Combinatorial Chemistry & High Throughput Screening, 1999, *2*, 91-104

Computational Ligand Design

J. Apostolakis† and A. Caflisch*

University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Abstract: A variety of computational tools that are used to assist drug design are reviewed. Particular emphasis is given to the limitations and merits of different methodologies. Recently, a number of general methods have been proposed for clustering compounds in classes of drug-like and non-drug-like molecules. The usefulness of this classification for drug design is discussed. The estimation of (relative) binding affinities is from a theoretical point of view the most challenging part of ligand design. We review three methods for the estimation of binding energies. Firstly, quantitative structure-activity relationships (QSAR) are presented. These have gained significantly from recent developments of experimental techniques for combinatorial synthesis and high-throughput screening as well as the use of powerful computational procedures like genetic algorithms and neural networks for the derivation of models. Secondly, empirical energy functions are shown to lead to more general models than standard QSAR, since they are fitted to a variety of complexes. They have been used recently with considerable success. Thirdly, we briefly outline free energy calculations based on molecular dynamics simulations, the method with the most sound theoretical foundation. Recent developments are reestablishing the interest in this approach. In the last part of this review structure-based ligand design programs are described. These are closely related to docking, with the difference that in design, unlike in most docking procedures, ligands are built on a fragment-by-fragment basis. Finally, a short description of our approach to computational combinatorial ligand design is given.

Introduction

The object of drug design is to suggest easily synthesizable molecules that act against the cause or merely the symptoms of a particular disease or disorder. Drug molecules usually bind at a target macromolecule of the host or the disease-causing agent and alter the macromolecules' function. Simple as this principle may sound it is complicated by the fact that molecular recognition and binding are not yet understood in terms of simple models that could lead the search for new drugs in an unerring way. Furthermore, there are a number of other issues that have to be addressed in the design of therapeutic compounds. The activity of a drug is seldomly absolutely specific. More often than not undesired side effects on the host are observed. The seriousness of these side effects reflects the toxicity of the corresponding compound. The determination of toxicity is a lengthy and costly process. It is therefore desirable to filter out problematic substances early on. Furthermore, as molecules foreign to the host, drugs are often degraded or simply washed out before they have a chance to bind to their target. Often the targets are even protected by physical barriers, e.g. the cell membrane. The pharmacokinetic properties together with the bioavailability of the drug determine to which extent drug-target interaction will take place. Finally, an ever increasing concern is the development of drug resistance through mutation of the target molecules. It is therefore of particular interest to find drugs that bind to evolutionary stable binding sites. The quality of a particular drug depends on the mentioned of properties in addition to its direct effectiveness. This very short description of the aim of drug design barely does justice to the work involved. Starting from the identification of a suitable molecular target and its structural characterization, it is not unusual for drug development to take a decade or more until a compound reaches the market. The problems that have to be addressed during development make the field extremely interdisciplinary and although it is natural for every specific branch to consider its own contributions of utmost importance it becomes ever clearer that drug design cannot be tackled within a single research field.

It is the object of computer aided drug design (CADD) to assist chemists and pharmacologists with this daunting task, making use of the immense

*Address correspondence to this author at University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; Phone: (41 1) 635 55 21; FAX: (41 1) 635 57 12; email: caflisch@bioc.unizh.ch

†GMD/SCAI Schloss Birlinghaven, D-53754, Sankt Augustin, Germany

1386-2073/99 $19.00+.00

computational and data handling power available through modern computers. The algorithms used are based on statistical or physical models. With statistical methods it is possible to study the correlations between different quantities, while physical models are useful for the understanding of the principles behind these correlations. This is important for predictions, since they usually correspond to some kind of interpolation or extrapolation. The quality of the physical model determines the reliability of the predictions. CADD has been a promising method for a long time. Recently, however, this promise has finally started to be fulfilled with the extensive use and successful application of computational methods in many pharmaceutical and biotechnology companies [1].

In this review we outline the methods most commonly used by molecular modelers and medicinal chemists in non-profit institutions and pharmaceutical companies. We first present the computational tools for predicting ligands for macromolecular targets whose three-dimensional structure is not known. A distinction is made between methods relying on the alignment of the molecules in the data set and approaches which overcome the alignment problem. Particular emphasis is given to the relations between the different approaches and methods to evaluate (relative) binding energies. The recently published structure-based methods are then reviewed.

General Properties of Known Drugs

As mentioned above, the biochemical function with respect to a specific target is only one of the necessary qualities of drugs. Properties like synthetic accessibility, stability, oral availability, good pharmacokinetic behavior, lack of toxicity and minor addictive potential are of supreme importance. Many of these depend on complex biochemical and physical phenomena that are not well, if at all, understood. For the determination of toxic side effects in humans, clinical trials are necessary and sometimes even not sufficient. Cumulative side effects may take years to make themselves noticed. One interesting approach to address all of these issues at once in the drug discovery process is to study the general features of known drugs. After all, drug databases contain substances that have been tested for exactly these properties. Therefore, the working hypothesis is that it is possible to derive relevant information by abstraction. Bemis and Murcko have performed a general analysis of the shapes of molecules found in a commercially available database of known drugs [2]. With the help of a simple graph theoretical approach that takes into account only the two-dimensional structure it is possible to decompose molecules into rings, side chains and linkers. The last two structural entities correspond to non-cyclical chains. Linkers join rings, while side chains end at atoms with only one neighbor. Linkers and rings together form the framework of the molecule. These definitions, which do not take into account the chemical nature of atoms and bonds show that the number of distinct frameworks (1179) in the database is roughly one fourth the number of compounds (5120). 783 frameworks are unique, i.e. they are found in only one compound. More interestingly, the shapes of half of the compounds in the database are described by only 32 frameworks. The most common ring structure is the six-membered ring. 306 acyclic compounds exist. Performing the analysis with criteria that include more information on the chemical structure (atom type, hybridization and bond order) naturally leads to higher diversity in the results. 2506 frameworks, 1908 unique and 41 common frameworks that account for roughly one fourth of the total were found. The identification of structural characteristics of known drugs is interesting for de novo drug design. Common structures could be used as initial fragments in drug design programs or combinatorial chemistry libraries.

One possible weakness of the described method is that to some extent chemical intuition was the guide for the definition of the frameworks. The search for common patterns is therefore limited by some more or less arbitrary *a priori* decisions. Assuming the chemical space of the database is not as intuitive, it is possible that some features are not recognized. Neural networks have often been used successfully for pattern recognition. Ajay *et al.* [3] and Sadowski and Kubinyi [4] have recently and independently from each other proposed the use of neural networks to distinguish between drugs and non-drugs. Ajay *et al.* [3] have used a Bayesian network for this purpose. In a first attempt the descriptor set contained one-dimensional descriptors that described the global properties of the molecules (molecular weight, number of hydrogen bond donors, number of hydrogen bond acceptors, log P, number of rotatable bonds, aromatic density and the parameter κ that specifies the degree of branching of the bonding pattern). With this set reasonable accuracy was obtained for the classification of the molecules into drug-like and non-drug-like. The second descriptor set they studied consisted of the ISIS fingerprint of the compounds. The ISIS fingerprint is a string of 166 bits, each of which indicates the presence or absence of a particular moiety. Also with this second set satisfactory results were obtained but it was shown that the combination of the two sets significantly reduces the error in classification (to approximately 10%), even with a reduced version of the ISIS fingerprint string with only 71 bits. Sadowski and Kubinyi [4] used descriptors based on the atom

types of Ghose and Crippen [5]. For every compound, the input consisted of the count of the atom types. A subset of 92 of the atom types which were populated at least 20 times in the training set of 10,000 molecules was used. In some ways this descriptor is similar to the ISIS fingerprint, since the atom types are correlated to the moieties described by the ISIS keys. Feed forward networks were used with 5 hidden neurons. The results seem comparable to those of Ajay *et al.* [3]. There is one particularly interesting experiment in this study, however. Whole sets of drugs with a particular indication area (hormones and antagonists, drugs acting on the nervous system, blood and cardiovascular system and drugs acting on the respiratory system) were successively removed from the training set. It was shown that neural networks trained without these sets were only slightly worse in predicting these compounds as drug-like. This indicates the existence of some very general rules or constraints that drugs satisfy.

The derivation of general rules for the classification of compounds as drug-like or not is a very ambitious project. Not only because such rules are expected to be complicated, but mainly because the question that is being addressed is not very well defined in the first place. In the best case a number of features are obtained that define the chemical space of a particular database. There are, however, many reasons why a compound is used as a drug and in extension why the composition of the database is such as it is. Assuming one does obtain an algorithm that can distinguish with great accuracy between molecules similar to the ones found in a drug database and those that are not, the question still remains as to what this information actually tells us. In fact it is conceivable that molecules that according to the derived rules are drug-like may even be poor candidates as drugs for a new target, since they may show higher cross-reactivity with the targets of the drugs in the database.

Ajay *et al.* [3] discuss the possibility that the application of such methods may be detrimental to the exploration of new structures. They argue against this on the basis of results according to which they could extrapolate with good accuracy to a database that was not used in training. Nevertheless, they indicate that other techniques should also be used for the choice of new compounds. In evaluating the results of the described methods one should bear in mind a number of historical facts that have formed the current drug repertoire and that may possibly be responsible for some of its characteristics.

a) Many currently used drugs are natural products or their derivatives. Information obtained from databases of drugs may to a large extent correspond to rules of what is synthesizable in nature.

b) A number of related drugs exist due either to patent reasons or simply because of the iterative process of drug finding and optimization. A neural network with 300 or more free parameters (the weights) might easily describe reliably a database of 5000 compounds if these are well clustered. This indicates a danger of overfitting in the sense that the classification in drugs and non-drugs is not based on general properties but on a relatively detailed description of the individual clusters.

c) Most drugs currently used target active sites of proteins with deep clefts. Surface binding is still very hard to achieve. It is however also very important for regulation cascades and signaling. It is quite probable that the qualities of drugs binding on relatively flat protein surfaces will differ from the qualities derived from current databases. Larger compounds might be necessary to prevent protein-protein association. The structural characteristics of such complexes are currently being analyzed [6].

The results obtained from neural network simulations are notoriously difficult to interpret. Since it has been shown that classification is possible it would be interesting to proceed also with more straightforward statistical methods. Once the principle of the classification is understood it will be more easy to determine to which extent it is useful to look for drug-like compounds. None of these points is considered a rigorous objection to the methods described in this section. These methods are very promising and address a very important and up to now widely neglected aspect of the drug finding process. Exactly because of the promise that these methods hold it is important to take them to their limits, but not further.

Prediction of Binding Affinities

The prediction of biological properties, especially of binding affinities, from structure is a particularly active and large field of research. The prediction of binding affinities in the form of a scoring function is an important part of any ligand design program and will be discussed here seperately. The methods used to predict binding affinities in combination with ligand design should ideally be fast and accurate. An accurate energy function should fulfill two requirements: For known ligands it should have its global minimum reasonably close to the experimentally obtained conformation of the complex, and it should rank different ligands according to their binding affinity. These two requirements can be described as correct ranking in conformational and chemical space, respectively. The former is easier to achieve, because the free energy of

the ligand in solution identically cancels. The approaches used can be subdivided in quantitative structure activity relationship (QSAR) methods, empirical energy functions, and knowledge-based potentials and free energy calculations based on more general physical methods such as molecular dynamics [7].

QSAR

The aim of a QSAR is to create models correlating physicochemical properties to some biologically relevant quantity [8]. In this section the discussion will be limited to binding affinities, while it is understood that the method can easily be applied to other properties, such as toxicity or reactivity. QSAR has been defined as the relationships derived primarily by empirical analysis of a data table whose columns are numerical properties and whose rows are compounds. The technique has evolved over the last thirty years to a large and important field with applications in many biochemical and pharmacological problems. In the field of drug design QSARs have been particularly important in situations where a number of known effective substances exist but little is known about the (structure of the) target. The simplest QSAR models are based on linear regression with a few variables. Nowadays neural networks, genetic algorithms, and powerful statistical methods like principal component analysis and partial least squares regression are routinely employed for the derivation of QSARs. The components necessary are a training set of molecules with measured biological activity, a set of physicochemical descriptors for these molecules and the mathematical framework for the model. The training set is typically relatively small, due to experimental constraints. On the other hand the number of descriptors is very large. In general, therefore, one has to deal with an underdetermined system. This leads to the particular problem of overfitting. Overfitting means that a good fit for the training set is obtained, although the actual predictive value of the model is very poor. The predictive value has to be assessed, for example, through the use of cross-validation procedures. The leave one out method is based on building the model on only a part of the training set (systematically leaving one data point out). The value that is not used in the training is then predicted. This is repeated for every data point leading to an average error of prediction. The measure most often used for the predictive quality of a model is the cross-validated r_{cv}^2:

$$r_{cv}^2 = 1 - \sum_{i=1}^{n}(y_{i,exp} - y_{i,pred})^2 \Big/ \sum_{i=1}^{n}(y_{i,exp} - \bar{y}_{exp})^2 \quad (1)$$

where n is the number of compounds. The values of r_{cv}^2 range between 1 and $-\infty$. Values above 0.5 indicate reasonable predictivity for the model, while a value around zero indicates that the predictions are no better than random. However, the leave-one-out method is more a test of internal consistency than of real predictive quality. For problems with enough data points it is more rigorous to divide the data set in two or more smaller sets that are used as training or test sets. The predictive quality is then assessed over the average error for the test set.

The task of building a QSAR model can be decomposed into two parts. First, the descriptors that contain the information necessary for the prediction have to be identified. It should be noted that the descriptors do not have to be and indeed rarely are independent. Second, the information contained in these descriptors has to be extracted in an optimal way. The second part is performed by one of the mathematical models already mentioned. To avoid overfitting both the number of descriptors and the number of adjustable parameters (e.g. the weights in a neural network) should be kept as low as possible, while still obtaining a good fit for the training set. In a sense the choice of the descriptors seems even more important than the mathematical model used. The prerequisite for solving any problem is to find the variables that describe the problem, to understand it in a mathematical sense. This corresponds to the first procedure, the choice of the best suited descriptors. This part of the work can also be understood as the interface between chemical intuition and the usually complex QSAR models. One can use chemical intuition to choose the descriptors and inversely when using an automatic procedure to find the best descriptors one can hope to understand something about the nature of the problem and even the quality of the solution by analyzing the set of (automatically) selected descriptors (if possible in combination with the weights they obtain in the mathematical model). This type of analysis can be very useful as is shown in some of the work described below.

Descriptors can be subdivided into two classes: Two-dimensional (2D) and three-dimensional (3D) descriptors. In the first group one finds features mainly dependent on the functional groups and the atomic properties of the compounds (octanol/water partition coefficient (logP), polarizability, van der Waals volume, surface area, molecular weight). Descriptors in the second group depend on the structure and the particular conformation of the molecules. A number of very interesting and powerful methods have been suggested recently for the determination of the best suited descriptors. The genetic function approximation (GFA) developed by Rogers is an elegant approach that can be used to determine the best descriptors and the functional form of the model at the same time [9]. The functional space of the models can be described as the set of all possible linear combinations $F(\mathbf{X})$ of M

basis set functions ϕ_k (**X**) of the variables (descriptors) **X** = {x_1, x_2,..., x_m}:

$$F(\mathbf{X}) = a_0 + \sum_{k=1}^{M} a_k \phi_k(\mathbf{X}) \quad (2)$$

The principle of the GFA is that a genetic algorithm is used to search the space of basis set functions. Genetic algorithms solve problems in an evolutionary way [10,11]. Possible solutions (individuals) are coded in bit-strings called chromosomes. An initial population of individuals is created usually by randomization of the strings. A fitness function is used to assess the quality of the solutions. Individuals with the highest fitness (best solutions) are more likely to be chosen as parents for offspring and thus propagate their genetic material, i.e. the information coded in the string, into the next generation. Offspring is created by random mutation of the parents string and/or crossover (combination of the information contained in two parent chromosomes). In the genetic function approximation the chromosomes code for different basis set functions of different features. For every chromosome the optimal weights for the different basis set functions are determined by least-squares regression. For the scoring function the inverse of the lack of fit (LOF) measure is used:

$$LOF = \frac{LSE}{\left(1 - \frac{c+dp}{M}\right)^2} \quad (3)$$

Here *LSE* is the least squares error, *c* is the number of basis functions used (other than the constant), *p* is the total number of features contained in all basis functions and *d* is a user-defined smoothing parameter. The advantage of the LOF is that it contains a bias against the inclusion of too many features in the model, and thus protects against overfitting. The strength of this bias is dependent on *d*. The GFA leads to a set of different highly predictive QSAR models for a number of ligand sets (e.g. r_{cv}^2=0.84 for the Selwood data set [12], which contains 31 compounds described by 56 features). The authors correctly point out the efficiency of genetic algorithms for sampling in large spaces, mainly comparing with incremental methods that neglect cooperative effects of the descriptors. The fact that for a single set different models (based on different descriptors) are very good indicates the possibility of using multiple QSARs to obtain averaged results and possibly also an estimation of the validity of predictions. Luke has used evolutionary programming to obtain a few additional solutions that were missed by GFA for the same systems [13]. Independently, Kubinyi has used another evolutionary algorithm to obtain similar results for the Selwood data set [14].

Evolutionary algorithms seem to be particularly interesting for the choice of features. For the choice of the actual functional form neural networks offer an extremely flexible possibility in particular for modeling nonlinear relationships and they have often been used for QSAR [15,16]. Wagener *et al.* have used Kohonen networks with spatial autocorrelation vectors as input [17]. A spatial autocorrelation vector has components corresponding to the value of the autocorrelation of some property at different distances. The property is sampled over the surface of the ligands. The autocorrelation values are therefore independent of the relative alignment of the compounds. It was shown that with this approach clustering of the compounds according to activity is obtained. This is particularly interesting, since Kohonen networks are self organizing, i.e. the activity data are not used during training of the network. The authors further showed that by using the same input with a feed-forward neural network trained by a back-propagation with momentum procedure they obtain a model that predicts all except one of the 31 corticosteroid binding globulin ligands very well r_{cv}^2=0.63 and r_{cv}^2=0.84 with all ligands and excluding the outlier, respectively). The outlier is the only molecule of the set with a substituent at position 9. This compound has been found to be problematic also in other studies [18,19].

So and Karplus have combined genetic algorithms for feature selection with neural networks for the regression [20]. The main idea is to use the GA to find features that optimize the predictive value of the model. This was achieved by using the r_{cv}^2 in the scoring function. Thus features that lead to a high predictivity are chosen. Application of this principle on the Selwood data set leads to highly predictive models with a r_{cv}^2=0.86. The authors suggest that with this method close to optimal QSARs are obtained. In agreement with previous studies [9,14] they find that logP is a very important feature for this data set. It is included in all their best models. Hydrophobicity is expected to play a role in many QSAR studies since in general it is considered to be the driving force of binding. The analysis of the other descriptors used in the best models shows that they also make sense chemically. Later the same authors combined their work with the molecular similarity matrices approach [19] to extend their method to 3D QSAR [21]. A molecular similarity matrix contains at every element a_{ij} the field similarity of compounds *i* and *j*. One can use steric or electrostatic fields. The GA was applied to find the columns of the similarity matrices that can be used as input to yield the highest predictivity. Neural networks with 1 to *N/5* input neurons were used for the regression, where *N* is the number of data points. On the basis of the r_{cv}^2 they find that, depending on the data set, 3-9 features are optimal, yielding r_{cv}^2-values

between 0.73 and 0.94 for nine different data sets [22].

3D descriptors depend on the molecular fields of the compounds and the overlap between them. One particularly interesting method based on molecular fields is the comparative molecular field analysis (CoMFA) [18]. In CoMFA the molecules are represented by their electrostatic and steric fields sampled over a grid that surrounds the molecule. The calculated fields at every point for every compound enter as a row in a matrix that is the basis of the model. This matrix has significantly less rows (compounds) than columns (grid points), thus describing an underdetermined system. However, because the field values are strongly correlated between neighboring points and similar compounds, it is possible to use partial least squares (PLS) analysis [23] to obtain models with few components and high predictivity. Initially a solution of low dimension is found and then the dimensionality is progressively increased, until the r_{cv}^2 stops increasing. In the original application of CoMFA to 21 ligands binding to two different steroid binding globulins satisfactory results were obtained (r_{cv}^2 was equal to 0.55 and 0.66 for the two proteins). The most important free parameter of the method (and of 3D methods in general) is the alignment of the molecules with respect to each other and the grid. The authors suggest a field fit procedure that maximizes the overlap of the fields, however as will be discussed shortly other possibilities are available. CoMFA has become a standard technique in QSAR, even though there are a number of potential difficulties with its use. Since it is a linear mapping technique and is based on the similarity of the compounds evaluated, it may fail when some of the compounds in the set are dissimilar to all the others. Furthermore, in spite of the use of PLS the system is still underdetermined and it is clear that a number of different solutions of approximately similar quality within the training set may exist [18]. These will in general differ in their predictions for the biological activity of designed drugs. As has already been mentioned this effect may be put to good use by estimating the quality of prediction over the standard deviation of the output of different models. Finally, the CoMFA predictions correspond to enthalpies rather than free energies, since the entropic contributions of the solute and solvation effects are generally neglected [24].

Alignment

3D QSAR methods are dependent on the relative alignment of the different molecules. One possibility for automatic alignment has already been mentioned with the field fit performed in CoMFA. Alignment methods are also important because they allow the identification of pharmacophores and the screening of compound databases for molecules similar to a known inhibitor or even the substrate. The application of alignment algorithms is based on two main assumptions which are not always given. The compounds should interact with the same groups in the active site and the binding of different molecules should not distort in different ways the binding site. Kearsley [25] has suggested the steric and electrostatic alignment method (SEAL), which optimizes overlap functions of the type:

$$A_F = -\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} exp\left(-\alpha r_{ij}^2\right) \quad (4)$$

where the subscript i and j run over the m atoms in the first molecule and the n atoms in the second molecule, respectively. The factor α determines the width of the Gaussians describing the overlap between two atoms. r_{ij} is the distance between the two atoms and the pre-exponential factors w_{ij} are functions of the partial charges q_i and q_j and the van der Waals radii R_i and R_j of the atoms:

$$w_{ij} = w_E q_i q_j + w_S R_i R_j \quad (5)$$

The weight factors w_E and w_S set the relative significance of the electrostatic and steric fit, respectively. The relative position of the molecules is described by the translation vector and four quaternion variables for rotation. Quaternions have the advantage that they do not show the singularities found for Euler angles. They also permit faster manipulations on the computer. The alignment is performed with a rational function optimization algorithm which expresses the function to be optimized as a Padé approximant and not as a Taylor series expansion as is usual for minimizers. Unlike the Taylor series the Padé approximation does not diverge for large distances from the point of expansion. The SEAL algorithm allows the exhaustive search of possible alignments of different molecules and has become a standard tool in drug design. Klebe *et al.* [26] have extended the SEAL approach to include conformational optimization in torsion space during the alignment. Since this optimization is local the method is combined with conformational search performed by the program MIMUMBA [27]. The approach is shown to lead to higher similarities than standard SEAL and to reproduce experimental binding conformations.

Alignment and docking are related problems. Therefore, it is not surprising that a number of algorithms originally developed for docking have been used also for alignment. Jones *et al.* have adapted their genetic algorithm [28] obtaining a procedure for the automatic and simultaneous alignment of a number of molecules with a template molecule [29]. The chromosome codes for the flexible torsion angles of

the molecules and an additional integer which maps features of the molecules with features of the template. A least squares procedure is used to obtain an optimal overlay corresponding to the coded mapping. The scoring function is a weighted average of the internal van der Waals energy of the molecules (to avoid non-physical structures), a volume integral for the common volume between each molecule and the template molecule, and a similarity score based on the features common to all molecules in the current overlay. In a similar spirit, Lemmen *et al.* have based their alignment procedure FLEXS [30,31] on the docking program FLEXX [32]. Starting from an initial base fragment the algorithm aligns the molecules adding one fragment after another. The method is very fast and shows good accuracy. On a test set of 284 experimentally given alignments 60% of the examples can be reproduced with an RMSD below 1.5Å . This compares well with the corresponding reproduction rate of 70% obtained with docking [33].

Empirical Functions and Knowledge-based Potentials

The main problem of QSAR models is their limited reliability and the fact that it is difficult to assess to which extent the model derived from a particular test set will be applicable for new compounds. It is therefore of interest to work with more general functions, sacrificing accuracy for the specific case, but gaining in robustness of predictions. Empirical functions and knowledge-based potentials can be understood as very general 3D-QSARs. They differ from the usual QSAR in two points: First, the training set contains many different ligand types that bind at different receptors. Second, the structure of the receptor itself is also used in the derivation of the model. First attempts to fit binding affinities of many different complexes to a single equation showed already promising results [34]. An equation of the following form was used:

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{hb} \sum_{h\text{-}bonds} f(\Delta R, \Delta\alpha) + \Delta G_{ionic} \sum_{ionic} f(\Delta R, \Delta\alpha) + \Delta G_{lipo} A_{lipo} + \Delta G_{rot} N_{rot} \quad (6)$$

ΔG_0 is a constant that can be interpreted as the loss of rotational and translational entropy, ΔG_{hb} corresponds to the contribution of an ideal hydrogen bond, ΔG_{ionic} corresponds to the contribution from an unperturbed ionic interaction, ΔG_{lipo} gives the contribution from lipophilic interactions and ΔG_{rot} is the entropy loss associated with the freezing of internal degrees of freedom in the ligand. A_{lipo} is the surface buried upon binding and N_{rot} the number of flexible torsion angles of the ligand that become frozen upon complex formation. The function f (ΔR, $\Delta\alpha$) takes into account the geometry of polar interaction in a very approximate way. The fitting of $\Delta G_{binding}$ with this equation leads to a root mean square error of approximately 1.9 kcal/mol for the training set of 45 protein ligand complexes. It has also been used very effectively in drug design [35] and docking programs [32]. On the other hand a number of problems are already evident from the functional form. For example polar interactions are dependent not only on the local geometry but also to a large extent on the screening by the solvent. An approximate screening has been included in later development of this energy function [36]. An even more general disadvantage of this approach is the fact that this equation does not contain any penalty terms for possible interactions that do not take place. For example the burial of a surface charge of a protein is connected with a significant loss of solvation energy. The complexes in the training set do not contain such situations, since they are extremely unfavorable and would destroy binding. A similar (but smaller) penalty is probably necessary for uncharged hydrogen acceptors or donors that are desolvated without satisfying their hydrogen bonding potential. The statistics on bad interactions is poor simply because when there are too many of them binding is destroyed and the structure of the complex cannot be determined experimentally. In docking or design, conformations are created that show unfavorable features that are not penalized, because their presence was never expected from the training data set (the complexes). Furthermore, the fact that the regression is performed on different complexes lends generality to this equation on the one hand but on the other it leads to lower precision for the specific case. Murray *et al.* have shown how to address this problem by using Bayesian regression [37]. They fit an equation, which is similar in spirit to that of Böhm, to 82 crystallographically determined protein-ligand complexes for which the binding affinity is known. This data set is treated as the prior information while the additional data consists of the affinities between a number of ligands and thrombin. The equation is optimized for the second data set, while still yielding reasonable results for the larger and more general data set. The obtained parameters for the equation lead to higher predictivity (more general models).

An interesting coarse grained knowledge-based potential has been derived by DeWitte and Shakhnovich [38]. In their approach very little physical intuition enters in the model, in contrast to the previously mentioned energy functions. In the previous models it was assumed that hydrogen bonds will be important, the training simply yields the weight for the corresponding term. In the model of DeWitte and Shakhnovich, the potential depends only on the

formation of atomic contacts (if two atoms lie within 5Å of each other) and was derived from a database of 106 complexes of ligands binding to proteins in pockets. A different potential was derived for surface binding from 17 complexes of non-peptidic ligands binding on the surface of proteins. As the authors point out, this distinction, although somewhat arbitrary, is necessary to account for the influence of solvation within the simple potential used. The derived potentials are interesting because they yield very low energies for the original complexes in comparison with designed ligands. This fact can also be interpreted as a weakness of the design algorithm, however, it is expected that known ligands should score well with the energy function.

COMBINE is an interesting approach which is formally similar to CoMFA [39]. The interaction energy between the ligand and the protein is decomposed in a residue-by-residue basis. With the help of partial least squares regression, a set of weights that represent the relative importance of each residue-ligand interaction is obtained. The method has been expanded to the use of screened interactions in solution and was shown to predict well ligands not included in the training set [40]. The advantage of the method is that it is based on molecular mechanics potentials and a fitting which can be understood to account for solvation or simply the difference to the unbound state. COMBINE is a smart combination of QSAR techniques with interaction energies calculated with the AMBER force field [41,42]. It is therefore expected and indeed shown to lead to highly predictive models.

Free Energy Calculations

Free energy calculations based on molecular dynamics can be used to calculate binding affinities between a ligand and the corresponding receptor. A number of excellent reviews on the subject exist [43-45]. Here we will briefly discuss approximate methods based on the statistical thermodynamics of binding. A short introduction to the formal approach for the calculation of free energies is given first. The dissociation constant (K_D) depends on the free energy of binding (ΔG) which is the free energy difference between the complex state and the state in which the receptor and ligand are free in solution.

$$\Delta G = RT\ln K_D \qquad (7)$$

where T is the temperature and R the gas constant. To calculate ΔG one would have to simulate the process of association or dissociation. Since the free energy is a state function it depends only on the end states. In the case of dissociation, it is only important that in the simulation the initial state corresponds to the complex and the final state to the compounds fully solvated. It is possible to take any path that connects these two states, even an unphysical path. For example the ligand can be gradually faded out from the binding site and - in a separate simulation - grown into the solvent. Although this is in general more efficient than the physical simulation of the dissociation it is still very time-consuming, due to the large size of the relevant conformational space. However, one can circumvent this problem by calculating ΔΔG, the difference in binding free energy between two different compounds A and B. The free energy change along any closed path is zero. Therefore, in the thermodynamic cycle shown in Fig. 1 under consideration of the direction of the arrows one obtains:

$$\Delta G_{12} + \Delta G_{24} - \Delta G_{34} - \Delta G_{13} = 0,$$
$$or$$
$$\Delta\Delta G = \Delta G_{12} - \Delta G_{34} = \Delta G_{13} - \Delta G_{24} \qquad (8)$$

The simulation necessary to perform the calculation of ΔG_{13} starts with the complex between the receptor and compound A. During the simulation compound A is gradually (usually discontinuously) mutated into compound B. The mutation is controlled over the coupling parameter λ, which reflects the chemical identity of the system, e.g. for the simulation from state 1 to state 3 in our cycle the system corresponds to the receptor complexed with compound A (λ=0) and compound B (λ=1). For intermediate values of λ the system corresponds to a chimera between states 1 and 3. If the difference between A and B is small enough the mutation can be performed in one step. The calculation of ΔG_{24} proceeds in the same way in the aqueous environment.

The free energy difference can be calculated from the simulation either by evaluating the derivative of the free energy over the coupling parameter λ and performing a numerical integration, according to the thermodynamic integration method [46]:

$$\Delta G = \int_0^1 \frac{\partial G}{\partial \lambda} d\lambda = \int_0^1 \langle \frac{\partial H(\mathbf{X}^N, \lambda)}{\partial \lambda} \rangle_\lambda d\lambda \qquad (9)$$

or with the help of the perturbation formula which evaluates the ratio of the partition sums at different values of λ [47]:

$$\Delta G = -k_B T \ln \langle e^{-\frac{\Delta H(\mathbf{X}^N)}{k_B T}} \rangle_0 \qquad (10)$$

ΔH is the difference Hamiltonian of the systems A and B, k_B is the Boltzmann factor, T the temperature and $\langle\rangle_\lambda$ is the ensemble average at λ. It has often been mentioned in the literature that both methods are formally exact and equivalent [43]. This is of course correct, however, there is a significant difference when

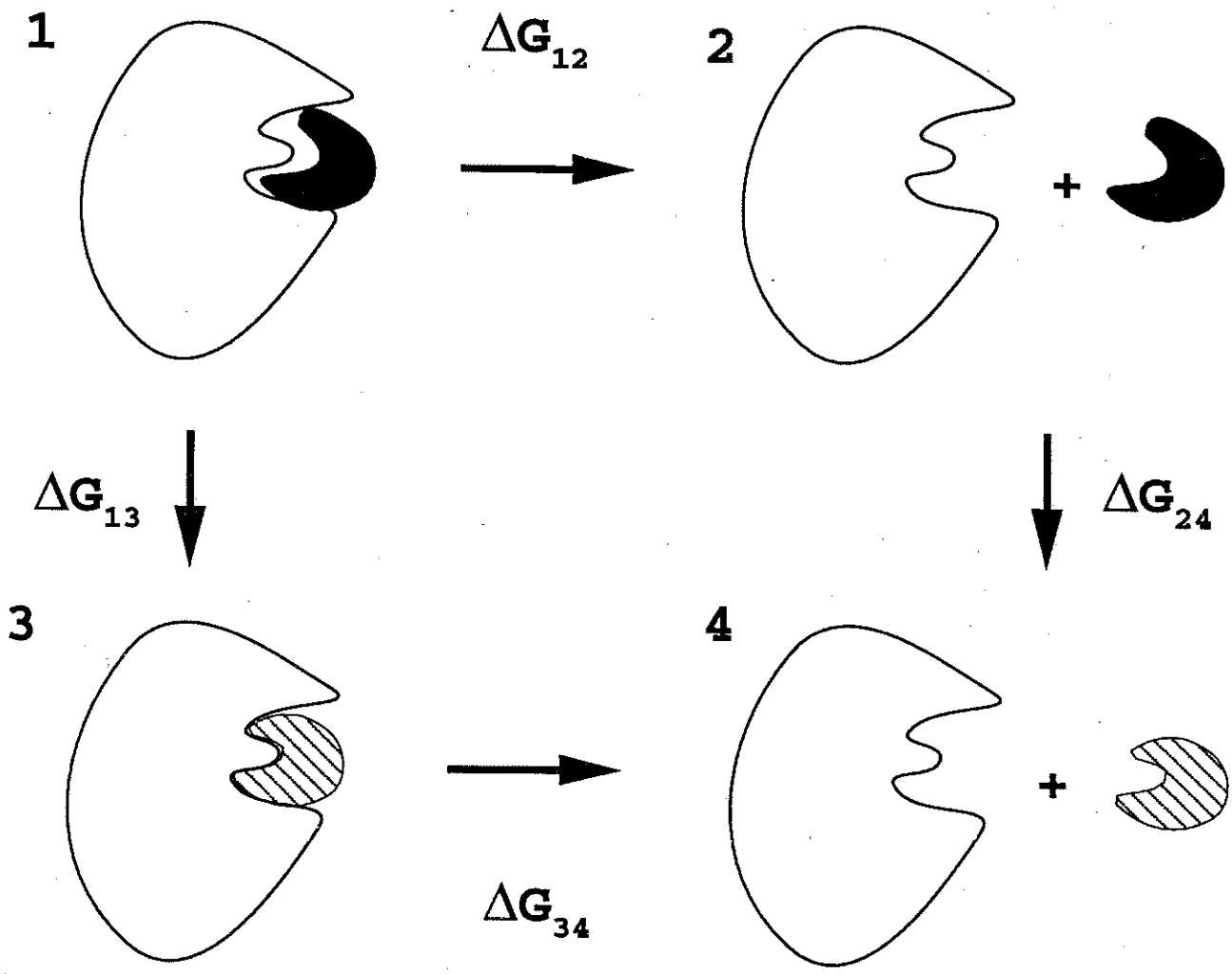

Fig. (1). Thermodynamic cycle for the mutation between different ligands binding to the same receptor.

it comes to practical applications. The perturbation method is also formally exact when performed in a single step, from a single ensemble, whereas in thermodynamic integration it is essential that one performs a number of simulations at different values of λ. This is important especially in the field of drug design since there is a clear trend towards fast and simple methods that need preferably only one simulation per compound. Gerber *et al.* have performed estimations of free energy changes in a single step with thermodynamic integration, but as expected this lead to poor results [48]. Liu *et al.* have argued that in order to perform a free energy calculation in one step it is necessary to include higher terms of the Taylor expansion of the free energy [49]. These, however, show increasingly worse convergence properties. Including the infinite number of terms of the complete Taylor series is equivalent to the perturbation method. Based on these observations the authors formulated a scheme with which the free energy of a number of different compounds could be estimated from a simulation of a single reference state. The scheme is based on standard perturbation, and the main contribution of the work lies in the definition of suitable "non-physical" reference states which use soft core potentials that allow better sampling with respect to the final states [50]. It is shown that acceptable estimations of the relative free energy of different substituted phenols in water can be obtained from a single reference state simulation.

Relative free energies of conformations of the same molecule can be calculated by simply sampling the probability of the conformational states of the molecule. However, conformational transitions are often hampered by energy barriers. In this case the statistics are poor and convergence is very slow. This problem is usually solved with the help of an umbrella potential which is applied to keep the system at a certain position along the reaction coordinate, which describes the conformational change. This umbrella potential is often a harmonic function of the reaction coordinate [43]. In the adaptive umbrella sampling it is the negative of the free energy calculated up to that point and is thus very

efficient at removing barriers [51]. These calculations then lead to the potential of mean force of the system with respect to the reaction coordinate. Kong and Brooks have developed a method for the calculation of free energy differences between different systems [52]. They treat the coupling parameter λ as a variable during dynamics. A fictitious mass for the λ degree of freedom is introduced, and λ is allowed to evolve according to Newtons law. Along with λ the chemical identity of the system changes. This so-called λ-dynamics allows the use of potential of mean force methods for the calculation of free energy differences between different molecules. The use of more than one coupling constant to describe the single terms of the changing Hamiltonian allows more efficient search for optimal pathways for the transition. What is much more interesting from the drug design point of view are the virtual competition experiments. The use of a number of coupling constants allows the system to choose among an equal number of molecular states. With this method the correct order of hydration free energies for C_2H_6, CH_3OH, CH_3SH, CH_3CN is obtained from a single simulation. Guo and Brooks have refined the method and applied it to trypsin inhibitors and obtained good qualitative agreement with more exact calculations. [53,54].

Semi-empirical methods based on linear approximations to the free energy have also been used with significant success [55-57]. They approximate the free energy of binding with an expression of the type [55]:

$$\Delta G_{bind} = 1/2(\langle V_{l-s}^{el}\rangle_{prot} - \langle V_{l-s}^{el}\rangle_{aq}) + \alpha\,(\langle V_{l-s}^{vdW}\rangle_{prot} - \langle V_{l-s}^{vdW}\rangle_{aq}) \quad (11)$$

where V_{l-s}^{el} and V_{l-s}^{vdW} are the electrostatic and van der Waals interaction energies between the ligand and its surroundings in protein (prot) or in aqueous solution (aq), respectively. The $\langle\rangle$ denotes an ensemble average sampled over a molecular dynamics or Monte Carlo trajectory [57]. The factor α is determined empirically. The method does not seem to offer significant advantages at first sight, however, it can be applied also for ligands with significant differences in their two-dimensional structure where standard free energy calculations usually fail to converge. The method was applied to five endothiapepsin inhibitors leading to an α value of 0.161 and a maximum error of 0.53 kcal/mol for the absolute binding free energies of the training set [55]. It has recently been used for studying the binding of 14 biotin analogues to avidin and yielded results that correlate well with experimental data for 10 of the ligands [58]. For the other four ligands the error of more than 7 kcal/mol originates mainly from conformational changes in the protein due to bulky substituents.

Structure-based Ligand Design Programs

Ligand design programs have been reviewed recently in great detail [59-62]. We will therefore limit the discussion on a few approaches implemented in the more widely distributed programs. Ligand design programs build ligands under the influence of some scoring function. The scoring function either corresponds to the potential in the binding site of the receptor or is a similarity function that biases the design toward known ligands. Ligands are built by connecting smaller molecular fragments or even atoms. Although atom-based approaches [63] have shown significant flexibility with respect to the structures that can be obtained, most methods build new compounds by combining predefined fragments mainly for two reasons: First, it is easier to control the synthesizability and the chemical stability of the designed molecules [64,65]. Second, fragments are more easily modeled, since the model parameters, like partial charges and force constants for the torsion angles, can often be assumed to depend mainly on the fragment and only to a lesser extent on the rest of the structure. With fragment-based approaches the number of newly created bonds is minimized.

The main ingredients of design programs are the docking procedure, the linking procedure and the energy function for the evaluation of the docked fragments and designed ligands. Since the problem of design is so closely related to docking the two fields have been developed in parallel. A short description of the most widespread approaches to docking with examples of corresponding programs will serve also as a short introduction to the concepts used in molecular design. DOCK, the earliest docking program was based on rigid docking and the use of geometrical criteria to judge the complementarity between receptor and ligand and was therefore fast enough to screen whole databases for leads [66]. DOCK uses spheres complementary to the receptor molecular surface to create a space filling negative image of the receptor site. Several atoms of the ligand are matched with receptor spheres to define the orientation of the ligand. Flexibility [67] and a more detailed force-field type of energy function for scoring [68] were included in later development of the program. DOCK has been able to find novel micromolar inhibitors of enzymes [69,70]. Recently it has been further developed to efficiently dock combinatorial libraries into a protein binding site by using a combinatorial branch and bound procedure [71].

FLEXX [32] is a very fast program for docking medium sized flexible ligands. It makes use of a number of methods that are typical of ligand design programs. It

docks molecules in a fragment-by-fragment basis and uses Böhms empirical function for scoring [34]. It is fast enough to allow screening of small databases of ligands and has been extended to predict water molecules in the binding site [72].

A number of genetic algorithms have been suggested for docking, lately [28,73,74]. They combine speed with simplicity of concept. GOLD [28] is based on a genetic algorithm that encodes the approximate conformation of the ligands in the chromosome and uses a simple least squares fitting procedure to obtain the final conformation. It also allows flexibility around bonds to hydrogen bond donors and acceptors in the receptor. The method has been validated on a large number of complexes leading to a success rate (identification of the experimental binding mode) of 70% [33]. Flexibility in the protein has been taken into account in various degrees from using approximate scoring functions [75] to allowing explicit flexibility for side chains [76,77] to methods that at least in principle allow arbitrary flexibility and are based on molecular dynamics [78] or Monte Carlo simulations [79]. However the inclusion of flexibility for the receptor generally leads to a very significant loss in efficiency.

In general two main approaches to the design problem can be distinguished. One is the grow approach as seen in programs such as GROW [80] and GroupBuild [81], where starting from a manually docked fragment the ligand is grown by adding fragment by fragment in the binding site. This has the advantage that chemical bonds are formed with correct geometry, that the intraligand interactions can be taken into account during the design, and that the number of ligands to be generated can easily be controlled. On the other hand, in the grow approach it is not straightforward to use existing information on energy minima and interesting binding pockets. For example programs based on the growing approach often face difficulties in creating ligands that bind to a number of independent pockets. Furthermore, the designed molecules depend crucially on the seed. That means that a number of runs with different seeds have to be performed. In these runs it is difficult to make use of information obtained in previous runs except for the trivial way of starting directly from a resulting conformation of a previous simulation [64]. In SMoG (Small Molecule Growth) the growing is implemented with a Monte Carlo type of approach [38,82]. A new fragment is grown from the previous structure and if the binding energy per atom decreases the new structure is accepted, while if it increases it is only accepted with a probability proportional to $exp(-\Delta g/T)$ where $g=G/N$ with G the free energy, N the number of atoms and T an algorithmic temperature [38,82].

In the combinatorial approaches fragments or probes are first docked in the binding site leading to functionality maps [83-85]. In the case of MCSS [85] these maps contain the different low energy conformations for each fragment type. These are clustered and the cluster representatives can subsequently be linked together with smaller [86,87] or larger (CAVEAT, [88] HOOK [89]) linkers. This approach has the advantage that the individual fragments are docked in optimal positions. On the other hand the geometry of the new bonds is not optimal and has to be accepted with a certain tolerance initially. LUDI [35,90,91], probably the most widely used *de novo* design program, makes extensive use of empirical information derived from structural databases. Interaction sites that indicate possible positions for functionalities complementary to the receptor are defined and used to dock fragments from a library. Alternatively the output of GRID can be used for the definition of interaction sites. The fragments are fitted on the interaction sites with the algorithm published by Kabsch [92] and are connected with small linkers. Interaction geometries were derived from structural data on small organic molecules. The energy function used is the empirical function by Böhm [34].

The computational combinatorial ligand design (CCLD) [86] approach was initially based on docking of functional groups with MCSS [85]. The fragments were ranked according to an approximated binding free energy whose solvation component was assumed to be the sum of electrostatic and non-polar contributions. The electrostatic term is obtained through the solution of the linearized Poisson Boltzmann equation [93-95], while the non-polar term is assumed to be proportional to the solvent accessible surface. CCLD creates two lists of fragment pairs, the first containing overlapping (i.e. mutually excluding) fragments and the second containing bonding fragments (i.e. fragments that can be bound by small linkers). Starting from every fragment in turn CCLD creates ligands by linking the docked fragments with the most favorable of small linkers. To avoid combinatorial explosion growing is discontinued when the average binding free energy of the fragments in the new ligand exceeds a user specified threshold. With the help of this approach novel compounds have been suggested for thrombin that show low micromolar binding affinity [96]. A new approach for docking fragments (SEED [97]) has been implemented recently. As mentioned previously, the main advantage of SEED over other docking programs is the comprehensive treatment of electrostatics in an efficient and accurate manner. Electrostatic solvation energy is decomposed in three contributions. The desolvation of the protein is calculated with the simple field approximation and integration of the energy density over the fragment volume. The interaction

energy between the receptor and the fragment and the desolvation of the fragment are both calculated with the help of the Generalized Born formula [98,99]. This scheme has been validated by comparison to the results obtained from finite difference solution of the Poisson equation [93-95]. Apart from the more accurate energy function, SEED has an additional advantage over MCSS, namely, that not only minima of the fragment positions are saved. Also other low energy conformations are kept for the creation of new ligands. This is based on the experience that in MCSS particularly good binding pockets will attract the fragments obscuring other low energy conformations. CCLD was used to create larger ligands from the fragments docked with SEED in the thrombin active site. Several of the *de novo* designed ligands resembled in structure and interactions a number of well known thrombin inhibitors [97].

At this stage a final note on the energy functions used seems necessary. Ligand design can be understood as the extension of the docking problem into chemical space. The degrees of freedom to be optimized are not only the positional and conformational variables of a particular compound, but, additionally, its chemical identity. This point of view makes one important problem in the field of ligand design particularly clear: The quality of the scoring or energy function used to evaluate the different solutions. When the search space is very limited as for example in the first programs that performed rigid docking [66] a very simple energy function based on geometrical criteria was sufficient to recognize the correctly docked structures. When flexibility in the ligand (and the protein) is allowed the effect of solvation has to be taken into account to avoid sampling irrelevant parts of the conformational space [79]. A simple example shows the higher quality requirements on the scoring function for design purposes: Assuming the charge on an atom in a designed ligand is a (continuous or discrete) variable of the optimization, any simple force-field-based energy function would tend to maximize its absolute value [100]. This is however in disagreement with empirical data. Although sometimes high affinity may be due to ionic interactions, more often than not the desolvation of full charges on the ligand and the protein is stronger than the direct interaction. Desolvation is the change in the solvation energy of the ligand and the receptor, upon complex formation. This further indicates that the scoring function should correspond to a difference between the free and the complexed states. The calculation of such differences is not necessary in docking because the term corresponding to the free state identically cancels. Accurate and reliable prediction of the absolute binding free energy for a medium-sized flexible ligand is currently beyond the limits of routine calculations, since it also includes finding the most probable conformations in water and averaging with the correct thermodynamic weights. Furthermore, in ligand design free energies are assumed to be additive, although of course it is clear that this is only a crude approximation [101]. The main task for a scoring function in a ligand design program is to find the conformations with the lowest energies for every chemical species (be that an atom, fragment or complete ligand) and in the case of different chemical entities (for example a benzene and a guanidinium docking in the same binding pocket) to decide which yields the lowest binding free energy. Both tasks and especially the latter are not straightforward and will most probably have to be addressed at different levels of accuracy during different stages of the design process.

Outlook

Drug design is a truly interdisciplinary subject. It is not expected that in the next few years any single method or approach will monopolize the discovery of new drugs. Different methods will support each other. Cross-fertilization between the disciplines is already evident when one compares computational and experimental approaches, such as MCSS [85] and SAR by NMR [102,103] and MSCS [104]. Also the combinatorial approach to design is more than reminiscent of the principle of combinatorial libraries [71,86]. One particularly interesting paper by Weber *et al.* exemplifies the synergies that one would expect in the future. A genetic algorithm was used in conjunction with a combinatorial library to obtain compounds with submicromolar affinity to thrombin [105]. The problem of the prediction of binding affinity was solved *a priori* by using experimentally determined binding constants as scoring function in the genetic algorithm. Computational methods will continue to play an increasingly important role in the drug design process since they are not only helpful for experimental approaches (e.g. refinement in X-ray crystallography [106]) but also contribute directly to the design and discovery of new drugs [1,107]. A particularly interesting new direction to follow is the development of methods to address toxicity and bioavailability from the very beginning of the design process. Further, the increasing amount of data on human gene sequences and the development of methods to reliably predict protein structures from their sequence [108] is spurring additional interest in new theoretical and computational approaches not only for drug design, but also for target finding and characterization.

Acknowledgment

This work was supported by the Swiss National Science Foundation.

References

[1] Kubinyi, H.; Curr. Opin. Drug Design Discov., **1998**, 1, 4.

[2] Bemis, G.; Murcko, M.A.; *J. Med. Chem.*, **1996**, *39*, 2887.

[3] Ajay, A.; Walters, W.P.; Murcko, M.A.; *J. Med. Chem.*, **1998**, *41*, 3314.

[4] Sadowski, J.; Kubinyi, H.; *J. Med. Chem.*, **1998**, *41*, 3325.

[5] Ghose, A.; Crippen, G.; *J. Chem. Inf. Comp. Sci.*, **1987**, *27*, 21.

[6] Lo Conte, L.; Chothia, C.; Janin, J.; *J. Mol. Biol.*, **1999**, *285*, 2177.

[7] van Gunsteren, W.F.; Berendsen, H. J.C.; *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 992.

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

[9] Rogers, D.; Hopfinger, A.J.; J. Chem. Inf. Comput. Sci., **1994**, *34*, 854.

[10] Davis, L.; *Handbook of Genetic Algorithms*; Van Nostrand Reinhold, New York NY, **1991**.

[11] Goldberg, D.E.; Genetic Algorithms in Search Optimization and Machine Learning; Addison-Wesley, Reading MA, **1989**.

[12] Selwood, D.; Livingstone, D.; Comley, J.; O'Dowd, A.; Hudson, A.; Jackson, P.; Jandu, K.; Rose, V.; Stables, J.; *J. Med. Chem.*, **1990**, *33*, 136.

[13] Luke, L.T.; *J. Chem. Inf. Comp. Sci.*, **1994**, *34*, 1279.

[14] Kubinyi, H.; Quant. *Struct.-Act. Relat.*, **1994**, *13*, 393.

[15] So, S.; Richards, W.; *J. Med. Chem.*, **1992**, *35*, 3201.

[16] Manallack, D.; Ellis, D.; Livingstone, D.; *J. Med. Chem.*, **1994**, *37*, 3758.

[17] Wagener, M.; Sadowski, J.; Gasteiger, J.; *J. Am. Chem. Soc.*, **1995**, *117*, 7769.

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

[19] Good, A.; So, S.; Richards, W.; *J. Med. Chem.*, **1993**, *36*, 433.

[20] So, S.; Karplus, M.; *J. Med. Chem.*, **1996**, *39*, 1521.

[21] So, S.; Karplus, M.; *J. Med. Chem.*, **1997**, *40*, 4347.

[22] So, S.; Karplus, M.; *J. Med. Chem.*, **1997**, *40*, 4360.

[23] Lindberg, W.; Persson, J.; Wold, S.; *Anal. Chem.*, **1983**, *55*, 643.

[24] Klebe, G.; Abraham, U.; *J. Med. Chem.*, **1993**, *36*, 70.

[25] Kearsley, S.; Smith, G.; *Tetrahedron Comput. Methodol.*, **1990**, *3*, 615.

[26] Klebe, G.; Mietzner, T.; Weber, F.; *J. Comput.-Aided Mol. Design*, **1994**, *8*, 751.

[27] Klebe, G.; Mietzner, T.; *J. Comput.-Aided Mol. Design*, **1994**, *8*, 583.

[28] Jones, G.; Willett, P.; Glen, R.C.; *J. Mol. Biol.*, **1995**, *245*, 43.

[29] Jones, G.; Willett, P.; Glen, R.; *J. Comput.-Aided Mol. Design*, **1995**, *9*, 532.

[30] Lemmen, C.; Lengauer, T.; *J. Comput.-Aided Mol. Design*, **1997**, *11*, 357.

[31] Lemmen, C.; Lengauer, T.; Klebe, G.; *J. Med. Chem.*, **1998**, *41*, 4502.

[32] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G.; *J. Mol. Biol.*, **1996**, *261*, 470.

[33] Jones, G.; Willett, P.; Glen, R.C.; Leach, A.; Taylor, R.; *J. Mol. Biol.*, **1997**, *267*, 727.

[34] Böhm, H.J.; *J. Comput.-Aided Mol. Design*, **1994**, *8*, 243.

[35] Böhm, H.J.; *J. Comput.-Aided Mol. Design*, **1992**, *6*, 61.

[36] Böhm, H.; J. Comput.-Aiderd Mol. Design, **1998**, 12, 309.

[37] Murray, C.; Auton, T.; Eldridge, M.; *J. Comput.-Aided Mol. Design*, **1998**, *12*, 503.

[38] DeWitte, R.; Shakhnovich, E.; *J. Am. Chem. Soc.*, **1996**, *118*, 11733.

[39] Ortiz, A.; Pisabarro, M.; Gago, F.; Wade, R.; *J. Med. Chem.*, **1995**, *38*, 2681.

[40] Perez, C.; Pastor, M.; Ortiz, A.; Gago, F.; *J. Med. Chem.*, **1998**, *41*, 836.

[41] Weiner, P.K.; Kollman, P.A.; *J. Comput. Chem.*, **1981**, *2*, 287.

[42] Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, Jr., S.; Weiner, P.; *J. Am. Chem. Soc.*, **1984**, *106*, 765.

[43] Beveridge, D.L.; DiCapua, F.M.; in *Computer simulation of biomolecular systems*, W.F. van Gunsteren,; P.K. Weiner, Eds.; **1989**, 1-26.

[44] Gilson, M.; Given, J.; Bush, B.; McCammon, J.; *Biophysical Journal*, **1997**, *72*, 1047.

[45] van Gunsteren, W.F.; in *Computer simulation of biomolecular systems*, W.F. van Gunsteren; P.K. Weiner, Eds.; **1989**, 27-59.

[46] Kirkwood, J.G.; *J. Chem. Phys.*, **1935**, *3*, 300.

[47] Zwanzig, R.; *J. Chem. Phys.*, **1954**, *31*, 11029.

[48] Gerber, P.R.; Mark, A.E.; van Gunsteren, W.F.; *J. Comput.-Aided Mol. Design*, **1993**, *7*, 305.

[49] Liu, H.; Mark, A.; van GunsterenW.F.; *J. Phys. Chem.*, **1996**, *100*, 9485.

[50] Beutler, T.C.; Mark, A.E.; van Schaik, R.C.; Gerber, P.R.; van Gunsteren, W.F.; *Chemical Physics Letters*, **1994**, *222*, 529.

[51] Hooft, R.; vanEijck, B.; Kroon, J.; *J. Chem. Phys.*, **1992**, *97*, 6690.

[52] Kong, X.andBrooks, C.; *J. Chem. Phys.*, **1996**, *105*, 2414.

[53] Guo, Z.; Brooks, C.; *J. Am. Chem. Soc.*, **1998**, *120*, 1920.

[54] Guo, Z.; Brooks, C.; Kong, X.; *J. Phys. Chem. B*, **1998**, *102*, 2032.

[55] Åqvist, J.; Medina, C.; Samuelsson, J.-E.; *Protein Engineering*, **1994**, *7*, 385.

[56] Hansson, T.; Å qvist, J.; *Protein Engineering*, **1995**, *8*, 1137.

[57] Jones-Hertzog, D.K.; Jorgensen, W.H.; *J. Med. Chem.*, **1996**, *40*, 1539.

[58] Wang, J.; Dixon, R.; Kollman, P.; *Proteins: Structure, Function and Genetics*, **1999**, *34*, 69.

[59] Kuntz, I.D.; *Science*, **1992**, *257*, 1078.

[60] Caflisch, A.; Ehrhardt, C.; in *Structure-based drug design*, P.Veerapandian, Ed.; **1997**, 541-558.

[61] Marrone, T.; Briggs, J.; McCammon, J.; *Ann. Rev. Pharmacol. Toxicol.*, **1997**, *37*, 71.

[62] Kubinyi, H.; *Curr. Opin. Drug Design Discov.*, **1998**, *1*, 16.

[63] Pearlman, D.; Murcko, M.; *J. Comput. Chem.*, **1993**, *14*, 1184.

[64] Pearlman, D.A.; Murcko, M.A.; *J. Med. Chem.*, **1996**, *39*, 1651.

[65] Bohacek, R.S.; McMartin, C.; *J. Am. Chem. Soc.*, **1994**, *116*, 5560.

[66] Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R.; Ferrin, T.E.; *J. Mol. Biol.*, **1982**, *161*, 269.

[67] Oshiro, C.M.; Kuntz, I.D.; Dixon, J.S.; J. *Comput.-Aided Mol. Design*, **1995**, *9*, 113.

[68] Meng, E.C.; Shoichet, B.K.; Kuntz, I.D.; *J. Comput. Chem.*, **1992**, *13*, 505.

[69] DesJarlais, R.L.; Seibel, G.L.; Kuntz, I.D.; Furth, P.S.; Alvarez, J.C.; Ortiz de Montellano, P.R.; DeCamp, D.L.; Babe, L.M.; Craik, C.S.; *Proc. Natl. Acad. Sci. USA.*, **1990**, *87*, 6644.

[70] Shoichet, B.K.; Stroud, R.M.; Santi, D.V.; Kuntz, I.D.; Perry, K.M.; *Science*, **1993**, *259*, 1445.

[71] Sun, Y.; Ewing, T.; Skillman, A.; Kuntz, I.D.; *J. Comput.-Aided Mol. Design*, **1998**, *12*, 597.

[72] Rarey, M.; Kramer, B.; Lengauer, T.; *Proteins: Structure, Function and Genetics*, **1999**, *34*, 17.

[73] Judson, R.S.; Tan, Y.T.; Mori, E.; Melius, C.; Jaeger, E.P.; Treasurywala, A.M.; Mathiowetz, A.; *J. Comput. Chem.*, **1995**, *16*, 1405.

[74] Clark, K.P.; Ajay; *J. Comput. Chem.*, **1995**, *16*, 1210.

[75] Jiang, F.; Kim, S.-H.; *J. Mol. Biol.*, **1991**, *219*, 79.

[76] Leach, A.R.; *J. Mol. Biol.*, **1994**, *235*, 345.

[77] Zacharias, M.; Luty, B.A.; Davis, M.; McCammon, J. A.,; *J. Mol. Biol.*, **1994**, *238*, 455.

[78] Elber, R.; Karplus, M.; *J. Am. Chem. Soc.*, **1990**, *112*, 9161.

[79] Apostolakis, J.; Plückthun, A.; Caflisch, A.; *J. Comput. Chem.*, **1998**, *19*, 21.

[80] Moon, J.B.; Howe, W.J.; Proteins: Structure, Function and Genetics, **1991**, 11, 314.

[81] Rotstein, S.H.; Murcko, M.A.; *J. Med. Chem.*, **1993**, 36, 1700.

[82] DeWitte, R.; Ishchenko, A.; Shakhnovich, E.; *J. Am. Chem. Soc.*, **1997**, *119*, 4608.

[83] Goodford, P.J.; *J. Med. Chem.*, **1985**, *28*, 849.

[84] Bobbyer, D. N.A.; Goodford, P.J.; McWhinnie, P.M.; Wade, R.C.; *J. Med. Chem.*, **1989**, *32*, 1083.

[85] Miranker, A.; Karplus, M.; *Proteins: Structure, Function and Genetics*, **1991**, *11*, 29.

[86] Caflisch, A.; *J. Comput.-Aided Mol. Design*, **1996**, *10*, 372.

[87] Joseph-McCarthy, D.; Hogle, J.; Karplus, M.; *Proteins: Structure, Function and Genetics*, **1997**, *29*, 32.

[88] Bartlett, P.A.; Shea, G.T.; Telfer, S.J.; Waterman, S.; in *Molecular Recognition in Chemical and Biological Problems*, R.C.S. London, Ed.; vol.*78*, **1989**, 182-196.

[89] Eisen, M.B.; Wiley, D.C.; Karplus, M.; Hubbard, R.E.; *Proteins: Structure, Function and Genetics*, **1994**, *19*, 199.

[90] Böhm, H.J.; *J. Comput.-Aided Mol. Design*, **1992**, *6*, 593.

[91] Böhm, H.J.; *J. Comput.-Aided Mol. Design*, **1994**, *8*, 623.

[92] Kabsch, W.; *Acta Cryst.*, **1976**, *A32*, 922.

[93] Davis, M.E.; McCammon, J.A.; *J. Comput. Chem.*, **1989**, *10*, 386.

[94] Davis, M.E.; McCammon, J.A.; *J. Comput. Chem.*, **1990**, *11*, 401.

[95] Davis, M.E.; McCammon, J.A.; *J. Comput. Chem.*, **1991**, *12*, 909.

[96] Caflisch, A.; Wälchli, R.; Ehrhardt, C.; *News Physiol. Sci.*, **1998**, *13*, 182.

[97] Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhard, C.; Caflisch, A., **1999**, submitted.

[98] Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T.; *J. Am. Chem. Soc.*, **1990**, *112*, 6127.

[99] Scarsi, M.; Apostolakis, J.; Caflisch, A.; *J. Phys. Chem. A*, **1997**, *101*, 8098.

[100] Shoichet, B.K.; Leach, A.R.; Kuntz, I.D.; *Proteins: Structure, Function and Genetics*, **1999**, *34*, 4.

[101] Mark, A.; van Gunsteren, W.; *J. Mol. Biol.*, **1994**, *240*, 167.

[102] Shuker, H.; Hajduk, P.; Meadows, R.; W., F.S.; *Science*, **1996**, *274*, 1531.

[103] Hajduk, P.; Meadows, R.; W., F.S.; *Science*, **1997**, *278*, 497.

[104] Mattos, C.; Ringe, D.; *Nature Biotechnology*, **1996**, *14*, 595.

[105] Weber, L.; Wallbaum, S.; Broger, C.; Gubernator, K.; *Angew. Chem. Int. Ed.*, **1995**, *34*, 2280.

[106] Brünger, A.T.; Kuriyan, J.; Karplus, M.; *Science*, **1987**, *235*, 458.

[107] Veerapandian, P.; *Structure-based drug design*; Marcel Dekker Inc., **1997**.

[108] Sanchez, R.; Šali, A.; *Proc. Natl. Acad. Sci. USA.*, **1998**, *95*, 13597.

Received 26 February, 1999 Accepted 26 March, 1999