertheless, ligand-based approaches and/or homology modeling and docking into a protein 3D model are in the foreground for receptors and ion channels, whereas docking into experimental 3D structures is preferentially applied for enzymes and other soluble proteins.

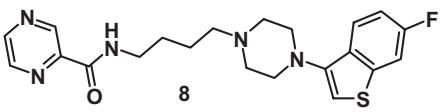
16.2 RECEPTORS

16.2.1 G Protein-Coupled Receptors

α_{1A} Adrenergic Receptor. A model of the α_{1A} adrenergic receptor was generated by ligand-supported homology modeling, based on the high-resolution X-ray structure of bovine rhodopsin and also using mutational and ligand SAR data. Virtual screening of the Aventis in-house compound repository was then performed in a stepwise manner. First, compounds with more than nine rotatable bonds and molecular weight >600 were eliminated; then 22,950 compounds were selected, using a α_{1A} receptor ligand pharmacophore hypothesis and the program Catalyst. These compounds were docked into the α_{1A} receptor homology model with the program GOLD and scored with PMF, after calibration of the docking procedure and evaluation of different scoring functions with a data set of 50 α_{1A} receptor antagonists and 990 druglike molecules from the MDDR database. The top-scoring 300 compounds were clustered according to their Unity fingerprint similarity, and a diverse set of 80 compounds was tested in a radioligand displacement assay. Of 37 compounds with K_i <10 μM, the most active hit was **compound 8** (K_i = 1.4 nM; Fig. 16.2) [78].

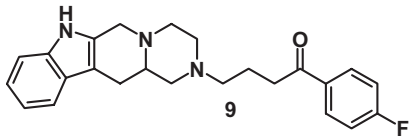
Dopamine D3 Receptor. The 3D structure of the dopamine 3 (D3) subtype receptor was also modeled from the X-ray structure of rhodopsin, with extensive structural refinement and validation using experimental data. A D3 pharmacophore model was derived from 10 potent and moderately selective known D3 receptor ligands. This pharmacophore model served to search 250,251 compounds from the National Cancer Institute (NCI) 3D database with the program Chem-X. The 6727 resulting hits were docked into four major conformational clusters of the D3 receptor, and ranking of the results was performed with the scoring function of the Cerius2 program. As an independent validation, 20 known D3 ligands were added to the set of 6727 compounds. The hit list of 2478 potential ligands was then filtered for known chemotypes. After removal of all compounds that are structurally similar to known D3 receptor ligands, 1314 candidates remained. Of 60 compounds requested from the NCI, only 20 were available in sufficient quantity. Eight of them had K_i values below 500 nM, for example, **compound 9** (K_i = 11 nM; Fig. 16.2) [79].

Endothelin A Receptor. A common pharmacophore for endothelin A (ET_A) receptor ligands was derived from a cyclic pentapeptide and a triterpene ester. The moderately selective lead structure 2,4-dibenzyloxybenzoic acid (IC_{50} ET_A = 9 μM, ET_B < 20% at 30 μM) was discovered by a 3D pharmacophore search in 60,000 compounds of the Rhone Poulenc Rorer UK corporate database with the ChemDBS-3D system [80]. The highly selective ET_A receptor **ligand 10** (IC_{50} ET_A = 5 nM, IC_{50} ET_B >10 μM; Fig. 16.2) resulted



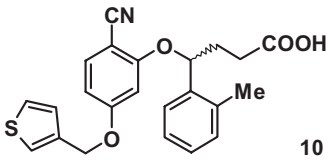
8

α_{1A} adrenergic receptor antagonist,
K_i = 1.4 nM



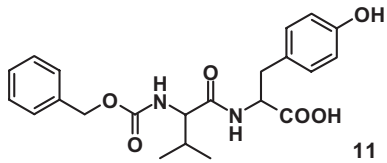
9

Dopamine D3 receptor antagonist,
K_i = 11 nM



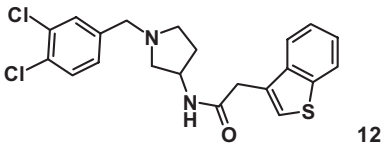
10

Endothelin A (ET_A) receptor antagonist,
(R)-isomer, IC_{50} = 5 nM



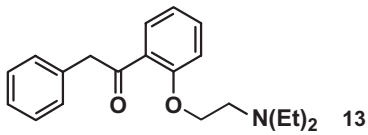
11

Endothelin A (ET_A) receptor antagonist
IC_{50} = 220 nM



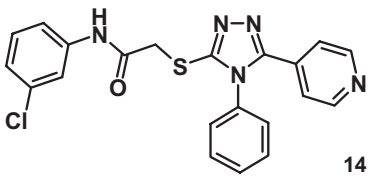
12

MCH-1 receptor ligand,
K_i = 360 nM



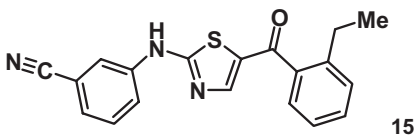
13

Muscarinic M3 receptor antagonist,
[A_{50}] M3≈0.2 μM (pA_2 = 6.67)



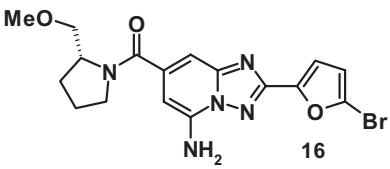
14

Neurokinin NK_1 antagonist,
K_i = 251 nM



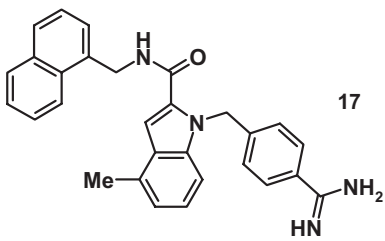
15

Neuropeptide Y (NPY5) receptor
antagonist, IC_{50} = 2.8 nM



16

Purinergic A_{2A} receptor antagonist,
K_i A_{2A} = 2.4 nM, K_i A_1 = 292 nM



17

Urotensin II receptor antagonist,
EC_{50} = 400 nM

Figure 16.2 GPCR ligands from virtual screening.

from a refined pharmacophore hypothesis and further chemical optimization [81, 82].

Another research group generated two pharmacophore models of ET_A-selective receptor antagonists from a training set of 18 ET_A antagonists by using the HypoGen and HipHop options of the program Catalyst. The best HypoGen hypothesis had five pharmacophoric features: two hydrophobic features, one aromatic ring, one acceptor, and one negative ionizable function. The highest-scoring HipHop model had six features: three hydrophobic features, one aromatic ring, one acceptor, and one negative ionizable group. The predictive power of the quantitative models was validated by their ability to extract a test set of 30 known ET_A antagonists from the World Drug Index. A 3D search of 55,000 compounds in the Maybridge database retrieved 498 hits from the HypoGen hypothesis and 5 hits from the HipHop hypothesis. After visual inspection, six hits from both analyses were selected for testing, of which four were biologically active, for example, **compound 11**, Z-Val-Tyr-OH (IC_{50} = 220 nM; Fig. 16.2) [83].

Melanin-Concentrating Hormone Type 1 Receptor. A "drug space" was defined by a BCUT metrics analysis of 81,560 drugs and druglike molecules. The resulting five-dimensional model (hydrogen bond donor and acceptor, two terms for polarizability, and charge) was used to locate the space for peptide G protein-coupled receptor (GPCR) ligands. Analysis of a virtual library of 9 million compounds, constructed from 19 predefined amine templates, yielded 2025 hits. After synthesis and biological testing, potent ligands of the GnRH, galanin, MC4, melanin-concentrating hormone (MCH), orexin, and other peptide GPCRs resulted, with a 4.5-fold (GnRH receptor) to 61-fold (MC4 receptor) enrichment of active analogs, as compared to a random selection of screening compounds from the Neurocrine Biosciences in-house compound repository. Out of several micromolar and submicromolar ligands of the MCH1 receptor, **compound 12** (K_i = 360 nM; Fig. 16.2) had the highest affinity [84].

Muscarinic M3 Receptor. A pharmacophore model was derived from known M3 receptor antagonists, using the program DISCO, and 3D searching was performed by Unity 3D in the Astra Charnwood in-house compound repository and the databases of several commercial suppliers. The 172 compounds that fitted the pharmacophore were screened for their M3-antagonistic potency. Several compounds with micromolar and even submicromolar activities resulted, for example, **compound 13** (A_{50} M3 antagonism $\approx$ 0.2 μM; pA_2 = 6.67; Fig. 16.2) [85].

Neurokinin-1 Receptor. A homology model of the neurokinin-1 (NK_1) receptor was built from the X-ray structure of rhodopsin, using the MOBILE (modeling binding sites including ligand information explicitly) approach. In this procedure, a preliminary model is generated, which is afterwards refined

by docking known ligands into the model. From this model a pharmacophore hypothesis was derived to search eight structural databases for molecules that fit this hypothesis. The workflow shows in an elegant manner how to perform stepwise virtual screening. From the 826,952 compounds of the various databases only 419,747 (51%) molecules passed a filter for molecular weight <450 and less than eight rotatable bonds; 131,967 molecules (16%) had the requested number of hydrophobic, donor, and acceptor properties, and 36,704 molecules (4.4%) fitted the pharmacophore hypothesis in 2D and 3D (database searches with Unity). On the basis of excluded volumes, 11,109 (1.34%) structures remained for docking into the modeled NK_1 receptor binding site, using FlexX-Pharm; the resulting docking poses were ranked with the knowledge-based scoring function DrugScore. The 1000 highest-scoring ligands were force field-minimized in the binding pocket and visually inspected for certain typical receptor-ligand interactions and features: (1) an amino-aromatic interaction between $His197^{5.39}$ and an aromatic ring; (2) a stacking between two aromatic rings; (3) a hydrogen bond between $Gln165^{4.60}$ and an acceptor group of the ligand; (4) similarity to known NK_1 receptor ligands in the β4-hairpin region; and (5) a small number of rotatable bonds. Of seven compounds for biochemical screening, **compound 14** (K_i = 251 nM; Fig. 16.2) showed submicromolar affinity [86, 87]. This result is especially remarkable because **compound 14** does not contain the "magic" 3,5-bis-trifluoromethyl substitution pattern of most highly active NK_1 receptor ligands.

NPY5 Receptor. A pharmacophore hypothesis for NPY5 receptor ligands was derived from three known ligands and used for a Catalyst 3D search in the Hoffmann-La Roche in-house compound repository. Of 632 retrieved molecules, 31 had IC_{50} values <10 μM. The most interesting compound was a substituted aminothiazole (IC_{50} = 40 nM), which after two cycles of chemical optimization resulted in some more nanomolar ligands, for example, **compound 15** (IC_{50} = 2.8 nM; Fig. 16.2) [88].

Purinergic A_{2A} Receptor. The CATS (chemically advanced template search) descriptor compares molecules by the topological pattern of their pharmacophore features [89]. Based on these descriptors, a self-organizing map (SOM) was generated from biologically active molecules, including purinergic receptor antagonists. Virtual libraries were designed from a triazolopyridine carboxylic acid and secondary amines. Projection of the resulting amides onto this map identified several hits with high affinity and selectivity, the most selective A_{2A} antagonist being **compound 16** (K_i A_{2A} = 2.4 nM, K_i A_1 = 292 nM; Fig. 16.2) [90].

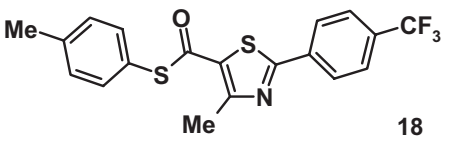
Urotensin II Receptor (GPR14). The vasoactive cyclic peptide urotensin II (U-II) is the endogenous ligand of the G protein-coupled orphan receptor GPR14. Structure-activity relationships from 25 peptide analogs, which mobilize intracellular calcium in GPR14-transfected CHO cells, and the

NMR 3D structure of the undecapeptide U-II generated a ligand pharmacophore hypothesis that served as query for the virtual screening of the Aventis in-house compound repository. Active leads from six different chemical classes could be identified by the 3D search, for example, **compound 17** (EC_{50} = 400 nM; Fig. 16.2) [91].

16.2.2 Nuclear Receptors

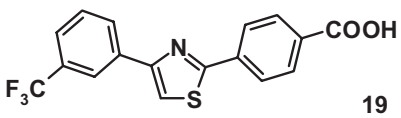
Retinoic Acid Receptor. A 3D structural model of the inactive conformation of the retinoic acid receptor (RAR) α-subtype (RARα) was developed from the RARγ 3D structure, bound to the agonist all-*trans*-retinoic acid, and the estrogen receptor α-subtype (ERα), bound to an antagonist. After validation of the method with known agonists and antagonists, 153,000 ACD compounds were docked into the RAR binding site with full flexibility of the ligand and the amino acid side chains of the protein, using the Molsoft Internal Coordinates Mechanics (ICM 2.7) program. Two novel RAR antagonists were discovered, for example, **compound 18** (55% inhibition at 20 μM; Fig. 16.3) [92]; comparable results were obtained with all three human isoforms: RARα, RARβ, and RARγ.

In a similar investigation, a model of the active RARα conformation was developed from the agonist-bound RARγ conformation. Docking of the ACD compounds as above but with a refined procedure, considering all atoms of the binding site, resulted in 5364 high-scoring hits. The 300 compounds with the lowest binding energy (i.e., highest predicted affinity) were visually inspected for shape complementarity, hydrogen bonding network, ligand conformations, and possible van der Waals clashes. Finally, 30 compounds were selected for biological testing. Despite the fact that an RARα 3D model was



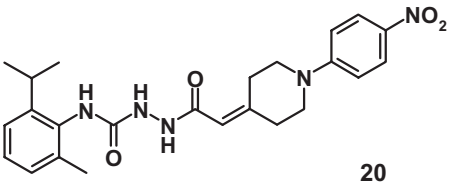
18

Retinoic acid receptor α (RARα) antagonist, 55% inhibition at 20 μM



19

Retinoic acid receptor β (RARβ) agonist, EC_{50} RARβ = 200 nM, EC_{50} RARα = 4 μM



20

Thyroid hormone receptor antagonist, IC_{50} = 0.75 μM

Figure 16.3 Nuclear receptor ligands from virtual screening.

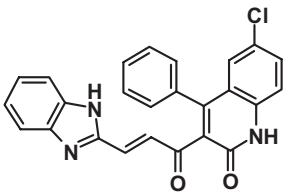
used for the docking, the two most active hits have a higher affinity to RARβ than to RARα, for example, **compound 19** (EC_{50} RARβ = 200 nM, EC_{50} RARα = 4 μM; Fig. 16.3) [93].

Thyroid Hormone Receptor. The 3D structure of the estrogen antagonist raloxifen, bound to the estrogen receptor α-subtype, was used to derive the "antagonist-binding" conformation of the thyroid receptor β-subtype (TRβ) from the 3D structure of an agonist complex of the TRβ ligand binding domain. Five grid potential representations of the binding site were generated by the Molsoft ICM virtual library screening module, accounting for hydrophobicity, van der Waals boundaries, hydrogen bonds, and electrostatic potential of the ligand binding site. The 190,000 rule of five-compatible compounds, out of 250,000 ACD compounds, were docked four times into the receptor grids by the ICM method, and the lowest score (i.e., best fit) of each ligand was retained. The geometry of the top 1000 ligand-protein complexes was refined, and the remaining 300 top-scoring complexes were visually inspected. Of 100 biologically tested compounds, 14 turned out to be TR antagonists. The most active hit (90% inhibition at 20 μM) served as the lead to construct a virtual library of a further 101 analogs. After docking, eight high-scoring compounds were synthesized and tested; all inhibited TR to 10–84% at 5 μM, the most active antagonist being **compound 20** (IC_{50} = 0.75 μM; Fig. 16.3) [94].

16.3 ENZYMES

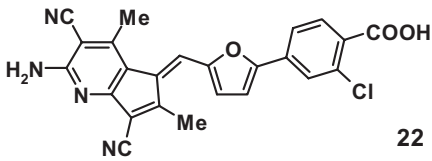
16.3.1 Kinases

Akt 1 (Protein Kinase Bα, PKBα). The three isoforms of protein kinase B are Akt 1 (PKBα), Akt 2 (PKBβ), and Akt 3 (PKBγ). A 3D structure of the binding site was extracted from the X-ray structure of a ternary complex of Akt1, a nonhydrolyzable ATP analog, and a peptide substrate derived from the binding sequence of glycogen synthase kinase 3β (GSK-3β). About 50,000 ChemBridge compounds were docked into this binding site in a flexible manner, using the program FlexX. The top 2000 compounds were ranked with the consensus scoring function CSORE; the top 100 compounds from the knowledge-based scoring function DrugScore, the top 200 compounds from GoldScore, and the top 200 compounds from ChemScore ranking were biologically tested. Only one hit, **compound 21** (IC_{50} = 4.5 μM, K_i = 3.9 μM; Fig. 16.4) resulted. To improve the result, the 4000 top-ranking compounds from FlexX and DrugScore were ranked according to GoldScore and ChemScore. Two hundred compounds were selected, which showed up within the top 700 rankings of both functions. From this set, 100 compounds were eliminated after visual inspection and 100 compounds were biologically tested. In addition to **compound 21** another low micromolar Akt1 inhibitor, **compound 22** (IC_{50} = 2.6 μM, K_i = 1.1 μM; Fig. 16.4), resulted [95].



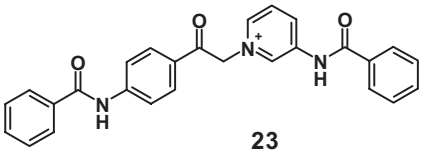
21

Akt 1 (protein kinase Bα, PKBα) inhibitor, IC_{50} = 4.5 μM, K_i = 3.9 μM



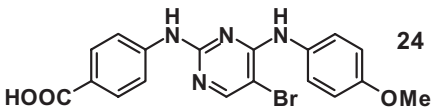
22

Akt 1 (protein kinase Bα, PKBα) inhibitor, IC_{50} = 2.6 μM, K_i = 1.1 μM



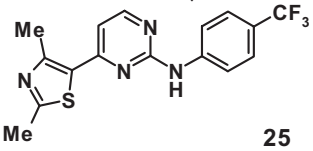
23

Bcr-abl tyrosine kinase inhibitor, IC_{50} = 24 μM



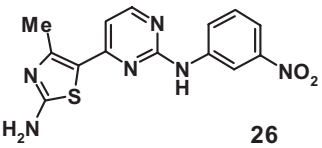
24

Checkpoint kinase 1 (Chk-1) inhibitor, IC_{50} = 450 nM



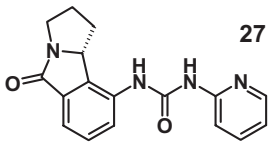
25

Cdk2 inhibitor, K_i Cdk2 = 0.29 μM, K_i Cdk4 > 20 μM



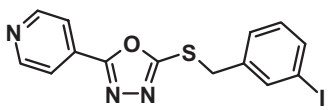
26

Cdk2 inhibitor, K_i Cdk2 = 2 nM, K_i Cdk4 = 53 nM



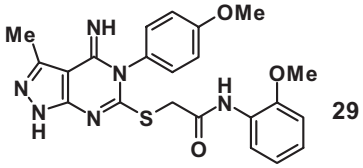
27

Cdk4 inhibitor, IC_{50} = 42 nM



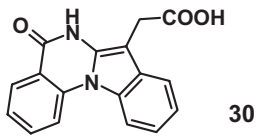
28

Glycogen synthase kinase 3 (GSK-3) inhibitor, IC_{50} = 0.39 μM



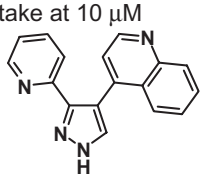
29

p56 lymphoid T cell tyrosine kinase (Lck) inhibitor, ≈75% inhibition of ^{3}H-thymidine uptake at 10 μM



30

Protein kinase CK2 (casein kinase II) inhibitor, IC_{50} = 80 nM



31

TGFβ receptor (TβRI) kinase inhibitor, IC_{50} = 27 nM, K_d = 5 nM

Figure 16.4 Kinase inhibitors from virtual screening.

Bcr-Abl Tyrosine Kinase. A database of 200,000 compounds of the ChemDiv compound collection was converted into 3D format and docked into the binding site of bcr-abl tyrosine kinase, using the program DOCK 4.0.1 for flexible docking. The 1000 top-scoring compounds were clustered by their molecular fingerprints. After filtering by the Lipinski rule of five, 15 compounds from diverse sets were selected for biological testing; eight of these compounds inhibited K562 tumor cell growth with IC_{50} values between 10 and 200 μM, for example, **compound 23** (IC_{50} = 24 μM; Fig. 16.4) [96].

Checkpoint Kinase 1. A subsection of the AstraZeneca in-house compound collection, containing 560,000 compounds, was used for virtual screening for checkpoint kinase 1 (Chk-1) inhibitors. Compounds with molecular weight >600 or with more than 10 rotatable bonds were removed, leaving about 400,000 compounds. Protonation and tautomeric states were corrected with the in-house program Leatherface. Then 3D structures were generated with Corina, with explicit enumeration of stereocenters (generating a maximum of 8 stereoisomers per molecule), and a multiconformer database was produced, using the program Omega. A 3D pharmacophore search was performed with the in-house program Plurality to eliminate compounds that do not have the typical binding motif for the kinase hinge region. The remaining 199,000 compounds (1 conformer per molecule) were flexibly docked into the ATP binding site of Chk-1, using the program FlexX-Pharm, which considers full flexibility of the ligand and demands certain interactions with the binding site, in this case to the backbone NH of Cys 87 and the backbone carbonyl of Glu 85. An enrichment study for known cyclin-dependent kinase 2 (Cdk2) inhibitors served to select the best consensus scoring, resulting in a combination of the FlexX and PMF scoring functions. Visual inspection of the 250 highest-scoring hits for unfavorable interactions with the binding site or compounds with unrealistic conformations resulted in a list of 103 compounds for biological testing. Inhibitory activities of 36 hits were in the range of 110 nM to 68 μM, for example, compound **24** (IC_{50} = 450 nM; Figure 16.4) [97].

Cyclin-Dependent Kinase 2. The flexible docking program LIDAEUS (developed from the program Sandock) was used to dock a database of about 50,000 commercially available compounds into the known 3D structure of the kinase Cdk2, to search for new chemotypes of Cdk inhibitors. Biochemical screening of 200 hits provided moderately active inhibitors. Structure-based modification led to the selective Cdk2 inhibitor **compound 25** (IC_{50} Cdk2/cyclin E = 0.9 μM; IC_{50} Cdk4/cyclin D1 = 5.5 μM [98]; K_i Cdk2 = 0.29 μM, K_i Cdk4 > 20 μM [99]; Fig. 16.4) with antiproliferative activity against tumor cells in vitro and in vivo. Further chemical optimization of **compound 25** produced the moderately selective nanomolar inhibitor **compound 26** (K_i Cdk2 = 2 nM, K_i Cdk4 = 53 nM; Fig. 16.4) [99].

Cyclin-Dependent Kinase 4. A homology model of the cyclin-dependent kinase Cdk4 was constructed from the X-ray structure of activated Cdk2, to perform a structure-based design of Cdk4 inhibitors. For this purpose the de novo design program LEGEND was combined with the program SEEDS (system for evaluation of availability of essential structures generated by de novo ligand design programs). LEGEND constructs ligands within the binding site of a protein on an atom-by-atom basis, and SEEDS extracts relevant scaffolds from the generated ligands to search databases for commercially available or synthetically feasible building blocks or analogs. On searching the ACD, 4884 compounds with molecular weight <350 were retrieved. After visual inspection, 382 compounds were purchased and tested in a cyclin D-Cdk4 complex assay. Eighteen compounds with IC_{50} values <500 μM were identified and clustered into four classes of new scaffolds. A diarylurea class could be further improved in biological activities by a dedicated library design. A docking study confirmed the binding mode of these ligands in the ATP binding pocket of the Cdk4 model. Further modifications led to the Cdk4 inhibitor **compound 27** (IC_{50} = 42 nM; Fig. 16.4) [100].

Glycogen Synthetase Kinase. Inhibitors of glycogen synthase kinase 3 (GSK-3), a serine protein kinase, may play a role in the treatment of diabetes. To search for potential ligands, 32 different virtual libraries with up to about 1.25 million compounds per library were generated. Then 47 hits from GSK-3 inhibitor screening were compared with up to 10,000 compounds from each of these libraries. CATS-2, a modification of the CATS descriptor, which compares molecules by the topological pattern of pharmacophore features assigned to atom pairs [89], was used for similarity search of each of the screening hits against 137,842 molecules that were randomly selected from the different virtual libraries. Whereas a classic 2D fingerprint similarity search did not provide any hits with a Tanimoto index >0.85, the CATS-2 search indicated that one of the virtual libraries had a high similarity to the screening hits. Filtering, library syntheses, and further optimization, including scaffold hopping, led to **compound 28** (IC_{50} = 0.39 μM; Fig. 16.4) [101].

p56 Lymphoid T Cell Tyrosine Kinase. The p56 lymphoid T cell tyrosine kinase (Lck) participates in protein-protein interactions through its Src homology-2 (SH2) domain. Virtual screening was performed, using the X-ray structure of the Lck SH2 domain complex with a pYEEI (pY+3) peptide. A 3D database of 2 million commercially available compounds was built with the 3D generator CORINA and docked into the pY+3 binding site with the program DOCK, using flexible ligands based on the anchored search method. Some further filters selected 25,000 compounds that were more rigorously docked by simultaneous energy minimization of the anchor fragment during the iterative build-up procedure. Two sets of 1000 compounds were selected on the basis of either the total interaction energy or a molecular weight-

normalized energy score, to account for the often observed overprediction of large molecules cf. 56]. Similarity clustering was performed for both sets, and compounds from the different clusters were selected according to their physicochemical properties. Thirty-four of 196 selected compounds, without a phosphotyrosine (pY) or a structurally related feature, inhibited Lck. Twenty-four of the active compounds were tested for their modulation of biological function: Thirteen showed inhibitory activity in a lymphocyte culture assay, for example, **compound 29** (~75% inhibition of ^{3}H-thymidine uptake at 10 μM; Fig. 16.4) [102].

Protein Kinase CK2 (Casein Kinase II). A homology model of human protein kinase CK2 (casein kinase II) was generated from the 3D structure of the highly homologous CK2 of *Zea mays*. Docking of 400,000 compounds of the in-house corporate collection of Novartis was performed with the program DOCK. The results were filtered according to the following criteria: Only compounds showing the typical hydrogen bond interaction to the hinge region of the kinase binding site were selected; results were ranked by a second scoring function and visually inspected for unrealistic conformations or unfavorable interactions. Four of twelve biologically tested compounds showed >50% inhibition at 10 μM, the most potent inhibitor being **compound 30** (IC_{50} = 80 nM; Fig. 16.4) [103].

TGFβ Receptor (TβRI) Kinase. TGFβ receptor (TβRI) kinase is activated by its association with the TGFβ type II receptor (TβRII). The activated kinase phosphorylates Smad substrates, which then induce TGFβ-dependent gene expression. The X-ray crystal structure of the unphosphorylated cytoplasmatic region of TβRI in complex with FKBP12, an inhibitor of the TGFβ pathway, served for a structure-based virtual screening to discover novel inhibitors. A starting point of the design was a pharmacophore hypothesis, derived from the experimental X-ray structure of the 2,4,5-triarylimidazole SB 203580 (IC_{50} = 30 μM) in the ATP binding site of TβRI. The pharmacophore search, which also considered shape constraints, identified 87 compounds from a commercially available database of 200,000 molecules, for example, **compound 31** (IC_{50} = 27 nM, K_d = 5 nM; Fig. 16.4) [104].

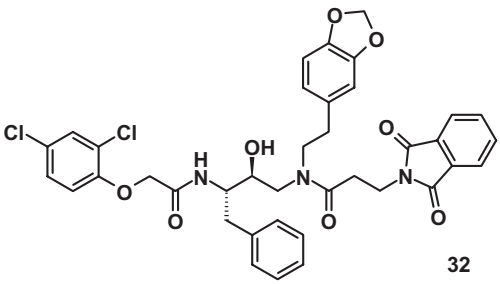
16.3.2 Proteases

Cathepsin D. The design of inhibitors of the aspartyl protease cathepsin D started from a virtual library of peptide analogs that contained the typical hydroxyethylamine isoster for the cleavable peptide bond. As the availability of starting materials would have generated a library of about 1 billion compounds, virtual screening was applied to reduce this multitude of candidate structures to a reasonable number. The backbone of a peptide

analog was docked into the active site of cathepsin D in the same pose as the natural product inhibitor pepstatin. Then fitting side chains for three different pockets of the binding site were selected by the program CombiBuild, which was developed from the program DOCK. A library of 1000 compounds resulted from this procedure, which in the following virtual screening was compared with a diversity-oriented library of peptide analogs. Whereas the directed library produced seven hits with IC_{50} values <100 nM, only one such hit resulted from the diversity-oriented library. In a further step, the best results from the directed library were the starting point for another directed library of 39 compounds. The inhibitor with the highest activity was **compound 32** (K_i = 9 nM, IC_{50} = 14 nM; Fig. 16.5) [105].

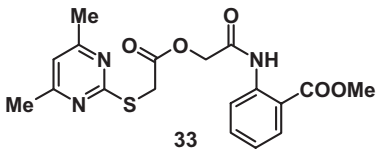
Falcipain-2. The 3D structures of the binding pockets of protozoal cysteine proteases are highly conserved. Homology models of the malarial cysteine proteases falcipain-2 and falcipain-3 were used for stepwise virtual screening of 241,000 compounds of the ChemBridge database. First, filters were applied to eliminate metal complexes and counterions, to neutralize charged compounds, and to eliminate compounds with inappropriate ADME properties, poor solubility, and violations of the Lipinski rule of five. 3D structures of the 60,000 compounds of this filtered database were generated and subjected to docking with the program GOLD, using three different protocols that were generated by docking a vinyl sulfone inhibitor into the cysteine protease cruzain. The first two rounds of docking, with 7–8 times and 2 times speed-up as compared to the standard protocol, were performed with the somewhat narrower binding pocket of falcipain-3. The remaining 1500 candidates were docked into both protein binding pockets, using the standard mode settings of GOLD. In both cases 10 known vinyl sulfone inhibitors were included, which showed up in the 20 highest-ranking ligands. The top 200 common hits for both proteins were visually inspected for reasonable geometry of the ligand, proximity of an electrophilic center (if present) to the SH group of the catalytic cysteine, and complementarity of the ligand and the protein. Of 100 selected compounds, 84 were biologically tested to identify 24 diverse inhibitors, of which 12 compounds are dual inhibitors of falcipain-2 and falcipain-3, with IC_{50} values between 1 and 62 μM; although many of these inhibitors are either Schiff bases or hydrazones, some of them have druglike structures, for example, **compound 33** (IC_{50} falcipain-2 = 6.2 μM, IC_{50} falcipain-3 = 12.0 μM; Fig. 16.5) [106]. Five compounds additionally inhibited *Leishmania donovani* cysteine protease, whereas four other, noninhibiting compounds showed strong antileishmanial activity in *L. donovani* promastigotes, obviously by a different mechanism of action.

HIV Protease. Docking of the 3D structures of the Cambridge Structural Database into the HIV protease binding site, by shape and to some extent by chemical complementarity, was performed with an early version of the



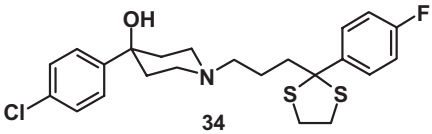
32

Cathepsin D inhibitor,
K_i = 9 nM, IC_{50} = 14 nM



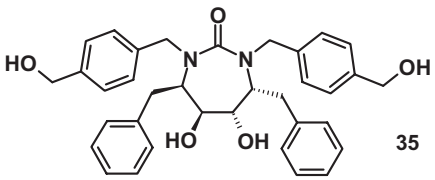
33

Falcipain inhibitor,
IC_{50} falcipain-2 = 6.2 μM,
IC_{50} falcipain-3 = 12.0 μM



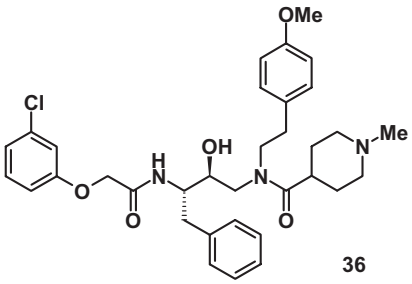
34

HIV protease inhibitor,
K_i HIV-1 protease = 15 μM,
K_i HIV-2 protease = 100 μM



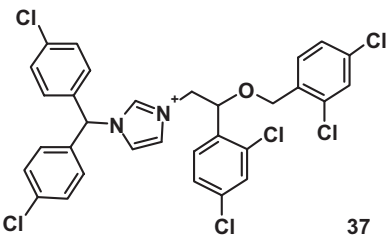
35

HIV protease inhibitor,
K_i = 0.27 nM, IC_{50} = 36 nM



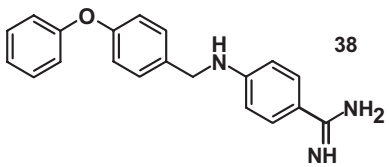
36

Plasmepsin II inhibitor,
K_i plasmepsin II = 2.0 nM,
K_i cathepsin D = 9.8 nM



37

SARS CoV 3C-like proteinase inhibitor,
K_i = 61 μM



38

Thrombin inhibitor,
K_i thrombin = 95 nM, K_i trypsin = 520 nM

Figure 16.5 Protease inhibitors from virtual screening.

program DOCK. Of the 200 top-scoring hits, 50 were commercially available and 15 were tested for their HIV protease inhibition. The neuroleptic haloperidol had an IC_{50} vs. HIV-1 protease of 100 μM but was toxic at high concentrations [107, 108]. Further chemical optimization resulted in the haloperidol derivative **compound 34** (K_i HIV-1 protease = 15 μM, K_i HIV-2 protease = 100 μM; Fig. 16.5) [108].

A pharmacophore hypothesis for HIV protease inhibitors was derived at Dupont from the X-ray structures of several inhibitor complexes and the modeled binding mode of vicinal diol inhibitors. A 3D database search yielded a substituted terphenyl compound, which suggested as starting point a six- or seven-membered ring, with a carbonyl group to position a structural water mimic and one or two hydroxy groups to interact with the catalytic aspartates. By extensive structural modification cyclic ureas, for example, **compound 35** (K_i = 0.27 nM, IC_{50} = 36 nM; Fig. 16.5) [109], resulted from which even more active but also poorly soluble inhibitors were derived e.g., [110].

Plasmepsin II. The malarial aspartyl protease plasmepsin II has a significant homology (35%) to cathepsin D. Correspondingly, the very same approach as for the cathepsin D inhibitors (see above) was followed. The best inhibitors have K_i values of 2–10 nM, a molecular weight <650, moderate selectivity vs. cathepsin D, the most closely related human protease, log P values <4.6, and no apparent binding to human serum albumin, for example, **compound 36** (K_i plasmepsin II = 2.0 nM, K_i cathepsin D = 9.8 nM; Fig. 16.5) [111].

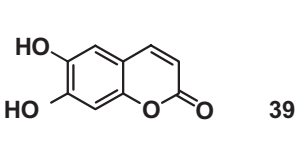
SARS CoV 3C-Like Proteinase. For the screening for SARS (severe acute respiratory syndrome) 3C-like proteinase A inhibitors, a "flexible" 3D model was built by homology modeling and molecular dynamics, starting from the known 3D structure of TGEV (transmissible-gastroenteritis virus) coronavirus 3C-like proteinase. Docking of 630,000 compounds from the ACD, MDDR, and NCI 3D databases was performed with the program DOCK 4.01. The docking hits were further ranked by a pharmacophore model, consensus scoring, and "drug-likeness" filters; 40 compounds were biologically tested. Three of these inhibited SARS 3C-like proteinase with K_i values below 200 μM, for example,. the known calmodulin antagonist **calmidazolium 37** (K_i = 61 μM; Fig. 16.5) [112].

Thrombin. New thrombin inhibitors were designed by a two-step procedure at Hoffmann-La Roche. *p*-Amino-benzamidine was the top-scoring ligand from a docking of 5300 commercially available amines into the recognition pocket of the serine protease thrombin. The link mode of the de novo design program LUDI connected this amine with 540 aldehydes by a reductive amination. Ten of the 100 top-scoring candidates were synthesized and tested; five bind with nanomolar affinities, for example, **compound 38** (K_i thrombin = 95 nM, K_i trypsin = 520 nM; Fig. 16.5) [113, 114].

16.3.3 Other Hydrolases

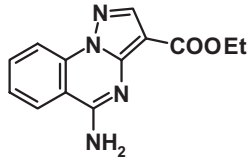
Acetylcholinesterase. The 3D structure of an acetylcholinesterase (AChE) complex with the natural product galanthamine was used to derive a Catalyst pharmacophore model with the program LigandScout. The pharmacophore, containing one donor, one acceptor, and two hydrophobic features, served to screen a 3D multiconformational database of more than 110,000 natural products. Among the observed hits were the coumarin **scopoletin 39** ($IC_{50} \approx 170\,\mu M$; Fig. 16.6) and its glucoside scopolin. In vivo, both compounds increase the extracellular acetylcholine concentration in rat brain to about 170% and 300% (intracerebrovascular application of 2 μmol compound), which is in the same range as the effect observed from galanthamine [115].

Adenylyl Cyclase (Edema Factor and CyaA). The adenylyl cyclases edema factor (EF) and CyaA are toxins of the pathogenic bacteria *Bacillus anthracis* and *B. pertussis*, which cause anthrax and whooping cough, respectively. The 3D structure of EF served to dock 205,226 ACD compounds into the catalytic site with a university version of the program DOCK. From 24 tested compounds two pyrazoloquinazolines could be identified as selective inhibitors of EF and CyaA, for example, **compound 40** (K_i EF ≈ 20 μM, IC_{50} EF = 90 μM; K_i CyaA ≈ 20 μM, IC_{50} CyaA = 80 μM; Fig. 16.6) [116].



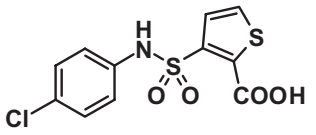
39

AChE inhibitor, $IC_{50} \approx 170\,\mu M$



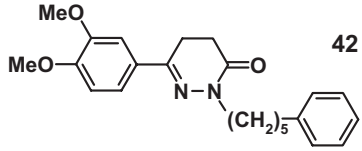
40

Edema factor (EF) adenylyl cyclase inhibitor,
K_i EF ≈ 20 μM, IC_{50} EF = 90 μM;
K_i CyaA ≈ 20 μM, IC_{50} CyaA = 80 μM



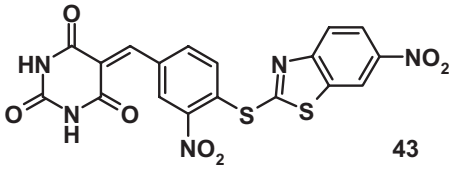
41

AmpC β-Lactamase noncovalent inhibitor,
K_i = 26 μM



42

Phosphodiesterase 4 inhibitor,
IC_{50} = 0.9 nM



43

Protein tyrosine
phosphatase 1B
(PTP1B) inhibitor,
IC_{50} = 4.1 μM

Figure 16.6 Other hydrolase inhibitors from virtual screening.

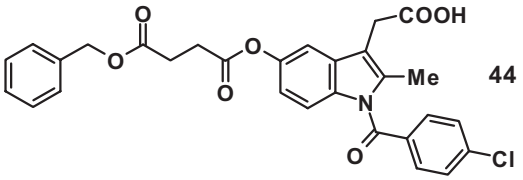
AmpC β-Lactamase. A map of "hot spots" was constructed from the X-ray structure of AmpC β-lactamase and a university version of the program DOCK was used to search for noncovalent inhibitors in 229,810 compounds of the ACD database. Of 56 tested compounds three had K_i values <650 μM, for example, **compound 41** (K_i = 26 μM; Fig. 16.6) [117]. The experimental X-ray structure of its complex with AmpC β-lactamase closely resembles the predicted binding mode.

Phosphodiesterase 4 Didier Rognan and his group used a "Scaffold-linker-functional group" (SCF) approach to design a virtual combinatorial library of analogs of the phosphodiesterase 4 (PDE4) inhibitor zardaverine (IC_{50} = 800 nM). All molecules were constructed from the invariable scaffold of zardaverine (with the exception of minor modifications) and a diverse set of variable linkers and building blocks. As the program FlexX produced the best results in the docking of zardaverine itself, this program was also used for the docking of all analogs. Nine molecules, out of 320 candidates, were synthesized and tested. **Compound 42** (IC_{50} = 0.9 nM; Fig. 16.6) was about 900 times more active than the original lead compound zardaverine [118].

Protein Tyrosine Phosphatase 1B. At Pharmacia, the in-house compound collection of 400,000 compounds was screened against protein tyrosine phosphatase 1B (PTP1B), resulting in 85 inhibitors (0.021%) with a validated IC_{50} < 100 μM; the most active compound had an IC_{50} = 4.2 μM. Shoichet and his group compared the efficacy of this high-throughput screening with docking and scoring [119]. Virtual screening was performed with 235,000 commercially available compounds from three different sources. After selection of only molecules with 17–60 nonhydrogen atoms, 165,581 compounds were docked into the 3D structure of PTB1B, using the program DOCK 3.5. Out of 365 high-scoring molecules, 127 (= 34.8%) inhibited PTP1B with an IC_{50} <100 μM, for example, **compound 43** (IC_{50} = 4.1 μM; Fig. 16.6) [74, 119]. The authors claim that the docking hits were more druglike than the screening hits, with respect to their physicochemical properties.

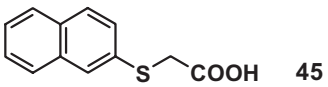
16.3.4 Oxidases/Reductases

Aldose Reductase. The ADAM&EVE docking program was used to screen about 120,000 structures of the ACD 3D database as potential aldose reductase inhibitors. Only one 3D conformation was generated for every molecule, but the ADAM&EVE program performed a systematic conformational search in the docking process, optimizing the conformation by a simplex method. After passing several filters (e.g., MW > 250, at least 1 ring system), total interaction energies were calculated and the resulting hits were visually inspected. An active hit served as a starting point for the dedicated design of analogs, resulting in **compound 44** (IC_{50} = 0.21 μM; Fig. 16.7) as the most potent inhibitor [120].



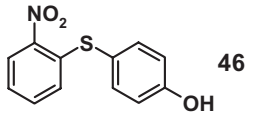
44

Aldose reductase inhibitor, IC_{50} = 0.21 μM



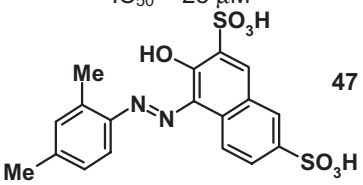
45

Aldose reductase inhibitor, IC_{50} = 28 μM



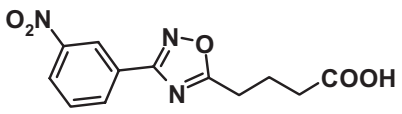
46

Aldose reductase inhibitor, IC_{50} = 4.8 μM



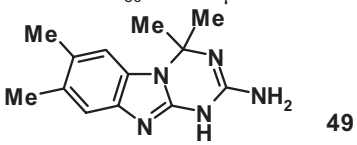
47

Aldose reductase inhibitor, IC_{50} = 0.58 μM



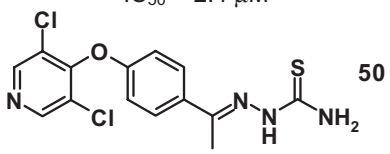
48

Aldose reductase inhibitor, IC_{50} = 2.4 μM



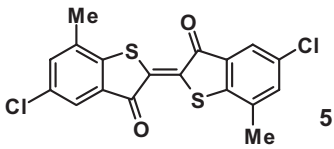
49

Plasmodium falciparum DHFR inhibitor, K_i = 0.54 μM, IC_{50} = 1.4 μM



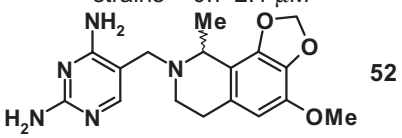
50

Plasmodium falciparum DHFR inhibitor, K_i = 0.9 μM, K_i vs. different resistant strains = 0.7-2.1 μM



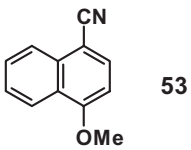
51

Pneumocystis carinii DHFR inhibitor, K_i = 6.9 μM (about 30-fold selectivity vs. human DHFR)



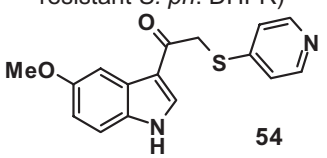
52

Streptococcus pneumoniae DHFR inhibitor, (*R*)-enantiomer, IC_{50} = 9.8 nM (*S. pn*. DHFR), IC_{50} = 2.8 nM (TMP-resistant *S. pn*. DHFR)



53

Inosine 5′-monophosphate dehydrogenase (IMPDH) inhibitor, IC_{50} = 31 μM



54

Inosine 5′-monophosphate dehydrogenase (IMPDH) inhibitor, IC_{50} = 32 μM

Figure 16.7 Oxidase and reductase inhibitors from virtual screening.

Among different 3D structures of aldose reductase, one with an additional open hydrophobic pocket was selected to search for ligands with additional aromatic rings. Molecular docking of 127,000 molecules of the NCI database into the binding pocket was performed with the program DOCK. The 1270 best-scoring compounds were clustered into chemical families, and a further selection was performed according to the interaction of the ligands with two key amino acids, Tyr48 and His110. Of the original 1270 molecules, 25 were selected; the most similar analogs in the ACD were taken for those that were commercially not available. Micromolar and submicromolar selective inhibitors resulted from biological testing (and chemical optimization of a nitro compound), for example, **compounds 45** (IC_{50} aldose reductase = 28 μM, aldehyde reductase inhibition: 6% at 45 μM; Fig. 16.7), **46** (IC_{50} aldose reductase = 4.8 μM, IC_{50} aldehyde reductase = 48 μM; Fig. 16.7). and **47** (IC_{50} aldose reductase = 0.58 μM, aldehyde reductase: 13% inhibition at 27 μM; Fig. 16.7) [121].

Gerhard Klebe and his group used a high-resolution 3D structure of aldose reductase (0.66-Å resolution) for a stepwise virtual screening for more drug-like inhibitors. First, 259,747 ACD compounds were filtered according to
2 certain properties: presence of a carboxylic group or its isoster and compliance with the Lipinski rule of five (but restricted to MW < 350 and a number of rotatable bonds <9). This resulted in 12,545 candidates that were filtered by a pharmacophore search, using the program Unity and a pharmacophore, which was derived from the aldose reductase binding site with the programs SuperStar and the knowledge-based scoring function DrugScore. The 1261 fitting compounds were flexibly docked with the program FlexX. In the scoring procedure, a correction had to be applied to avoid overprediction of the affinity of large, flexible molecules [cf. 56]. The highest-scoring 216 compounds were clustered and visually inspected for the binding conformation, the surface complementarity of the ligand and the protein, and for unfilled space along the protein-ligand interface. A subset of nine carboxylic acids was selected for acquisition and biological testing. The most active hit was **compound 48** (IC_{50} = 2.4 μM; Fig. 16.7) [122].

Dihydrofolate Reductase. A 3D model of the dihydrofolate reductase (DHFR) domain of the bifunctional DHFR-thymidylate synthase of the malaria parasite *Plasmodium falciparum* was derived from the experimental 3D structures of human, chicken, *Escherichia coli*, and *Lactobacillus casei* DHFRs. Compounds with bifunctional basic groups, like amidines and guanidines, were extracted from the ACD, and the program GREEN was used to dock these compounds into the substrate binding site of the DHFR domain, under the constraint of an interaction of their basic group with Asp54. Among 32 candidates from docking and scoring, 21 were purchased and tested. Two compounds showed significant inhibitory activity, for example, **compound 49** (K_i = 0.54 μM, IC_{50} = 1.4 μM; Fig. 16.7) [123].

In malaria chemotherapy, resistant parasites have significantly reduced the efficiency of classic antifolate drugs. In the search for novel inhibitors of

P. falciparum dihydrofolate reductase (PfDHFR), first 3D pharmacophores and other filters were used to reduce the number of potential candidates in a database of 230,000 ACD compounds to 4061 molecules. Docking of these "focused" compounds was performed with the program DOCK 3.5. Twelve compounds were identified that are structurally unrelated to known antifolates; they inhibit not only wild-type PfDHFR but also different resistant mutants at micromolar concentrations. The most potent inhibitor was **compound 50** (K_i = 0.9 μM, K_i vs. the antifolate-resistant strains A16V, S108T, A16V+S108T, C59R+S108N+I164L, and N51I+C59R+S108N+I164L = 0.6–2.1 μM; Fig. 16.7) [124].

An opportunistic infection with the fungus *Pneumocystis carinii* is the principal cause of mortality in HIV-infected patients. Inhibitors of *P. carinii* DHFR with selectivity against human DHFR were identified by docking 53,328 compounds of the FCD (fine chemicals directory, a precursor of the ACD) into an unpublished 3D structure of the ternary complex of *P. carinii* DHFR with folate and NADPH, using the program DOCK. Of 2700 fitting compounds, 1266 were eliminated by energetic considerations. After two steps of chemical diversity selection the number of candidates was reduced to 89 compounds, of which 40 were ordered for biological testing. The most potent inhibitor was **compound 51** (K_i = 6.9 μM; Fig. 16.7), with about 30-fold selectivity vs. human DHFR [125].

At Hoffmann-La Roche 5-*N*,*N*-disubstituted aminomethyl-2,4-diaminopyrimidines were designed and tested as *Streptococcus pneumoniae* DHFR inhibitors. A virtual library was generated by substituting 2,4-diaminopyrimidine with 9448 secondary amines, and two approaches were followed: (1) a diversity-oriented selection and (1) virtual screening by docking and scoring, using the program FlexX and a homology model that was constructed from the 3D structure of the closely related *S. aureus* DHFR. The FlexX scoring function was modified to penalize hydrogen bonds that are formed at the surface of the protein.

Significantly more hits and more active compounds were obtained from the structure-based library design than from diversity-based design (21% vs. 3% hit rate). In general, the compounds showed high activity against trimethoprim (TMP)-sensitive and TMP-resistant *S. pneumoniae* DHFR. Some compounds were highly selective for the bacterial enzyme, as compared to the inhibition of the human enzyme, for example, the (*R*)-enantiomer of **compound 52** (IC_{50} *S. pn.* DHFR = 9.8 nM, IC_{50} TMP-resistant *S. pn.* DHFR = 2.8 nM, IC_{50} human DHFR = 1.2 μM; Fig. 16.7) [126].

Inosine 5′-Monophosphate Dehydrogenase Inhibitors. A series of 21 known inosine 5′-monophosphate dehydrogenase (IMPDH) inhibitors was used to validate a virtual screening protocol. By application of a molecular weight filter (80 < MW < 400), 3425 compounds were extracted from an in-house reagent inventory system. Docking of these compounds into a substrate-IMPDH complex 3D structure was performed with the program FlexX; three

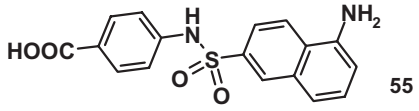
different scoring functions were tested, with and without conserved water molecules in the NAD cofactor binding site. The resulting 74 compounds gave a hit rate of 10% active compounds of diverse chemistry, for example, **compounds 53** (IC_{50} = 31 μM; Fig. 16.7) and **54** (IC_{50} = 32 μM; Fig. 16.7) [127].

16.3.5 Other Enzymes

5-Aminoimidazole-4-Carboxamide Ribonucleotide Transformylase. The NCI diversity library, a set of 1990 compounds with nonredundant pharmacophore profiles, was used for virtual screening of the human 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase active site with the program AutoDock. Biological testing of 16 soluble compounds, out of 44 potential inhibitors, revealed eight micromolar inhibitors with novel scaffolds, for example, **compound 55** (IC_{50} = 4.1 μM; Fig. 16.8). Docking of all compounds with similar scaffolds, from the entire NCI 3D database, yielded another 11 inhibitors, for example, **compound 56** (K_i = 154 nM, IC_{50} = 600 nM; Fig. 16.8) [128].

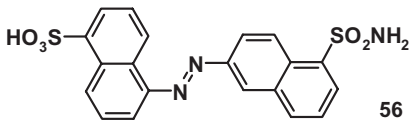
Carbonic Anhydrase II. Carbonic anhydrases are metalloenzymes with a catalytically active Zn^{2+} ion in the catalytic center. Aromatic and other acidic sulfonamides bind as anions that form the warhead group of all carbonic anhydrase inhibitors. The experimental X-ray structure of carbonic anhydrase II was used for virtual screening of potential inhibitors. 3D structures were generated for 98,850 compounds of the Maybridge and LeadQuest compound collections. The binding pocket of carbonic anhydrase was investigated by the computer programs GRID, SuperStar, LUDI, and DrugScore. Hot spots obtained from these programs were converted into a pharmacophore model, and 2D and 3D searches were performed with the program Unity. The resulting 3314 structures were flexibly superimposed on the highly potent inhibitor dorzolamide, using the program FlexS. The best hits were docked as flexible ligands with the program FlexX. Binding affinities to carbonic anhydrase were estimated with the knowledge-based scoring function DrugScore, and the top ranking 13 molecules were biologically tested. Three inhibitors exhibited subnanomolar activity, for example, compounds **57** (IC_{50} = 0.6 nM; Fig. 16.8) and **58** (IC_{50} = 0.8 nM; Fig. 16.8) [129, 130].

A search for even more potent carbonic anhydrase II inhibitors, by a Harvard University group in cooperation with Concurrent Pharmaceuticals (now Vitae Pharmaceuticals), started from the nanomolar inhibitor *p*-H_2NCO-C_6H_4-SO_2NH_2 (K_d = 120 nM). Derivatives of this base fragment, substituted at the carboxamido nitrogen atom, were generated in the binding site of the protein from 100 different small organic groups. The growth algorithm CombiSMoG (combinatorial small molecule generator) randomly selected fragments from this library and attached them to the growing ligand. The affinity of the generated ligands was estimated by the knowledge-based



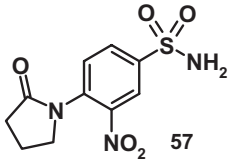
55

5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase inhibitor, IC_{50} = 4.1 μM

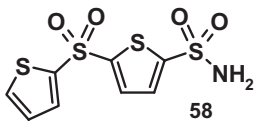


56

5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase inhibitor, K_i = 154 nM, IC_{50} = 600 nM

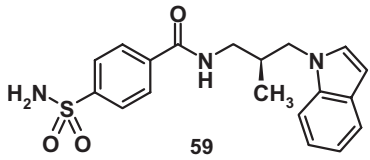


57



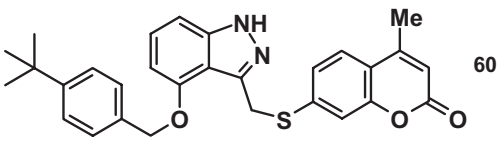
58

Carbonic anhydrase II inhibitors, IC_{50} = 0.6 nM (left) and 0.8 nM (right)



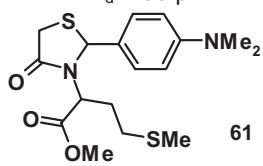
59

Carbonic anhydrase II inhibitor, K_d = 30 pM



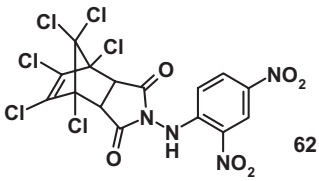
60

DNA gyrase inhibitor MNEC = 0.03 μg/ml



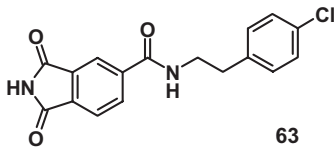
61

dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase (RmlC) inhibitor, 100% inhibition at 20 μM



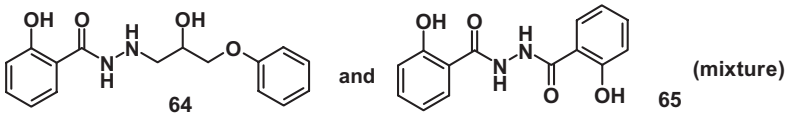
62

Farnesyl transferase inhibitor IC_{50} = 25 μM



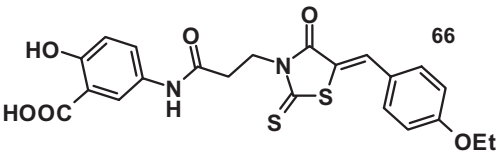
63

Guanine-phosphoribosyl transferase inhibitor, K_i = 23 μM, IC_{50} = 52 μM



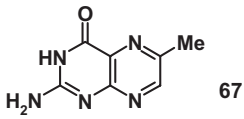
64 65

HIV-1 integrase inhibitors, IC_{50} 3′-processing = 0.6 μg/ml, IC_{50} strand transfer = 0.46 μg/ml



66

HIV-1 integrase inhibitor, IC_{50} 3′-processing = 17 μM, IC_{50} strand transfer = 11 μM



67

tRNA-guanine transglycosylase (TGT) inhibitor, K_i = 0.25 μM

Figure 16.8 Other enzyme inhibitors from virtual screening.

CombiSMoG potential function, which was derived from 1000 protein-ligand complex 3D structures. After inspection of 100,000 candidates, the five best hits were ranked by a force field calculation. The (*R*)-isomer of the indole **compound 59** (K_d = 30 pM; Fig. 16.8) is the highest-scoring compound and has the highest affinity of all synthesized molecules, whereas the (*S*)-isomer has only a K_d = 230 pM [131, 132].

DNA Gyrase. After the failure to find DNA gyrase inhibitors by conventional screening of the Hoffmann-La Roche compound collection, Böhm et al. performed a virtual screening procedure, called "needle screening." First, small "needle-type" molecules were selected from about 350,000 compounds of the ACD and part of the Roche compound inventory and docked into the DNA gyrase active site with the de novo design program LUDI. The resulting hits were then analyzed for their binding site interactions. About 200 compounds were tested for DNA gyrase inhibition, and activities in the range of 5–64 μg/ml were obtained. X-ray structure analysis verified the proposed binding modes of an indazole, an aminotriazine, and a pyrrolopyrimidine lead structure. **Compound 60** [maximal noneffective concentration (MNEC) = 0.03 μg/ml; Fig. 16.8] resulted after 3D structure-guided optimization [133].

dTDP-6-Deoxy-D-Xylo-4-Hexulose3,5-Epimerase (RmlC). dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase (RmlC) has been selected as a new promising target in the fight against tuberculosis. A virtual library of 2,3,5-trisubstituted thiazolidin-4-ones was generated from 24 amino acids, 27 aldehydes, and 2 thioacids with the program CombiLibmaker, and the resulting 3888 structures (containing all possible stereoisomers) were docked into the active site of the enzyme with the program FlexX. After consensus scoring with the CScore module, the top 5% (= 144 compounds) were selected for synthesis and biological tests; 30 of 94 compounds had biological activities >50% at 20 μM, for example, **compound 61** (100% inhibition at 20 μM; Fig. 16.8) [134].

Farnesyl Transferase. A rigid docking of 219,390 ACD compounds into the binding site of farnesyl transferase was performed with the program EUDOC. Of 21 hits, four inhibited the enzyme with IC_{50} values in the range 25–100 μM. The most potent inhibitor, **compound 62** (IC_{50} = 25 μM; Fig. 16.8), inhibited farnesyl transferase also in human lung cancer cells [135a]. A Catalyst 3D pharmacophore search of a Schering-Plough corporate database yielded five compounds with IC_{50} values smaller or equal to 5 μM, representing three different structural classes [135b].

Guanine Phosphoribosyl Transferase. Guanine phosphoribosyl transferase (GPRT) is one of the enzymes of the purine salvage pathway, which is needed by protozoa because they lack the ability to synthesize purine nucleotides.

Two micromolar phthalimide GPRT inhibitors were identified by screening the in-house phthalimide library. On the basis of this result, a virtual library of substituted phthalimides was constructed and docked into the binding sites of six GPRTs from different sources, including *Giardia lamblia*, various trypanosomes, *E. coli*, and human with the program DOCK 4.01. Several micromolar inhibitors resulted, for example, **compound 63** (K_i = 23 μM, IC_{50} = 52 μM; Fig. 16.8) [136].

HIV-1 Integrase. Several 3D pharmacophore models were derived from known HIV-1 integrase inhibitors. These models were validated with a 3D database of 152 compounds with known integrase inhibitory activities. The most probable pharmacophore model was used as query for a 3D search of 206,876 compounds of the NCI 3D database. From 340 hits 29 compounds were selected for biological tests, resulting in 10 novel, structurally diverse HIV-1 integrase inhibitors. Four of these had IC_{50} values <30 μM, for example, a salicylic acid derivative, which later turned out to be a mixture of two salicylic acid hydrazides, **compounds 64** and **65** (IC_{50} 3′-processing ~2.0 μM, IC_{50} strand transfer ~1.5 μM; Fig. 16.8) [137].

A pharmacophore hypothesis for HIV-1 integrase inhibitors was derived from four isosteric β-diketo integrase inhibitors by the HipHop module of Catalyst. A 3D search in a multiconformer Catalyst database of 150,000 ChemBridge compounds yielded 1700 molecules that fitted a four-point pharmacophore. Subsequently, the program GOLD 1.2 was used to dock the structures into the integrase binding site. Afterwards, the 200 top-scoring hits were visually inspected for their ability to chelate a metal ion, for structural novelty, and for compliance with the Lipinski rule of five. Finally, 110 molecules were biologically tested, yielding 48 compounds with IC_{50} values from 7 to 100 μM, for example, **compound 66** (IC_{50} 3′-processing = 17 μM, IC_{50} strand transfer = 11 μM; Fig. 16.8). The most active compounds had a salicylic acid substituent and a 2-thioxo-thiazolidinone (rhodanine) scaffold. On the basis of a 2D substructure search for these moieties, another 22 compounds were selected and tested, resulting in some more micromolar integrase inhibitors [138].

tRNA-Guanine Transglycosylase. In the search for tRNA-guanine transglycosylase (TGT) inhibitors, 800,000 molecules from eight different databases were screened in a stepwise manner, using the programs Selector (to eliminate molecules with more than 7 rotatable bonds and a MW > 450), Unity for 3D pharmacophore search, and FlexX for flexible docking. About 50% of all molecules were eliminated by the Selector procedure. Three different binding site-derived pharmacophore hypotheses were applied to perform 3D pharmacophore searches. This filter reduced the set of compounds to 20% of the original size. In the next step, volume constraints defined the shape of the binding site, producing a hit list of 872 compounds. After flexible docking into two different conformations of the enzyme, some other criteria

were applied, to end up with 9 compounds that were biologically tested. All had micromolar to submicromolar activities, for example, **compound 67** (K_i = 0.25 μM; Fig. 16.8) [139, 140].

16.4 ION CHANNELS

16.4.1 T-Type Selective Ca^{2+} Channel

The T-type selective channel blocker lead structure mibefradil (IC_{50} = 1.2 μM) served as a template for virtual screening of the Hoffmann-La Roche in-house compound collection. Several filters were applied, and the similarity of the candidates to the lead structure was compared by the CATS descriptor [89]. Because only pharmacophoric features and their topological distances describe the molecules, the CATS descriptor enables a "scaffold hopping"; that is, molecules with different scaffolds but comparable biological properties result from this approach. The 12 highest-ranking molecules were biologically tested; nine of them showed T-channel blocking activities in the same range as the lead structure mibefradil. Whereas one highly active compound was the known neuroleptic **clopimozide 68** (IC_{50} < 1 μM; Fig. 16.9) [89], several other active hits, for example, **compounds 69** (IC_{50} = 2.4 μM; Fig. 16.9) [63,141] and **70** (IC_{50} = 0.8 μM; Fig. 16.9) [141], are new chemotypes. Despite the topological pharmacophore similarity, the scaffolds of all compounds are significantly different from mibefradil.

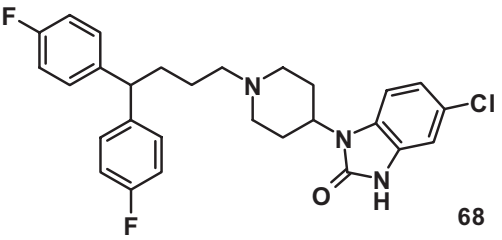
16.4.2 Kv1.5 Potassium Channel

A potent hKv1.5 potassium channel blocker from literature served as template for a TOPAS (topology assigning system) de novo design [142]. The "scaffold hopping" program TOPAS starts from a collection of building blocks that are generated by retrosynthetic fragmentation of the World Drug Index (WDI). By using 11 chemical reactions of the RECAP procedure [143], 24,563 unique building blocks were generated. After assembling new structures from various scaffolds and building blocks, an evolutionary algorithm selects the "fittest" molecules, that is, the ones that are most similar to the original template. Although the "most similar" **compound 71a** (R = OMe, IC_{50} = 7.34 μM; Fig. 16.9) is much less active than the template (IC_{50} = 0.11 μM), a close analog, **compound 71b** (R = H, IC_{50} = 0.47 μM) [142], has about the same order of biological activity (wrong substitution pattern in Scheme 1 of Ref. 142; see Scheme 2).

The same template as for **compound 71** was used by Peukert et al. for a 2D similarity search in the Aventis in-house compound collection [144]. 75 Compounds with a similarity index >0.80 were biologically tested. The moderately active 1-carboxy,8-sulfonamido-naphthalene (IC_{50} = 9.5 μM), with insufficient chemical stability, was the starting point for the design of substi-

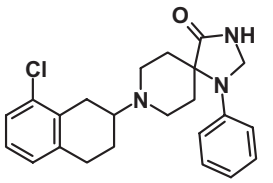
tuted biphenyls, which after further optimization produced **compound 72** (IC_{50} = 0.16 μM; Fig. 16.9) [144].

An improved 3D pharmacophore, considering all results obtained so far, and a new 3D search in the Aventis compound collection with the program



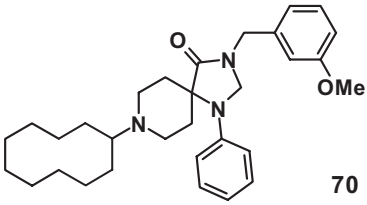
68

Clopimozid, Ca^{2+} antagonist,
IC_{50} < 1 μM



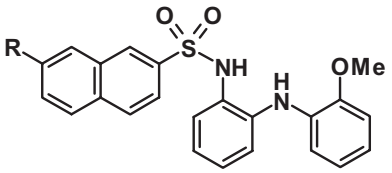
69

Ca^{2+} antagonist (T channel blocker),
IC_{50} = 2.4 μM



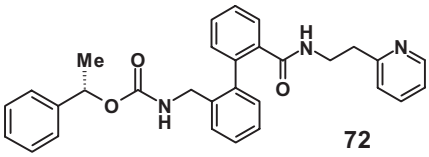
70

Ca^{2+} antagonist (T channel blocker),
IC_{50} = 0.8 μM



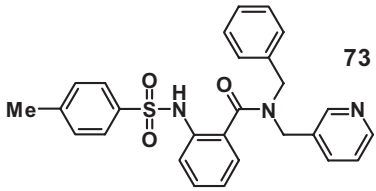
71

hKv1.5 Potassium channel inhibitor,
a R = OMe, IC_{50} = 7.34 μM
b R = H, IC_{50} = 0.47 μM



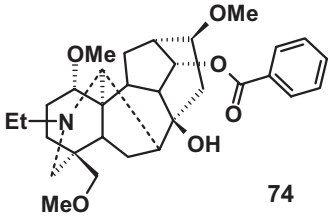
72

hKv1.5 Potassium channel inhibitor,
IC_{50} = 0.16 μM



73

hKv1.5 Potassium channel inhibitor,
IC_{50} = 0.7 μM



74

Shaker K^+ channel blocker,
IC_{50} = 3.8 μM

Figure 16.9 Ion channel blockers from virtual screening.

Unity resulted in 4234 hits. After application of several filters and clustering of the remaining 1975 molecules, compounds from 18 of the 27 clusters were screened in *Xenopus* oocytes. One compound with an IC_{50} of 5.6 μM belonged to a new class of Kv1.5 blockers and exhibited a favorable pharmacokinetic profile. After further optimization, **compound 73** (IC_{50} = 0.7 μM; Fig. 16.9) resulted, with good oral bioavailability in rats [145].

In a further investigation, the most interesting hits of the prior work were used together with other reference Kv1.5 channel blockers to perform virtual screening for new chemotypes. A protein-based pharmacophore for a 3D search was derived from a homology model of the potassium channel. The five most active hits from the corporate database had IC_{50} values between 0.9 and 7.9 μM (structures not given) [146]. Whereas chemical similarity between these compounds, as measured by pairwise Tanimoto similarity based on Unity fingerprints, was low, feature tree similarity values, which measure pharmacophore similarity across chemically diverse classes, are high.

16.4.3 Shaker K^+ Channel

Although a large number of drugs have their origin in natural products [147], databases of natural products are rarely used for virtual screening. A 3D homology model of the eukaryotic Shaker K^+ channel was built from the known 3D structure of the KcsA potassium channel. The refined 3D model was used to dock more than 50,000 compounds of the China Natural Product Database (Shanghai Institute of Materia Medica, Chinese Academy of Sciences, and Neotrident Technology Ltd.) with the program DOCK 4.0 into the extracellular tetraethylammonium (TEA) binding site. Of 14 hits, only four diterpenoid alkaloids from *Aconitum leucostonum* were accessible. Extracellular application of the four compounds inhibited the delayed rectifier current (I_K) at micromolar concentration, for example, **14-benzoyl-talatisamine 74** (IC_{50} = 3.8 μM; Fig. 16.9) [148].

16.5 OTHER TARGETS; PROTEIN-PROTEIN AND PROTEIN-RNA INTERACTIONS

16.5.1 Bcl-2 Protein-Protein Interaction

Bcl-2 is one of the many factors that control apoptosis, and overexpression of Bcl-2 has been observed in many different cancers. A homology model of Bcl-2 was derived from the NMR 3D structure of the Bcl-XL complex with a Bak BH3 peptide. This model served to search the NCI 3D database of 206,876 organic compounds for potential Bcl-2 inhibitors, which bind to the Bak BH3 binding site of Bcl-2. Full conformational flexibility of the ligands was taken into account in the program DOCK. Thirty-five potential inhibitors were tested, and seven of them had IC_{50} values from 1.6 to 14.0 μM. One of

the hits, **compound 75** (Fig. 16.10), had the highest antiproliferative activity (IC_{50} = 10.4 μM) in the human myeloid leukemia cell line HL-60. Whereas **compound 75** induced apoptosis in cancer cells with high Bcl-2 expression, it had only little effect on cancer cells with low or undetectable levels of Bcl-2 [149].

16.5.2 Cyclophilin A

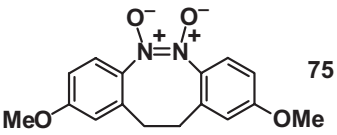
The immunophilins cyclophilin A [CyPA; binds cyclosporin A (CsA)] and FK506-binding protein (FKBP12; binds FK506 and rapamycin) are peptidylprolyl isomerases (PPIases, rotamases). However, it is the interaction of the drug-immunophilin complexes with the calcium/calmodulin-dependent protein phosphatase calcineurin (CsA/CyPA and FK506/FKBP12 complexes) and the serine/threonine kinase FRAP (rapamycin/FKBP12 complex) that is responsible for their immunosuppressive effects. A pharmacophore model for potential cyclophilin ligands was derived from cyclosporin and dipeptides that bind to CyPA. Compounds of the ACD, WDI, and Chapman-Hall Dictionary of Organic Compounds were filtered to remove molecules with MW >700 and reactive compounds. Then 3D structures were generated with the program Concord, and a Unity 3D search was performed, using the cyclophilin ligand 3D pharmacophore. In the resulting hits a lead structure with IC_{50} = 6 μM was identified. It served as the starting point for further chemical optimization, from which several submicromolar CyPA inhibitors resulted, for example, **compound 76** (IC_{50} = 930 nM; Fig. 16.10) [150].

16.5.3 FK 506-Binding Protein (FKBP12)

FK506-binding proteins (FKBP) belong to the family of immunophilins. Together with their ligand FK506 and the serine/threonine phosphatase calcineurin, they form ternary complexes that block signal transduction in T cells. A 3D version of the ACD and the 3D structures of the Cambridge Crystallographic Database (CCD) were docked into the binding pocket of FKBP with the program Sandock. Several hits bound with micromolar affinities, for example, the steroid **compound 77** (K_d = 7 μM) and the spiro **compound 78** (K_d = 11 μM); the dipeptide **Z-L-Pro-L-Pro 79** had even submicromolar affinity (K_d = 0.8 μM) (Fig. 16.10) [151].

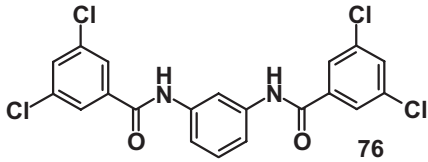
16.5.4 HIV-1 RNA Transactivation Response Element

The binding of the HIV-1 transactivating regulatory protein (tat) to the RNA transactivation response element (TAR) is an essential step for HIV-1 replication. The ACD was screened for inhibitors of the tat-TAR protein-RNA interaction. A four-step procedure was used: Rigid docking was followed by three steps of flexible docking, using a stochastic torsional angle modification of the ligands. The procedure was validated by docking ligands of five RNA



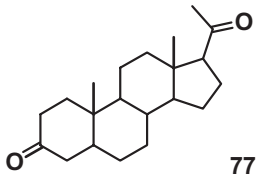
75

Bcl-2 protein-protein interaction inhibition, IC_{50} = 10.4 μM



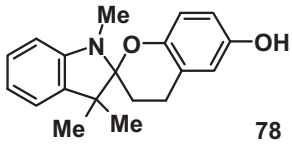
76

Cyclophilin A (CyPA) inhibitor, IC_{50} = 930 nM



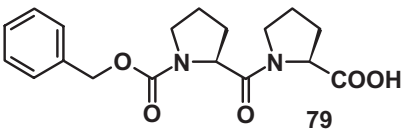
77

FK-506 binding protein (FKBP) ligand, K_d = 7 μM



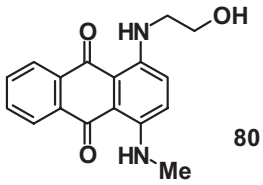
78

FK-506 binding protein (FKBP) ligand, K_d = 11 μM



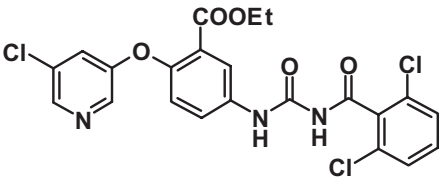
79

FK-506 binding protein (FKBP) ligand, K_d = 0.8 μM



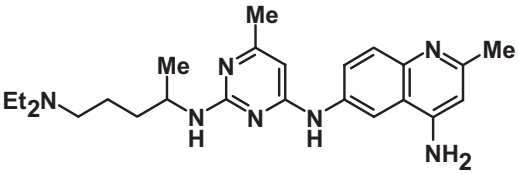
80

HIV-1 RNA Transactivation response element (TAR) inhibitor, $CD_{50} \approx 1$ μM



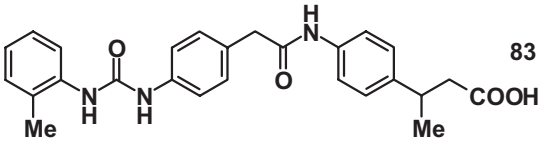
81

Mesangial cell (MC) proliferation inhibitor, 90% inhibition at 100 nM



82

Inhibitor of Rac1 protein-protein interaction; $IC_{50} \approx 50$ μM



83

VLA-4 (= α4β1 antigen) antagonist, IC_{50} = 1.3 nM

Figure 16.10 Inhibitors of protein-protein and protein-RNA interactions from virtual screening.

complexes of known structure and scoring them by an empirical function, which was derived from ligand-RNA complexes with known structure and affinity, accounting also for solvation and changes in conformational entropy. Screening of about 153,000 ACD compounds yielded high-ranking known TAR ligands, as well as new structures, for example, **compound 80** (CD_{50} ≈ 1 μM; CD_{50} = competitive dose, concentration of compound required to reduce the binding of the tat protein to TAR to 50%; Fig. 16.10) [152].

16.5.5 Mesangial Cell Proliferation

Mesangial cell (MC) proliferation inhibitors were searched, using the HipHop module of the Catalyst software. A 3D pharmacophore model, consisting of two hydrophobic regions, two hydrophobic aromatic regions, and three hydrogen bond acceptors, was generated from a training set of heterocyclic phosphonic acid diethyl esters, using the Catalyst HipHop option. This model served as a 3D query to search 47,045 compounds of the Maybridge 3D database. Among 41 structurally novel inhibitors with >50% inhibitory activity at 100 nM, the most potent hit was **compound 81** (90 % MC proliferation inhibition at 100 nM; Fig. 16.10) [153].

16.5.6 Rac1 Protein-Protein Interaction

Rac GTPase is involved in one of several signaling pathways mediated by Rho family GTPases. The 3D structure of a Rac1-Tiam1 complex was used to specify the binding pocket for inhibitors, and a flexible 3D search was performed in 140,000 compounds of the NCI database with the program Unity. The hits of this search were flexibly docked with the program FlexX and ranked by the consensus scoring function CScore. By visual inspection, 58 of the 100 highest-scoring hits were eliminated because they did not show an interaction of the ligand with Trp56. Considering solubility and availability of the remaining compounds, finally 15 compounds were tested for their ability to inhibit the Rac1-binding interaction with its guanine nucleotide exchange factor (GEF) TrioN. **Compound 82** (IC_{50} ≈ 50 μM; Fig. 16.10) was the only active and selective compound, significantly inhibiting TrioN binding to Rac1 but not interfering with Cdc42 binding to Intersectin. Also in cells it effectively inhibited Rac1 binding and activation, and in human prostate cancer PC-3 cells it inhibited proliferation [154].

16.5.7 VLA-4 ($\alpha 4\beta 1$ Antigen)

A 3D model of the fibrinogen-derived (very late antigen-4, VLA-4) inhibitor 4-[N′-(2-methylphenyl)ureido]phenylacetyl-Leu-Asp-Val was derived from the X-ray structure of the related integrin-binding region of the vascular cell adhesion molecule-1 (VCAM-1). A 3D pharmacophore was generated with the program Catalyst, and a 3D search was performed in 8624 molecules from

the ACD, containing either a free amino or a nitro group and a carboxyl group, in order to replace the tripeptide part of the inhibitor. All 12 selected molecules that passed additional filters inhibited the association of the α4β1 antigen with VCAM-1. The most potent analog, **compound 83,** had an IC_{50} = 1.3 nM (Fig. 16.10) [155].

16.6 SUMMARY AND CONCLUSIONS

Virtual screening comprises several computational techniques that have already shown their efficiency in delivering interesting lead structures. In the most effective application, cascades of different steps serve to reduce very rapidly the number of potential candidates from hundreds of thousands or even millions of structures to a manageable size, for example, by first applying simple filters (molecular weight, polar surface area, number of rotatable bonds, Lipinski rule of five, lead-likeness rules, drug-likeness neural nets), followed by pharmacophore generation and pharmacophore searches. The number of potential candidates can be reduced by a filter that checks the presence of all necessary pharmacophoric features. The generation of a pharmacophore hypothesis can be ligand based or may be derived from the protein 3D structure (if available) by a hot spot analysis (programs GRID [11, 156], LUDI [35], DrugScore [49, 157, 158]). Ligand-based pharmacophore generation is most often performed with the HipHop and HypoGen options of the program Catalyst [159–163]. LigandScout is a new program for the automated generation of pharmacophore hypotheses from 3D structures of protein-ligand complexes [164]. Finally, a 2D (topological) or 3D pharmacophore search is performed. The CATS descriptor [89] and the feature trees [165, 166] are extremely fast and effective search tools for pharmacophore similarity, very often producing active hits with new scaffolds. For 3D searches the programs Catalyst [e.g., 78, 81, 82, 88, 91, 104, 115, 138, 153, 155] and Unity [e.g., 78, 122, 139, 140, 145, 146, 150, 154, 167] are most often used.

If a 3D structure of the target is available from protein crystallography or NMR studies, or can be modeled by homology, the last step, using flexible docking and scoring, is more time demanding. For docking, the programs most commonly used in the success stories described in this review are DOCK [34, 168], in several different versions [96, 102, 103, 107, 112, 116, 117, 119, 121, 124, 125, 136, 148, 149], and FlexX and FlexX-Pharm [86, 87, 95, 97, 118, 122, 126, 127, 129, 130, 134, 139, 140, 154, 169–171]. Stepwise virtual screening protocols have been applied in several examples described in this review [e.g., 78, 86, 87, 95, 97, 106, 122, 124, 129, 130, 139, 140]. A surprisingly large number of successful docking studies used a homology model of the respective protein [78, 79, 86, 87, 92–94, 100, 103, 106, 112, 123, 148, 149].

As discussed in the introduction, scoring functions still pose problems (see also Chapter 14). Some of these problems arise from insufficient consideration of details of favorable and unfavorable protein-ligand interactions,

whereas others are more systematic in their nature, for example, the overprediction of the affinity of large molecules [56, 102, 122] and the overprediction of the affinity contribution of hydrogen bonds at the solvent-accessible surface of the protein [126]. Thus a visual inspection of the docking results [e.g., 93–95, 97, 100, 103, 106, 120, 122, 138, 154] is of utmost importance, to check for unfavorable geometry of the docked ligand, geometric complementarity, for example, space filling of hydrophobic pockets, key interactions with the protein (which is the key option of the program Flex-Pharm), and unfavorable electrostatic interactions, for example, oxygen-oxygen repulsion. Of course, synthetic accessibility or commercial availability and certain physicochemical properties, such as solubility, are also critically important for the selection of candidates for biological screening.

Although some virtual screening hits described in this review, do not look very druglike, for example, **compounds 47**, **50**, **51** (Fig. 16.7), **56**, **62** (Fig. 16.8), and **75** (Fig. 16.10), several other compounds have already been optimized to interesting candidates for further development. It must be kept in mind that ligand-based and 3D structure-based approaches enable only ligand design, not drug design. In the future, computer programs for virtual screening not only should aim at the further improvement of the scoring functions but should consider also synthetic accessibility and allow the construction of ligands with chemically reasonable fragment-based approaches.

REFERENCES

1. Goodford PJ. Drug design by the method of receptor fit. *J Med Chem* 1984;27:557–64.
2. Beddell CR, editor. *The design of drugs to macromolecular targets*. Chichester: John Wiley & Sons, 1992.
3. Beddell CR, Goodford PJ, Norrington FE, Wilkinson S, Wootton R. Compounds designed to fit a site of known structure in human haemoglobin. *Br J Pharmacol* 1976;57:201–9.
4. Kuyper LF, Roth B, Baccanari DP, Farone R, Beddell CR, Champness JN, Stammers DK, Dann JG, Norrington FEA, Baker D, Goodford PJ. Receptor-based design of dihydrofolate reductase inhibitors: comparison of crystallographically determined enzyme binding with enzyme affinity in a series of carboxy-substituted trimethoprim analogues. *J Med Chem* 1982;25:1120–2.
5. Cushman DW, Cheung HS, Sabo EF, Ondetti MA. Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. *Biochemistry* 1977;16:5484–91.
6. Redshaw S. Angiotensin-converting enzyme (ACE) inhibitors and the design of captopril. In: Ganellin CR, Roberts SM, editors, *Medicinal chemistry. The role of organic chemistry in drug research*, 2nd edition. London: Academic Press, 1993. p. 163–85.
7. Baldwin JJ, Ponticello GS, Anderson PS, Christy ME, Murcko MA, Randall WC, Schwam H, Sugrue MF, Springer JP, Gautheron P, Grove J, Mallorga P,

Viader MP, McKeever BM, Navia MA. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J Med Chem* 1989;32:2510–13.

8. Vacca JP. Clinically effective HIV-1 protease inhibitors. *Drug Discov Today* 1997;2:261–72.
9. Wlodaver A, Vondrasek J. Inhibitors of HIV-1 protease—a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct* 1998; 27:249–84.
10. von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, Phan TV, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* 1993;363: 418–23.
11. Goodford PJ. A computational procedure for determining energetically favourable binding sites on biologically important molecules. *J Med Chem* 1985; 28:849–57.
12. Hillisch A, Pineda LF, Hilgenfeld R. Utility of homology models in the drug discovery process. *Drug Discov Today* 2004;9:659–69.
13. Hillisch A, Peters O, Kosemund D, Müller G, Walter A, Schneider B, Reddersen G, Elger W, Fritzemeier KH. Dissecting physiological roles of estrogen receptor α and β with potent selective ligands from structure-based design. *Mol Endocrinol* 2004;18:1599–609.
14. Hillisch A, Peters O, Kosemund D, Müller G, Walter A, Elger W, Fritzemeier KH. Protein structure-based design, synthesis strategy and in vitro pharmacological characterization of estrogen receptor α and β selective compounds. *Ernst Schering Res Found Workshop* 2004;46:47–62.
15. Veerapandian P, editor. *Structure-based drug design*. New York: Marcel Dekker, 1997.
16. Gubernator K, Böhm HJ, editors. *Structure-based ligand design* (Vol. 6 of: Mannhold R, Kubinyi H, Timmerman H, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 1998.
17. Babine RE, Abdel-Meguid SS. *Protein crystallography in drug discovery* (Vol. 20 of: Mannhold R, Kubinyi H, Folkers G, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 2004.
18. Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* 1996;16:3–50.
19. Böhm HJ, Klebe G. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs? *Angew Chem Int Ed Engl* 1996;35:2589–614.
20. Babine RE, Bender SL. Molecular recognition of protein-ligand complexes: applications to drug design. *Chem Rev* 1997;97:1359–472.
21. Kubinyi H. Structure-based design of enzyme inhibitors and receptor ligands. *Curr Opin Drug Discov Dev* 1998;1:4–15.
22. Kubinyi H. Combinatorial and computational approaches in structure-based drug design. *Curr Opin Drug Discov Dev* 1998;1:16–27.

23. Davis AM, Teague SJ, Kleywegt GJ. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew Chem Int Ed Engl* 2003;42:2718–2736; *Angew Chem* 2003;115:2822–41.
24. Congreve M, Murray CW, Blundell TL. Structural biology and drug discovery. *Drug Discov Today* 2005;13:895–907.
25. Lahana R. How many leads from HTS? *Drug Discov Today* 1999;4:447–8.
26. Ramesha CS. Comment: How many leads from HTS? *Drug Discov Today* 2000;5:43–4.
27. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 1997;46:3–26.
28. Ajay A, Walters WP, Murcko MA. Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? *J Med Chem* 1998;41:3314–24.
29. Sadowski J, Kubinyi H. A scoring scheme for discriminating between drugs and non-drugs. *J Med Chem* 1998;41:3325–9 (1998).
30. Güner OF, editor. *Pharmacophore perception, development and use in drug design*. La Jolla: International University Line, 2000.
31. Mason JS, Good AC, Martin EJ. 3-D pharmacophores in drug discovery. *Curr Pharm Des* 2001;7:567–97.
32. van Drie J. Pharmacophore discovery: a critical review. In: Tollenaere J, de Winter H, Langenaeker W, Bultinck P, editors, *Computational medicinal chemistry and drug discovery*. New York: Marcel Dekker, 2004. p. 437–60.
33. Langer T, Hoffmann R. *Pharmacophores and pharmacophore searches* (Vol. ■■ of Mannhold R, Kubinyi H, Folkers G, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 2006.
34. Meng EC, Shoichet B, Kuntz ID. Automated docking with grid-based energy evaluation. *J Comput Chem* 1992;13:505–24.
35. Böhm HJ. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aided Mol Design* 1992;6:61–78.
36. Lengauer T, Rarey M. Computational methods for biomolecular docking. *Curr Opin Struct Biol* 1996;6:402–6.
37. Dixon JS. Evaluation of the CASP2 docking section. *Proteins Struct Funct Genet* 1997;Suppl 1:198–204.
38. Abagyan R. Totrov M. High-throughput docking for lead generation. *Curr Opin Chem Biol* 2001;5:375–82.
39. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins Struct Funct Genet* 2002;47:409–43.
40. Böhm HJ, Schneider G, editors. *Protein-ligand interactions. From molecular recognition to drug design* (Vol. 19 of Mannhold R, Kubinyi H, Folkers G, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 2003.
41. Brooijmans N, Kuntz ID. Molecular recognition and docking algorithms. *Annu Rev Biophys Biomol Struct* 2003;32:335–73.

42. Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M. Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. *J Med Chem* 2004;47:45–55.
43. Alvarez JC. High-throughput docking as a source of novel drug leads. *Curr Opin Chem Biol* 2004;8:365–70.
44. Muegge I, Enyedy I. Docking and scoring. In: Tollenaere J, de Winter H, Langenaeker W, Bultinck P, editors, *Computational medicinal chemistry and drug discovery*. New York: Marcel Dekker. 2004. p. 405–36.
45. Krovat EM, Steindl T, Langer T. Recent advances in docking and scoring. *Curr Comput Aided Drug Des* 2005;1:93–102.
46. Böhm HJ. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J Comput Aided Mol Des* 1994;8:243–56.
47. Ajay A, Murcko MA. Computational methods to predict binding free energy in ligand-receptor complexes. *J Med Chem* 1995;38:4953–67.
48. Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *J Med Chem* 1999;42:791–804.
49. Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. *J Mol Biol* 2000;295:337–56.
50. Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J Med Chem* 2000;43:4759–67.
51. Stahl M, Rarey M. Detailed analysis of scoring functions for virtual screening. *J Med Chem* 2001;44:1035–42.
52. Schulz-Gasch T, Stahl M. Binding site characteristics in structure-based virtual screening: evaluation of current docking tools. *J Mol Model* 2003;9:47–57.
53. Wang R, Lu Y, Wang S. Comparative evaluation of 11 scoring functions for molecular docking. *J Med Chem* 2003;46:2287–303.
54. Charifson PS, Corkery JJ, Murcko MA, Walters,WP. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J Med Chem* 1999;42:5100–9.
55. Wang RX, Wang SM. How does consensus scoring work for virtual library screening? An idealized computer experiment. *J Chem Inf Comput Sci* 2001;41:1422–6.
56. Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. *Proc Natl Acad Sci USA* 1999;96:9997–10002.
57. Böhm HJ, Schneider G, editors. *Virtual screening for bioactive molecules* (Vol. 10 of: Mannhold R, Kubinyi H, Timmerman H, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 2000.
58. Klebe G, editor. *Virtual screening: an alternative or complement to high throughput screening*. Dordrecht: Kluwer Academic Publishers, 2000; also published in *Persp Drug Discov Design* 2000;20:1–287.
59. Alvarez J, Shoichet B, editors. *Virtual screening in drug discovery*. Boca Raton: CRC Press, Taylor & Francis Group, 2005.
60. Walters WP, Stahl MT, Murcko MA. Virtual screening—an overview. *Drug Discov Today* 1998;3:160–78.

61. Good A. Structure-based virtual screening protocols. *Curr Opin Drug Discov Dev* 2001;4:301–7.
62. Langer T, Hoffmann RD. Virtual screening: an effective tool for lead structure discovery? *Curr Pharm Des* 2001;7:509–27.
63. Schneider G, Böhm HJ. Virtual screening and fast automated docking methods. *Drug Discov Today* 2002;7:64–70.
64. Waszkowycz B. Structure-based approaches to drug design and virtual screening. *Curr Opin Drug Discov Dev* 2002;5:407–13.
65. Bajorath, J. Integration of virtual and high throughput screening. *Nat Rev Drug Discov* 2002;1:882–94.
66. Lyne PD. Structure-based virtual screening: an overview. *Drug Discov Today* 2002;7:1047–55.
67. Xu H, Agrafiotis DK: Retrospect and prospect of virtual screening in drug discovery. *Curr Top Med Chem* 2002;2:1305–20.
68. Bissantz C, Bernard P, Hibert M, Rognan D. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins Struct Funct Genet* 2003;50:5–25.
69. Bleicher KH, Böhm HJ, Müller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. *Nature Rev Drug Discov* 2003;2:369–78.
70. Shen J, Xu X, Cheng F, Liu H, Luo X, Shen J, Chen K, Zhao W, Shen X, Jiang H. Virtual screening on natural products for discovering active compounds and target information. *Curr Med Chem* 2003;10:1241–53.
71. Lengauer T, Lemmen C, Rarey M, Zimmermann M. Novel technologies for virtual screening. *Drug Discov Today* 2004;9:27–34.
72. Klebe G. Lead identification in post-genomics: computers as a complementary alternative. *Drug Discov Today Technol* 2004;1:225–30.
73. Oprea TI, Matter H. Integrating virtual screening in lead discovery. *Curr Opin Chem Biol* 2004;8:349–358.
74. Shoichet BK. Virtual screening of chemical libraries. *Nature* 2004;432: 862–5.
75. Langer T, Wolber G. Virtual combinatorial chemistry and *in silico* screening: efficient tools for lead structure discovery? *Pure Appl Chem* 2004;76:991–6.
76. Rarey M, Lemmen C, Matter H. Algorithmic engines in virtual screening. In: Oprea TI, editor, *Chemoinformatics in drug discovery* (Vol. 23 of Mannhold R, Kubinyi H, Folkers G, editors, ?????). Weinheim: Wiley-VCH, 2005. p. 59–115.
77. Anderson AC, Wright DL. The design and docking of virtual compound libraries to structures of drug targets. *Curr Comp Aided Drug Des* 2005;1:103–27.
78. Evers A, Klabunde T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J Med Chem* 2005;48:1088–97.
79. Varady J, Wu X, Fang X, Min J, Hu Z, Levant B, Wang S. Molecular modeling of the three-dimensional structure of dopamine 3 (D3) subtype receptor: discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J Med Chem* 2003;46: 4377–92.

80. Astles PC, Brown TJ, Handscombe CM, Harper MF, Harris NV, Lewis RA, Lockey PM, McCarthy C, McLay IM, Porter B, Roach AG, Smith C, Walsh RJA. Selective endothelin A receptor antagonists. 1. Discovery and structure-activity of 2,4-disubstituted benzoic acid derivatives. *Eur J Med Chem* 1997;32:409–23.

81. Astles PC, Brown TJ, Harper MF, Harris NV, McCarthy C, Porter B, Smith C, Walsh RJA. Selective endothelin A receptor antagonists. 2. Discovery and structure-activity of 5-ketopentanoic acid derivatives. *Eur J Med Chem* 1997; 32:515–22.

82. Astles PC, Brealey C, Brown TJ, Facchini V, Handscombe C, Harris NV, McCarthy C, McLay IM, Porter B, Roach AG, Sargent C, Smith C, Walsh RJA. Selective endothelin A receptor antagonists. 3. Discovery and structure-activity relationships of a series of 4-phenoxybutanoic acid derivatives. *J Med Chem* 1998;41:2732–44.

83. Funk OF, Kettmann V, Drimal J, Langer T. Chemical function based pharmacophore generation of endothelin-A selective receptor antagonists. *J Med Chem* 2004;47:2750–60.

84. Lavrador K, Murphy B, Saunders J, Struthers S, Wang X, Williams J. A screening library for peptide activated G-protein coupled receptors. 1. The test set. *J Med Chem* 2004;47:6864–74.

85. Marriott DP, Dougall IG, Meghani P, Liu YJ, Flower DR. Lead generation using pharmacophore mapping and three-dimensional database searching: application to muscarinic M3 receptor antagonists. *J Med Chem* 1999;42:3210–16.

86. Evers A, Klebe G. Ligand-supported homology modelling of G-protein coupled receptor sites: models sufficient for successful virtual screening. *Angew Chem Int Ed* 2004;43:248–251; *Angew Chem* 2004;116:250–3.

87. Evers A, Klebe G. Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J Med Chem* 2004;47:5381–92.

88. Guba W, Neidhart W, Nettekoven M. Novel and potent NPY5 receptor antagonists derived from virtual screening and iterative parallel chemistry design. *Bioorg Med Chem* Lett 2005;15:1599–603.

89. Schneider G, Neidhart W, Giller T, Schmidt G. ''Scaffold-hopping'' by topological pharmacophore search: a contribution to virtual screening. *Angew Chem Int Ed Engl* 1999;38:2894–2896; *Angew Chem* 1999;111:3068–70.

90. Schneider G, Nettekoven M. Ligand-based combinatorial design of selective purinergic receptor (A_{2A}) antagonists using self-organizing maps. *J Comb Chem* 2003;5:233–7.

91. Flohr S, Kurz M, Kostenis E, Brkovich A, Fournier A, Klabunde T. Identification of nonpeptidic urotensin II receptor antagonists by virtual screening based on a pharmacophore model derived from structure-activity relationships and nuclear magnetic resonance studies on urotensin II. *J Med Chem* 2002;45:1799–805.

92. Schapira M, Raaka BM, Samuels HH, Abagyan R. Rational discovery of novel nuclear hormone receptor antagonists. *Proc Natl Acad Sci USA* 2000; 97:1008–13.

93. Schapira M, Raaka BM, Samuels HH, Abagyan R. *In silico* discovery of novel retinoic acid receptor agonist structures. *BMC Struct Biol* 2001;1:1 (7 pages, www.biomedcentral.com/1472-6807/1/1).
94. Schapira M, Raaka BM, Das S, Fan L, Totrov M, Zhou ZU, Wilson SR, Abagyan R, Samuels HH. Discovery of diverse thyroid hormone receptor antagonists by high-throughput docking. *Proc Natl Acad Sci USA* 2003;100:7354–9.
95. Forino M, Jung D, Easton JB, Houghton PJ, Pellecchia M. Virtual docking approaches to protein kinase B inhibition. *J Med Chem* 2005;48:2278–81.
96. Peng H, Huang N, Qi J, Xie P, Xu C, Wang J, Yang C. Identification of novel inhibitors of BCR-ABL tyrosine kinase via virtual screening. *Bioorg Med Chem Lett* 2003;13:3693–9.
97. Lyne PD, Kenny PW, Cosgrove DA, Deng C, Zabludoff S, Wendoloski JJ, Ashwell S. Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening. *J Med Chem* 2004;47:1962–8.
98. Wu SY, McNae I, Kontopidis G, McClue SJ, McInnes C, Stewart KJ, Wang SD, Zheleva DI, Marriage H, Lane DP, Taylor P, Fischer PM, Walkinshaw MD. Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop. *Structure* 2003;11:399–410.
99. Wang SD, Meades C, Wood G, Osnowski A, Anderson S, Yuill R, Thomas M, Mezna M, Jackson W, Midgley C, Griffiths G, Fleming I, Green S, McNae I, Wu SY, McInnes C, Zheleva D, Walkinshaw MD, Fischer PM. 2-Anilino-4-(thiazol-5-yl)pyrimidine CDK inhibitors: synthesis, SAR analysis, X-ray crystallography, and biological activity. *J Med Chem* 2004;47:1662–75.
100. Honma T, Hayashi K, Aoyama T, Hashimoto N, Machida T, Fukasawa K, Iwama T, Ikeura C, Ikuta M, Suzuki-Takahashi I, Iwasawa Y, Hayama T, Nishimura S, Morishima H. Structure-based generation of a new class of potent Cdk4 inhibitors: new de novo design strategy and library design. *J Med Chem* 2001; 44:4615–27.
101. Nærum L, Nørskov-Lauritsen L, Olesen PH. Scaffold hopping and optimization towards libraries of glycogen synthase kinase-3 inhibitors. *Bioorg Med Chem Lett* 2002;12:1525–8.
102. Huang N, Nagarsekar A, Xia GJ, Hayashi J, MacKerell AD Jr. Identification of non-phosphate-containing small molecular weight inhibitors of the tyrosine kinase p56 Lck SH2 domain via *in silico* screening against the pY+3 binding site. *J Med Chem* 2004;47:3502–11.
103. Vangrevelinghe E, Zimmermann K, Schoepfer J, Portmann R, Fabbro D, Furet P. Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking. *J Med Chem* 2003;46:2656–62.
104. Singh J, Chuaqui CE, Boriack-Sjodin PA, Wen-Cherng Lee, Pontz T, Corbley MJ, Cheung HK, Arduini RM, Mead JN, Newman MN, Papadatos JL, Bowes S, Josiah S, Ling LE. Successful shape-based virtual screening: the discovery of a potent inhibitor of the Type I TGFβreceptor kinase (TβRI). *Bioorg Med Chem Lett* 2003;13:4355–4359; cf. *Bioorg Med Chem Lett* 2004;14:2991.
105. Kick EK, Roe DC, Skillman G, Liu AG, Ewing TJA, Sun Y, Kuntz ID, Ellman JA. Structure-based design and combinatorial chemistry yield low nanomolar inhibitors of cathepsin D. *Chem Biol* 1997;4:297–307.

106. Desai PV, Patny A, Sabnis Y, Tekwani B, Gut J, Rosenthal P, Srivastava A, Avery M. Identification of novel parasitic cysteine protease inhibitors using virtual screening. 1. The ChemBridge database. *J Med Chem* 2004;47: 6609–15.

107. DesJarlais RL, Seibel GL, Kuntz ID, Furth PS, Alvarez JC, Ortiz de Montellano PR, DeCamp DL, Babé LM, Craik CS. Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proc Natl Acad Sci USA* 1990;87:6644–8.

108. Rutenber E, Fauman EB, Keenan RJ, Ortiz de Montellano PR, Meng E, Kuntz ID, DeCamp DL, Salto R, Rosé JR, Craik CS, Stroud RM. Structure of a non-peptide inhibitor complexed with HIV-1 protease. Developing a cycle of structure-based drug design. *J Biol Chem* 1993;268:15343–6

109. Lam PYS, Jadhav PK, Eyermann CJ, Hodge CN, Ru Y, Bacheler LT, Meek JL, Otto MJ, Rayner MM, Wong YN, Chang CH, Weber PC, Jackson DA, Sharpe TR, Erickson-Viitanen S. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* 1994;263:380–4.

110. De Lucca GV, Erickson-Viitanen S, Lam PYS. Cyclic HIV protease inhibitors capable of displacing the active site structural water molecule. *Drug Discov Today* 1997;2:6–18.

111. Haque TS, Skillman AG, Lee CE, Habashita H, Gluzman IY, Ewing TJA, Goldberg DE, Kuntz ID, Ellman JA. Potent, low-molecular-weight non-peptide inhibitors of malarial aspartyl protease plasmepsin II. *J Med Chem* 1999;42: 1428–40.

112. Liu Z, Huang C, Fan K, Wei P, Chen H, Liu S, Pei J, Shi L, Li B, Yang K, Liu Y, Lai L. Virtual screening of novel noncovalent inhibitors for SARS-CoV 3C-like proteinase. *J Chem Inf Model* 2005;45:10–17.

113. Böhm HJ, Banner DW, Weber L. Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitors. *J Comput Aided Mol Design* 1999;13:51–6.

114. Stahl M. Structure-based library design. In: Böhm HJ, Schneider G, editors, Virtual screening for bioactive molecules (Vol. 10 of Mannhold R, Kubinyi H, Timmerman H, editors, *Methods and principles in medicinal chemistry*). Weinheim: Wiley-VCH, 2000, pp. 229–64.

115. Rollinger JM, Hornick A, Langer T, Stuppner H, Prast H. Acetylcholinesterase inhibitory activity of scopolin and scopoletin discovered by virtual screening of natural products. *J Med Chem* 2004;47:6248–54.

116. Soelaiman S, Wei BQ, Bergson P, Lee YS, Shen Y, Mrksich M, Shoichet BK, Tang WJ. Structure-based inhibitor discovery against adenylyl cyclase toxins from pathogenic bacteria that cause anthrax and whooping cough. *J Biol Chem* 2003;278:25990–7.

117. Powers RA, Morandi F, Shoichet BK. Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. *Structure* 2002;10:1013–23.

118. Krier M, de Araújo-Júnior JX, Schmitt M, Duranton J, Justiano-Basaran H, Lugnier C, Bourguignon JJ, Rognan D. Design of small-sized libraries by combinatorial assembly of linkers and functional groups to a given scaffold: application to the structure-based optimization of a phosphodiesterase 4 inhibitor. *J Med Chem* 2005;48:3816–22.

119. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, Connolly DT, Shoichet BK. Molecular docking and high throughput screening for novel inhibitors of protein tyrosine phosphatase 1B. *J Med Chem* 2002;45:2213–21.
120. Iwata Y, Arisawa M, Hamada R, Kita Y, Mizutani MY, Tomioka N, Itai A, Miyamoto S. Discovery of novel aldose reductase inhibitors using a protein structure-based approach: 3D-database search followed by design and synthesis. *J Med Chem* 2001;44:1718–28.
121. Rastelli G, Ferrari AM, Constantino L, Gamberini MC. Discovery of new inhibitors of aldose reductase from molecular docking and database screening. *Bioorg Med Chem* 2002;10:1437–50.
122. Krämer O, Hazemann I, Podjarny AD, Klebe G. Virtual screening for inhibitors of human aldose reductase. *Proteins Struct Funct Genet* 2004;55:814–23.
123. Toyoda T, Brobey RKB, Sano G, Horii T, Tomioka N, Itai Akiko. Lead discovery of inhibitors of the dihydrofolate reductase domain of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Biochem Biophys Res Commun* 1997;235:515–19.
124. Rastelli G, Pacchioni S, Sirawaraporn W, Sirawaraporn R, Parenti MD, Ferrari AM. Docking and database screening reveal new classes of *Plasmodium falciparum* dihydrofolate reductase inhibitors. *J Med Chem* 2003;46:2834–45.
125. Gschwend DA, Sirawaraporn W, Santi DV, Kuntz ID. Specificity in structure-based drug design: identification of a novel, selective inhibitor of *Pneumocystis carinii* dihydrofolate reductase. *Proteins Struct Funct Genet* 1997;29:59–67.
126. Wyss PC, Gerber P, Hartman PG, Hubschwerlen C, Locher H, Marty HP, Stahl M. Novel dihydrofolate reductase inhibitors. Structure-based versus diversity-based library design and high-throughput synthesis and screening. *J Med Chem* 2003;46:2304–12.
127. Pickett SD, Sherborne BS, Wilkinson T, Bennett J, Borkakoti N, Broadhurst M, Hurst D, Kilford I, McKinnell M, Jones PS. Discovery of novel low molecular weight inhibitors of IMPDH via virtual needle screening. *Bioorg Med Chem Lett* 2003;13:1691–4.
128. Li C, Xu L, Wolan DW, Wilson IA, Olson AJ. Virtual screening of human 5-aminoimidazole-4-carboxamide ribonucleotide transformylase against the NCI diversity set by use of AutoDock to identify novel nonfolate inhibitors. *J Med Chem* 2004;47:6681–90.
129. Grüneberg S, Stubbs MT, Klebe G. Successful virtual screening for novel inhibitors of human carbonic anhydrase: strategy and experimental confirmation. *J Med Chem* 2002;45:3588–602.
130. Grüneberg S, Wendt B, Klebe G. Subnanomolar inhibitors from computer screening: a model study using human carbonic anhydrase II. *Angew Chem Int Ed Engl* 2001;40:389–393; *Angew Chem* 2001;113:404–8.
131. Grzybowski BA, Ishchenko AV, Kim CY, Topalov G, Chapman R, Christianson DW, Whitesides GM, Shakhnovich EI. Combinatorial computational method gives new picomolar ligands for a known enzyme. *Proc Natl Acad Sci* 2002;99:1270–3.
132. Grzybowski BA, Ishchenko AV, Shimada J, Shakhnovich EI. From knowledge-based potentials to combinatorial lead design *in silico*. *Acc Chem Res* 2002;35:261–9.

133. Boehm HJ, Boehringer M, Bur D, Gmuender H, Huber W, Klaus W, Kostrewa D, Kuehne H, Luebbers T, Meunier-Keller N, Mueller F. Novel inhibitors of DNA gyrase: 3-D structure based biased needle screening. Hit validation by biophysical methods, and 3-D guided optimization. A promising alternative to random screening. *J Med Chem* 2000;43:2664–74.

134. Babaoglu K, Page MA, Jones VC, McNeil MR, Dong C, Naismith JH, Lee RE. Novel inhibitors of an emerging target in *Mycobacterium tuberculosis*; substituted thiazolidinones as inhibitors of dTDP-rhamnose synthesis. *Bioorg Med Chem Lett* 2003;13:3227–30.

135a. Perola E, Xu K, Kollmeyer TM, Kaufmann SH, Prendergast FG, Pang YP. Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads. *J Med Chem* 2000;43:401–8.

135b. Kaminski JJ, Rane DF, Snow ME, Weber L, Rothofsky ML, Anderson SD, Lin SL. Identification of novel farnesyl protein transferase inhibitors using three-dimensional database searching methods. *J Med Chem* 1997;40:4103–12.

136. Aronov AM, Munagala NR, Kuntz ID, Wang CC. Virtual screening of combinatorial libraries across a gene family: in search of inhibitors of *Giardia lamblia* guanine phosphoribosyltransferase. *Antimicr Agents Chemother* 2001;45:2571–6.

137. Hong H, Neamati N, Wang S, Nicklaus MC, Mazumder A, Zhao H, Burke Jr TR, Pommier Y, Milne GWA. Discovery of HIV-1 integrase inhibitors by pharmacophore searching. *J Med Chem* 1997;40:930–6.

138. Dayam R, Sanchez T, Clement O, Shoemaker R, Sei S, Neamati N. β-Diketo acid pharmacophore hypothesis. 1. Discovery of a novel class of HIV-1 integrase inhibitors. *J Med Chem* 2005;48:111–20.

139. Brenk R, Naerum L, Gradler U, Gerber HD, Garcia GA, Reuter K, Stubbs MT, Klebe G. Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis. *J Med Chem* 2003;46:1133–43.

140. Brenk R, Meyer EA, Reuter K, Stubbs MT, Garcia GA, Diederich F, Klebe G. Crystallographic study of inhibitors of tRNA-guanine transglycosylase suggests a new structure-based pharmacophore for virtual screening. *J Mol Biol* 2004;338:55–75.

141. Schneider G, So SS. *Adaptive Systems in Drug Design* (Biotechnology Intelligence Unit 5). Georgetown: Landes Bioscience 2002. p.45–7.

142. Schneider G, Clément-Chomienne O, Hilfiger L, Schneider P, Kirsch S, Boehm HJ, Neidhart W. Virtual screening for bioactive molecules by evolutionary de novo design. *Angew Chem Int Ed Engl* 2000;39:4130–3; *Angew Chem* 2000;112:4305–9.

143. Lewell XQ, Judd DB, Watson SP, Hann MM. 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;38:511–22.

144. Peukert S, Brendel J, Pirard B, Brüggemann A, Below P, Kleemann HW, Hemmerle H, Schmidt W. Identification, synthesis, and activity of novel blockers of the voltage-gated potassium channel Kv1.5. *J Med Chem* 2003;46:486–98.

145. Peukert S, Brendel J, Pirard B, Strübing C, Kleemann HW, Böhme T, Hemmerle H. Pharmacophore-based search, synthesis, and biological evaluation of anthranilic amides as novel blockers of the Kv1.5 channel. *Bioorg Med Chem Lett* 2004;14:2823–7.

146. Pirard B, Brendel J, Peukert S. The discovery of Kv1.5 blockers as a case study for the application of virtual screening approaches. *J Chem Inf Model* 2005;45:477–85.

147. Newman DJ, Cragg GM, Snader KM. Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod* 2003;66:1022–37.

148. Liu H, Li Y, Song M, Tan X, Cheng F, Zheng S, Shen J, Luo X, Ji R, Yue J, Hu G, Jiang H, Chen K. Structure-based discovery of potassium channel blockers from natural products: virtual screening and electrophysiological assay testing. *Chem Biol* 2003;10:1103–13.

149. Enyedy IJ, Ling Y, Nacro K, Tomita Y, Wu X, Cao Y, Guo R, Li B, Zhu X, Huang Y, Long YQ, Roller PP, Yang D, Wang S. Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *J Med Chem* 2001;44:4313–24.

150. Wu YQ, Belyakov S, Chi Choi, Limburg D, Thomas BE IV, Vaal M, Wei L, Wilkinson DE, Holmes A, Fuller M, McCormick J, Connolly M, Moeller T, Steiner J, Hamilton, GS. Synthesis and biological evaluation of non-peptidic cyclophilin ligands. *J Med Chem* 2003;46:1112–15.

151. Burkhard P, Hommell U, Sanner M, Walkinshaw MD. The discovery of steroids and other novel FKBP inhibitors using a molecular docking program. *J Mol Biol* 1999;287:853–8.

152. Filikov AV, Mohan V, Vickers TA, Griffey RH, Cook PD, Abagyan RA, James TL. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J Comput Aided Mol Design* 2000;14:593–610.

153. Kurogi Y, Miyata K, Okamura T, Hashimoto K, Tsutsumi K, Nasu M, Moriyasu M. Discovery of novel mesangial cell proliferation inhibitors using a three-dimensional database searching method. *J Med Chem* 2001;44:2304–7.

154. Gao Y, Dickerson JB, Guo F, Zheng J, Zheng Y. Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc Natl Acad Sci USA* 2004;101:7618–23.

155. Singh J, van Vlijmen H, Liao Y, Lee WC, Cornebise M, Harris M, Shu I, Gill A, Cuervo JH, Abraham WM, Adams SP. Identification of potent and novel α4β1 antagonists using *in silico* screening. *J Med Chem* 2002;45:2988–93.

156. Cruciani G, editor. *Molecular interaction fields* (Vol. 27 of Mannhold R, Kubinyi H, Folkers G, editors, *Methods and principles in medicinal chemistry*), Weinheim: Wiley-VCH, 2005.

157. Gohlke H, Hendlich M, Klebe G. Predicting binding modes, binding affinities and 'hot spots' for protein-ligand complexes using a knowledge-based scoring function. In: Klebe G, editor, *Virtual screening: an alternative or complement to high throughput screening.* Dordrecht: Kluwer Academic Publishers, 2000; also published in *Persp Drug Discov Des* 2000;20:115–44.

158. Sotriffer CA, Gohlke H, Klebe G. Docking into knowledge-based potential fields: a comparative evaluation of DrugScore. *J Med Chem* 2002;45:1967–70.

159. Catalyst software. Accelrys Inc, San Diego.
160. Li H, Sutter J, Hoffmann R. HypoGen: an automated system for generating 3D predictive pharmacophore models. In: Güner OF, editor, *Pharmacophore perception, development and use in drug design*., La Jolla: International University Line, 2000. p. 173–89.
161. Clement OO, Mehl AT. HipHop: pharmacophores based on multiple common-feature alignments. In: Güner OF, editor, *Pharmacophore perception, development and use in drug design*. La Jolla: International University Line, 2000. p. 69–84.
162. Kurogi Y, Güner OF. Pharmacophore modeling and three-dimensional database searching for drug design using Catalyst. *Curr Med Chem* 2001;8:1035–155.
163. Güner O, Clement O, Kurogi Y. Pharmacophore modeling and three dimensional database searching for drug design using Catalyst: recent advances. *Curr Med Chem* 2004;11:763–71.
164. Wolber G, Langer T LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 2005;45:160–9.
165. Rarey M, Dixon JS. Feature trees: a new molecular similarity measure based on tree matching. *J Comput Aided Mol Des* 1998;12:471–90.
166. Rarey M, Stahl M. Similarity searches in large combinatorial chemistry spaces. *J Comput Aided Mol Des* 2001;15:497–520.
167. Unity software. Tripos Associates Inc, St. Louis.
168. Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 2001;15:411–28.
169. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 1996;261:470–89.
170. Kramer B, Rarey M, Lengauer T. Evaluation of the FlexX incremental construction algorithm for protein ligand docking. *Proteins Struct Funct Genet* 1999;37:228–41.
171. Hindle SA, Rarey M, Buning C, Lengauer T. Flexible docking under pharmacophore type constraints. *J Comput Aided Mol Des* 2002;16:129–49.