DrugScore Meets CoMFA: Adaptation of Fields for Molecular Comparison (AFMoC) or How to Tailor Knowledge-Based Pair-Potentials to a Particular Protein

Holger Gohlke and Gerhard Klebe*

Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

Received January 3, 2002

The development of a new tailor-made scoring function to predict binding affinities of protein−ligand complexes is described. Knowledge-based pair-potentials are specifically adapted to a particular protein by considering additional ligand-based information. The formalism applied to derive the new function is similar to the well-known CoMFA approach, however, the fields used in the approach originate from the protein environment (and not from the aligned ligands as in CoMFA, thus, a "reverse" CoMFA (= AFMoC) named Adaptation of Fields for Molecular Comparison is performed). A regular-spaced grid is placed into the binding site and knowledge-based pair-potentials between protein atoms and ligand atom probes are mapped onto the grid intersections resulting in "potential fields". By multiplying distance-dependent atom-type properties of actual ligands docked into the binding site with the neighboring grid values, "interaction fields" are produced from the original "potential fields". In a PLS analysis, these atom-type specific interaction fields are correlated to the actual binding affinities of the embedded ligands, resulting in individual weighting factors for each field value. As in CoMFA, the results of the analysis can be interpreted in graphical terms by contribution maps, and binding affinities of novel ligands are predicted by applying the derived 3D QSAR equation. The scope of the new method is demonstrated using thermolysin and glycogen phosphorylase *b* as test examples. Impressive improvements of the predictive power for affinity prediction can be achieved compared to the application of the original knowledge-based potentials by considering a sample set of only 15 known training ligands. Thus, with growing information about the drug target studied, the new method allows one to move gradually from generally valid to protein-specifically adapted pair-potentials, depending on the amount of training information available and its degree of structural diversity. In addition, convincing predictive power is also achieved for ligand poses generated by automatic docking tools.

Introduction

As direct consequence of the ongoing post-sequencing efforts,[1] a tremendous increase of structurally characterized drug targets will be discovered. Accordingly, structure-based drug design techniques, in particular virtual screening, will be of steadily increasing importance. As a major advantage compared to experimental high-throughput screening, they are, by far, both less time-consuming and expensive. In addition, new insights acquired during an ongoing drug design project can be steadily fed back into every new design cycle, allowing one to exploit as much information as is available on each single step. However, to be successful as a strategy for lead finding and optimization, in silico structure-based approaches must generate and identify relevant binding modes and predict accurately binding affinities of potential drug candidates. Both aspects strongly depend on our understanding of the determinants of protein−ligand binding and how we translate this knowledge back into computer methods to score protein−ligand interactions.[2−6]

Over recent years, a broad spectrum of competitive methods for scoring protein−ligand interactions has emerged.[7] Established approaches have been further improved, e.g., in the area of the regression-based scoring functions[8−10] or methods based on first principles.[11−13] Furthermore, well-known techniques have been applied to protein−ligand scoring by using atom−atom contact potentials to develop knowledge-based scoring functions.[14−16] As a pragmatic alternative, existing scoring functions have been combined for intersection-based consensus scoring that allows one either to enhance enrichment rates in virtual screening[17,18] or to improve binding affinity predictions.[19−21]

Besides solely exploiting the structural information of a particular protein−ligand complex, a more convincing approach would be to include additional structural *and* energetic information about already known ligands. As a first step toward such a strategy, the consecutive or iterative application of tools for molecular comparison along with docking has been suggested.[22] A set of putative candidate molecules has been prescreened for similarity with known inhibitors of carbonic anhydrase II using the small molecule superpositioning program FlexS.[23] Subsequently, the candidates ranked most similar to the references have been docked using

* To whom correspondence should be addressed. Tel.: (+49) 6421 2821312. Fax: (+49) 6421 2828994. E-mail: klebe@mailer.uni-marburg.de.

10.1021/jm020808p CCC: $22.00 © 2002 American Chemical Society
Published on Web 08/15/2002

FlexX.[24] Using this procedure, subnanomolar inhibitors have been discovered. The idea of considering ligand information to improve the model building of proteins by homology has recently been described.[25] In turn, such information can also be used for protein−ligand docking.[26−28]

To develop so-called tailor-made scoring functions specifically adapted to predict binding affinities with respect to one particular protein, *energetic* information about already known ligands is evaluated in addition to *structural* information as described in the preceding paragraph. An early approach along these lines has been introduced by Holloway et al.[29] who correlated intermolecular force field energies calculated for HIV-1 protease inhibitor complexes using the MM2X force field with observed pIC_{50} values. Instead of correlating the *total* computed binding energy to the experimentally determined binding free energy, Wade et al. used selected interaction energy *components*. They are derived from interactions between inhibitors and individual protein residues (including waters) and correlated to binding free energies via PLS analyses (COMBINE method).[30,31] Similar approaches have also been described.[32−36] Considering only energetic components and neglecting entropic contributions can only be applied in the case of ligands where entropic changes upon binding do not matter.[37,38] To overcome this limitation, several attempts to combine force field derived contributions with terms to describe changes in solvation and conformational degrees of freedom upon binding have emerged. In these studies, either a linear regression or PLS analysis has been used to correlate experimental binding affinities with computed properties.[34,39−45]

The methods described in the last paragraph suffer however from two shortcomings. First, to obtain high-quality correlations, a large data set of structurally diverse training compounds spreading over a sufficient range of binding affinities is necessary. This prerequisite is difficult to accomplish at the beginning of a drug design project. Second, the translation of the results of a regression model into suggestions for how to modify a particular ligand structure to improve its binding affinity is by no means straightforward. Usually, rather general information such as "decrease the van der Waals repulsion" or "increase the solvent accessible surface buried upon ligand binding" is not very conclusive with respect to the design process. Murray et al.[46] introduced an interesting alternative to partially circumvent the first limitation. They adapted a previously published general scoring function[47] to a set of thrombin complexes under the restraints of the earlier (diverse) training set using Bayesian regression. This allows for fine-tuning of the function to either general or specific purposes. COMBINE and related approaches provide at least some insight for how to translate scoring results into structural design aspects.[30,33,40]

Interestingly, the field of drug design knows alternative techniques to score and predict binding affinities within a set of ligands usually with a surprisingly high predictive power. These are the different variants[48] of comparative molecular field analyses (CoMFA).[49] Furthermore, the results of these analyses can be translated rather intuitively into ligand design aspects. Once a particular alignment of ligands is given, the approach maps binding affinity differences in terms of molecular field differences calculated by the use of various probes placed at the intersections of a regularly spaced grid. The actual correlation and predictive QSAR equation is extracted using PLS analysis. Originally, 3D QSAR was developed for design problems where the structure of the target protein is unknown. However, increasingly, it is also applied to cases where the corresponding receptor structure has been determined. There, usually, the protein information is only considered to suggest a meaningful ligand alignment; however, in the potential field calculations, information about the surrounding protein environment is, rather unreasonably, neglected.

Recently, we developed the knowledge-based scoring function DrugScore[14] by extracting structural information from crystallographically determined protein−ligand complexes and converting them into distance-dependent pair-potentials and solvent accessible surface-dependent singlet-potentials. This function reliably recognizes correct binding modes and predicts binding affinities. Furthermore, it can be used to reliably predict which type of ligand atom would most favorably bind at a given site in a protein binding pocket ("hot spot" analysis).[50] Thus, the pair-potentials used for scoring ligand affinities can also be applied to support the actual design and ligand optimization process.

Regarding the strength of both methods, the knowledge-based potentials to score binding modes using protein information and the comparative molecular field analyses to map differences in ligand binding data prompted us to develop a novel approach to adapt knowledge-based potentials specifically to one protein by considering ligand-based information in a CoMFA-type approach. As a result, both a tailor-made scoring function and a protein-based CoMFA are developed. Since the interaction fields used in our approach have their origin in the protein environment, i.e., *reverse* to CoMFA-type fields, we anticipate to call our method AFMoC (Adaptation of Fields for Molecular Comparison). The approach allows us to gradually move from general knowledge-based potentials to protein-specifically adapted ones regarding the amount of ligand data available for training. It thus overcomes the prerequisite to involve complete ligand training sets. In addition, since interaction fields based on atom types are used, interpreting the PLS results in terms of ligand structure optimization to achieve better binding affinities is straightforward. Also, the information contained in the different atom-type based interaction fields is mutually orthogonal. This is an important advantage over the information comprised by generic fields such as "van der Waals" or "electrostatic" interaction used in CoMFA. Finally, since structural information of experimentally determined complexes is converted into statistical pair-potentials, the latter (and the derived interaction fields) are expected to contain besides enthalpic also entropic effects, resulting from (de-)solvation.[14]

First, we will describe the theoretical background of the new AFMoC method followed by the application to two data sets composed by thermolysin and glycogen phosphorylase *b* inhibitors. Finally, the impact of an improved scoring function together with suggestions for an improved ligand design will be given.

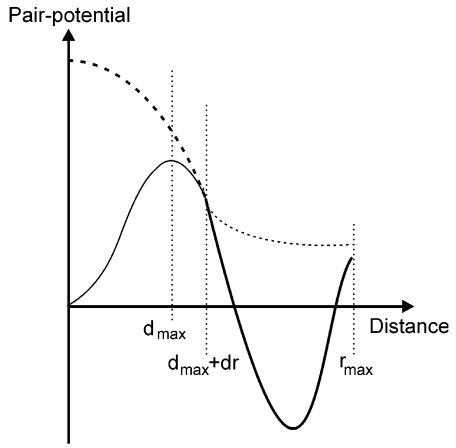


Figure 1. Schematic representation of the Gaussian-type repulsion term (broken line) added to the DrugScore pair-potentials (solid line). For distances between 0 and $d_{max} + dr$, values from the Gaussian are used whereas for distances between $d_{max} + dr$ and r_{max}, values of the pair-potentials are applied (bold lines). Regions of both functions not considered in functional dependency are drawn as thin lines.

Theory

The distance-dependent pair-potentials of mutual atom−atom contacts used in DrugScore to score protein−ligand interactions were derived from a database of structurally known protein−ligand complexes.[14] In case of a *sufficient* number of *different* complexes, it is anticipated that the contact pair distributions observed in the crystalline state also represent these interactions under general conditions. Similarly, the pair-potentials derived from these distributions contain generally valid information about protein−ligand interactions.[51,52] However, as a consequence of their origin, the pair-potentials only depict an *averaged* representation of these interactions. Situations that deviate from this "average" are possibly not treated adequately. In addition, the pair-potentials are used assuming that the distributions of intermolecular distances (and hence the pair-potentials) are similar even for molecular environments deviating from a pair of atoms under investigation. Using additional *structural* and *energetic* information about the special conditions present in one particular protein would enable us to transform these averaged (yet generally valid) potentials into *protein-specifically* adapted ones.

In the original CoMFA, a unique probe is placed to all intersections of a regularly spaced grid. Next the ligands to be analyzed are embedded into this grid and, by using some functional form, interaction energies are calculated between probe atoms and ligand atoms. In CoMFA Lennard-Jones and Coulomb interactions are computed. Equally well-known other functional forms can be used, e.g., potentials taken from GRID[53] or Gaussian-type similarity indices as in CoMSIA,[54] or, as in the present study, knowledge-based pair-potentials. Here, however, instead of placing a uniform probe to all intersections, the protein environment is considered by placing the grid into the binding site and mapping the pair-potentials between protein and ligand atom probes onto the grid intersections (Figure 2), resulting in "potential fields". By "multiplying" atom-type properties of ligands docked into the binding site with neighboring grid values, "interaction fields" are produced from the "potential fields". As in usual CoMFA in a second step, a PLS analysis[55,56] is performed to correlate these fields with the actual binding affinities of the embedded ligands. This results in individual weighting factors for each field value. The information contained in these tailor-made fields can be translated into ideas for how to modify a particular ligand to optimize its binding. Finally, the original "interaction fields" as well as the specifically adapted ones can be used to score, e.g., the binding modes of novel ligands. These three steps will be described in more detail in the following paragraphs.

Calculation of Interaction Fields. Inside the binding pocket of a protein a cubic grid G is constructed. The interaction ΔP between a *probe atom* of ligand atom type $t \in T^{PLS}$ (where T^{PLS} is a subset of all available ligand atom types T parametrized in the DrugScore pair-potentials[14]) located at grid point g and the surrounding protein atoms $p \in P$ of type T(p) is calculated using the distance-dependent pair-potentials of DrugScore $\Delta W_{t,T(p)}(r_{g,p})$ (eq 1). Here, $r_{g,p}$ is the distance between g and p.

$$\forall\, t \in T^{PLS}, \forall\, g \in G:\ \Delta P_t^g = \sum_{p \in P} \Delta W_{t,T}(p)(r_{g,p}) \quad (1)$$

So far, this procedure is identical to the previously described "hot spots" analysis.[50] In addition, to be able to calculate interactions even in the vicinity of protein atoms where the knowledge-based potentials cannot be defined due to missing experimental contact data, an artificial repulsion term is added to each pair-potential. It is computed using a Gaussian function centered at the origin of the pair-potential (Figure 1). The half-width of this function has been adjusted such that its value and gradient coincide with those of the potential at a distance $d_{max} + dr$, with $dr = 0.1$ Å.[14] d_{max} corresponds to the position of the first maximum of a pair-potential with respect to the origin. Thus, for distances $r_{g,p}$ between 0 and $d_{max} + dr$, the values of the repulsion function are used, whereas for distances larger than $d_{max} + dr$, the values of the original pair-potentials are applied. Fields calculated following this protocol will be referred to as "potential fields".

In the following step, the interactions of the protein with a docked ligand in the binding pocket are mapped onto neighboring grid points by multiplying the potential field value ΔP_t^g for ligand atom type t at grid point g with the sum of distance-dependent contributions $B(r_{g,l})$ of ligand atoms $l \in L$ of type T(l) = t (eq 2).

$$\forall\, t \in T^{PLS}, \forall\, g \in G:\ \Delta W_t^g = \Delta P_t^g \sum_{l \in L:T(l)=t} B(r_{g,l}) \quad (2)$$

The distance-dependent contribution $B(r_{g,l})$ is calculated using a Gaussian function, and it is normalized with respect to the sum of the contributions to all grid points (eq 3). We selected a Gaussian-type distance dependence similarly to the CoMSIA approach since the Gaussian-type functional form is well suited to adjust and distribute the contribution of ligand atoms to the neighboring grid points.

Thus, for k different ligands, $k \cdot ||T^{PLS}||$ different interactions fields are obtained. Here, $||T^{PLS}||$ represents the number of distinct atom types in the set T^{PLS}. By multiplying "potential field" values with "properties"

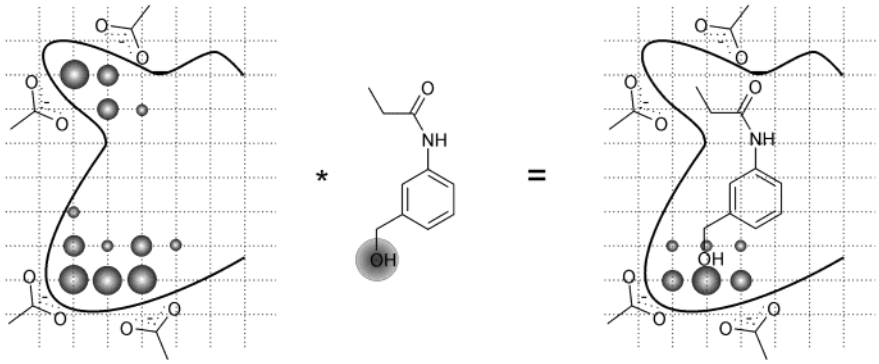


Figure 2. Scheme to explain the calculation of "interaction fields". Multiplying the "potential field" values (eq 1) for ligand atom type "hydroxyl oxygen" (left) with values of a Gaussian function centered at the position of a ligand atom (eq 3) (center), pair interactions between a ligand and a protein atom are mapped onto the neighboring grid intersections (eq 2) (right). For clarity, the binding pocket (left) and the ligand (center) are depicted separately.

$$B(r_{g,l}) = \frac{\exp\left\{-\frac{r_{g,l}^2}{2\sigma^2}\right\}}{\sum_{g \in G} \exp\left\{-\frac{r_{g,l}^2}{2\sigma^2}\right\}} \qquad (3)$$

$\sum_{l \in L:\mathrm{T}(l)=t} B(r_{g,l})$, fields are produced that may be considered as "interaction fields" or "energy fields". Together with eq 1, the sum of ΔW_t^g over all grid points $g \in G$ results in the pair-potential contributions of all ligand atoms of type t with the surrounding protein atoms. They are obtained as distance-dependent weighted average of the potential values ΔP_t^g at grid points surrounding the ligand atoms. For the extreme case of a grid with zero spacing and a Delta function instead of the Gaussian function (i.e., $\sigma = 0$ in eq 3), this average value becomes equal to the sum over the appropriate contributions of the pair-potentials directly evaluated between ligand and protein atoms.

Correlating Interaction Field Values with Binding Affinities. The values of the "interaction fields" calculated for a set of ligands docked into the binding pocket of one protein could, in principle, be correlated with experimentally determined binding affinities using PLS[55,56] similar to CoMFA.[49] However, it has to be considered that the interaction fields are atom-type specific, rather than generic as in the related CoMFA or CoMSIA[54] approaches. As a consequence, "interaction fields" for atom types $t \in T^{\mathrm{Pair}}$ (which are contained in T but not in T^{PLS}) rarely present in the ligand training set have to be excluded for statistical insignificance or their minor contributions as "two-/ three-/ four-/ .../ level variables" in the subsequent regression analysis. Yet, contributions of these atoms with respect to binding affinity must be taken into account.

Recently, we have shown that scoring values obtained by DrugScore for protein−ligand complexes can be scaled to experimentally determined pK_i values.[50] In addition, the obtained scaling factor c_S^{Pair} can be transferred reliably between sets of different protein−ligand complexes. Hence, the contribution $\mathrm{p}K_i^{\mathrm{Pair}}$ to the binding affinity due to all ligand atoms of type $t \in T^{\mathrm{Pair}}$ can be determined. Subtracting $\mathrm{p}K_i^{\mathrm{Pair}}$ from the experimental $\mathrm{p}K_i$ value reveals the affinity contribution arising from all ligand atoms of type $t \in T^{\mathrm{PLS}}$, denoted as $\mathrm{p}K_i^{\mathrm{PLS}}$. The latter is then correlated to the "interaction field" values by PLS, which results in weighting factors c_t^g (termed "coefficients" in the QSAR equation) for every grid point g and every ligand atom type $t \in T^{\mathrm{PLS}}$.

This concept is based on two assumptions: (i) the contribution $\mathrm{p}K_i^{\mathrm{Pair}}$ to the overall $\mathrm{p}K_i$ value will be of minor importance if the number of ligand atoms with type $t \in T^{\mathrm{Pair}}$ is small compared to the number of ligand atoms with atom type $t \in T^{\mathrm{PLS}}$; (ii) estimating $\mathrm{p}K_i^{\mathrm{Pair}}$ using the DrugScore pair-potentials applying a common c_S^{Pair} will be sufficiently accurate.

Predicting Binding Affinities while Gradually Switching between Original and Adapted Interaction Fields. Multiplication of the "potential field" values ΔP_t^g with the weighting factors c_t^g at every grid point g yields the protein-specifically adapted potential fields ΔP_t^{*g}. They are used to calculate the binding affinity contribution $\mathrm{p}K_i^{\mathrm{PLS}}$ of a ligand by summing over all grid points and ligand atoms with types $t \in T^{\mathrm{PLS}}$ (eqs 2 and 3). Of course, for the ΔP_t^g values the same centering and scaling has to be applied as in the PLS prior to multiplication with the weighting factors c_t^g and the intercept of the QSAR equation needs to be considered. Finally, the contribution $\mathrm{p}K_i^{\mathrm{Pair}}$ of atoms with types $t \in T^{\mathrm{Pair}}$ must be calculated (see above). Due to the latter step, also ligands with atom types not contained in the training set can be considered in the affinity prediction.

In the case of small data sets of closely related ligands or affinity data used for adaptation not well spread over at least 3 orders of magnitude,[57] the obtained PLS results will inevitably suffer from large prediction errors for statistical reasons. In such situations, likely the original pair-potentials in DrugScore reveal more reliable predictions. With increasing importance and coverage of the added structural and energetic information comprised in the training set, this situation will change. The gradually enhanced consideration of specifically adapted potentials replacing the original ones appears to be the strategy of choice. Optionally, we "mix" linearly the original (ΔP_t^g) and adapted (and centered and scaled) potential fields (ΔP_t^{*g}) to reveal better affinity predictions (eq 4). The "mixing coefficient" θ is allowed to vary between 0 and 1. As a rule of thumb, small values for θ should be selected if molecules in the

training set reflect little additional information or the molecules to be predicted significantly deviate from those in the original training set.

$$pK_i^{PLS} = \sum_{g \in G} \sum_{\substack{l \in L: \\ T(l) \in T^{PLS}}} B(r_{g,l})[\theta \Delta P^{*g}_{T(l)} + (1 - \theta) c_S^{Pair} \Delta P^g_{T(l)}] + \theta \langle pK_{i,train}^{PLS} \rangle \quad (4)$$

Here, $\langle pK_{i,train}^{PLS} \rangle$ is the average of all pK_i^{PLS} values of the training set and equals the intercept of the QSAR equation obtained by PLS.

Novel compounds that are structurally similar to the members of the training set are docked into the binding pocket following the same protocol as applied to the training molecules. Equation 4, eventually complemented by a pK_i^{Pair} contribution, then yields an estimate of the binding affinity. However, if the prediction for molecules with structurally strongly deviating skeletons is attempted (e.g., in virtual screening), two limitations must be reflected. First, ligand portions that exceed beyond the scope of the "trained" grid will be scored based on the original DrugScore pair-potentials. Similarly, if ligand portions coincide with regions of the grids where the variance of information used for training has been too scarce to be considered in PLS (see Methods), we switch back to the original potential field values ΔP_t^g for scoring, assuming that original DrugScore performs better than simply ignoring such situations.

Methods

Data Sets and Alignments. The AFMoC approach has been validated using two data sets of thermolysin and glycogen phosphorylase *b* inhibitors, respectively. The thermolysin set comprises 61 compounds for training and 15 compounds for testing, taken from Klebe et al. (Table 4 in ref 54; see also Supporting Information). The experimentally determined inhibition constants are used as pK_i values and spread over a satisfactorily large range of 9.7 log units for the training compounds. The data set of glycogen phosphorylase inhibitors was compiled from studies of Cruciani et al.,[58] Pastor et al.,[59] and Hopfinger et al.[40,60] Comparing pK_i values of compounds that are listed in more than one reference ensured the consistency of this data set. It contains 58 compounds for training and 8 compounds for testing. The structures of these glucose derivatives are summarized in Table 1 and Table 2 together with the experimentally determined pK_i values. They cover a range of 4.7 log units in case of the training compounds. For all compounds, standard protonation states are assumed, i.e., carboxylate and phosphate groups are considered to be deprotonated, and aliphatic amino groups are considered to be protonated.

In both cases, a protein-based alignment was applied to obtain a consistent superimposition of all inhibitor molecules. The binding modes of 8 of the 76 thermolysin inhibitors have been determined crystallographically (Protein Data Base (PDB)[61] codes: 1tlp, 1tmn, 2tmn, 4tln, 4tmn, 5tln, 5tmn, 6tmn). Superimposing the remaining seven protein chains onto 1tlp yielded a mean *rmsd* of only 0.14 ± 0.05 Å. A similar value was obtained considering only the binding site residues. These crystal structures were used as templates to construct and superimpose the remaining thermolysin inhibitors. The protein environment, including Zn, as given in 1tlp was used to calculate the potential fields (see below). Invariant molecular portions or interaction patterns present in the various ligands of the set were superimposed onto the crystallographic references followed by a minimization inside the rigid binding pocket of 1tlp (Figure 3). The MAB force field as implemented in MOLOC was used.[62] For the glycogen phosphorylase inhibitors, a similar protocol has been applied. Here, the protein−ligand complexes 1a8i, 1axr, 1b4d, 1e1y, 1noj, 1nok, 2gpb, 2prj, 3gpb, 4gpb, 5gpb, and 6gpb were taken as templates. Superimposing all protein chains onto 2gpb indicated intrinsic rigidity of this binding pocket (mean *rmsd* = 0.43 ± 0.16 Å) except for Lys574, which exhibits slight induced fit motions. Once constructed all ligands were minimized in the rigid binding pocket of 2gpb with the MAB force field (Figure 3). The protein structure of 2gpb was also used for the calculation of potential fields (see below).

AFMoC Analysis. Interaction fields for the superimposed thermolysin and glycogen phosphorylase *b* inhibitors were calculated as described (Theory section (eqs 1−3)) using the protein structures 1tlp and 2gbp as reference, respectively. The size of the grid boxes was selected in a way that all inhibitor molecules were sufficiently embedded with a margin of at least 4 Å. To study the influence of grid spacing on PLS performance, values of 0.5, 1.0, 1.5, and 2.0 Å were tested. However, if not stated differently, results are reported for 1.0 Å spacing. Table 3 summarizes details about the grids used in the analyses.

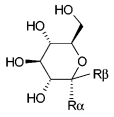
The repulsive properties of the function added to the pair-potentials at short atom−atom distances influences interaction field values at grid points close to the protein surface. A (dimensionless) value of 10 has been selected for the height of the Gaussian repulsion function at the origin of an atom−atom contact. This value has been adjusted to be of the same order of magnitude than the largest values of the pair-potentials. Variations of the steepness of the repulsion function did not significantly alter the results.

To estimate pK_i^{Pair} values by scaling scoring values calculated with DrugScore pair-potentials, the coefficient $c_S^{Pair} = -3.11 \times 10^{-2}$ was determined as average over scaling factors obtained for data sets of serine proteases, metallo proteases, endothiapepsines, and two mixed data sets ("others" and "Böhm1998"; see ref 50 for details). Of course, experimentally determined pK_i values of these complexes were correlated with scoring values computed only with pair-potentials.

To correlate the individual interaction fields with binding affinities, statistical analyses were performed using our own in-house implementation of the (SAM)PLS algorithms.[63] According to the described fitting algorithm, the results obtained from eq 12′ and 16 in the cited study have also been centered in our case. Furthermore, the columns of the descriptor matrix have been centered on input. However, since these ΔW_t^g values are obtained applying DrugScore pair-potentials (eq 2), which were already mutually weighted,[14] neither scaling of the columns nor of the interaction fields in total is performed. Test calculations applying auto scaling or block scaling resulted in AFMoC models of reduced predictive power. To check statistical significance of the PLS models, cross-validation runs were performed by means of the "leave-one-out" (LOO) procedure using the enhanced SAMPLS method. Thereby, following recommendations of Wold,[56] Kubinyi et al.,[64] and Bush et al.,[63] the columns of the descriptor matrix were recentered for every new cross-validation run. The optimal number of components has been determined by selecting the smallest s_{PRESS} value. The same number of components was subsequently used to derive the final QSAR models. For all conventional analyses (no cross-validation), the "minimum-sigma" standard deviation threshold was set to 5×10^{-4}. This value was adjusted in a way that approximately 90% of all columns of the descriptor matrix were eliminated during PLS analysis. The statistical results are summarized in Table 4. The q^2, s_{PRESS}, r^2, S, and *contribution* values were computed as defined in ref 49. Plots of predicted versus actual binding affinities for the fitted PLS analyses are shown in Figure 4.

To apply an even more rigorous statistical test, several runs of a "leave-five-out" procedure have been performed using models based on the 1.0 Å grid spacing. For this, five

Table 1. Glucose Analog Inihibitors[a] of Glycogen Phosphorylase *b* Used as Training Set[40,58−60]



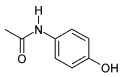
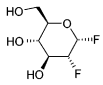
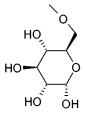
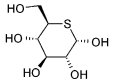
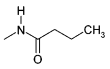
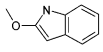
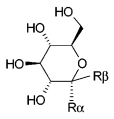
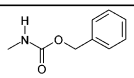
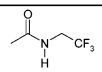
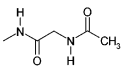
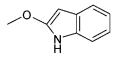
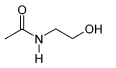
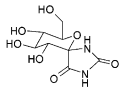
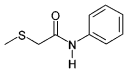
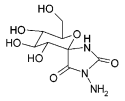
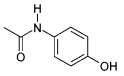
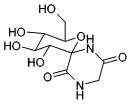
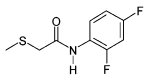
No.	Ref.	Substituent at C1 position Rα	Rβ	pK_i	No.	Ref.	Substituent at C1 position Rα	Rβ	pK_i
1	58		H	2.8	**27**	58	H		3.4
2	58		H	1.3					
3	58		H	1.9	**28**	58		H	1.9
4	58	H		1.6					
5	58	H		1.6	**29**	58		H	1.4
6	58		H	1.6	**30**	58		H	2.3
7	58	H		1.8					
8	58			2.1	**31**	58		H	1.6
9	58		H	2.8					
10	58	H		1.7	**32**[a)]	58			3.7
11	58			1.8					
12	58	H		2.3	**33**	58	H		2.6
13	58			2.4	**34**	58	H		3.7
14	58		H	1.5	**35**	58	H		3.4
15	58	H		1.8	**36**	58	H		4.4
16	58	H		2.3	**37**	40	H		4.5
17	58	H		2.0					
18	58			2.1	**38**	40	H		4.4
19	58		H	1.8	**39**	40	H		4.4
20	58		H	3.4	**40**	40	H		4.1
21[a)]	58			2.7	**41**	40	H		4.0
22	58	H		3.0	**42**	40	H		3.9
23	58		H	2.1	**43**	40	H		3.4
24	58		H	2.6	**44**	40	H		2.3
25	58		H	2.5	**45**	60			4.8
26	58	H		1.7					

Table 1. (Continued)



No.	Ref.	Substituent at C1 position Rα	Rβ	pK_i	No.	Ref.	Substituent at C1 position Rα	Rβ	pK_i
46	60	H		3.5	**53**	59	H		2.1
47	60	H		3.0	**54**	59	H		2.9
48	60	H		2.7	**55**	59	H		2.6
49	60	H		2.6	**56**[a)]	59			5.5
50	60	H		2.4	**57**[a)]	59			6.8
51	60	H		2.4	**58**[a)]	59			4.2
52	60	H		1.7					

[a] The whole ligand is depicted.

arbitrarily selected compounds were discarded from the training set and used for prediction. Statistical results are shown in Table 5.

Furthermore, to detect possible chance correlations, the biological data were scrambled randomly and model calculations were repeated. Only negative q^2 values were obtained in such cases (Table 6), thus supporting the assumption that proper correlations with the original biological data are given.

To distribute the atomic protein-to-ligand interactions over neighboring grid points, a Gaussian function has been introduced (eqs 2 and 3). The half-width of this function determines the "local smearing" of these interactions: smaller σ values result in a more locally restricted mapping, whereas larger σ values tend to average adjacent interactions. To study the importance of this feature, σ values of 0.55, 0.7, ..., 1.3 Å were used to calculate interaction fields (Figure 5). In all other calculations, a σ value of 0.7 (thermolysin data) and 0.85 (glycogen phosphorylase *b* data) has been applied.

Since the interaction fields obtained by mapping protein−ligand interactions onto neighboring grid points are (ligand) atom-type specific, various combinations can be used to build up the descriptor matrix as input for PLS analysis. To check how much the individual fields contribute to the derived model, we started with a combination of five fields representing to our best knowledge interaction differences of the considered ligands. We subsequently added further fields until no significant improvement in q^2 values could be achieved. Table 7 displays the results for thermolysin and glycogen phosphorylase *b*, respectively. In bold, the field combinations are indicated which were selected for all remaining calculations throughout this study.

To investigate by how much the statistical results depend on the relative orientation of the molecules of the data set with respect to the lattice,[65] the entire data sets were translated with respect to the grid origin. The original grid boxes (1.0 and 2.0 Å lattice spacing) were extended by 1.0 and 2.0 Å in *x*, *y*, and *z* directions, respectively, to guarantee a sufficiently large margin embedding all molecules during translation. The data sets were then translated along the *xyz* diagonal from 0 to 1.0 or 2.0 Å in steps of $^1/_4$, $^1/_2$, and $^3/_4$ of the grid spacing. For all orientations produced, interaction fields were calculated and PLS analyses were performed. The obtained q^2 values are shown in Figure 6 for both data sets.

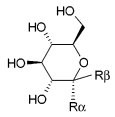
To interpret the AFMoC results graphically, at every grid point the product of the standard deviation of the considered interaction field values times the coefficient of the QSAR equation has been calculated ("STDEV*COEFF" field) and stored for contouring within SYBYL.[66] In an iterative manner, appropriate contour levels were selected to display best interpretable contour maps (Figure 9 and Figure 10).

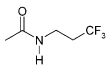
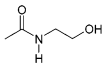
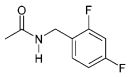
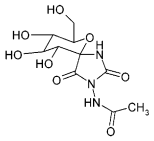
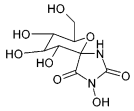
For comparison, CoMFA analyses were performed as implemented in SYBYL[66] for both data sets using steric and electrostatic fields. Partial atomic charges were computed using the AM1 Hamiltonian[67] within MOPAC 6.0.[68] All CoMFA calculations were done with SYBYL standard parameters using a sp^3 carbon probe atom with a charge of +1.0.

Prediction of Binding Affinities. To estimate the predictive power of the AFMoC models, 15 thermolysin and 8 glycogen phosphorylase inhibitors were selected as test sets (see above). Following the same protocol as for the training compounds, the molecules were manually docked into the binding pockets. Predictions were then performed using the AFMoC models based on the 1 Å grid spacing and compared to the results based on a prediction using the original DrugScore pair-potentials. Statistical results are given in Table 8, and plots of predicted versus experimentally determined binding affinities are shown in Figure 7.

To avoid large prediction errors in case of training sets comprised by only a small number of compounds or data sets composed by closely related molecules, contributions to binding affinity based on the original DrugScore potentials can be mixed with those from the individually adapted AFMoC fields (eq 4). To estimate the predictive power of our approach in situations where only a reduced number of reference compounds is available, 100 times a subset of 5, 10, 15, 30, and 45 compounds were randomly selected from the training set

Table 2. Glucose Analog Inihibitors[a] of Glycogen Phosphorylase *b* Used as Test Set[40,59,60]



No.	Ref.	Substituent at C1 position Rα	Rβ	pK_i
59	40	H		2.1
60	60	H		2.1
61	60		H	1.8
62	60		H	1.8
63	60		H	1.6
64	59	H		4.1
65[a)]	59			3.3
66[a)]	59			4.4

[a] The whole ligand is depicted.

of thermolysin inhibitors. For these 5 × 100 selections, PLS analyses were performed. Finally, the 500 AFMoC models obtained were used together with the original DrugScore potential fields to predict binding affinities for the 15 test set compounds. To achieve a mixing of the fields, the parameter θ (eq 4) was varied between 0 (solely DrugScore pair-potential fields) and 1 (solely AFMoC fields) in steps of 0.1; r^2 values were averaged over all 100 runs, respectively. In addition, the model based on all 61 training compounds has been regarded. Results are given in Figure 8 (left). With the glycogen phosphorylase inhibitors, an analogous procedure has been performed (Figure 8, right). Ligand coordinates of all molecules used in the analyses can be obtained from the authors in SYBYL molfile format upon request.

To use specifically adapted AFMoC fields in virtual screening, the method must be robust enough not only to predict binding affinities reliably for manually placed ligands but also to handle correctly ligand binding modes generated by automatic docking. Therefore, we generated up to 500 ligand poses for each test set compound with FlexX. During the incremental build-up procedure and the final scoring FlexX used DrugScore pair-potentials. For the glycogen phosphorylase inhibitors, default parameters were applied, whereas the higher flexibility of the thermolysin ligands afforded enhanced sampling. Finally, all generated ligand binding modes were rescored using the AFMoC models based on 61 or 58 training compounds. Again, the mixing parameter θ was varied between 0.0 and 1.0 in steps of 0.1. To investigate whether a subsequent geometry optimization of the generated binding modes of thermolysin ligands improves the affinity predictions, all docking solutions were submitted to an energy minimization inside the (rigid) enzyme binding pocket using the MAB force field, restricting the number of iterations to 300. The results are summarized in Table 9.

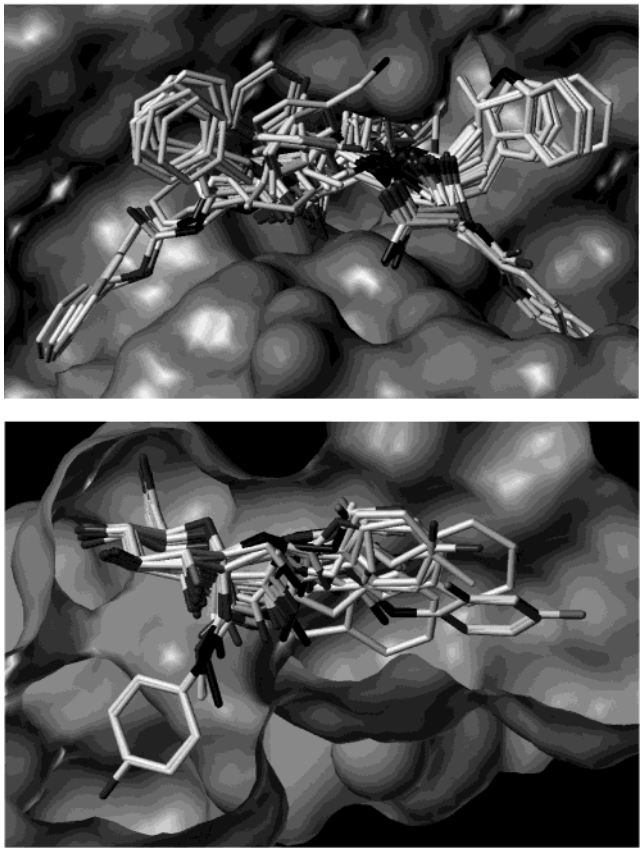


Figure 3. Alignment of 61 thermolysin inhibitors (top) and 58 glycogen phosphorylase *b* inhibitors (bottom) within the binding sites of 1tlp and 2gpb, respectively. The solvent-accessible surface is indicated as a solid surface.

Results and Discussion

Data Sets and Alignment Procedure. To validate the AFMoC approach, two data sets comprising thermolysin and glycogen phosphorylase *b* inhibitors, respectively, were selected. All of the 61 thermolysin compounds applied for training and at least 11 of the 15 test compounds have also been used in the study of Waller and Marshall,[69] De Priest et al.,[70] and Klebe et al.[54,71] for the validation of 3D QSAR approaches. In case of the glycogen phosphorylase *b* inhibitors, combining data from several studies[40,58−60] lead to a data set of 58 training (Table 1) and 8 test compounds (Table 2). In both cases, the pK_i values of the training compounds spread over a range of at least 3 logarithmic units.[57] The analysis of the presently determined crystal structures of protein−ligand complexes showed only minor deviations in the protein structures, demonstrating their rigidity upon ligand binding. Furthermore, all ligands accommodate the same region of the binding pocket.

For the data sets of thermolysin inhibitors, we already reported the prediction of binding affinities. The results showed a moderate predictive power applying DrugScore.[50] In the case of glycogen phosphorylase *b* inhibitors, DrugScore even failed to predict binding affinities (see below). Unsatisfactory predictions have also been reported in other studies for sugar-binding protein−ligand complexes (see Table 5 in ref 15). Thus, both cases are challenging to see whether any tailoring of the DrugScore pair-potentials increases the predictive power.

Table 3. Parameters of the Grid Sizes Used for AFMoC Analyses

	thermolysin				glycogen phosphorylase *b*			
grid spacing[a]	0.5	1.0	1.5	2.0	0.5	1.0	1.5	2.0
X[b]	−7.19 (46)	−7.19 (24)	−7.19 (16)	−7.19 (13)	−10.11 (40)	−10.11 (21)	−10.11 (14)	−10.11 (11)
Y[b]	−12.72 (45)	−12.72 (23)	−12.72 (16)	−12.72 (12)	−6.44 (32)	−6.44 (17)	−6.44 (12)	−6.44 (9)
Z[b]	−10.69 (37)	−10.69 (19)	−10.69 (13)	−10.69 (10)	−6.27 (41)	−6.27 (21)	−6.27 (15)	−6.27 (11)
no. of grid points	76590	10488	3328	1560	52480	7497	2520	1089

[a] In Å. [b] Given is the coordinate of the lower, left, front corner. In parentheses, the number of grid points in each direction is shown.

Table 4. Statistical Results for AFMoC Analyses

	thermolysin				glycogen phosphorylase *b*			
spacing[a]	0.5	1.0	1.5	2.0	0.5	1.0	1.5	2.0
σ	0.70	0.70	0.70	0.70	0.85	0.85	0.85	0.85
q^{2} [b,c]	0.58 (0.61)	0.59 (0.62)	0.52 (0.56)	0.46 (0.50)	0.53 (0.47)	0.58 (0.52)	0.57 (0.52)	0.48 (0.42)
s_{press} [d,e]	1.41	1.38	1.50	1.51	0.78	0.75	0.78	0.82
r^{2} [b,f]	0.97 (0.97)	0.97 (0.97)	0.95 (0.95)	0.73 (0.75)	0.71 (0.67)	0.79 (0.77)	0.91 (0.90)	0.72 (0.68)
S [d,g]	0.35	0.34	0.46	1.03	0.60	0.50	0.33	0.59
F [b,h]	159.1 (173.4)	165.1 (180.0)	99.2 (108.3)	37.4 (41.9)	66.4 (55.9)	51.5 (44.2)	74.0 (64.9)	45.9 (38.7)
comp[i]	10	10	9	4	2	4	7	3
fraction	0.35	0.36	0.37	0.36	0.36	0.36	0.40	0.35
C.3								
C.2					0.24	0.21	0.13	0.22
C.ar	0.31	0.31	0.32	0.35	0.13	0.18	0.17	0.13
O.3	0.08	0.08	0.09	0.10	0.08	0.07	0.10	0.11
O.2	0.08	0.07	0.07	0.07	0.06	0.06	0.08	0.06
O.co2	0.07	0.06	0.07	0.09				
N.3					0.01	0.01	0.02	0.01
N.am	0.09	0.09	0.07	0.02	0.12	0.12	0.10	0.11
S.3	0.02	0.02	0.01	0.01				

[a] In Å. [b] Values are given considering only the part of binding affinity (pK_i^{PLS}) used in PLS analysis or considering the total binding affinity (values in parentheses). [c] $q^2 = 1 - \text{PRESS/SSD}$ as obtained by "leave-one-out" cross-validation. PRESS equals the sum of squared differences between predicted and experimentally determined binding affinities, SSD is the sum of the squared differences between experimentally determined binding affinities and the mean of the training set binding affinities. [d] In logarithmic units. [e] $s_{PRESS} = \sqrt{\text{PRESS}/(n-h-1)}$ as obtained by "leave-one-out" cross-validation. n equals the number of data points, h is the number of components. [f] Correlation coefficient. [g] $S = \sqrt{\text{RSS}/(n-h-1)}$. RSS equals the sum of squared differences between fitted and experimentally determined binding affinities. [h] Fisher's F-value. [i] Number of components.

Relative mutual superposition of all ligands has been obtained using the protein binding-pocket as a constraint. As recently shown, this produces superior models in 3D QSAR approaches.[72] As a final step, geometry optimization has been applied to relief unfavorable inter- (and intra-) molecular interactions.

In the glycogen phosphorylase *b* case, for several of the considered ligands, complex structures have been determined crystallographically.[73] These were used as a starting point for the mutual alignment, complemented by manually docked ligands following the same procedure as in the thermolysin case.

Significance and Robustness of AFMoC Analyses. Table 4 summarizes the statistical results for both data sets as a function of the applied grid spacing. Models with $q^2 > 0.5$ are generally accepted as significant.[74] In the present study, $q^2 > 0.5$ has been obtained with a lattice spacing of 0.5, 1.0, and 1.5 Å, either for the part of pK_i being adapted during PLS analysis (pK_i^{PLS}, s. Theory section) and for the consideration of the total binding affinity (Figure 4). For both data sets, CoMFA analyses using a 1 Å grid spacing were performed for comparison. In case of thermolysin, a q^2 value of 0.63 ($s_{PRESS} = 1.33$; 8 components) was obtained which is comparable to those reported in previous studies.[54,71] For the glycogen phosphorylase *b* data set, a q^2 value of 0.62 ($s_{PRESS} = 0.68$; six components) was calculated. Hence, in both cases, q^2 values obtained by CoMFA are comparable to those computed by AFMoC.

Table 5 summarizes for 10 runs of "leave-five-out" cross-validation with AFMoC that also q^2 values > 0.5 were obtained. Further indication for the significance of the obtained models is provided by the randomly scrambled data that do not allow one to obtain reasonable models (Table 6).

Table 4 shows that the dependency on the grid spacing is small up to 1.5 Å, beyond this spacing q^2 drops below 0.5. The latter has to be seen with regard to the setting of σ that governs the degree of "smearing" information about protein−ligand interactions across neighboring grid points (eq 3). Figure 5 shows for a 1 Å grid spacing that the interdependence of q^2 with σ is negligible for glycogen phosphorylase *b*, whereas an inverse correlation of q^2 with σ can be observed for the thermolysin data. Hence, values of 0.7 and 0.85 Å have been chosen for subsequent calculations using thermolysin and glycogen phosphorylase *b* inhibitors, respectively. A selection of $\sigma = 0.7$ Å means that a particular interaction declines across a distance of 1 Å (2 Å) to 21% (1%) of its original value. This shows that particularly local interactions are taken into account by our approach.

Another reason for selecting rather small σ values results from the fact that ligands approach the protein surface to van der Waals distance. Larger σ values involve also distant grid points in the calculation that can easily coincide with the space occupied by the protein atoms. The interaction information at grid points next to the protein surface is mainly determined by the repulsion term that is rather unspecific with respect to the ligand atom type. On the other hand, PLS models using interaction fields neglecting such repulsion

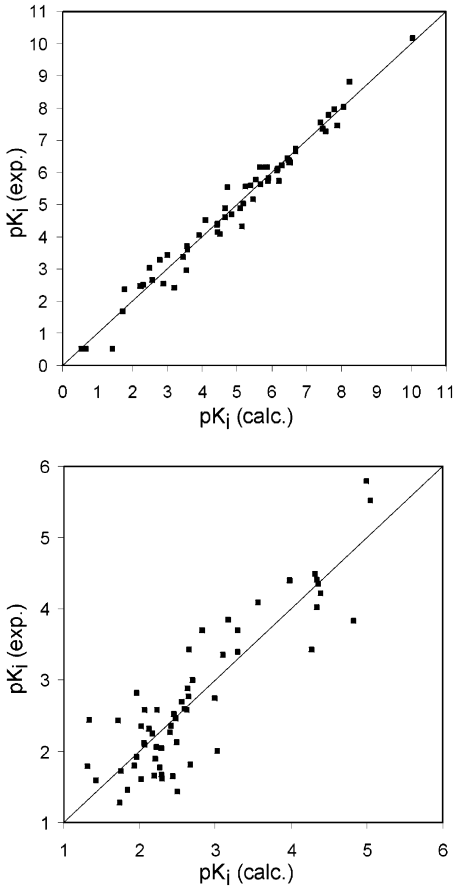


Figure 4. Experimentally determined binding affinities versus fitted predictions for the training sets of 61 thermolysin inhibitors (top) and 58 glycogen phosphorylase inhibitors (bottom).

Table 5. Statistical Results for 10 Runs of "Leave-Five-Out" Cross-Validation

	thermolysin			glycogen phosphorylase *b*		
no. of run	q^2 [a]	S_{press} [b]	components	q^2 [a]	s_{press} [b]	components
1	0.61 (0.64)	1.34	9	0.54 (0.48)	0.78	4
2	0.60 (0.63)	1.37	10	0.59 (0.54)	0.74	4
3	0.66 (0.69)	1.25	10	0.59 (0.54)	0.74	3
4	0.60 (0.63)	1.36	10	0.67 (0.63)	0.66	4
5	0.61 (0.64)	1.35	10	0.58 (0.53)	0.74	3
6	0.51 (0.55)	1.51	10	0.52 (0.46)	0.79	2
7	0.61 (0.64)	1.35	10	0.55 (0.49)	0.76	2
8	0.52 (0.56)	1.49	10	0.58 (0.52)	0.74	3
9	0.53 (0.57)	1.48	10	0.56 (0.50)	0.76	3
10	0.57 (0.61)	1.40	9	0.62 (0.58)	0.70	3
LOO[c]	0.59 (0.62)	1.38	10	0.58 (0.52)	0.75	4

[a] Values are given considering only the part of binding affinity (pK_i^{PLS}) used in PLS analysis or considering the total binding affinity (values in parentheses). [b] In logarithmic units. [c] For comparison, results of the "leave-one-out" cross-validation are shown.

terms resulted in significantly lower q^2 values (data not shown). This demonstrates the importance of including steric repulsion. Apart from higher q^2 values for 1 Å compared to 2 Å spacing, Figure 6 illustrates that translations along the grid do not affect the statistical results of the models in the 1 Å case, in contrast to results often reported for CoMFA.[65] Accordingly, for all subsequent calculations, a 1 Å grid spacing has been used.

Table 4 also depicts the single interaction field contributions to explain binding affinity differences.

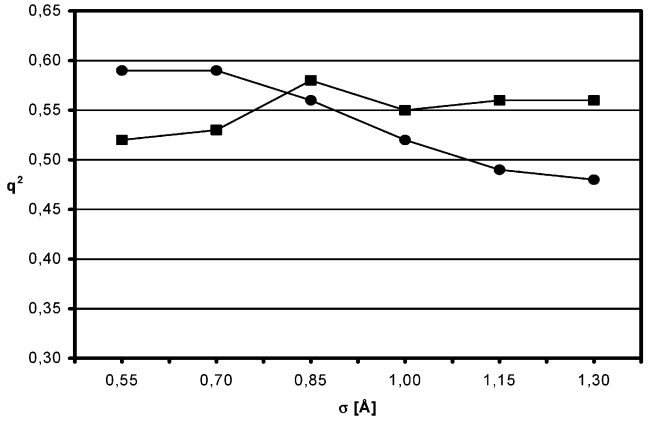


Figure 5. Dependence of q^2 values for AFMoC models on the size of σ values (eq 3). The analyses were performed for the training sets of glycogen phosphorylase (■) and thermolysin (●) inhibitors.

Table 6. q^2 Values for AFMoC Models Obtained with Randomly Scrambled Affinity Data

	q^2 [a]	
no. of components	thermolysin	glycogen phosphorylase *b*
1	−0.17 (−0.08)	−0.35 (−0.52)
2	−0.48 (−0.36)	−0.75 (−0.97)
3	−0.52 (−0.40)	−0.74 (−0.96)
4	−0.62 (−0.49)	−0.77 (−0.99)
5	−0.85 (−0.70)	−0.57 (−0.77)
6	−0.66 (−0.53)	−0.36 (−0.53)
7	−0.69 (−0.55)	−0.57 (−0.77)
8	−0.65 (−0.52)	−0.57 (−0.77)
9	−0.69 (−0.55)	−0.55 (−0.75)
10	−0.68 (−0.55)	−0.61 (−0.81)

[a] Values are given considering only the part of binding affinity (pK_i^{PLS}) used in PLS analysis or considering the total binding affinity (values in parentheses).

Contributions from C.3, C.ar, and C.2 fields are noticeably larger than those from fields calculated for polar ligand atoms. This trend corresponds to the frequency distribution of atom types in the ligand data sets. Table 7 shows how the q^2 values depend on the interaction field combinations used to derive AFMoC models. Starting with five initial fields representative for the interactions experienced by the various ligands, the inclusion of two further fields increased q^2 by 31 and 14%, respectively. The inclusion of further fields or the replacement by others resulted in reduced q^2 values. Hence, the final models were derived using seven different interaction fields (bold in Table 7).

At this point, an important difference between CoMFA- or CoMSIA-type and AFMoC-type interaction fields has to be mentioned. For CoMFA or CoMSIA applications, the increase in statistical significance of PLS results, due to the addition of further fields, has been discussed.[48,75] Including fields describing hydrogen bonding properties in addition to steric, electrostatic and hydrophobic property fields in CoMSIA did not reveal models of significantly higher q^2 values.[71] However, it has to be emphasized that these interaction or property descriptions are not atom-type specific but "generic". Adding for example a hydroxyl group to a molecule will influence *all* of the above-mentioned CoMSIA fields; hence, these fields are not orthogonal to each other. In contrast, the AFMoC fields are independent per definition, i.e., an additional hydroxyl group will only influ-

Table 7. q^2 Values Obtained by AFMoC Using Different Combinations of Interaction Fields

thermolysin		glycogen phosphorylase *b*	
combination of fields[a]	$q^{2\,b}$	combination of fields[a]	$q^{2\,b}$
C.3/O.3/O.2/O.co2/N.am	0.45 (0.50)	C.3/C.2/C.ar/O.3/N.am	0.51 (0.48)
C.3/C.ar/O.3/O.2/O.co2/N.am	0.56 (0.60)	C.3/C.2/C.ar/O.3/O.2/N.am	0.56 (0.50)
C.3/C.ar/O.3/O.2/O.co2/N.am/S.3	**0.59 (0.62)**	**C.3/C.2/C.ar/O.3/O.2/N.3/N.am**	**0.58 (0.52)**
C.3/C.ar/O.3/O.2/O.co2/N.am/N.3	0.58 (0.61)	C.3/C.2/C.ar/O.3/O.2/N.3/N.am/S.3	0.56 (0.52)
C.3/C.ar/O.3/O.2/O.co2/N.am/N.ar	0.56 (0.60)		
C.3/C.ar/O.3/O.2/O.co2/N.am/C.2	0.54 (0.57)		

[a] The denotation of the combination of interaction fields follows the atom type convention of SYBYL;[66] further details are given in ref 14. [b] Values are given considering only the part of binding affinity (pK_i^{PLS}) considered in PLS analysis or considering the total binding affinity (values in parentheses).

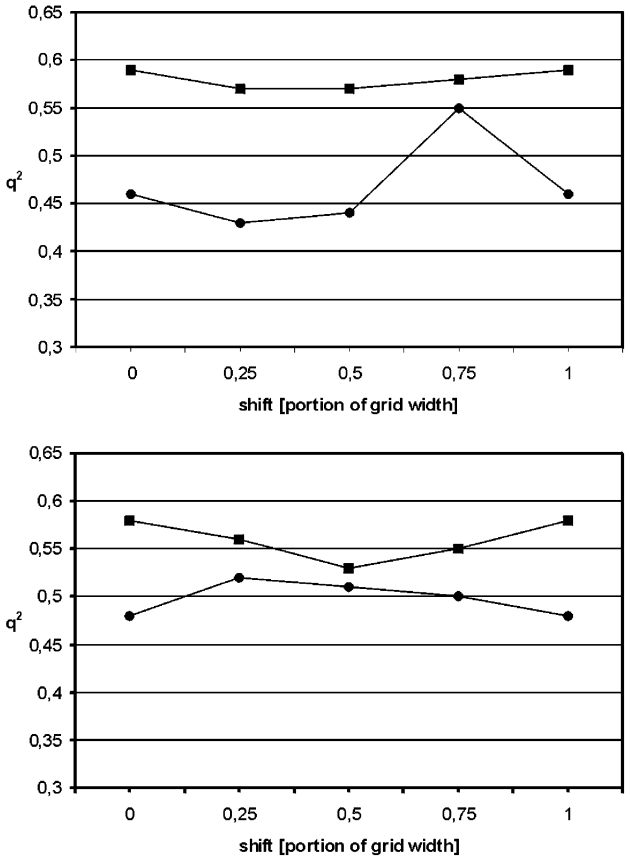


Figure 6. Dependence of q^2 values for AFMoC models on the translation of the training compound sets (top: thermolysin; bottom: glycogen phosphorylase) with respect to the grid origin. Translations were performed in steps of $^1/_4$, $^1/_2$, and $^3/_4$ of the grid spacing (■: 1.0 Å; ●: 2.0 Å) along the *xyz* diagonal.

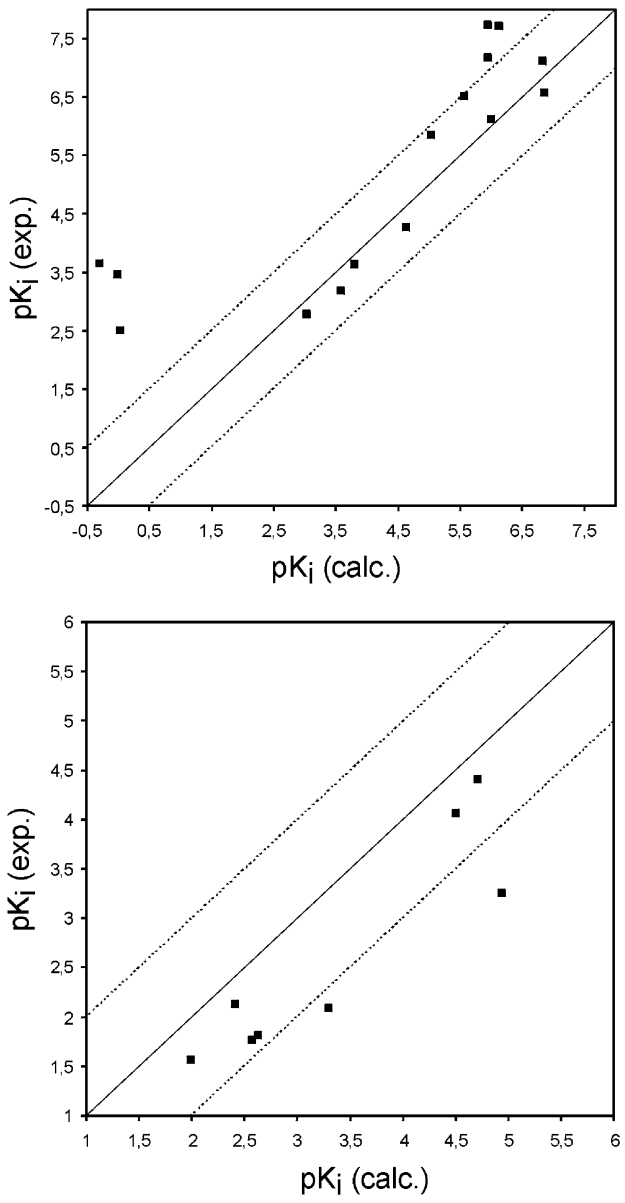


Figure 7. Experimentally determined binding affinities versus predicted values for the test set compounds not included in the training set (top: thermolysin inhibitors; bottom: glycogen phosphorylase inhibitors b). For prediction, AFMoC models generated using all 61 and 58 training compounds, respectively, were considered. In addition to the line of ideal correlation, dashed lines are given to indicate deviations of one logarithmic unit from ideal prediction.

ence the O.3 atom-type field. The inclusion of additional field types is thus only limited by an adequate representation of atoms of the respective types in the training set. This, however, does not restrain the range of ligand atom types permitted in the compounds of the test set. For those ligand atoms not represented by the model fields, the contributions from the original pair-potentials are used (see Theory section).

In view of these arguments, the numbers of PLS components (Table 4) have to be considered with respect to the number of descriptors used as input for PLS analyses. Especially in the thermolysin case, these figures are higher for PLS models based on smaller grid spacing. Similarly, the number of components selected in terms of lowest s_{PRESS}, also produced thermolysin models with rather high values using CoMFA and CoMSIA.[54,70,71] The difference in considered components between the glycogen phosphorylase *b* and thermolysin data presumably arises from the fact that the compounds of the former set are all derivatives of glucose

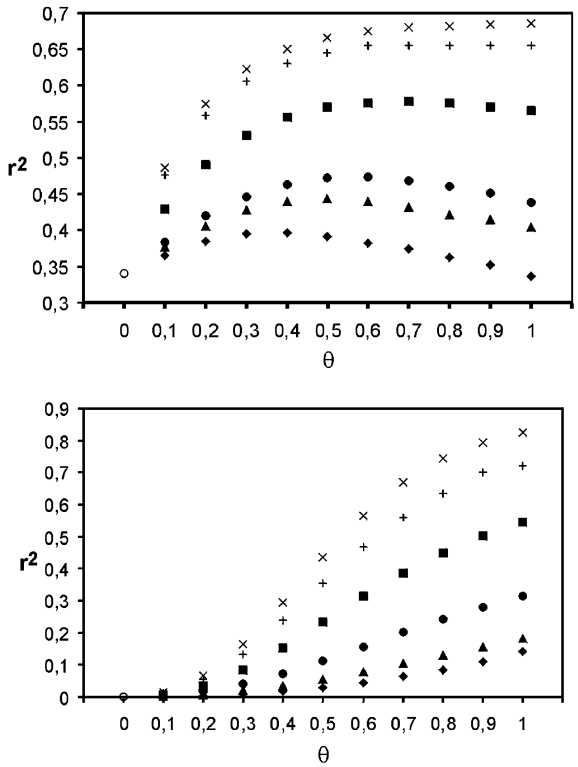


Figure 8. Dependence of r^2 values for the binding affinity prediction of the test sets of thermolysin inhibitors (top) and glycogen phosphorylase *b* (bottom) on the mixing parameter θ (eq 4). In addition, the influence of the size of the training sets used to derive AFMoC models is depicted (number of compounds in the training sets: 5 (◆), 10 (▲), 15 (●), 30 (■), 45 (+), all (x)). For comparison, the r^2 value obtained using the original DrugScore pair-potentials is shown (○).

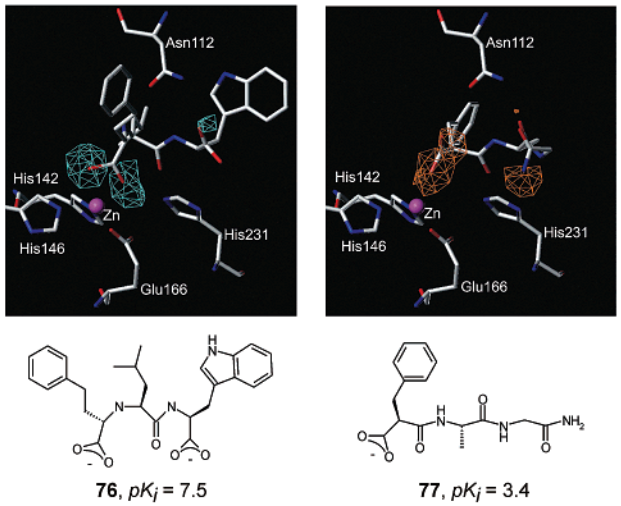


Figure 9. AFMoC STDEV*COEFF contour plots elucidating regions in the binding pocket of thermolysin where the presence of carboxylate oxygens in different regions will enhance (left) or reduce (right) binding. Contour levels are −0.01 and 0.007, respectively.

being modified at the C1 atom (exceptions are **21** and **32**). Thus, restricted structural (and hence energetic) variations in a smaller region of the interaction grids are experienced. The thermolysin inhibitors cover structural variations that range from the P1 to P2′ site, thus involving a much larger grid space. Consequently, to extract all information from the latter (higher dimensional) case, more components are necessary. The

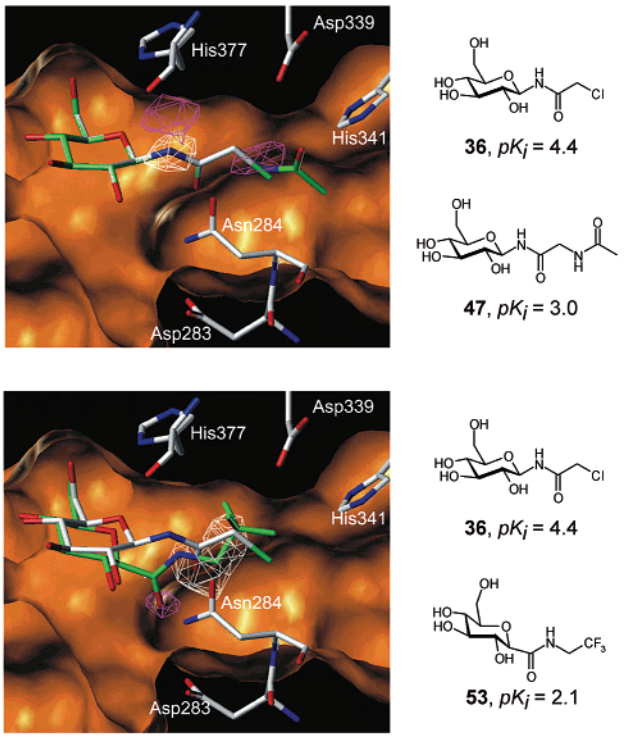


Figure 10. AFMoC STDEV*COEFF contour maps showing regions in the glycogen phosphorylase *b* binding pocket where the presence (white isopleths) or absence (magenta isopleths) of amide nitrogen (top) or carbonyl oxygen (bottom) atoms in a ligand will enhance binding. Contour levels are −0.015 and 0.004 for the N.am contours and −0.006 and 0.0022 for the O.2 contours, respectively. In both cases, ligands with weak and strong binding affinities orient such atoms into the contoured regions. In addition to the solvent-accessible surface (solid), adjacent binding site residues are shown.

Table 8. Statistical Parameters for the Prediction of Binding Affinities of the Test Set Compounds

	thermolysin	glycogen phosphorylase *b*
$r^2_{\text{DrugScore}}$	0.34	0.00
$SD_{\text{DrugScore}}$[a]	1.62	3.11
r^2_{AFMoC}	0.69	0.83
SD_{AFMoC}[a]	1.76	0.93
r^2_{CoMFA}	0.63	0.67
SD_{CoMFA}[a]	1.50	1.04

[a] In logarithmic units.

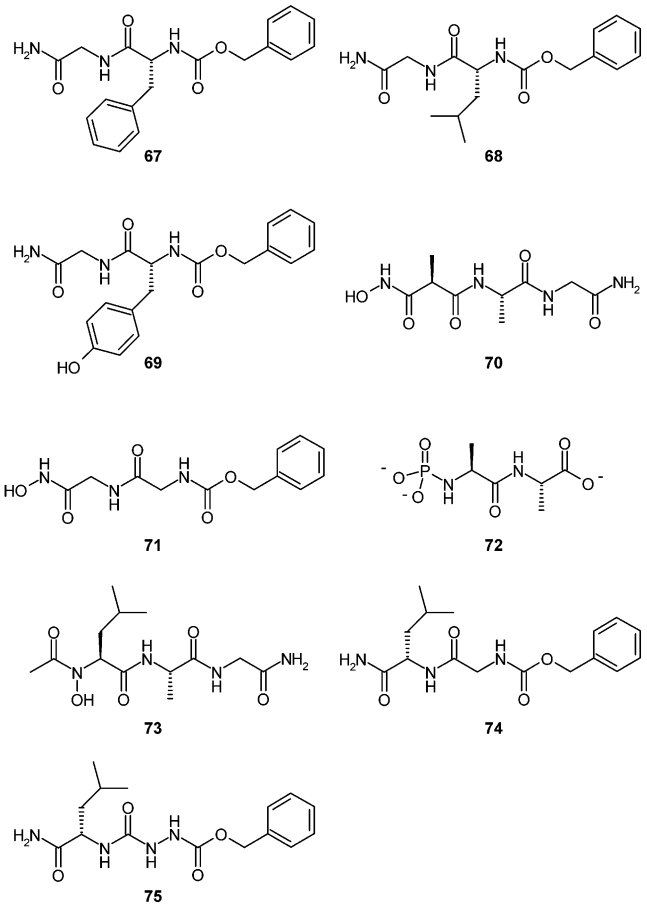
Table 9. Statistical Results for the Prediction of Binding Affinities Using Docked Ligand Geometries for the Test Set Compounds

	thermolysin	thermolysin[a]	glycogen phosphorylase *b*
rmsd[b]	11	40	100
$r^2_{\text{DrugScore}}$	0.12	0.24	0.12
r^2_{Mixed}[c]	0.21 (0.6)	0.44 (0.7)	0.48 (0.9)
r^2_{AFMoC}	0.15	0.36	0.48

[a] All docking solutions were submitted to 300 cycles of energy minimization with the MAB force field inside the (rigid) binding pocket of thermolysin. [b] Percentage of cases with rmsd <2.0 Å of automatically docked solutions with respect to the manually placed geometries. [c] In parentheses, the mixing parameter θ is given.

robustness of statistical results (e.g., leave-five-out analysis and a convincing predictive power for the test set, see below) makes us confident, however, that no overfitting of the data occurred. This anticipation is further corroborated by the fact that the q^2 values still rise by more than 5% on adding the last components to both the thermolysin and glycogen phosphorylase *b* models.

Scheme 1. Thermolysin Inhibitors



It has to be emphasized that all statistical analyses were performed without applying variable selection techniques that have been developed to distinguish explanatory information contained in the descriptor matrix from noise. Although usually improved models of lower dimensionality result for the training sets, it is still a matter of debate whether an improved predictive power with respect to (new) molecules, not included in the training set, is achieved (see refs 48 and 75 for a critical summary). In addition, especially if selection techniques on single variables are applied, noncontiguous graphical contribution maps are obtained, that hamper the interpretation of PLS results (see below). The q^2 values close to 0.6 for the 1.0 Å grid models indicate in our case that PLS itself has successfully extracted the relevant information from the descriptor matrixes.

Predictive Power of AFMoC Models. To assess the predictive power of the derived AFMoC models, binding affinities of additional, manually docked ligands, not contained in the training set, have been calculated (eq 4) exploiting the spatially adapted interaction fields.

For 9 of the 15 compounds in the thermolysin case and 6 of the 8 ligands in the glycogen phosphorylase *b* case, the deviations with respect to the experimental binding affinity are less than one logarithmic unit (Figure 7). However, the binding constants of thermolysin inhibitors **67**−**69** (Scheme 1) are underestimated by more than 2.5 orders of magnitude. All three compounds do not possess a substituent accommodating the hydrophobic S1′ pocket and their Zn binding is performed via an amide group. Among the training compounds, nonoccupancy of S1′ is only experienced by **70**−**72** and zinc coordination through an amide group (or the carbonyl oxygen of an hydroxamate group) occurs for **73**−**75**. All six compounds are weak inhibitors with $pK_i \leq 4$. Assuming that such weak inhibition can be attributed to either a missing S1′ substituent or the feeble amide zinc coordination, simultaneous presence of both features in the three "outliers" explains why the prediction of their binding affinities fails substantially.

In case of the glycogen phosphorylase *b* inhibitors, the predicted binding affinity of compound **65** deviates by 1.6 logarithmic units. Here, an additional acetamido substituent at the five-membered ring of the *spiro* compound occupies a region of the binding pocket previously not experienced by any of the training compounds.

Table 8 shows the squared correlation coefficients and standard deviations of the test set compounds derived with the AFMoC models to those calculated by DrugScore. In the thermolysin case, the latter values are very similar to those reported previously for docked geometries,[50] indicating that neither the process of manual docking nor the grid representation of pair-potentials for scoring influences prediction results. For both sets of test compounds, a remarkable increase in the r^2 values is apparent if protein-specifically adapted interaction grids are applied. This is in particular the case for glycogen phosphorylase *b* inhibitors. Using DrugScore pair-potentials only, no reliable prediction of binding affinities is achieved. The increased standard deviations, despite improved r^2, for thermolysin can be attributed to the three outliers discussed above. Excluding these three compounds reveals $r^2 = 0.84$ and SD = 0.92 logarithmic units. Interestingly, "mixing" AFMoC and DrugScore fields with a mixing parameter $\theta = 0.7$ yields a very similar r^2 value of 0.68, while the SD value drops to 1.36 logarithmic units. Here, the deviation of each of the three outliers is reduced by more than one logarithmic unit. In addition, Table 8 shows the statistical parameters for the test set predictions applying CoMFA models. While in the thermolysin case, AFMoC already shows a slightly superior predictive power compared to CoMFA, our new method performs significantly better for glycogen phosphorylase *b*. Together with the comparison of q^2 value reported above, this latter case indicates that high *predictive* power of 3D QSAR models cannot be assessed solely considering q^2 values.[76,77]

The linear mixing of DrugScore and AFMoC interaction fields for the binding affinity prediction is depicted in Figure 8. Each data point with $\theta \geq 0.1$ corresponds to an average over 100 PLS models obtained with 5, 10, 15, 30, 45, or all training compounds (see Methods). The maximum in predictive power reflected by the r^2 value with respect to a variation of the number of thermolysin inhibitors used for training gradually shifts from $\theta = 0.4$ for five compounds to $\theta = 1.0$ for all 61 compounds. Hence, if only a small amount of training information is available, best predictive power is achieved by combining generally valid and specifically adapted scoring information. In contrary, for the glycogen phosphorylase *b* case, the highest r^2 values are obtained using only the adapted AFMoC models ($\theta = 1.0$), irrespective of the amount of information used for training. Certainly,

the lacking predictive power reflected by the sole usage of pair-potentials determines these results. Furthermore, the more pronounced structural variance of the thermolysin test compounds compared to those of the glycogen phosphorylase *b* test set requires a more distinct consideration of the general pair-potentials in the thermolysin case. For both examples, considering a set of only 15 (approximately 25% of all) training compounds, r^2 values are already obtained that mirror a significant improvement in predictive power compared to the sole consideration of the original DrugScore pair-potentials (thermolysin case: $\theta = 0.6$, $r^2 = 0.48 \pm 0.08$; glycogen phosphorylase *b* case: $\theta = 1.0$, $r^2 = 0.31 \pm 0.22$). In contrast, increasing the number of training compounds from 45 to 61 or 58, respectively, yields only a slight ascent in predictive power.

The observed performance suggests using AFMoC as a tunable or tailor-made scoring function in structure-based drug design. While for initial lead finding a generally valid scoring function such as DrugScore is required, subsequently collected experimental evidence (in structural and energetic terms) for the biological target of interest can be fed back into AFMoC models. In response to the amount of new and target-specific information provided and the structural similarity of the compounds to be scored, predictions based on a protein-specifically adapted scoring function are achieved depending on the "mixing" parameter θ. In an iterative process, a hierarchical procedure is followed. However, in contrast to other approaches,[78,79] the underlying methodology is not *changed* but steadily *adjusted* to the increasing knowledge about the target under investigation.

Predictions for FlexX-Docked Ligand Geometries. For CoMFA-type calculations, the mutual relative alignment of the data set molecules is crucial. However, to use specifically adapted AFMoC fields in virtual screening, a reliable prediction of binding affinities must also be provided for ligand binding modes generated by automatic docking tools. Hence, we generated with FlexX up to 500 ligand poses for all molecules of both test sets and rescored them with appropriate AFMoC models obtained using all training compounds. Table 9 shows that in case of glycogen phosphorylase *b* a significantly improved prediction of binding affinities is achieved using the AFMoC model ($r^2 = 0.48$) compared to the sole scoring with DrugScore pair-potentials ($r^2 = 0.12$). For the thermolysin inhibitors, however, no improvement is found using the FlexX generated binding modes by either methods ($r^2 = 0.12$ compared to $r^2 = 0.15$). In the former case, all inhibitors were docked within less than 2.0 Å *rmsd* compared to the orientations obtained by manual docking. However, in the thermolysin case, this only holds for 11% of the molecules, although here a more elaborate sampling has already been performed during binding mode generation. To assess whether improved binding geometries would result in better affinity predictions, all binding modes of thermolysin inhibitors were subjected to an energy minimization with MAB.[62] As a consequence, 6 of the 15 test set molecules (40%) adopt binding modes deviating by less than 2.0 Å *rmsd* from the manually docked geometries. More important, a remarkable improvement of r^2 to 0.36 is obtained using AFMoC fields. These enhanced predictions based on minimized ligand poses strongly support the fact that automatic docking tools have to produce geometrically improved ligand orientations. Nevertheless, one should keep in mind that the set of thermolysin inhibitors with up to 15 rotatable bonds per molecule is a challenging test for every docking algorithm. Interestingly, combining DrugScore and AFMoC interaction fields using $\theta = 0.7$ even yields $r^2 = 0.44$. Hence, while for the manually docked test ligands solely using AFMoC fields revealed the best results, the more "blurred" ligand placements produced by a docking tool requires an increasing consideration of the generally valid DrugScore pair-potentials for ligand scoring. Similarly, using AFMoC fields, which were generated using an increased σ value of 0.85 and 1.00 (instead of $\theta = 0.7$ as "default" for thermolysin while keeping all other parameters identical), r^2 values (with $\theta = 0.6$) of 0.33 and 0.36, respectively, are obtained even with the ligand poses calculated directly with FlexX. More widely distributing the information of the protein−ligand interactions to neighboring grid points during the calculation of interaction fields hence results in larger robustness and tolerance with respect to deviations in ligand placement for the prediction of binding affinities. In any case, the results of these calculations suggest an enhanced scope of our new method in the field of drug development.

Graphical Interpretation of AFMoC Results. In drug design, first of all reliable predictions of binding affinities for new compounds are important. Furthermore, an intuitive graphical interpretation of the models used to predict affinities support the proposal how structural modifications of already known ligands could improve binding. In this context, AFMoC reveals important advantages. Figures 9 and 10 show correlation results mapped back onto molecular structures in terms of "STDEV*COEFF" fields.[49] Contour diagrams of field contributions for the different ligand atom types are given together with important binding-site residues and exemplary inhibitors. The obtained contour diagrams possess a contiguous and smoothly connected shape, which facilitates their interpretation. In addition, AFMoC field contributions occur at the location of ligand atoms, clearly denoting regions where it is either more favorable or disfavorable to place ligand atoms of a given type with respect to binding affinity. These characteristics are similar to those of CoMSIA fields.[54,80]

However, there are more advantages. First, atom-type specific interaction fields in contrast to generic property fields enable one to propose structural modifications in terms of favorable ligand atom types. CoMFA or CoMSIA only allow suggestions such as "place a more bulky group with more negative electrostatic potential in this binding pocket site". Second, since only protein−ligand interactions up to a distance of 6 Å are mapped onto neighboring grid points, those parts of the ligands which are solvent exposed will not be taken into account. Hence, in contrast CoMFA or CoMSIA examine interactions to *all* parts of the ligands by the different probes. Finally, as implicitly provided by the methodology, map features can be discussed (and validated) in terms of the protein environment.

Figure 9 elucidates regions in the binding pocket of thermolysin where the presence of carboxylate oxygens

enhances (left) or reduces (right) binding affinity. Obviously, a -COO^- chelating the Zn is favored over a single coordination. In addition, the second carboxylate of **76** interacts with the amide nitrogen of Asn112, which according to the contour map also contributes favorably to binding. However, given the rather pronounced differences between both ligands, the observed binding affinities (**76:** $pK_i = 7.5$; **77:** $pK_i = 3.4$) cannot solely be attributed to deviations in carboxylate oxygen interactions. Rather, additional contour maps must be consulted to detect further explanations.

The upper part of Figure 10 shows the contours of amide nitrogen atoms in the binding pocket of glycogen phosphorylase *b*. Both ligands **36** and **47** possess an amide nitrogen in the region indicated as favorable (white contour). Here, an hydrogen bond to the carbonyl group of His377 enhances binding.[73,81] Interestingly, next to this contour a disfavorable region (magenta) is indicated for amide nitrogens that originates from close contact with the protein. One has to recall, that the *product* of the coefficient of the QSAR equation times the value of the interaction field at this grid point determines the contribution to binding affinity. The values of the interaction field for N.am calculated with DrugScore potentials obviously possess a negative sign in this region. It should be noted, however, that they are calculated by considering *all* protein atoms within a cutoff of 6 Å. In addition, test calculations have shown that a rather "soft" repulsion term (operating between amide nitrogen and carbonyl oxygen in this case) yields models with best predictive power. Within the protein, interaction fields have of course positive signs. Since, however, the standard deviation in interaction field values is close to zero in these regions, they are excluded from input into PLS, hence no contours are revealed.

The disfavored region for amide nitrogens on the right is more difficult to interpret. One has to keep in mind that disfavorable contributions to binding affinity do not only result from the *presence* of *disfavorable* interaction partners in the protein but also from the *absence* of *favorable* ones. The latter holds in this case, even so, the carboxylate group of Asp339 is more than 3.5 Å apart. However, a different orientation of the substituent of **47** is hardly possible due to the rather stringent fixation of the sugar moiety in its subpocket and the conformationally rigid amide side chain. Supposedly, an unfavorable desolvation of the terminal amide nitrogen that does not find a reasonable interaction partner in the protein reduces binding affinity.

Finally, the lower part of Figure 10 shows regions of the glycogen phosphorylase *b* binding pocket where carbonyl-oxygens are (dis-)favorable. In the white contoured area, binding is favored due to hydrogen bond formation with the amide nitrogen of Asn284 (3.0 Å). Furthermore, the adjacent amide nitrogen of **36** forms an hydrogen bond to the carbonyl oxygen of His377 as discussed above. Ligand **53** exhibits an amide group of reversed orientation. It positions its carbonyl oxygen to an unfavorable region (magenta contour), which supposedly suffers from the electrostatic repulsion with the carboxylate group of Asp283 (3.2 Å). The reverse orientation of the amide group in **53** compared to **36** reinforces a ligand pose that also affects the position of the sugar moiety. While part of the low binding affinity ($pK_i = 2.1$) can be attributed to the formation of a nonoptimal hydrogen-bonding network to the protein (also indicated by contour maps for hydroxyl-oxygens not shown here), the unfavorable positioning of the carbonyl oxygen is also detrimental for binding.

The examples discussed show that differences in binding affinities can be interpreted in structural terms in particular if the protein environment is considered and information from contribution maps of various atom types are consulted.

Current Limitations and Future Enhancements. In the current study, only interactions between the protein environment (and a Zn ion in the case of thermolysin) and the ligands are taken into account to calculate interaction fields. In the glycogen phosphorylase *b* case, the protein structure of the T state is used. Here, the distance between the cofactor pyridoxal phosphate and the C1-substituents of the glucose in the modeled inhibitors is larger or close to 6 Å. This distance coincides with the cutoff of the pair-potentials. Assuming similar cofactor-to-sugar distances across the ligands of the data set, the influence on binding affinity *differences* due to the presence of pyridoxal phosphate should be negligible. However, cofactors can in principle be included by considering them as part of the protein. In the compilation of atom−atom pair-distributions, such consideration has been performed and information is contained in the pair-potentials.[14]

In addition, convincing results are obtained although water molecules were not included into the analyses. Several aspects could explain this finding. Converting structural database information of experimentally determined complexes into pair-potentials *implicitly* includes entropy-driven effects arising from (de-)solvation.[7] Although water molecules possibly mediate protein−ligand interactions and likely are a major player in receptor plasticity,[82] their actual contribution to binding affinity depends on the molecular system considered. Finally, if all ligands in a series interact similarly with active site waters, the water influence will cancel out during PLS.

Recently, Pastor et al.[83] incorporated binding site waters present in glycogen phosphorylase *b* into a GRID/GOLPE analyses. Apart from lower dimensionality of the models including water compared to those missing water, the increase in predictive power with respect to q^2 was not significant. For thermolysin, no QSAR study with explicit consideration of water molecules has been performed. The inclusion of (implicit) desolvation energy fields in CoMFA[69] did not improve the statistical quality of the models. Recently, we developed pair-potentials for water-to-ligand or water-to-protein interactions following the DrugScore formalism (unpublished results). The explicit consideration of water molecules as part of the receptor or as part of the ligands during interaction field calculations can hence be performed.

In contrast to COMBINE[30] and FEFF 3D QSAR,[84] we neglect internal ligand strain energies. In our case, minimizing the ligand structures inside the binding pocket releases internal ligand strain. In addition, recent studies have shown that energy differences between ligand conformations in solution and protein-bound state are rather small.[85,86] One may anticipate,

however, that one reason for the improved predictive power found after minimization with respect to thermolysin inhibitors docked with FlexX is the relief of strain energy. Hence, the generated ligand configurations better adapt to regions where training molecules were located.

A current limitation is the assumption of a rigid protein structure, although repulsive interactions are reduced by energy minimization. Possibly distorted ligand conformations can result from inappropriate considerations of protein induced-fit effects. Presently, such effects have not been reported for the two case studies used in this paper. To account for protein plasticity, a possible alternative would be "floating independent reference frames". Here, every adjacent grid point is allowed to move with the protein region.[87]

Conclusion

In the present study, we developed protein-specifically adapted knowledge-based pair-potentials tailored to one particular protein by considering additional ligand-based information in a CoMFA-type approach. This results in either a tailor-made scoring function or a "reverse" protein-based CoMFA (= AFMoC) method. Thus, we combine the benefits of CoMFA (spatial adaptation of interaction fields, graphical evaluation of results) with those of knowledge-based pair-potentials. Furthermore, descriptors are used that produce mutually orthogonal interaction fields and are atom-type specific. Finally, by gradually moving from generally valid to protein-specifically adapted pair-potentials one permits to reflect the amount and the degree of structural diversity available in the ligand training data.

We validated the new approach using two data sets of inhibitors for thermolysin and glycogen phosphorylase *b*. We obtained convincing predictive power in both cases, based on leave-one- and leave-five-out procedures as well as the prediction of external compounds. Most remarkably, we have found significant increase in predictive power compared to the application of the original pair-potentials when additional information of only 15 ligands is incorporated. Furthermore, convincing binding affinity predictions have been achieved not only for manually docked ligand configurations but also for those generated by automatic docking. This enhances the scope of our new method considerably with respect to lead finding. Since the method can be steadily adjusted to the increasing information in ongoing drug development projects, the continuous feedback of novel experimental information will enhance the predictive power of each subsequently computed model.

Acknowledgment. The help of B. Wegscheid and C. Gerlach (University Marburg) in developing the glycogen phosphorylase *b* model is gratefully acknowledged. We thank T. Mietzner (BASF, Ludwigshafen) for helpful advice in implementing the SAMPLS algorithm. The present study has been supported by the German Ministry of Education, Science, Research and Technology (BMBF, Grant Number 0311619).

Supporting Information Available: Two tables with thermolysin inhibitors used as training and test sets together with the pK_i values. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Drews, J. Drug Discovery: A Historical Perspective. *Science* **2000**, *287*, 1960−1964.
(2) Müller-Dethlefs, K.; Hobza, P. Noncovalent Interactions: A Challenge for Experiment and Theory. *Chem. Rev.* **2000**, *100*, 143−167.
(3) Davis, A. M.; Teague, S. J. Hydrogen Bonding, Hydrophobic Interactions, and Failure of the Rigid Receptor Hypothesis. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 736−749.
(4) Hirst, J. D. Predicting ligand binding energies. *Curr. Opin. Drug. Discov. Dev.* **1998**, *1*, 28−33.
(5) Tame, J. R. H. Scoring functions: A view from the bench. *J. Comput. Aided Mol. Des.* **1999**, *13*, 99−108.
(6) Böhm, H.-J.; Stahl, M. Rapid empirical scoring functions in virtual screening applications. *Med. Chem. Res.* **1999**, *9*, 445−462.
(7) Gohlke, H.; Klebe, G. Statistical potentials and scoring functions applied to protein−ligand binding. *Curr. Opin. Struct. Biol.* **2001**, *11*, 231−235.
(8) Wang, R.; Liu, L.; Lai, L.; Tang, Y. SCORE: A New Empirical Method for Estimating the Binding Affinity of a Protein−Ligand Complex. *J. Mol. Model.* **1998**, *4*, 379−394.
(9) Stahl, M.; Böhm, H.-J. Development of filter functions for protein−ligand docking − Fast, fully automated docking of flexible ligands to protein binding sites. *J. Mol. Graph. Model.* **1998**, *16*, 121−132.
(10) Böhm, H. J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput. Aided Mol. Des.* **1998**, *12*, 309−323.
(11) Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. Ligand Solvation in Molecular Docking. *Proteins* **1999**, *34*, 4−16.
(12) Zou, X.; Sun, Y.; Kuntz, I. D. Inclusion of Solvation in Ligand Binding Free Energy Calculations Using the Generalized-Born Model. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.
(13) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−97.
(14) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based Scoring Function to Predict Protein−Ligand Interactions. *J. Mol. Biol.* **2000**, *295*, 337−356.
(15) Muegge, I.; Martin, Y. C. A general and fast scoring function for protein−ligand interactions: a simplified potential approach. *J. Med. Chem.* **1999**, *42*, 791−804.
(16) Mitchell, J. B. O.; Laskowski, R. A.; Alex, A.; Thornton, J. M. BLEEP-potential of mean force describing protein−ligand interactions: I. Generating potential. *J. Comput. Chem.* **1999**, *20*, 1165−1176.
(17) Charifson, P. S.; Corkerey, J. J.; Murcko, M. A.; Walters, W. P. Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins. *J. Med. Chem.* **1999**, *42*, 5100−5109.
(18) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(19) So, S.-S.; Karplus, M. A comparative study of ligand−receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors. *J. Comput. Aided Mol. Des.* **1999**, *13*, 243−258.
(20) Terp, G. E.; Johansen, B. N.; Christensen, I. T.; Jorgensen, F. S. A new concept for multidimensional selection of ligand conformations (MultiSelect) and multidimensional scoring (MultiScore) of protein-ligand binding affinities. *J. Med. Chem.* **2001**, *44*, 2333−2343.
(21) Wang, R.; Wang, S. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422−1426.
(22) Grüneberg, S.; Wendt, B.; Klebe, G. Sub-nanomolar Inhibitors from Computer Screening: A Modell Study using Human Carbonic Anhydrase II. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 389−393.
(23) Lemmen, C.; Lengauer, T.; Klebe, G. FLEXS: a method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502−4520.
(24) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.
(25) Schafferhans, A.; Klebe, G. Docking ligands onto binding site representations derived from proteins built by homology modelling. *J. Mol. Biol.* **2001**, *307*, 407−427.
(26) Makino, S.; Kuntz, I. D. Automated Flexible Ligand Docking Method and Its Application for Database Search. *J. Comput. Chem.* **1997**, *18*, 1812−1825.
(27) Mestres, J.; Rohrer, D. C.; Maggiora, G. M. A molecular field-based similarity approach to pharmacophoric pattern recognition. *J. Mol. Graph. Model.* **1997**, *15*, 114−21, 103−106.

(28) Fradera, X.; Knegtel, R. M.; Mestres, J. Similarity-driven flexible ligand docking. *Proteins* **2000**, *40*, 623−636.
(29) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; de-Solms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. A priori prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J. Med. Chem.* **1995**, *38*, 305−317.
(30) Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. Prediction of drug binding affinities by comparative binding energy analysis. *J. Med. Chem.* **1995**, *38*, 2681−2691.
(31) Wang, T.; Wade, R. C. Comparative binding energy (COMBINE) analysis of influenza neuraminidase-inhibitor complexes. *J. Med. Chem.* **2001**, *44*, 961−971.
(32) Kurinov, I. V.; Harrison, R. W. Prediction of new serine proteinase inhibitors. *Nat. Struct. Biol.* **1994**, *1*, 735−743.
(33) Kulkarni, S. S.; Kulkarni, V. M. Structure Based Prediction of Binding Affinity of Human Immunodeficiency Virus-1 Protease Inhibitors. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1128−1140.
(34) Rognan, D.; Lauemoller, S. L.; Holm, A.; Buus, S.; Tschinke, V. Predicting Binding Affinities of Protein Ligands from Three-Dimensional Models: Application to Peptide Binding to Class I Major Histocompatibility Proteins. *J. Med. Chem.* **1999**, *42*, 4650−4658.
(35) Grootenhuis, P. D. J.; van Galen, P. J. M. Correlation of binding affinities with nonbonded interaction energies of thrombin-inhibitor complexes. *Acta Crystallogr. Sect. D* **1995**, *51*, 560−566.
(36) Joseph-McCarthy, D.; Hogle, J. M.; Karplus, M. Use of the multiple copy simultaneous search (MCSS) method to design a new class of picornavirus capsid binding drugs. *Proteins* **1997**, *29*, 32−58.
(37) Böhm, H.-J.; Klebe, G. What Can We Learn from Molecular Recognition in Protein−Ligand Complexes for the Design of New Drugs? *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2566−2587.
(38) Ajay; Murcko, M. A. Computational Methods to Predict Binding Free Energy in Ligand−Receptor Complexes. *J. Med. Chem.* **1995**, *38*, 4953−4967.
(39) Takamatsu, Y.; Itai, A. A New Method for Predicting Binding Free Energy Between Receptor and Ligand. *Proteins* **1998**, *33*, 62−73.
(40) Venkatarangan, P.; Hopfinger, A. J. Prediction of Ligand−Receptor Binding Thermodynamics by Free Energy Force Field Three-Dimensional Quantitative Structure−Activity Relationship Analysis: Application to a Set of Glucose Analogue Inhibitors of Glycogen Phosphorylase. *J. Med. Chem.* **1999**, *42*, 2169−2179.
(41) Viswanadhan, V. N.; Reddy, M. R.; Wlodawer, A.; Varney, M. D.; Weinstein, J. N. An Approach to Rapid Estimation of Relative Binding Affinities of Enzyme Inhibitors: Application to Peptidomimetic Inhibitors of the Human Immunodeficiency Virus Type 1 Protease. *J. Med. Chem.* **1996**, *39*, 705−712.
(42) Bohacek, R. S.; McMartin, C. Multiple Highly Diverse Structures Complementary to Enzyme Binding Sites: Results of Extensive Application of a *de Novo* Design Method Incorporating Combinatorial Growth. *J. Am. Chem. Soc.* **1994**, *116*, 5560−5571.
(43) Kasper, P.; Christen, P.; Gehring, H. Empirical Calculation of the Relative Free Energies of Peptide Binding to the Molecular Chaperone DnaK. *Proteins* **2000**, *40*, 185−192.
(44) Pierce, A. C.; Jorgensen, W. L. Estimation of binding affinities for selective thrombin inhibitors via Monte Carlo simulations. *J. Med. Chem.* **2001**, *44*, 1043−1050.
(45) Rizzo, R. C.; Tirado-Rives, J.; Jorgensen, W. L. Estimation of Binding Affinities for HEPT and Nevirapine Analogues with HIV-1 Reverse Transcriptase via Monte Carlo Simulations. *J. Med. Chem.* **2001**, *44*, 145−154.
(46) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand−receptor binding affinities and the use of Bayesian regression to improve the quality of the model. *J. Comput. Aided Mol. Des.* **1998**, *12*, 503−519.
(47) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425−445.
(48) Norinder, U. Recent Progress in CoMFA Methodology and Related Techniques. In *3D QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer Academic Publisher: Dordrecht, 1998; Vol. 3, pp 25−39.
(49) Cramer III, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). I. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.
(50) Gohlke, H.; Hendlich, M.; Klebe, G. Predicting Binding Modes, Binding Affinities and "Hot Spots" for Protein−Ligand Complexes using a Knowledge-based Scoring Function. *Persp. Drug. Discov. Design.* **2000**, *20*, 115−144.
(51) Sippl, M. J. Knowledge-based potentials for proteins. *Curr. Opin. Struct. Biol.* **1995**, *5*, 229−35.
(52) Jernigan, R. L.; Bahar, I. Structure-derived potentials and protein simulations. *Curr. Opin. Struct. Biol.* **1996**, *6*, 195−209.
(53) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Am. Chem. Soc.* **1985**, *28*, 849−857.
(54) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.
(55) Wold, S.; Ruhe, A.; Wold, H.; Dunn III, W. J. The Collinearity Problem in Linear Regression. The Partial Least Squares Approach to Generalized Inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735−743.
(56) Wold, S.; Johansson, E.; Cocchi, M. PLS − Partial Least Squares Projections to Latent Structures. In *3D QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Eds.; ESCOM: Leiden, 1993.
(57) Thibaut, U.; Folkers, G.; Klebe, G.; Kubinyi, H.; Merz, A.; Rognan, D. Recommendations to CoMFA Studies and 3D QSAR Publications. In *3D QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Eds.; ESCOM: Leiden, 1993; pp 711−716.
(58) Cruciani, G.; Watson, K. A. Comparative molecular field analysis using GRID force-field and GOLPE variable selection methods in a study of inhibitors of glycogen phosphorylase b. *J. Med. Chem.* **1994**, *37*, 2589−2601.
(59) Pastor, M.; Cruciani, G.; Clementi, S. Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships. *J. Med. Chem.* **1997**, *40*, 1455−1464.
(60) Hopfinger, A. J.; Reaka, A.; Venkatarangan, P.; Duca, J. S.; Wang, S. Construction of a Virtual High Throughput Screen by 4D-QSAR Analysis: Application to a Combinatorial Library of Glucose Inhibitors of Glycogen Phosphorylase b. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1151−1160.
(61) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. E., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112*, 535−542.
(62) Gerber, P. R.; Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des.* **1995**, *9*, 251−268.
(63) Bush, B. L.; Nachbar, R. B. Sample-distance Partial Least Squares: PLS optimized for many variables, with application to CoMFA. *J. Comput. Aided Mol. Des.* **1993**, *7*, 587−619.
(64) Kubinyi, H.; Abraham, U. Practical Problems in PLS Analyses. In *3D QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Eds.; ESCOM: Leiden, 1993; pp 717−728.
(65) Cho, S. J.; Tropsha, A. Cross-validated R2-guided region selection for comparative molecular field analysis: a simple method to achieve consistent results. *J. Med. Chem.* **1995**, *38*, 1060−1066.
(66) SYBYL *Molecular Modeling Software*; Version 6.6. Tripos Inc., St. Louis, MO.
(67) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(68) Stewart, J. J. MOPAC: a semiempirical molecular orbital program. *J. Comput. Aided Mol. Des.* **1990**, *4*, 1−45.
(69) Waller, C. L.; Marshall, G. R. Three-Dimensional Quantitative Structure−Activity Relationship of Angiotensin-Converting Enzyme and Thermolysin Inhibitors. II. A Comparison of CoMFA Models Incorporating Molecular Orbital Fields and Desolvation Free Energies Based on Active-Analogue and Complementary-Receptor-Field Alignement Rules. *J. Med. Chem.* **1993**, *36*, 2390−2403.
(70) De Priest, S. A.; Mayer, D.; Naylor, C. B.; Marshall, G. R. 3D-QSAR of Angiotensin-Converting Enzyme and Thermolysin Inhibitors: A Comparison of CoMFA Models Based on Deduced and Experimentally Determined Active Site Geometries. *J. Am. Chem. Soc.* **1993**, *115*, 5372−5384.
(71) Klebe, G.; Abraham, U. Comparative Molecular Similarity Index Analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *J. Comput. Aided Mol. Des.* **1999**, *13*, 1−10.
(72) Golbraikh, A.; Bernard, P.; Chretien, J. R. Validation of protein-based alignment in 3D quantitative structure- activity relationships with CoMFA models. *Eur. J. Med. Chem.* **2000**, *35*, 123−136.

(73) Watson, K. A.; Mitchell, E. P.; Johnson, L. N. Glucose Analogue Inhibitors of Glycogen Phosphorylase: from Crystallographic Analysis to Drug Prediction using GRID Force-Field and GOLPE Variable Selection. *Acta Crystallogr. Sect. D* **1995**, *51*, 458−472.
(74) Cramer III, R. D.; DePriest, S. A.; Patterson, D. E.; Hecht, P. The Developing Practice of Comparative Molecular Field Analysisl. In *3D QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Eds.; ESCOM: Leiden, 1993.
(75) Kim, K. H.; Greco, G.; Novellino, E. A critical review of recent CoMFA applications. *Persp. Drug. Discov. Design*. **1998**, *12*, 257−315.
(76) Kubinyi, H.; Hamprecht, F. A.; Mietzner, T. Three-Dimensional Quantitative Similarity-Activity Relationships (3D QSiAR) from SEAL Similarity Matrixes. *J. Med. Chem*. **1998**, *41*, 2553−2564.
(77) Golbraikh, A.; Tropsha, A. Beware of q^2! *J. Mol. Graph. Model.* **2002**, *20*, 269−276.
(78) Eriksson, M. A.; Pitera, J.; Kollman, P. A. Prediction of the binding free energies of new TIBO-like HIV-1 reverse transcriptase inhibitors using a combination of PROFEC, PB/SA, CMC/MD, and free energy calculations. *J. Med. Chem*. **1999**, *42*, 868−881.
(79) Pearlman, D. A.; Charifson, P. S. Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System. *J. Med. Chem*. **2001**, *44*, 3417−3423.
(80) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-Dimensional Quantitative Structure−Activity Relationship Analyses Using Comparative Molecular Field Analysis and Comparative Molecular Similarity Indics Analysis To Elucidate Selectivity Differences of Inhibitors Binding to Trypsin, Thrombin, and Factor Xa. *J. Med. Chem*. **1999**, *42*, 458−477.
(81) Gregoriou, M.; Noble, M. E.; Watson, K. A.; Garman, E. F.; Krulle, T. M.; de la Fuente, C.; Fleet, G. W.; Oikonomakos, N. G.; Johnson, L. N. The structure of a glycogen phosphorylase glucopyranose spirohydantoin complex at 1.8 A resolution and 100 K: the role of the water structure and its contribution to binding. *Protein Sci.* **1998**, *7*, 915−927.
(82) Ladbury, J. E. Just add water! The effect of water on the specifity of protein−ligand binding sites and its potential application to drug design. *Chem. Biol.* **1996**, *3*, 973−980.
(83) Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem*. **1997**, *40*, 4089−4102.
(84) Tokarski, J. S.; Hopfinger, A. J. Prediction of Ligand−Receptor Binding Thermodynamics by Free Energy Force Field (FEFF) 3D-QSAR Analysis: Application to a Set of Peptidomimetic Renin Inhibitors. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 792−811.
(85) Bostrom, J.; Norrby, P. O.; Liljefors, T. Conformational energy penalties of protein-bound ligands. *J. Comput. Aided Mol. Des.* **1998**, *12*, 383−396.
(86) Vieth, M.; Hirst, J. D.; Brooks, C. L., III. Do active site conformations of small ligands correspond to low free-energy solution structures? *J. Comput. Aided Mol. Des.* **1998**, *12*, 563−572.
(87) Pearlman, D. A. Free Energy Grids: A Practical Qualitative Application of Free Energy Perturbation to Ligand Design Using the OWFEG Methodol. *J. Med. Chem*. **1999**, *42*, 4313−4324.

JM020808P