Computational Chemistry at Novartis

Richard Lewis[a]*, Peter Ertl[a], Edgar Jacoby[a], Marina Tintelnot-Blomley[a], Peter Gedeck[b],
Romain M. Wolf[a], and Manuel C. Peitsch[a]

Abstract: Computational approaches have become an integral part of modern drug discovery and medicinal chemistry. These approaches can be roughly classified into data/information mining (or filtering) and modelling/simulation methods. Taken together, they represent an ever growing source of hypotheses used to guide experimental approaches and hence drug discovery decisions. Therefore, it is not only important to optimally understand and apply existing methods, but also invest in the development of new algorithms to further improve our selection of drug candidate. The present contribution will describe a few approaches which have become routine at Novartis.

Keywords: Computational chemistry · *In silico* Screening · Molecular informatics · QSAR ·
Structure-based drug design

The Renaissance of 2D-QSAR

In every drug discovery project, a thorough understanding of the structure–activity relationships (SAR) is pivotal for fast and efficient development of drug candidates. By using a variety of computational techniques qualitative or quantitative SAR (QSAR) models can be derived that allow either the classification of compounds as active or inactive (classification) or the prediction of binding activities of compounds quantitatively (regression). In all cases, the first step is to describe the chemical structure with a set of numbers, so-called descriptors. Since the early work of Hansch and Fujita [1] a large number of different descriptors were developed for QSAR model generation [2]. Fundamentally, these descriptors can be distinguished by the dimensionality of the structural representation. For a long time, mainly 2D structure representations were used to derive numerical descriptors.

However, following the development of the CoMFA (Comparative molecular field analysis) method by Cramer and coworkers [3] 3D structure descriptors became more and more popular based on the belief that 3D descriptors capture the effects relevant for binding more efficiently than the simpler 2D descriptors. As described below, 3D-QSAR models provide easily interpretable hints on where and how active compounds should be modified. But to make 3D methods really work, it is crucial to get the binding conformation and molecular alignments right. As good methods for this are still lacking, this often requires tedious manual manipulation of the different structures. In an industrial research environment there is often not the time to do this; for large datasets 2D-QSAR methods are therefore the method of choice

Over the last years 2D-QSAR methods experienced a renaissance, due to several changes in the pharmaceutical industry shifting the focus of computational chemistry from single compounds to larger datasets. The introduction of combinatorial chemistry techniques and improved purification methods to the medicinal chemist's lab, paired with new screening methods produced a considerable increase of data for lead optimisation projects. The volume of data is even higher in the analysis of high-throughput screening (HTS) campaigns and valuable SAR information can be extracted, if the results are mined with adequate techniques. Another exciting development is the appearance of workflow management tools that are capable of processing structural data (see *e.g.* SciTegic [4] and InforSense [5]). This simplifies working with extremely large datasets and leads to an increasing automation of computational chemistry workflows. Finally, new statistical approaches have been introduced that are robust enough to cope with large and/or noisy datasets. Naïve Bayesian classifiers, decision trees, random forests, and support vector machines are just some of the methods that have been added to the computational chemist's toolkit [6–9].

A few applications should help demonstrate the value of 2D-QSAR approaches. While high-throughput screening is often perceived as the method of choice for hit finding, not every assay can be formatted in a way suitable for HTS. Also, it is sometimes desirable to extract subsets from the screening selection focused on a specific biological target for tool finding. It has been shown that 2D descriptors combined with Binary Kernel Discrimination classifiers are useful approaches to develop such focused screening sets [10]. Another interesting application developed by Klon *et al.* [6][11][12] is to use a naïve Bayesian classifier to improve the enrichment of HTD results.

In a recent study, we compared the performance of 2D and 3D descriptors in regression models [13] using almost 1000 datasets. In practically all cases, models using 2D descriptors were far superior to models using the alignment-free 3D Grind descriptors [14]. The use of applying QSAR models during a lead optimisation study was demonstrated in a retrospective study by Hirons *et al.* [15]. The advantage of using iteratively refined QSAR models was clearly demonstrated. The 2D-descriptors used in this work were specifically de-

Correspondence: Dr. R. Lewis[a]
Tel.: +41 61 696 2449
Fax: +41 61 696 8672
E-Mail: Richard.lewis@novartis.com
[a]Novartis Institutes for BioMedical Research
CH-4002 Basel
[b]Novartis Institutes for BioMedical Research
Novartis Horsham Research Center
Horsham, RH12 5AB, UK

veloped for lead optimisation datasets and are particularly useful for identifying bio-isosteric replacements and combinatorial library design [16].

In summary, 2D-QSAR approaches can successfully be applied in all stages of the drug development process. The speed in which they can be calculated makes them especially suitable for handling large datasets. 2D-QSAR methods are also attractive approaches for iterative, automatic generation of QSAR models during lead optimisation.

Indirect Drug Design

Indirect methods for drug design are used when there are no high-quality structures of the target from which a structure-based drug design effort can be launched. Such methods can also be used in combination with homology models to provide supporting evidence for the model. This is particularly useful for targets where the X-ray structure is of low resolution, or there is very low homology to known structures. Indirect methods start with the structures of small molecules and their activity at a target, and then attempt to deduce information about the bioactive conformation and binding mode of the active compounds. The principal methods, pharmacophore analysis and 3D-QSAR, work best if the activity data is taken from a binding assay rather than an *ex vivo* or *in vivo* assay, to reduce the influence of bioavailability effects. It is also assumed in the first instance that the molecules have a single common binding mode. This is however a questionable assumption in many cases, but in absence of any contrary evidence, it is the starting point.

A pharmacophore is a model defined by the features of a molecule (hydrogen bonding groups, hydrophobic regions, charged groups or aromatic ring centres) thought to be essential for biological activity and by the geometric relationships of the features, for instance the distances between them. From these features, a 3D pharmacophore model for bioactivity can be proposed. These models are then used to score novel compounds before synthesis or to search corporate databases for structurally novel active chemotypes [17]. The first stage of the process is to sample the low-energy conformations of all the molecules, followed by an analysis of the pharmacophores present in the active molecules but not in the inactive ones. The key difficulties are to ensure that the bioactive conformation is sampled, and that the features encode the right information (for example that carboxylate, hydroxamate and tetrazole groups are recognised as bioisosteres. The strengths and weaknesses of commercial programs for pharmacophore

analysis have been examined by Patel and coworkers [18].

3D-QSAR methods try to build a statistical model which describes the biological activity as a function of the steric and electrostatic fields around the molecules, simulating potential receptor interactions [19]. The model may be used for activity prediction and the design of novel, more potent molecules during lead optimisation. Visualisation of the results, in terms of favourable and unfavourable regions around the molecules, greatly facilitates the design of new structures. The key to the success of the method is to have a good way to superimpose the molecules [20]. A good example of the application of the method to a relevant target is in the work of Cavalli and coworkers [21], looking at the requirements of hERG binding. Pharmacophore and 3D-QSAR models are used routinely as *in silico* screens by chemists, to help set synthetic priorities, and to assist chemists to design better compounds.

Structure-based Drug Design (SBDD)

Structure-based drug design, as drug discovery in general, is not a linear, stand-alone, single shot procedure, but an iterative and integrated approach, which follows – in close collaboration with all the different disciplines involved – a circular process of improving all parameters describing a final drug. In this respect, earlier structure-based ligand design has matured towards an approach which became an integral part of most industrial drug discovery projects where structural information of the target is available or deducible from related targets. Development, growth and appreciation of SBDD are firmly linked to the rapid progress in protein crystallography during the last 20 years. Starting with the revolutionary ideas for understanding protein–ligand interactions [22] and the first success stories of SBDD in HIV protease taking advantage of multiple co-crystal structures, this situation has become a standard scenario for many projects. In addition to structural biology, the explosion in computer power and graphics capabilities should not be underestimated as factors leading to the current success.

A target with a reasonably well-defined binding pocket, *i.e.* possibly no surface interaction or unspecific binding, combined with high-resolution structural information, is the favourable starting point, if not the prerequisite, for SBDD. A thorough analysis and characterisation of this binding site – by graphical visualisation and using appropriate software (*e.g.* for calculating and storing molecular fields as grids) – is a first crucial step at the beginning of a project.

Knowledge of 3D requirements, functional groups possibly involved in interactions and a physicochemical description of the binding site are the keys to the second step: to optimise the complementarity between a putative ligand and its binding site by making use of insight into the forces which make a small molecule fit at the active site (favourable interaction geometries and chemical principles), and accordingly by designing molecules which match the requirements of their binding site [23].

A plethora of methods and strategies can contribute here depending on the current level of knowledge. Visual inspection and modeller's and chemist's intuition, *i.e.* the human brain, creativity and experience (of course supported by steadily improving means to analyse situation and needs) is still of utmost importance. Conformational preferences or patterns of non-bonded interactions might still be better judged by a modeller than a simplified calculation. Virtual screening, especially for small fragments or building blocks fitting a well defined pocket, the design of focused libraries, or *de novo* methods can be part of the SBDD approach. In order to apply these methods most efficiently, the limited knowledge at the start of a program is expanded progressively. Major limitations such as protein flexibility and the effect of solvent are beyond high-throughput methods (and challenges for every approach) and have to be borne in mind. Starting from different levels of approximations, and cycle by cycle, more details can be included.

At the same time synthetic accessibility and the potential for modifications that lead to drug-like properties are taken into account. The inability to include such criteria is a major drawback of most *de novo* design programs.

Structure-based methods are applied in all phases of a project. Whilst one has to be lucky to identify a good hit as starting point from the screening of in-house databases, a co-crystal structure of the target with a substrate, cofactor, known lead *etc.* is a solid starting point. Based on such an entry point, fragments that match particular parts of the identified binding site can be searched for. New scaffolds are explored, which represent characteristic 3D or physicochemical features and are open to wider chemical variation. The identification and optimisation of a declared lead is always guided by the firm knowledge gained from the binding site analysis.

In the field of oncology research at Novartis – to take an example – SBDD was successful at different stages of medicinal chemistry projects. Virtual screening of a subset of the corporate compound collection in a homology model of human CK2 protein kinase – combined with filtering based on known pharmacophore informa-

tion – identified a potent lead among only a dozen compounds selected for testing [24]. The design of a new lead series for CDK2 inhibition exemplifies how target family information can be utilised for the design of specificity [25]. The work on inhibitors of the chymotrypsin-like activity of the 20S proteasome shows a structure-based optimisation of a high-throughput screening hit [26].

The first example shows that the lines between SBDD and HTD are flexible. Depending on the level of input from existing knowledge, HTD can be just one more method in SBDD making use of the additional throughput of ideas. But on the other hand, HTD can also be a very basic screening tool, focusing on speed and throughput, as opposed to SBDD's focus on accuracy or on guidance from the scientist's knowledge.

HTD and Scoring

HTD (High-throughput docking), where candidate ligand molecules are docked *in silico* into a 3D molecular structure model of the target protein to generate physically reasonable models of the ligand–target complex and to assess their binding potential [27], is today successfully integrated in various scenarios of the target-based lead finding process [28]. An impressive number of success stories have been reported for lead finding applications with typical experimental validation hit rates of 1–10% for various target families providing specific micromolar compounds [29].

With the progress within the field of structural biology, HTD approaches hold an enormous potential for the early drug discovery process. These expectations and the availability of cheap computing hardware, including Linux clusters and more recently GRID computing platforms [24][30], motivated in the last decade the development of around 20 docking software packages (*e.g.* DOCK, Autodock, Gold, FlexX, Glide, ICM, Ligfit) [27][31].

The HTD algorithms are either based on incremental build-up procedures and/or systematic or random sampling techniques to generate coordinates or poses – candidate models – of the ligand 3D conformation and molecular interactions within the target binding site. This geometric probing samples, ideally, the entire conformational space of the fully flexible candidate ligand by variations of the dihedral angles and by all possible placements and orientations of the entire molecule, a combinatorial complex problem. The existing algorithms are fast and robust within the rigid or limited side-chain flexibility receptor approximation and some of the guided docking methods allow the inclusion of constraints in

form of desired pharmacophore contacts (*e.g.* GOLD) [32]. Conversely, the computational assessment of the binding potential is a very challenging scientific problem [26][31][33]. In theory, the computed binding energy should correspond to the experimental binding energy and the predicted pose should correspond to the experimental 3D structure of the ligand–receptor complex. While in general, the docking methods are surprisingly efficient to predict accurate binding poses for high-affinity drug-like small molecular ligands when the used 3D model of the receptor is of relevance, accurate computational prediction of absolute binding energies still do not exist [31][33].

The final selection of compounds must rely on approximate scoring techniques and careful visual inspection by a computational and medicinal chemist to integrate complex non-computer amenable expertise. Among the diverse scoring methods, basically three categories can be distinguished. Molecular mechanics force-field based scoring functions (*e.g.* G-Score, D-Score, Gold, DOCK, AutoDock) include explicit Lennard-Jones and Coulomb terms to compute, respectively, the van der Waals and electrostatic terms of the binding energy and take into account the conformational strain energy. Empirical scoring functions (*e.g.* LUDI, F-Score, ChemScore) describe the binding energy as a sum of empirically based parameters for specific interaction contacts (*e.g.* hydrophobic contact, hydrogen-bond, *etc.*) derived from regression analysis for the binding energy and/or the binding poses of selected reference sets of receptor-ligand complexes with known experimental 3D structures and binding energies. Finally, knowledge-based scoring functions (*e.g.* PMF, DrugScore, SMoG) are based on categories of simple atom pair contact parameters and were derived to reproduce binding poses. All of these scoring functions have significant imperfections and depending on the nature of target family or the binding site, characterised by the shape and polarity, the one or other method might perform better than another one [26][31][33][34]. Most essentially, the quality allows the distinction of active and inactive molecules and a number of computational protocols are currently developed to boost the enrichment. As for instance, it is *a priori* not clear which scoring function is the most appropriate for a newly investigated target, consensus scoring functions (*e.g.* X-CSCORE, C-Score) combining the output of individual scores or ranks can be used and were sometimes shown to perform better than any single score alone [33][34]. More recently, the application of multivariate analysis and machine learning methods were shown to be useful [35]. For instance, the application of the *naive* Bayes classifier method after consensus scoring was shown

capable to provide further enrichment [17]. Other knowledge-based approaches emerge which apply ligand specific transferable multiple active site correction terms derived from docking the candidate ligands on multiple targets in parallel [36].

There are several limitations on the physically-based binding energy [37–39]. First, it is that ligand binding is a much more complex phenomenon than the rigid key and lock mechanism. It is a dynamic process, including structural reorganisation of the ligand, the protein and the hydrating water environment [40], which is better described as a transient binding site occupation. In the bound state the ligand executes restricted back and forward motions. Hence, not all specific atom contacts defining the binding interactions exist persistently and providing a rationale why straight forward summing-up of the interactions terms might be an imperfect methodology [37]. In addition, effects linked to spatial confinement in the binding site could require quantum theory based methods [38]. Ligand binding integrates cooperative and frustrating compensatory effects of multiple weak interactions which result in the observed time averaged total binding energy [38][39]. To achieve further progress with the first principle methods, future developments will probably include molecular dynamics simulations in explicit water and refined force field parameters as recently shown by the case study of the progesterone receptor [41].

Web-based Cheminformatics Tools

Computational methods are becoming increasingly part of the workflow of modern medicinal chemists. Typical tasks involved in this process include fast calculation of molecular properties and ADMET characteristics for large data sets, analysis of molecular diversity or design of targeted combinatorial libraries. Although these tasks are traditionally a domain of expert molecular modellers, bench chemists are increasingly interested in performing these analyses themselves. At Novartis we are supporting this effort by providing a collection of cheminformatics tools for medicinal chemists available on their desktops through an easy-to-use web-based interface.

Tools based on the web technology are indeed ideal for interactive processing of chemical information. Such systems provide easy access to relevant data, allow visualisation, processing and analysis of the data and development of models which help users to understand complex relationships within the data. The web tools support medicinal chemists in their daily tasks, as well as more specialised activities, such as the design of targeted combinatorial libraries,

bioisosteric design or virtual screening.

Ciba-Geigy (which merged in 1996 with Sandoz to form Novartis) was one of the first large pharmaceutical companies to recognise the advantages of web technology for processing molecular information. By 1995 its chemists were already able to use a web-based cheminformatics system through the company intranet [42–44]. The Novartis web-based cheminformatics system has been continually updated and currently consists of more than 20 modules, supporting a broad range of various cheminformatics and molecular processing tasks including:

- calculation of molecular properties and drug transport characteristics;
- sophisticated toxicology alerting;
- molecular visualisation, including visualisation of surface molecular properties;
- support for diversity analysis and enumeration of combinatorial libraries;
- drug design based on bioisosteric principle;
- various helper tools (file format conversion, molecule depiction, 2D to 3D conversion, *etc.*).

The most popular tool on the Novartis intranet is a service for calculation of various molecular properties called ***In silico Profiling*** (Fig.). It offers access to properties which may be calculated in a straightforward manner from the molecular connectivity (SMILES string), are easily interpretable and provide good correlation with molecular bioavailability. Available properties include octanol–water partition coefficient logP, molar refractivity, flexibility index, hydrogen bonding characteristics and molecular polar surface area calculated ac-

cording to the protocol developed in-house [45]. Various drug transport characteristics, such as intestinal absorption, blood-brain-barrier penetration or plasma-protein binding calculated based on in-house models are also provided. Novartis cheminformatics web system features also a sophisticated toxicity alerting system – ***In silico Tox Check*** [46] using a database of nearly 200 complex substructures which may cause toxicity, collected from literature or based on results of Novartis in-house toxicity testing. The system allows toxicity alerting on about 20 endpoints, with main focus on carcinogenicity and mutagenicity.

Novartis web-based molecular processing tools installed on the corporate intranet bring easy-to-use cheminformatics and molecular modelling capabilities directly to the desks of synthetic chemists enabling them comfortable access to data and their visualisation and analysis, considerably improving efficiency of the drug design and development process.

In silico Drug Discovery Pipeline

As exemplified in the previous section, bench scientists have an ever increasing need for simple-to-use tools to conduct *in silico* calculations in drug discovery. Our vision is to further extend the portfolio of approaches offered through Web-based interfaces and thereby enable scientists to rapidly test their hypothesis before contacting domain experts who would then focus on applying more advanced methods and spend more time to analysis specific project-related problems. In this context, we intend to further develop several Web-

based applications that can be chained in a pipeline to address the computational steps from target identification, protein modelling to high throughput docking and hit selection. The basic concept of this vision is derived from previously implemented automated systems such as GeneQuiz [47] and SwissModel [48], both of which focus on a portion of the pipeline. These can be seen as proof of concept systems for our vision.

Detailed Atomistic Models of GPCRs

G Protein-coupled receptors (GPCRs) are the most prominent target family for pharmaceutical compounds. But with only one X-ray structure available so far (bovine rhodopsin [49]), detailed structural information for GPCRs is sparse. Computer-aided drug design for GPCRs has to rely either on QSARs (without a relation to the structure of the receptor) or on homology models based on the rhodopsin template (Building models *ab initio* based on receptor primary structure alone is not feasible at this point). The quality of these models can be assessed experimentally. However, despite the fact they may explain a limited set of experimental findings, they must be considered with care.

To study the interactions between GPCRs and low-molecular-weight compounds, the most interesting part of GPCRs is the predominantly hydrophobic seven-helices trans-membrane (7TM) region, including in addition the (more polar) extracellular loops when these are supposed to contribute significantly to interactions with ligands.

The standard way to generate homology models of the 7TM part of a GPCR is to
- find its location in the primary sequence (*e.g.* by hydrophobicity plots);
- align the sequences of the seven individual helices to those in the X-ray structure of rhodopsin;
- substitute the side chains of rhodopsin with the respective ones of the GPCR under construction;
- refine the crude model by techniques like energy refinement and molecular dynamics, mostly applying some constraints to preserve the overall helix arrangement as in rhodopsin.

Many pre-built models can be found at *www.gpcr.org* [50]. Procedures to generate and use GPCR models have been described throughout recent literature [51]. Vriend *et. al.* [52] list the problems encountered in building GPCR models, summarising the situation both before the publication of the rhodopsin crystal structure and afterwards.

In general, models are generated to explain or predict experimental findings. This is possible with GPCRs where for example
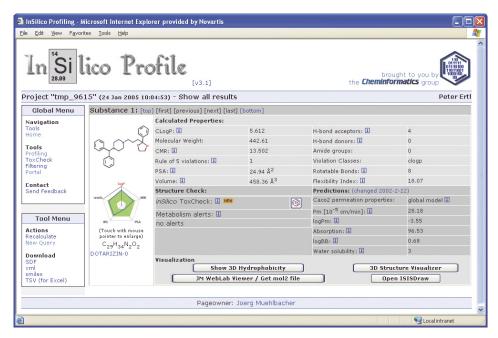

Fig. Cheminformatics *In Silico* Profiling web service

explicit 3D models may serve to guide mutation studies, *e.g.* to determine the residues responsible for a specific function [53]. Models can also be used to dock ligands, giving hints on the complexation mode and on favourable modifications to drug candidate molecules [54]. In such cases, an iterative process leads to the refinement of the crude model which explains initial experimental data, suggests additional experiments (point mutations and/or changes to the chemical structure of the ligands) which then require further adjustment of the model until experimental data and model are congruent.

In summary, homology models of GPCRs are useful in drug design. Like homology models for other receptor classes, they yield a working hypothesis, which should be tested by experimentation.

Received: May 19, 2005

[1] C. Hansch, T. Fujita, *J. Am. Chem. Soc.* **1964**, *86*, 1616.

[2] R. Todeschini, V. Consonni, 'Handbook of Molecular Descriptors', Wiley-VCH, Weinheim, **2000**.

[3] R.D. Cramer III, D.E. Patterson, J.D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959.

[4] SciTegic, San Diego, *http://www.scitegic.com*.

[5] InforSense Ltd., London, *http://www.inforsense.com*.

[6] A.E. Klon, M. Glick, M. Thoma, P. Acklin, J.W. Davies, *J. Med. Chem.* **2004**, *47*, 2743.

[7] A. Rusinko III, M.W. Farmen, C.G. Lambert, P.L. Brown, S.S. Young, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1017.

[8] V. Svetnik, A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947.

[9] D. Wilton, P. Willett, K. Lawson, G. Mullier, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 469.

[10] J. Hert, P. Willett, D.J. Wilton, P. Acklin, K. Azzaoui, E. Jacoby, A. Schuffenhauer, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1177.

[11] A.E. Klon, M. Glick, J.W. Davies, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2216.

[12] A.E. Klon, M. Glick, J.W. Davies, *J. Med. Chem.* **2004**, *47*, 4356.

[13] P. Gedeck, C. Bartels, B. Rohde, to be published.

[14] M. Pastor, G. Cruciani, I. McLay, S. Pickett, S. Clementi, *J. Med. Chem.* **2000**, *43*, 3233.

[15] L. Hirons, J.D. Holliday, S.P. Jelfs, P. Willett, P. Gedeck, *QSAR,* to be published.

[16] J.D. Holliday, S.P. Jelfs, P. Willett, P. Gedeck, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 406–411.

[17] D.P. Marriott, I.G. Dougall, P. Meghani, Y.J. Liu, D.R. Flower, *J. Med Chem.* **2002**, *42*, 3210.

[18] Y. Patel, V.J. Gillet, G. Bravi, A.R. Leach, *J. Comp.-Aid. Mol. Des.* **2002**, *16*, 653.

[19] Y.C. Martin, '3D QSAR: Current State, Scope, and Limitations', in 'Perspectives in Drug Discovery and Design', **1998**, p. 3.

[20] C. Lemmen, T.J. Lengauer, *J. Comp.-Aid. Mol. Des.* **2000**, *14*, 215.

[21] Cavalli, E. Poluzzi, F. De Ponti, M. Recanatini, *J. Med. Chem.* **2002**, *45*, 3844.

[22] P.J.Goodford, *J. Med. Chem.* **1984**, *27*, 557.

[23] H. Gohlke, G. Klebe, *Angew. Chem. Int. Ed.* **2002**, *41*, 2644.

[24] E. Vangrevelinghe, K. Zimmermann, J. Schoepfer, R. Portmann, D. Fabbro, P. Furet, *J. Med. Chem.* **2003**, *46*, 2656.

[25] J. Schoepfer, H. Fretz, B. Chaudhuri, L. Muller, E. Seeber, L. Meijer, O. Lozach, E. Vangrevelinghe, P. Furet, *J. Med. Chem.* **2002**, *45*, 1741.

[26] P. Furet, P. Imbach, P. Fuerst, M. Lang, M. Noorani, J. Zimmermann, C. García-Echeverría, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1331.

[27] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, *Nature Drug Discov. Rev.* **2004**, *3*, 935.

[28] E. Jacoby, A. Schuffenhauer, M. Popov, K. Azzaoui, E. Vangrevelinghe, J. Priestle, P. Ferrara, B. Faller, P. Acklin, *Chimia* **2004**, *58*, 577.

[29] J.C. Alvarez, *Curr. Opin. Chem. Biol.* **2004**, *8*, 365.

[30] P. Afflard, B. Almeida, J. Basse-Welker, M.C. Peitsch, *Digma* **2005**, *5*, 14.

[31] N. Brooijmans, I.D. Kuntz, *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335.

[32] X. Fradera, J. Mestres, *Curr. Top. Med. Chem.* **2004**, *4*, 687.

[33] H.-J. Böhm, M. Stahl, *Rev. Comput. Chem.* **2002**, *18*, 41.

[34] R. Wang, Y. Lu, S. Wang, *J. Med. Chem.* **2003**, *46*, 2287.

[35] M. Jacobsson, P. Liden, E. Stjernschantz, H. Bostrom, U. Norinder, *J. Med. Chem.* **2003**, *46*, 5781.

[36] G.P.A. Vigers, J.P. Rizzi, *J. Med. Chem.* **2004**, *47*, 80.

[37] D.A. Leigh, *Chem. Biol.* **2003**, *10*, 1143.

[38] A. Cooper, C.M. Johnson, J.H. Lakey, M. Nöllmann, *Biophys. Chem.* **2001**, *93*, 215.

[39] W. Murray, M.L. Verdonk, *J. Comput. Aided Mol. Des.* **2002**, *16*, 151.

[40] S. J. Teague, *Nat. Rev. Drug Discov.* **2003**, *2*, 527.

[41] T. Mordasini, A. Curioni, R. Brusi, W. Andreoni, *ChemBioChem* **2003**, *4*, 155.

[42] P. Ertl, *Chimia* **1998**, *52*, 673.

[43] P. Ertl, J. Mühlbacher, B. Rohde, P. Selzer, S*AR and QSAR Env. Res.* **2003**, *14*, 321.

[44] P. Ertl, P. Selzer, J. Mühlbacher, *Drug Disc. Today BIOSILICO* **2004**, *2*, 201.

[45] P. Ertl, B. Rohde, P. Selzer, *J. Med. Chem.* **2000**, *43*, 3714.

[46] J. Mühlbacher, P. Ertl, P. Selzer, L. Müller, S. Glowienke, A. Schuffenhauer, *Nachr. Chem.* **2004**, *52*, 162.

[47] S. Hoersch, C. Leroy, N.P. Brown, M.A. Andrade, C. Sander, *Trends Biochem Sci.* **2000**, *25*, 33.

[48] N. Guex, M.C. Peitsch, *Electrophoresis* **1997**, *18*, 2714.

[49] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I.L. Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289*, 739.

[50] F. Horn, E. Bettler, L. Oliveira, F. Campagne, F.E. Cohen, G. Vriend, *Nucleic Acids Res.* **2003**, *31*, 294.

[51] I. Visiers, J.A. Ballesteros, H. Weinstein, *Meth. Enzymol.* **2002**, 329.

[52] L. Oliveira, T. Hulsen, D.L. Hulsik, A.C.M. Paiva, G. Vriend, *FEBS Letts.* **2004**, *564*, 269.

[53] M.-G Ludwig, M. Vanek, D. Guirini, J.A. Gasser, C.E. Jones, U. Junker, H. Hofstetter, R.M. Wolf, K. Seuwen, *Natur*e **2003**, *425*, 93.

[54] S.U. Miedlich, L. Gama, K. Seuwen, R.M. Wolf, G.E. Breitwieser, *J. Biol. Chem.* **2004**, *279*, 7254.