

Recent Advances in Docking and Scoring

E. M. Krovat*, T. Steindl and T. Langer

Institute of Pharmacy, Department of Pharmaceutical Chemistry, Innrain 52, University of Innsbruck, A-6020 Innsbruck, Austria

Abstract: This review is focused on recent advances and new aspects in the field of molecular docking and scoring, and it covers multiple applications and case studies. Basic requirements and different algorithms for docking are briefly discussed. Moreover, parameters that influence docking results, combination of different docking algorithms and scoring functions, performance of scoring functions, docking using homology models, and ligand and protein flexibility are examined to give an overview of the state-of-the-art methods and a survey of innovative approaches in molecular docking and scoring. Regarding the enormous amount of literature in this field we restrict ourselves on an overview of several important advances in docking and scoring techniques published within the last two years, i.e. we considered publications ranging from 2002 to 2004.

Keywords: Molecular docking, scoring functions, virtual screening, docking algorithm, homology modelling, protein flexibility, high-throughput screening, tautomerism.

INTRODUCTION

Docking procedures aim to identify correct poses of ligands in the binding pocket of a protein and to predict the affinity between the ligand and the protein. In other words, docking describes a process by which two molecules fit together in three-dimensional space. Molecular docking has contributed important proceedings to drug discovery for many years. One main motivation in drug discovery is the identification of innovative small molecular scaffolds exhibiting high binding affinity and selectivity for the target together with a reasonable ADME (absorption, distribution, metabolism, excretion) profile, lead and/or drug likeness. Such chemical entities are likely to be able to enter higher phases in the further drug development process. Molecular docking, compared to the fast and successful method of three-dimensional pharmacophore modelling [1,2] is a rather complex and computer-intensive approach to find new compounds by virtual screening. The crystalline structure of ligands bound to their target receptor is one of the most important sources to gain information about the basic mechanisms of interaction between the parts constituting the three-dimensional complex structure. The interactions result in a cascade of events, e.g. a catalytic reaction, like the cleavage of the substrate, or the stabilisation of a transition state, e.g. by peptidomimetic enzyme inhibitors, or the blockage of the protein's active site due to the tight binding of an inhibitor. Furthermore, these events lead to a sequence of steps that present the molecular basis of pharmacological effects. The behaviour of small molecules in the binding pockets of target proteins can be described by molecular docking. This technique is widely used in lead discovery and optimisation. A survey of the impact of docking, scoring, and scoring functions is given elsewhere [3-9]. In our work the recent advances of the last few years in docking

and scoring are summarised based on multiple application cases.

BASIC REQUIREMENTS FOR MOLECULAR DOCKING

The setup for a ligand docking approach requires the following components: A target protein structure with or without a bound ligand, the molecules of interest or a database containing existing or virtual compounds for the docking process, and a computational framework that allows the implementation of the desired docking and scoring procedures. The three-dimensional structure of the protein-ligand complex has to be detailed at atomic resolution. In many cases only the unbound (ligand-free, apo) form of the protein is determined, without the bioactive conformation of the ligand. Most docking algorithms assume the protein to be rigid, according to the high computational cost that the demand of flexibility implicates. The ligand is mostly regarded as flexible. Beside the conformational degrees of freedom the binding pose in the protein's binding pocket must be taken into consideration. More details regarding protein and ligand flexibility are given later in this article. Docking can be performed by placing rigid molecules or fragments into the protein's active site using different approaches like the clique-search, geometric hashing, or pose clustering. In the clique search matches are searched to describe the compatible characteristics (shape or interaction pattern) of ligand and protein by means of a distance compatibility graph. This approach is implemented in the program DOCK [10]. A geometric hashing function is created to describe geometric features like distances in two steps, the preprocessing phase and the recognition phase. This approach is attractive regarding its time-efficacy and the option of partial matching of the ligand in the protein pocket [11,12]. Pose clustering is an algorithm based on the matching of triplets of features of the ligand with a triplet of features of the protein. The features represent the interaction zones of ligand and receptor. This approach is implemented in the program LUDI [13,14]. A more detailed description of

*Address correspondence to this author at the Institute of Pharmacy, Department of Pharmaceutical Chemistry, Innrain 52, University of Innsbruck, A-6020 Innsbruck, Austria; Tel.: +43 512 507 5264; Fax +43 512 507 5269; E-mail: eva.krovat@uibk.ac.at

the briefly summarised methods of the basic rigid docking algorithms of pattern recognition are reviewed in the work of Muegge *et al.* [3]. The flexibility of the ligand can be represented by a set of conformers covering the conformational space in an exhaustive way. According to the size of the conformational model, a larger time frame is necessary to perform the docking of a single molecule. This means that a compromise between the coverage of conformational space and computational cost (time) must be effected. Different approaches show that the conformational model can vary between 20 and a few hundred conformers per molecule of the ensemble [15,16]. A more sophisticated approach is the implementation of algorithms that simulate flexible docking, like incremental fragmentation (FLEXX), [17,18] molecular simulation (simulated-annealing in AutoDOCK [19], molecular dynamics) [20]. Also Monte Carlo-based algorithms, e.g. implemented in the program PRODOCK [21-23] are used to handle the flexibility of ligands in the docking process. A docking technique based on genetic algorithms is implemented in GOLD [24,25]. As shown by the listed examples, numerous approaches have been introduced in the last years to handle the definition of binding geometry within the docking problem. Besides the accurate docking of the ligand, the preparation of the protein is of big importance. The assessing of the validity of a structure model is useful to prevent errors in the proximate docking procedure. The atomic resolution of a complex must be attended in the very beginning of the preparation of the protein. Especially the water molecules of a PDB file must be regarded cautiously, since this information is a subjective interpretation of a crystallographer from an observed electron density map. Temperature factors, as indicators for the reliability of the atomic coordinates and presence of the atom on defined positions, can help to interpret the presence of atoms in distinct positions. Sometimes remarks are found within the PDB files that indicate special problems in the interpretation process. Also the electron density maps can be inspected together with the PDB files [4]. This data may help to prevent failures in the basic material and increase the success of docking studies and structure-based design. The number of available crystal structures in high resolution has enormously increased in the last years due to the efforts of the human genome project and data collected in high-throughput (HT) crystallography laboratories. This becomes obvious regarding the growth of the Protein Data Bank that exceeds 27000 three-dimensional biological macromolecular structures in end of September 2004 [26]. The complexity of docking algorithms ranges many different procedures developed in the last two decades, which have their advantages and disadvantages. Docking programs as DOCK, FLEXX, and GOLD are frequently used platforms which address the complex issues of docking and scoring in different ways. One bottle-neck in the docking algorithms is the time required for docking a single ligand into the binding pocket of the protein. On modern hardware this procedure still takes a time range of several seconds up to minutes per molecule on a single node. Depending on the purpose of the docking experiment, this time restricting component deserves increasing importance. For large scale experiments, like HT docking, quite rapid algorithms must be used to ensure a reasonable time frame. This can be reached by the use of parallel computing, e.g. in Linux

clusters, but in the first instance, by a fast algorithm that produces accurate docking results. Parallelisation using grid computing has also been proposed [27].

VIRTUAL SCREENING TECHNIQUES AND DOCKING

Virtual screening is a widely accepted method in lead discovery because it is advantageous in the elimination of undesired molecules from compound libraries and the reduction of cost and time in drug discovery projects. In structure-based design ligands are modelled regarding the demands of the protein binding pocket. Docking may help in this case to investigate the active site in detail and detect uncovered binding pockets or interaction points. This approach is carried out by several de novo design tools, e.g. ligand construction and docking in GROWMOL [28]. These applications provide new scaffolds and can result in new synthesis strategies for the medicinal chemist. Docking of virtual combinatorial libraries may yield innovative ligands as well [29]. The aspect of privileged motive design [30] can be implemented by using the innovative software tool *ilib diverse* to generate focused virtual libraries for the desired target [31,32]. This program builds libraries of drug-like organic molecules for rational lead structure discovery. Compounds are generated by combining user-defined fragments according to state-of-the-art chemical knowledge. The technique of virtual screening is used to prioritise molecules for biological assays. It is cost and time efficient and has contributed important advances to lead discovery programs in many pharmaceutical companies. Often virtual screening techniques are used in combination with HT screening lead discovery tools [8]. The docking of a huge molecule database against a specific target may yield new candidates for further lead development. The combination of a prior pharmacophore search to restrict a large database to a smaller hitlist, and the subsequent docking of the hitlist in order to distinguish binders from non-binders, or even high affinity binders from low affinity binders, presents another combination of docking with a virtual screening technique. The selection of compounds for experimental validation can be justified by this approach. These combinations of different techniques seem to reduce the time scale to yield success in the identification of a novel lead structure. Furthermore, docking can help to correlate the experimentally determined biological activities, ligand poses, and predicted binding affinities by the docking program, to evaluate the scoring functions and to identify a good score of the target protein. A frequently used method is the re-docking of a complexed ligand in order to verify the validity of the docking and scoring algorithms.

SCORING METHODS

Scoring of docked poses is still regarded as one of the major challenges in the field of molecular docking. The purpose of the scoring procedure is the identification of the correct binding pose by its lowest energy value, and the ranking of protein-ligand complexes according to their binding affinities [3]. Though much effort has been invested in the development of accurate scoring functions, still no single scoring function can correctly rank every protein-

ligand complex today. Two approaches may help to overcome to this problem. First the identification of a good score for the target protein or second the application of a consensus score. The relative contribution of different protein-ligand interactions may vary between structural families, hence not a single scoring function can be determined to represent this. A consensus score is the combination of several scoring function. The common top rankers get a higher consensus rank than single outliers. False positives can be detected easier than with one singular scoring function. However, it is advisable to use a well-suited scoring function, if possible, because the consensus score always presents an average value and will never outperform the best scoring function. Another problem is the occurrence of induced-fit binding modes of ligands, as observed in several cases [7,33]. Compared to the flexibility of small molecules the tackling of the flexibility issue in macromolecules is still a remaining problem. Hence, the induced-fit protein movements must be regarded as a crucial task at the moment and in the near future to develop docking and scoring algorithms. Scoring functions can be divided in empirical scoring functions, scoring functions derived from force fields, and knowledge-based scoring functions. Scoring functions derived from force fields handle the ligand binding prediction with the use of potential energies (non-bonded interaction terms) and sometimes in combination with solvations and entropy contributions. Knowledge-based scoring functions are based on atom pair potentials derived from structural databases. Forces and potentials are collected from known protein-ligand complexes to get a score for their binding affinities (e.g. PMF). Empirical scoring functions derive from training sets of protein-ligand complexes with determined affinity data. One general aspect in the finding of an accurate empirical scoring function is the assumption that each occurrence of an individual interaction is considered as equivalent. In the case of hydrogen bonding it was shown that corresponding to the chemical environment the strength of hydrogen bonds can vary from 2 kJ/mol to 15 kJ/mol, in extreme cases up to 42 kJ/mol [34]. Hence, an average value may not always be representative, especially in multiple hydrogen bonded systems. An extremely important issue is the pH value. Changes in the pH value can cause the protonation or deprotonation of the interaction partners, ligand and protein, and result in different interaction patterns. This factor must also be included to some extent in scoring functions. The correct affinity ranking of ligands in a protein binding site and the optimisation of the interactions with the target are demands for a scoring function [6]. The description of non-covalent molecular associations in water is highly complex. A quantitative agreement between estimated affinity values and experimental results yet can not be guaranteed for every application case because the contribution of the individual interactions (hydrogen bonding, hydrophobicity, van der Waals forces, and electrostatic interactions) is faced differently in the different scoring systems [6]. A rigorous distinction of active and inactive compounds is often not possible with one single scoring function, hence the enrichment is a way to count the number of active compounds in the top ranked candidates. A database search can be considered as successful, if an enrichment is reached by a scoring function compared to the random screening procedure.

PARAMETERS INFLUENCING THE DOCKING RESULTS

A critical issue is the multi-factorial dependence of docking results. Much effort has been devoted to investigate, how and to what extent the outcome and accuracy of a docking approach are influenced by different parameters. Aside from docking algorithm and scoring function, binding site definition and the use of additional pharmacophore constraints are decisive [35]. Also the nature of the biological target, the properties of the active site, crystallographic resolution, as well as ligand flexibility were found to influence docking reliability [36]. A study carried out with CDOCKER, a molecular dynamics simulated-annealing-based algorithm, focuses on the role of a grid representation of protein-ligand interactions in comparison with a full force field approach [37]. It uses a test set of 41 diverse complexes, the ligand structures created by CORINA. The protein-ligand interactions for 50 replicas of each ligand are computed by grid approaches with equally spaced as well as discrete probe radii, or full force field, in combination with a final minimisation step. It is demonstrated that full force field runs give the highest success rate - success meaning a root mean square deviation (rmsd) between top ranked docking position and X-ray structure of less than 2 Å. Equally good results follow from the combination of grid methods with a full force field final minimisation step. Another discussed parameter is ligand flexibility. A lower number of rotatable bonds can clearly be connected with higher docking success, whereas molecular weight or binding affinity seem less differentiating for docking accuracy. Increasing the ligand sampling rate leads to improved accuracy, but not beyond a level of 50 replicas. Looking at the starting structures better outcomes are observed for the X-ray geometries as initial conformations than for the CORINA structures. The importance of interaction representation is also discussed in an article presenting a novel, knowledge-based method for the identification of the most favorable points between proteins and ligands for hydrophobic and hydrogen bond interactions [38], which has been incorporated in the SLIDE program. Thereby, knowledge on chemical groups providing the most energetically favorable contacts and on optimal geometries for hydrogen bonds around protein groups was included. Especially for apo structures of proteins, where the side chain conformation is different from that in complex with a ligand, promising results can be obtained. The knowledge-based method proves superior over grid-based sampling creating dockings close to the X-ray structure, discriminating between active and inactive structures in HT virtual screening, and improving scoring accuracy. Evaluation of the three docking tools GLIDE, GOLD, and ICM together with a detailed analysis of parameters influencing their performance is presented in an article by Perola, Walters, and Charifson [39]. The authors focus on the critical issue of test set selection to study docking accuracy. A set of 200 protein-ligand complexes from the PDB is chosen showing high pharmaceutical relevance, diversity, and drug-likeness of the ligands. Docking accuracy, i.e. top ranking poses within 2 Å of the corresponding crystal structure, is correlated with ligand flexibility (number of rotatable bonds), predominant nature of interactions between the ligand and the receptor (hydrophobic contacts versus

hydrogen bonds including metal interactions), and the degree of solvent exposure of the binding pocket. GLIDE is found to perform most consistently with respect to these influences. Its multistage systematic algorithm provides more extensive coverage in conformational space than both the genetic algorithm and the stochastic search implemented in GOLD and ICM. Besides, the latter two are far more binding site-dependent. Therefore, GLIDE provides the most accurate results on this wide and diverse set of systems: Correct identification of crystallographic pose within the defined rmsd cut-off in 61% of the cases in comparison to 48% for GOLD and 45% for ICM. Furthermore, the impact of energy minimisation and re-ranking of the top poses are investigated and suggested as beneficial post-processing steps to overcome some limitations of a given docking function.

DOCKING ALGORITHMS AND COMBINATIONS WITH SCORING FUNCTIONS

Nowadays we face a vast pool of different docking and scoring methodologies. Often a selection of an appropriate combination for a particular target becomes necessary since no truly generally applicable tool that offers robust and accurate solutions to a majority of various docking problems has been found so far. Advantages and disadvantages of these methods and the targets they are best suited for are critical and intensively investigated issues. Several studies comparing docking approaches in terms of accuracy in virtual screening, correct prediction of binding modes, and CPU time consumption have been published [36,40]. In an article by Schulz-Gasch and Stahl [35] the authors compare the new docking programs FRED (OpenEye Scientific Software) and GLIDE (Schrödinger Inc.) with results obtained from previous studies using FLEXX (Tripos Inc.). Two widely used algorithms are discussed: Multiconformer docking (FRED, GLIDE), which separates conformational search of a molecule from its placement in the binding site, and incremental construction process (FLEXX). The latter docks a small, rigid fragment, capable of specific directed interactions with the protein, at various favourable positions in the active site. These algorithms are combined with different scoring functions (empirical, knowledge-based, or based on molecular force fields). Functions not including directional terms are referred to as "soft" and emphasise lipophilic contacts and general sterical fit (i.e. PMF, PLP, DrugScore). "Hard" functions contain angular terms for defining hydrogen bonds (ChemScore, FLEXX scoring function). The study distinguishes between functions, called "objective", for estimation of receptor-ligand interaction energy and minimisation during docking, and "scoring" functions as such used for rank ordering of ligands. For seven protein targets docking studies with a subset of the Derwent World Drug Index [41] seeded with known inhibitors are investigated with different combinations of docking algorithms, objective functions, and scoring functions. Studies of the binding site characteristics and knowledge of the qualities of the used methodologies allow interpretation of the results. Although other influences, like binding site definition, are also crucial, Schulz-Gasch and Stahl arrive at a number of general guidelines to select the best combination for a particular virtual screening problem:

For lipophilic binding sites multiconformer docking with a soft objective and a harder scoring function is suggested. If polar groups also play a certain role incremental construction algorithms, which are always combined with hard objective functions, should be applied. Very polar sites with a dense network of directed interactions require incremental construction and hard scoring functions, or alternatively multiconformer docking in combination with a hard scoring function. Overall ChemScore is considered to be the most generally applicable and robust function. GLIDE performs well, where electrostatically mismatching poses have to be filtered out, because of its intermediate force field optimisation step. On the other hand, CPU time consumption is quite high, which makes it disadvantageous compared to the other two programs. Kontoyianni and co-workers [36] investigated five common docking programs (FLEXX, GOLD, DOCK, LigandFit, and GLIDE) against 14 protein families (69 targets) in order to compare the strength and limitations of these tools. Remarkable here is the large diversity of methods, upon which these programs are based (incremental construction in FLEXX, shape-based and genetic algorithm in DOCK and GOLD, respectively, systematic search in GLIDE, and Monte Carlo simulations in LigandFit). Furthermore, great diversity is achieved in protein selection providing a wide range of crystallographic resolutions, receptor families, ligand flexibility, ligand affinity, and active site topology including the presence of metal ions. Special attention is paid to the ionisation state of the ligands, active site definition (12 Å radius from the bound ligand or an amino acid central to the binding site), and identification of suitable parameter settings for the different docking algorithms. Docking accuracy is detected *via* visual inspection and rmsd calculation between the docked and experimentally determined ligand conformations. Thereby, GOLD clearly outperforms the other programs. However, none of them performs particularly well in ranking these conformers (in accordance with the crystal structures) at top positions. Here, best results, i.e. a rate of 67%, are achieved by GLIDE. Possible correlations between the obtained docking results and ligand flexibility, active site topology, and crystallographic resolution are investigated: Crystallographic resolution equal or less than 2.5 Å and ligands with a maximum of 10 rotatable bonds enable good results for all of the programs, whereby LigandFit seems to be more sensitive than the others. In terms of binding site polarity, DOCK is found preferable for rather hydrophilic sites. GLIDE on the other hand does not seem discriminatory at all in this context. Despite the highly favourable over-all performance of GLIDE, it is the slowest program (LigandFit being the fastest) and is therefore not advisable for usage in docking large databases without prior filtering. It also becomes clear that particular targets are best dealt with particular programs, e.g. GLIDE and LigandFit seem the best choice for COX2 and thymidine kinases, whereas GOLD should be preferred for neuraminidases and thermolysins. An evaluation study on eight docking programs (DOCK, FLEXX, FRED, GOLD, SLIDE, SURFLEX, and QXP) under comparable conditions has been published recently by Kellenberger *et al.* [42]. After protein and ligand pre-processing of 100 complexes from the PDB using the software tools Sybyl and Unity, the structures are imported into the mentioned docking programs. Docking results are discussed in the light of: (i)

docking accuracy, i.e. the ability to reproduce the X-ray pose of the ligand defining a 1 Å rmsd cut-off for success, (ii) ranking accuracy, i.e. discrimination between 10 known inhibitors of an enzyme (thymidine kinase) from 990 randomly chosen drug-like molecules, and (iii) speed. Accurate prediction of ligand binding positions is achieved with GLIDE, QXP, and GOLD for 61-63% of the cases and with FLEXX and SURFLEX for 48% and 54% of the cases, respectively. Best enrichment of active compounds results from SURFLEX (eight true hits among 50 compounds), and also GLIDE, GOLD, and FLEXX perform fairly well. The programs vary considerably in CPU time consumption: FRED, the fastest tool, attractive for ultra-HT docking, shows a mean docking time of 18 seconds. DOCK and FLEXX are also remarkably fast whereas GLIDE proves inferior with an average docking time of about four minutes. Additionally, the docking performance of these eight tools is discussed in terms of dependency on different functional and steric features of the ligand and the protein cavity, like the size and the shape of the binding site or ligand flexibility and polarity. Systematic docking errors occur in all programs and originate from insufficient conformational sampling of highly flexible molecules, open and shallow active sites, or unusual binding modes. Low ranking performance is observed in cases, where either no or very few lipophilic or electrostatic interactions occur between the protein and the ligand. Furthermore, mismatches between hydrophobic/ electrostatic potentials of the ligand and those of the binding site are challenging in this context. Highest docking and ranking accuracy is stated for GLIDE, GOLD, and SURFLEX.

PERFORMANCE OF SCORING FUNCTIONS

Prediction of interaction energies between ligands and their receptors remains a major challenge for structure-based inhibitor discovery. Though much effort [43-47] has been devoted to developing robust scoring schemes that can successfully rank the affinities of possible ligands to a binding site, an all-purpose scoring function has not been found. Trends to overcome this problem emerged: (i) On the one hand the design of target/ligand-directed scoring systems, for example for metalloproteinases [48] or for protein-carbohydrate complexes [49], can provide valuable scoring tools for special applications. Another approach to overcome the inability of empirical, force field, and knowledge-based scoring functions to deal with metal ions in the active site is a quantum mechanics method presented by Raha and Merz [50]. In this study quantum mechanics, which has until recently only been used to study smaller systems because of the enormous computational cost associated with it, is applied on a set of 18 carbonic anhydrase and five carboxy peptidase inhibitors. The method performs well across these two different classes of metalloenzymes recognising significant and variable charge-transfer interactions between the metal and the ligand. (ii) On the other hand, consensus scoring is often the concept of choice to improve hit rates, since it makes use of the merits of different scoring functions and therefore makes up for the low over-all predictive abilities of the single functions. Several studies find that reliability and interpretability of results can be improved by combining results from different

functions into a consensus score [48,51,52]. Thereby various possibilities for consensus score calculation exist: Rank-by-vote gives each compound a vote from a particular scoring function, if it appears in the top n% of the database according to that scoring function. The compounds are then ranked after the number of votes they received. Rank-by-vote often lacks sufficient significance especially, when dealing with few scoring functions and large data sets. Rank-by-number aims to average the score of several different scoring functions and is valid only in cases, where the different scoring functions calculate measurements on a comparable scale. A third approach, rank-by-rank, averages the ranks given to a single compound for each scoring function. A study by Klon, Glick, and Davis describes the application of consensus scoring subsequent to HT docking and prior to a naive Bayes categorisation on the target of tyrosine phosphatase 1B and protein kinase B/Akt in order to rescue poor docking results [53]. Much emphasis has been laid on comparing the performance of different scoring functions in accurate library ranking [36,40,43,54]. Wang *et al.* [55] present a study on 11 popular scoring functions, which have been tested on 100 protein-ligand complexes to evaluate their ability to reproduce experimentally determined structures and binding affinities. They include four scoring functions implemented in the LigandFit module in Cerius² (LigScore, PLP, PMF, and LUDI), four scoring functions implemented in the CScore module in Sybyl (F-Score, G-Score, D-Score, and ChemScore), the scoring function implemented in the AutoDOCK program, and two stand-alone scoring functions (DrugScore and X-Score). They can be roughly grouped into three categories: (i) force field-based methods, i.e. AutoDOCK, G-Score, and D-Score, (ii) empirical scoring functions, i.e. LigScore, PLP, LUDI, F-Score, ChemScore and X-Score, and (iii) knowledge-based potentials, i.e. PMF and DrugScore. The most decisive challenge in this study is that these scoring functions should be tested and compared independently of a particular docking program. Therefore, conformational sampling and scoring are separated into two consecutive steps. An ensemble of docked conformations for each ligand molecule is generated with the AutoDOCK program, focusing on exhaustive coverage of conformational space. In a first step the scoring functions are evaluated in terms of docking accuracy. Judging from the rmsd between best-scored docked conformation and the X-ray structure PLP, F-Score, LigScore, DrugScore, LUDI, and X-Score outperform the other scoring functions. When considering not only the best but rather the three best scored yet non-duplicate conformations in the ensemble, the success rates can be further increased. Moreover, a consensus scoring approach to identify the correct bound conformation of a given ligand is carried out. All possible double and triple combinations of these six best scoring functions are tested. The success rates escalate, for triple scoring schemes even higher than for double scoring schemes. The next aim of the study is to find out, whether the different scoring functions have sufficient consideration of hydrogen bonding as well as of hydrophobic interactions, which are much more difficult to characterise. According to the chemical nature of their protein-ligand interactions the 100 complexes are classified as "hydrophilic", "hydrophobic", and "mixed" type. Only PLP and F-Score are able to maintain their success rates across all three subsets. Typically a well-balanced consideration of polar and non-polar, enthalpic and entropic

factors can be expected from empirical scoring functions. In force field equitation hydrophobic effects cannot adequately be formalised. Another critical issue is the prediction of binding affinity. Therefore the correlation between the scores from the 11 scoring functions and the experimentally measured binding affinities of the 100 complexes are examined. Best agreement can be seen for X-Score, followed by PLP, DrugScore, and G-Score. Other scoring functions give rather poor results. Finally, all scoring functions are inspected by their abilities to construct a descriptive, funnel-shaped energy surface for protein-ligand complexation. All docked conformations of a ligand present spots on the protein-ligand complexation energy surface, whose shape is defined by the scoring function. Despite the multidimensional nature of the energy surface the score is clearly associated with the rmsd value. X-Score and DrugScore give the best score-rmsd correlation throughout the entire conformational ensemble. The other scoring functions perform worse in constructing a descriptive, funnel-shaped energy surface. An assessment of nine scoring functions (CHARMm potential, DrugScore, the AutoDOCK scoring function, PMF, GOLD, ChemScore, and the three scoring functions used in DOCK) implemented in widely applied docking programs has been carried out by Ferrara *et al.* [56]. Protein-ligand interactions are studied for a database of 189 complexes. In a first step, nearly 100 decoys are constructed for each complex, whose deviations from the crystal structure present a continuous spectrum in the neighbourhood of the binding site. To investigate the applicability of scoring functions in binding site definition decoys on the protein surface are generated. The nine scoring functions are evaluated for their ability to discriminate near-native ligand poses from misdocked conformations. They perform well, especially CHARMm, DOCK-energy, DrugScore, ChemScore, and AutoDOCK, which yield recognition rates of approximately 80%. The analyses reveal that steric complementarity between the ligand and the receptor is more important than electrostatics to identify near-native poses. In this context, the impact of assigning protonation states and the role of solvation models are discussed, as well as the critical issue of using a rigid receptor representation in docking. The second question in this study addresses the problem of ligand ranking accuracy: The scoring functions are compared with respect to their ability to correctly rank known ligands by estimating binding affinities, which is of primary interest in lead optimisation. The more broadly parameterised ChemScore achieves the highest correlation with an R^2 value of 0.51 for the 189 test set complexes and an R^2 value of 0.43 for a subset of 116 complexes which was deprived of structures used to calibrate this scoring function. Other scoring functions prove suitable only for the serine protease and the metalloprotease families but provide unsatisfactory results for the remaining data sets. The authors stress that binding affinity prediction still remains a major challenge in the docking and scoring process. Constantly new or improved scoring functions are developed to account for the various effects and interaction possibilities that occur upon ligand binding. Novel scoring function generation built with several statistical methods (principal component analysis, neural networks, genetic algorithms) is described by Giordanetto and colleagues [57]. The functions combine descriptors generated by the QXP docking program as well

as new descriptors. The latter are based on solvent accessible surface areas and account for conformational entropy changes and desolvation effects of both ligand and receptor upon binding, which are crucial but nevertheless often neglected aspects of the binding process. A training set of 100 and a test set of 24 protein-ligand complexes are used and confirm accurate affinity prediction and therefore applicability of the new functions as filters to guide compound selection after docking.

DOCKING USING HOMOLGY MODELS

Despite the constantly increasing number of proteins, whose three-dimensional structures have been determined, we still face the problem nowadays that many proteins, like membrane receptors, are not yet crystallisable. Among them is the enormously important group of GPCRs (G-protein coupled receptors). When no experimental structure is available for a potential drug target, it is often possible to generate an approximate three-dimensional structure based on a template protein of a similar sequence and a known structure. Several studies on the use of homology models for docking and scoring have been published: To determine the dependence of docking screens on the conformation of the binding site and to evaluate the information loss that occurs as the active site conformation becomes less defined, is the aim of a study by Susan McGovern and co-workers [58]. Therefore, a small molecule database is docked against the holo (ligand-bound), apo (ligand-free), and homology-modelled structures of 10 different enzyme binding sites using the Northwestern version of DOCK3.5. The database contains at least 35 ligands for each of the 10 systems. For better comparability of the influence of the protein conformation the holo, apo, and modelled structures for each system are used. The structures are analysed for their ability to enrich the known ligands in that system over random selection. This desired enrichment can be observed in all systems for each of the three structures. However, it does not always reach the level that justifies the effort of a docking approach. When demanding a 20-fold enrichment, holo conformation meets this requirement in eight of 10 systems, apo in only two, and modelled structures in three. The best overall discrimination between known ligands and decoy molecules comes from holo structures (seven systems), followed by apo (two systems), and homology models (one system, which is the purine nucleoside phosphorylase). The study clearly shows that the performance of docking calculations is dependent of the representation of the protein. Holo structures are most likely to yield the best results, but important exceptions also emerge from this study. The question of protein flexibility seems particularly important in this context. The creation of a fast algorithm, which also takes hinge-bending motions of protein receptor domains into account, has been presented. Experimental data show that docking ligands into a rigid receptor homology model yields poor results, whereas computational mobilisation of the receptor can clearly improve accuracy [59]. Bissantz *et al.* [60] present homology models based on bovine rhodopsin for the important family of GPCRs and test their applicability in virtual screening of databases. "Antagonist-bound" models of a dopamine, a muscarinic, and a vasopressin receptor are constructed. GOLD, FLEXX, and

DOCK in combination with seven frequently used scoring functions are applied and succeed in discriminating known inhibitors from randomly chosen molecules of a database. In this study rhodopsin-based homology models prove suitable for screening databases for antagonists, but not sufficiently accurate for agonists. Another rhodopsin-derived homology model for the α_{1A} adrenergic GPCR has been published recently [61]. It offers insights into the binding site key residues by docking an endogenous agonist (norepinephrine) as well as an antagonist (WB-4101) with the BioDock program. Thereby, mutagenesis data can be confirmed and suggestions for future ligand design can be given. Recently, a retrospective study on factor VIIa inhibitors and CDK2-antagonists has been published, which seeks to provide some guidelines for the successful usage of homology models in docking [62]. Homology models of these two target proteins, of which reference crystal structures are available, are built with the program MOE based on template serine proteases and on different kinases. Sequence identity of these templates near the binding site ranges from approximately 40% to 80%. Large data sets seeded with known active ligands are docked using the software UCSF DOCK, whereby conformational analysis (10 different conformations per stereoisomer) is outsourced. The DOCK energy scoring function is applied for ranking. Analysis of the enrichment, i.e. the ratio of active compounds in the top ranked 100 hits to active compounds in 100 randomly picked structures, allows systematic evaluation of the performance of molecular docking applied to homology models in comparison to crystal structures. Although an enrichment better than random can be observed in all experiments, tremendous improvement only becomes obvious in cases where the sequence identity exceeds 50%. This cut-off can be correlated with an rmsd of less than 2 Å between the binding site residues of the homology models and the target crystal structures. The authors conclude that especially homology models generated from templates with sequence identity greater than 50% are highly useful for docking applications, since they provide enrichments similar to those from the target crystal structures (approximately five times better than random).

APPLICATIONS OF DOCKING

Docking has become an essential tool in structure-based ligand design. It is widely applied and meets very heterogeneous demands. Of course a major task remains the identification of new active compounds for a particular target protein. Docking proves a reliable and fast filter in HT virtual screening [63,64], thereby providing a pool of ideas for novel lead structures, and can show several success stories [65-68]. Many research groups apply docking and scoring methods, when synthesis and experimental testing have already been performed, in order to correlate the scores with the biological activity [69,70]. Often the explanation or affirmation of a binding mode for a (structurally new) class of compounds is desired [71-75]. Many docking approaches aim to find out, whether a specific docking methodology and/or which scoring functions are best suited for a particular target system. Therefore it is tested, if correct binding orientations can be reproduced [75]. Furthermore, databases seeded with active molecules are screened to investigate, if

and which docking/scoring methods successfully discriminate between active and inactive compounds [64]. Some critical targets demand the creation of protein models or the customisation of docking processes. In such cases docking studies with known ligands are essential to test the new systems [60,76].

LIGAND AND PROTEIN FLEXIBILITY, OVERESTIMATION OF MOLECULAR SIZE, AND TAUTOMERISM

Several interesting topics related to flexibility of proteins, to overestimation of docking scores due to high molecular size, and to issues with respect to tautomeric equilibria were recently discussed in the scientific community and will be mentioned briefly in the following part. First, most docking methodologies treat the ligand as flexible, whereas the protein is usually kept rigid. Although it is well established that the ligand binding event is dynamic, this simplification is often necessary in order to prevent not practicable, too time-intensive calculations. An examination of the determinants of docking accuracy, which focuses on the influence of ligand and protein flexibility, has been presented by Erickson *et al.* [77]. As criterion for successful protein-ligand binding mode prediction and sufficiently accurate results an rmsd of less than 2 Å between the docked and the X-ray position of the ligand is defined. In a first approach a test set of 41 ligand-protein complexes is used to study the dependence of docking accuracy on ligand properties, rigid versus flexible docking, force field parameters, and starting conformation, applying the four docking algorithms DOCK, FLEXX, GOLD, and CDOCKER. The usage of a rigid ligand provides accurate docking results for the X-ray conformations as starting structures. Starting from CORINA generated ligand geometries however the success rate drops by almost half, which emphasises the importance of the ability to treat ligands flexibly. For flexible ligand docking experiments only small differences between results starting from CORINA and X-ray structures respectively can be seen. A marked difference in flexible ligand docking accuracy between the four programs becomes obvious and demonstrates the decisiveness of docking methodology selection. Two trends emerge from both the rigid and the flexible approach: Average success rates increase, when a softer non-bonded potential is used. On the other hand, force field parameter variation, and therefore the variation of ligand atomic charge and dielectric constant, only has a very small effect. Furthermore, the molecular properties of the ligands, i.e. the size (molecular weight), flexibility (number of rotatable bonds), and polarity (number of polar atoms) are studied. For DOCK, FLEXX, and GOLD docking accuracy decreases substantially for ligands with eight or more rotatable bonds, whereas CDOCKER produces correct results for most of those ligands. To examine the effects of protein flexibility on docking accuracy a test set of protein structures bound to several inhibitors and also the unbound form is selected. It consists of three proteases, whose degree of conformational change upon ligand binding varies significantly. The ligands are docked back to the structure they were taken from and to the apo form of the enzyme. A third docking experiment uses an "average" structure.

Thereby, a structure is selected whose binding site coordinates are the closest to the average position of all the structures. The results clearly show, that it is important, if and to what extent the protein changes in response to ligand binding and which structure is used. A drop in docking accuracy occurs when the ligand is docked to a structure other than its own. This drop off is correlated with the degree of protein movement in the active site, i.e. the higher the amount of movement upon ligand binding the lower is the success rate. If the rmsd between the protein used for docking and the one from the initial complex is greater than 1.5 Å, almost 90% of the docking accuracy is lost. Therefore, apo structures perform even worse than the average structures. Erickson *et al.* point out, what implications follow from these results in structure-based lead design, virtual screening, and especially for the utilisation of homology models constructed from templates with low sequence similarity to the target. An interesting approach to account for receptor rearrangement upon ligand binding (induced-fit) has been presented by Cavasotto and Abagyan [78]. Therein the novel algorithm IFREDA is used to allow protein flexibility in virtual screening: A discrete set of receptor conformations is created, which are then used for flexible ligand-rigid receptor docking and scoring followed by a merging and shrinking process. Aliquot sampling of receptor conformational space is guaranteed by taking both side chain rearrangements and essential backbone movements into consideration. The results of cross-docking experiments with 33 crystal structures of four protein kinase sub-families suggest that this methodology is especially advantageous in cases, where few or no experimental data of complexes are available, while some binders are known. Second, the overestimating of large molecules by actual scoring methods and the usage of correction terms have been considered in the work of Verdonk *et al.* and Pan *et al.*, respectively [79, 80]. In the study of Verdonk *et al.*, GOLD is validated against four targets (neuraminidase, CDK2, PTP1b, and estrogen receptor). According to the authors, one critical factor in the validation of scoring functions is the design of a focused rather than a random library. To handle the issue of overestimation of large compounds the validation is performed against a focused library containing compounds with similar one-dimensional properties to those of the actives. This approach avoids biases like enrichment based on a difference in molecular size between actives and random compounds. Molecules with a high molecular weight implicate a larger surface area, thus resulting in a larger number of unspecific interactions to the protein. The scoring functions GoldScore, DrugScore, and a modified version of ChemScore are extended by a pharmacophore thereby including structural knowledge about the way ligands bind to their corresponding proteins. The pharmacophore restraints increase the enrichment rates against CDK2 and neuraminidase. This approach shows that actives need to be screened against focused libraries to circumvent the generation of artificial enrichments [79]. A score normalisation technique is presented in the work of Pan *et al.* The goal is to suppress the tendency for the selection of large molecules in virtual database screening according to the van der Waals attractive contribution. Molecules with larger size may be ranked high, but not necessarily be as structurally complementary to the binding site as smaller compounds. The normalisation strategy based on the number

of heavy atoms, N, can control the bias of the distribution of docked compounds' molecular weight towards a higher level than the full database distribution. The molecular size bias can be overcome by scaling the scores according to the size of the compounds. The score can be divided by N^x , meaning N is the number of heavy atoms in a compound and x is a constant. In this method no other one-dimensional property than molecular size is considered. Further, x can be changed for every set of actives, because the molecular size distribution of actives varies between test sets [80]. Third, problems regarding the correct protonation state and tautomerism have been highlighted in the review of Pospisil *et al.* [81]. Molecular modelling application, e.g. ligand docking, often disregard the issues related to tautomeric equilibria of molecules. A change in shape, functional groups, surface, and hydrogen bonds can be observed between tautomers of a single compound. Structural databases rarely included tautomeric forms of compounds. The complementary pattern of a ligand to the protein's binding site, which is of special interest in docking simulations, may be met by one tautomer, but disregarded by another. Hence, it is recommendable to cover the dimension of multiple tautomers as completely as possible in virtual screening methods. Two reasons favor the use of tautomers in virtual screening: The increase of the chemical space in a database on the one hand, and consideration of the binding of a compound in its tautomeric state on the other hand. Tautomeric compounds can be classified as hits in virtual database screening, but would not have been detected in a database disregarding tautomeric forms. Another topic in ligand docking is the accurate description of the protonation state of ligands and amino acids in the binding pocket. The correct protonation state of a ligand has significant impact. It depends on the pH value of the system and should be checked precisely for the amino acids of the binding pocket as well as for the docked molecules.

CONCLUSION

In this review we focused on recent progressions in molecular docking and scoring by the description of several applications and case studies. Successful usage of these techniques has still severe limitations. The aim of a docking procedure is often the discovery of new lead candidates. The identification of an overall reliable and robust scoring function seems to be one of the main challenges to be addressed in the near future. Yet, the combination of scoring functions in a consensus score presents a well-established, however, questionable approach to overcome this problem. Nevertheless, novel algorithms will arise to find new solutions to the docking problem and overcome the limitations of recently developed scoring functions. Especially the issue of protein flexibility and induced-fit motions of the protein will gain in importance over the coming years in the design and discovery of novel lead candidates by means of protein-ligand docking and scoring.

ABBREVIATIONS

- ADME = Absorption, distribution, metabolism, excretion
GPCR = G-protein coupled receptor

HT = High-throughput
 PDB = Protein Data Bank
 rmsd = Root mean square deviation

REFERENCES

- [1] Kurogi, Y.; Güner, O. F. *Curr. Med. Chem.*, **2001**, *8*, 1035-1055.
 [2] Langer, T.; Krovat, E. M. *Curr. Opin. Drug Discov. Devel.*, **2003**, *6*, 370-376.
 [3] Muegge, I.; Rarey, M. *Rev. Comput. Chem.*, **2001**, *17*, 1-60.
 [4] Davis, A. M.; Teague, S. J.; Kleywegt, G. J. *Angew. Chem. Int. Ed.*, **2003**, *42*, 2718-2736.
 [5] Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins*, **2002**, *47*, 409-443.
 [6] Tame, J. R. H. *J. Comput. Aided Mol. Des.*, **1999**, *13*, 99-108.
 [7] Toledo-Sherman, L. M.; Chen, D. *Curr. Opin. Drug Discov. Devel.*, **2002**, *5*, 414-421.
 [8] Waszkowycz, B. *Curr. Opin. Drug Discov. Devel.*, **2002**, *5*, 407-13.
 [9] Schneidman-Duhovny, D.; Nussinov, R.; Wolfson, H. *J. Curr. Med. Chem.*, **2004**, *11*, 91-107.
 [10] Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langrince, R.; Ferrin, T. E. *J. Mol. Biol.*, **1982**, *161*, 269-288.
 [11] Fischer, D.; Norel, R.; Wolfson, H.; Nussinov, R. *Proteins*, **1993**, *16*, 278-292.
 [12] Fischer, D.; Lin, S. L.; Wolfson, H. L.; Nussinov, R. *J. Mol. Biol.*, **1995**, *248*, 459-477.
 [13] Boehm, H. J. *J. Comput. Aided Mol. Des.*, **1992**, *6*, 593-606.
 [14] Boehm, H. J. *J. Comput. Aided Mol. Des.*, **1992**, *6*, 61-78.
 [15] Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P.; Miller, M. D. *J. Comput. Aided Mol. Des.*, **1994**, *8*, 565-482.
 [16] Lorber, D. M.; Shoichet, B. K. *Protein Science*, **1998**, *7*, 938-950.
 [17] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.*, **1996**, *261*, 470-489.
 [18] Rarey, M.; Kramer, B.; Lengauer, T. *J. Comput. Aided Mol. Des.*, **1997**, *11*, 369-384.
 [19] Goodsell, D. S.; Olson, A. J. *Proteins*, **1990**, *8*, 195-202.
 [20] Given, J. A.; Gilson, M. K. *Proteins*, **1998**, *33*, 475-495.
 [21] Trosset, J. Y.; Scheraga, H. A. *J. Comput. Chem.*, **1999**, *20*, 412-427.
 [22] Trosset, J. Y.; Scheraga, H. A. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 8011-8015.
 [23] Trosset, J. Y.; Scheraga, H. A. *J. Comput. Chem.*, **1999**, *20*, 244-252.
 [24] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.*, **1997**, *267*, 727-748.
 [25] Jones, G.; Willett, P.; Glen, R. C. *J. Comput. Aided Mol. Des.*, **1995**, *9*, 532-549.
 [26] Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. *Nucleic Acids Res.*, **2000**, *28*, 235-242.
 [27] Richards, W. G. *Nat. Rev. Drug Discov.*, **2002**, *1*, 551-555.
 [28] Bohacek, R.; McMartin, C.; Glunz, P.; Rich, D. H. *IMA Volumes in Mathematics and Its Applications*, **1999**, *108*(Rational Drug Design), 103-114.
 [29] Krovat, E. M.; Langer, T. *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 1123-1129.
 [30] Bleicher, K. H.; Böhm, H.-J.; Müller, K.; Alaine, A. I. *Nat. Rev. Drug Discov.*, **2003**, *2*, 369-378.
 [31] Wolber, G.; Langer, T. **2004**, <http://www.inteligand.com>.
 [32] Wolber, G.; Langer, T. Rational Approaches to Drug Design, Proceedings of the European Symposium on Quantitative Structure-Activity Relationships, **2000**, 390-399.
 [33] Teague, S. J. *Nat. Rev. Drug Discov.*, **2003**, *2*, 527-541.
 [34] Boehm, H. J.; Klebe, G.; Kubinyi, H.: *Wirkstoffdesign. Protein-Ligand Wechselwirkungen* Spektrum Akademischer Verlag: Heidelberg, **1996**.
 [35] Schulz-Gasch, T.; Stahl, M. *J. Mol. Model.*, **2003**, *9*, 47-57.
 [36] Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J. Med. Chem.*, **2004**, *47*, 558-565.
 [37] Wu, G.; Robertson, D. H.; Brooks, C. L. 3rd; Vieth, M. *J. Comput. Chem.*, **2003**, *24*, 1549-1562.
 [38] Zawadzki, M. I.; Sanschagrin, P. C.; Korde, R. S.; Kuhn, L. A. *J. Comput. Aided Mol. Des.*, **2002**, *16*, 883-902.
 [39] Perola, E.; Walters, W. P.; Charifson, P. S. *Proteins*, **2004**, *56*, 235-249.
 [40] Turchi, I. In *Middle Atlantic Regional Meeting of the American Chemical Society*; American Chemical Society, Washington, D. C: Princeton, NJ, United States, **2003**.
 [41] WDI *Derwent World Drug Index*, <http://www.derwent.com>.
 [42] Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. *Proteins*, **2004**, *57*, 225-242.
 [43] Chen, R.; Li, L.; Weng, Z. *Proteins*, **2003**, *52*, 80-87.
 [44] Smith, R.; Hubbard, R. E.; Gschwend, D. A.; Leach, A. R.; Good, A. C. *J. Mol. Graph. Model.*, **2003**, *22*, 41-53.
 [45] Feher, M.; Deretey, E.; Roy, S. *J. Chem. Inf. Comput. Sci.*, **2003**, *43*, 1328-1334.
 [46] Wang, R.; Lai, L.; Wang, S. *J. Comput. Aided Mol. Des.*, **2002**, *16*, 11-26.
 [47] McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. *Biopolymers*, **2003**, *68*, 76-90.
 [48] Hu, X.; Shelver, W. H.; Pan, F. In *225th ACS National Meeting*; American Chemical Society, Washington, D. C: New Orleans, LA, United States, **2003**.
 [49] Laederach, A.; Reilly, P. J. *J. Comput. Chem.*, **2003**, *24*, 1748-1757.
 [50] Raha, K.; Merz, K.M. Jr. *J. Am. Chem. Soc.*, **2004**, *126*, 1020-1021.
 [51] Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. *J. Mol. Graph. Model.*, **2002**, *20*, 281-295.
 [52] Nicodeme, P.; Didier, R. *Proteins*, **2002**, *47*, 521-533.
 [53] Klön, A. E.; Glick, M.; Davies, J. W. *J. Med. Chem.*, **2004**, *47*, 4356-4359.
 [54] Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins*, **2003**, *52*, 609-623.
 [55] Wang, R.; Lu, Y.; Wang, S. *J. Med. Chem.*, **2003**, *46*, 2287-2303.
 [56] Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. *J. Med. Chem.*, **2004**, *47*, 3032-3047.
 [57] Giordanetto, F.; Cotesta, S.; Catana, C.; Trosset, J.-Y.; Vulpetti, A.; Stouten, P. F. W.; Kroemer, R. T. *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 882-893.
 [58] McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.*, **2003**, *46*, 2895-2907.
 [59] Gilson, M. K.; Kairys, V. In *ACS National Meeting*; American Chemical Society, Washington, D. C: New Orleans, LA, United States, **2003**.
 [60] Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. *Proteins*, **2002**, *50*, 5-25.
 [61] Pedretti, A.; Elena, S. M.; Villa, L.; Vistoli, G. *Biochem. Biophys. Res. Commun.*, **2004**, *319*, 493-500.
 [62] Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lancot, J. K.; Putta, S.; Stanton, R.; Grootenhuys, P. D. *J. Med. Chem.*, **2004**, *47*, 764-767.
 [63] Jenkins, J. L.; Kao, R. Y. T.; Shapiro, R. *Proteins*, **2003**, *51*, 81-93.
 [64] Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. *J. Mol. Graph. Model.*, **2003**, *21*, 289-307.
 [65] Park, H. J.; Yoo, J.; Kim, Y. K.; Han, J. W.; Lee, H. W. In *224th ACS National Meeting*; American Chemical Society, Washington, D. C: Boston, MA, United States, **2002**.
 [66] Grueneberg, S.; Stubbs, M. T.; Klebe, G. *J. Med. Chem.*, **2002**, *45*, 3588-3602.
 [67] Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. *J. Med. Chem.*, **2002**, *45*, 2213-2221.
 [68] Rastelli, G.; Ferrari, A. M.; Constantino, L.; Gamberini, M. C. *Bioorg. Med. Chem.*, **2002**, *10*, 1437-1450.
 [69] Batra, S.; Sabnis, Y. A.; Yogesh, A.; Rosenthal, P. J.; Avery, M. A. *Bioorg. Med. Chem.*, **2003**, *11*, 2293-2299.
 [70] Vicker, N.; Ho, Y.; Robinson, J.; Woo, L. L. W.; Purohit, A.; Reed, M. J.; Potter, B. V. L. *Bioorg. Med. Chem. Lett.*, **2003**, *15*, 863-865.
 [71] Cozzini, P.; Dottorini, T. *Eur. J. Med. Chem.*, **2004**, *39*, 601-609.
 [72] Hou, T.; Zhang, W.; Xu, X. *J. Comput. Aided Mol. Des.*, **2002**, *16*, 27-41.
 [73] Shiao, Y. S.; Huang, P. T.; Liou, H. H.; Liaw, Y. C.; Shiao, Y. Y.; Lou, K. L. *Chem. Res. Toxicol.*, **2003**, *16*, 1217-1225.
 [74] Chen, H. F.; Li, Q.; Yao, X.; B.T.Fan; Yuan, S. G.; Panaye, A.; Doucet, J. P. *QSAR Comb. Sci.*, **2003**, *22*, 604-613.
 [75] Illapakurthy, A. C.; Ashok, C.; Yogesh, Y. A.; Avery, B. A.; Avery, M. A.; Wyandt, C. M. *J. Pharm. Sci.*, **2003**, *92*, 649-655.

- [76] Cavasotto, C. N.; Orry, A. J. W.; Abagyan, R. A. *Proteins*, **2003**, *51*, 423-433.
- [77] Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. *J. Med. Chem.*, **2004**, *47*, 45-55.
- [78] Cavasotto, C. N.; Abagyan, R. A. *J. Mol. Biol.*, **2004**, *337*, 209-225.
- [79] Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 793-806.
- [80] Pan, Y.; Huang, N.; Cho, S.; MacKerell, A. D. J. *J. Chem. Inf. Comput. Sci.*, **2003**, *43*, 267-272.
- [81] Pospisil, P.; Ballmer, P.; Scapozza, L.; Folkers, G. *J. Recept. Signal Transduct. Res.*, **2003**, *23*, 361-371.

Received: October 6, 2004

Accepted: November 10, 2004