Journal of Medicinal Chemistry Article

J. Med. Chem. **2009,** *52,* 5673–5684 **5673**
DOI: 10.1021/jm8016464

Structural Artifacts in Protein−Ligand X-ray Structures: Implications for the Development of Docking Scoring Functions

Chresten R. Søndergaard,[†] Alison Elizabeth Garrett,[†] Tommy Carstensen,[†] Gianluca Pollastri,[‡] and Jens Erik Nielsen*,[†]

[†]*School of Biomolecular and Biomedical Science, Centre for Synthesis and Chemical Biology, UCD Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland, and* [‡]*School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland*

Received December 27, 2008

The development of docking scoring functions requires high-resolution 3D structures of protein−ligand complexes for which the binding affinity of the ligand has been measured experimentally. Protein−ligand binding affinities are measured in solution experiments, and high resolution protein−ligand structures can be determined only by X-ray crystallography. Protein−ligand scoring functions must therefore reproduce solution binding energies using analyses of proteins in a crystal environment. We present an analysis of the prevalence of crystal-induced artifacts and water-mediated contacts in protein−ligand complexes and demonstrate the effect that these can have on the performance of protein−ligand scoring functions. We find 36% of ligands in the PDBBind 2007 refined data set to be influenced by crystal contacts and find the performance of a scoring function to be affected by these. A Web server for detecting crystal contacts in protein−ligand complexes is available at http://enzyme.ucd.ie/LIGCRYST.

Introduction

Rational design of novel proteins and small-molecule drugs is playing an increasing role in addressing the challenges arising in modern medicine and chemical industrial processes. In the past decade computational procedures for protein design and virtual drug screening have matured and are now routinely used for engineering novel biocatalysts and for developing small-molecule drugs. Both protein design (PD[a]) and virtual drug screening (VDS) depend on algorithms that can rapidly and accurately predict the energy of a given protein or protein−ligand complex. In VDS, for example, it is often necessary to predict the binding energies of hundreds of thousands of protein−ligand complexes to arrive at a set of candidate drugs. In PD energies must be evaluated for a similar high number of protein structure models to identify mutations that can confer some new characteristic on an existing or a de novo designed protein.[1−4] Both VDS and PD has been used extensively in modern research, and novel enzymes constructed using PD techniques are able to catalyze exotic reactions.[5,6] Similar successes have been achieved for VDS, with the identification of promising drug candidates in big pharma drug discovery projects[7,8] as striking examples.

Despite the apparent successes of PD and VDS, it is clear that these methods can be improved in two ways: first, by developing faster and more accurate energy calculation algorithms and, second, by developing techniques that sample the conformational space of molecules better.

In the present article we address the former objective and examine the protein X-ray structures that are used to calibrate the energy calculation algorithms for the small-molecule docking programs used in VDS. In drug ligand docking, the energy calculation algorithms are often referred to as "scoring functions". The task of a scoring function is primarily to identify the most realistic protein−ligand complex structure from the thousands of structures generated by a small-molecule docking program[9−14] and, second, to calculate an accurate binding energy for that protein−ligand complex.

The performance of a scoring function is measured using different tests. These typically include a "binding energy test" that tests the ability of the scoring function to predict accurate ligand binding affinities, and a "discrimination test", which tests the ability of the scoring function to identify the correct protein−ligand complex from a pool of so-called "decoy structures". Both the binding energy test and the discrimination test are performed using a well-characterized set of protein−ligand X-ray structures (the "test set"), which in the case of the discrimination test has been supplemented by a set of artificially created decoy protein−ligand structures. During development, the scoring function is calibrated using a different set of protein−ligand complex X-ray structures (the "training set") to avoid overfitting of the scoring function to the test set.

Both the training set and the test set are essential to the development of the scoring function, and the accuracy and applicability of a scoring function are therefore dependent on the quality of the data in these two data sets. If the data sets contain incorrect or distorted X-ray structures, then the scoring function will be unduly influenced by these. It is therefore essential to only use "well-behaved" protein−ligand X-ray structures in the training set and the test set.

*To whom correspondence should be addressed. Phone: +353 1 716 6724. Fax: +353 1 716 6898. E-mail: Jens.Nielsen@ucd.ie.

[a] Abbreviations: PD, protein design; VDS, virtual drug screening; PDB, Protein Data Bank.

Published on Web 08/27/2009

Here, we examine the characteristics of a large number of protein−ligand complexes with respect to the prevalence of crystal symmetry contacts (biological and nonbiological) and the prevalence of water-mediated protein−ligand contacts. We study the perturbations that crystallization is likely to induce in protein−ligand complexes and provide a Web-based service that allows researchers to determine if a specific protein−ligand complex is influenced by crystal artifacts. Finally we discuss the implications of our findings for the development of scoring functions for VDS and drug docking.

Scoring Functions

The development of scoring functions has been an active area of research for at least 2 decades with countless algorithms being released. In general scoring functions can be divided into three categories: force field based scoring functions, knowledge based scoring functions, and empirical scoring functions.[15]

Force Field Based Scoring Functions. Force field based scoring functions are based on well-established force fields used in molecular mechanics simulations. Force field based scoring functions have been developed by truncating molecular mechanics force fields, leaving only the nonbonded terms to predict protein−ligand binding affinities. Often the molecular mechanics force field is supplemented by a term describing the desolvation component of the binding energy using implicit solvent models based on the Poisson−Boltzmann equation[16] or the generalized Born framework.[17]

Knowledge-Based Scoring Functions. Knowledge-based scoring functions are based on statistical analysis of a large data set of protein−ligand complex structures. The frequency distributions of specific contact types are derived from experimental 3D protein structures and translated into knowledge-based potentials using the inverse Boltzmann law.[18−20] This approach is based on the assumption that the interactions between the protein and the ligand have been optimized during evolution, and interactions that are observed frequently therefore must be favorable. Interactions that rarely occur are considered to be unfavorable. Popular scoring functions in this category include PMF,[21] BLEEP,[22,23] and DrugScore.[18]

Empirical Scoring Functions. Empirical scoring functions are based on the assumption that the binding free energy of a protein−ligand interaction can be divided into nonbonded interactions that are physically intuitive: electrostatic interactions, hydrogen bonds, van der Waals interactions, entropic effects, and hydrophobic effects.

Generally, empirical scoring functions are based on the principles first implemented by Böhm.[24] It is assumed that the change in free energy (from here on "the binding energy") when a ligand binds to a protein can be decomposed into a set of independent energy terms that each model a specific aspect of the energetics of the protein−ligand complex. The individual energy terms are calibrated to reproduce experimentally measured binding energies for a set of protein−ligand complexes with known binding affinities (the training set). Calibration of the individual energy terms is typically done using linear regression. Scoring functions in this category include LigScore,[25] ChemScore,[26] X-Score,[27] and the scoring function in eHITS.[14]

Improving Scoring Functions

During development empirical scoring functions typically undergo an optimization procedure that aims to minimize the difference between the predicted and experimentally measured binding energies. Several data sets of experimentally measured binding energies and corresponding protein−ligand complex X-ray structures have been compiled[28−31] (see Table 1) and are routinely used in the development and benchmarking of scoring funcions.[15]

Much effort has been devoted to developing efficient optimization procedures and to constructing fast and sophisticated energy calculation algorithms. Significantly less effort has been devoted to ensuring that the experimental data are appropriate for use in the training set and test set used during scoring function development. This is partly because there are many problems associated with capturing and standardizing binding energy measurements. Buffer components, pH, temperature, the methods used for preparing the protein, and the exact protocol applied are known to influence ligand binding, and these parameters are often not reported in sufficient detail in original publications. This fact makes it impossible, or very difficult, to construct accurate electronic compilations of K_d values containing these parameters. To the best of our knowledge no scoring function can predict binding energies as a function of any of the above-mentioned experimental parameters, thus illustrating the impact that lack of detail in experimental data can have on the development of computational methods. In contrast, crystal symmetry data are reported quite consistently because of the data standardization efforts of the Protein Data Bank and can therefore be investigated. The proper processing and detection of symmetry-induced effects that we concern ourselves with in this work are therefore fairly easy to implement but have, to our knowledge, been ignored for the majority of scoring functions. However, the Astex diverse set[32] has been constructed to exclude protein−ligand complexes with crystal contacts.

Bridging the Crystal−Solution Gap. A fundamental problem in the calibration of structure-based scoring functions

Table 1. Databases Containing Protein−Ligand Complexes with Known 3D-Structure and Experimentally Determined Binding Affinity

name	number of binding affinities as of March 2008	remarks
BindingDB[30]	binding affinities for 1514 structures in the PDB based on 85% sequence identity	K_i, $\Delta G°$, IC_{50}, K_d, and EC_{50} data
		link to the PDB but no designated structures associated with data
AffinDB[31]	748 binding affinity values for 474 protein−ligand complexes	K_d, K_a, K_i, and IC_{50} data.
		link to the PDB but no designated structures associated with data
K_iBank[44]	>16000 inhibition constants for 50 protein targets and 5900 chemical structures	inhibition constants
Binding MOAD[45]	2964 structures with binding data	K_d, K_i, and IC_{50} data
PDBBind[28,29]	3124 structures with binding data in the 'general' set	K_d, K_i, and IC_{50} data. Using biological units

is that the protein 3D structure and the binding energies are determined in different experiments and under different conditions. The experiments determining the 3D structure (e.g., X-ray crystallography) and the binding affinity (e.g., isothermal titration calorimetry, enzyme kinetic experiments, surface plasmon resonance measurements, etc.) should ideally be carried out under identical conditions to ensure that the measured binding energy reflects the ligand environment observed in the X-ray structure. However, in practice, this is typically not possible, since the conditions that influence ligand binding are the very conditions that a protein crystallographer changes to induce crystallization. (It is mentioned that in a single case ligand binding energies have been measured in a protein crystal.[33])

In the calculation of solution K_d values from an X-ray structure, a scoring function thus attempts to bridge the gap between a set of binding energy measurement conditions and a set of crystallization conditions. While this is a problem that, to our knowledge, is not addressed in any scoring function, it is possible to minimize the number of structures in the training set that contain significant crystallization-induced artifacts. In this study we concern ourselves with quantifying the prevalence of crystallization-induced artifacts in protein−ligand complexes and present a Web server that allows researchers to quickly assess if a specific protein−ligand structure is likely to contain any crystal artifacts.

Water. The treatment of water also presents a problem in scoring functions. It is well-known that the release of water upon protein−ligand complexation plays a large role in determining the enthalphy−entropy balance of the system. However, since it is hard to model this term, most scoring functions include a "fudge factor" that is supposed to account for the change in water entropy. Complicating the issue further is the fact that some water molecules form "water bridges" between a protein atom and a ligand atom, between two ligand atoms, or between two protein atoms. Since this happens both in the apo state and in the holo state, we must in principle detect these bridges, estimate how often the bridges exists, and add the energetic contribution to the relevant state. For example, a water bridge between two protein atoms in the binding site that occurs only in the apo state would thus increase the K_d value for a ligand, while a protein−ligand water bridge would decrease the K_d value.

While we can construct any number of scenarios where the energy from water bridges completely dwarfs the contributions from the protein−ligand interactions, it is prudent to establish how often each type of water bridge occurs. In doing this, we must also establish which water bridges in a protein−ligand X-ray structure are likely to be crystallization artifacts in order to identify the water bridges that also will occur in solution. The statistics from this analysis will enable us to decide how much effort we should devote to developing accurate energy terms for water bridges.

Constructing a Biologically Realistic Training Set or Test Set

Before studying the prevalence of crystal-induced artifacts in protein−ligand complexes, it is useful to review the two main types of distortions to protein−ligand complex X-ray structures that are likely to be observed.

Oligomeric State of the Protein (Biological Symmetry). The asymmetric unit structure determined in X-ray diffraction experiments can consist of more than one biological unit, or more frequently the biological unit consists of several asymmetric units. Simply downloading a PDB file will typically yield a file containing only the atoms in the asymmetric unit, and in many cases this is not a realistic representation of the biologically relevant state of the protein. However, it is possible to construct the biological unit by referencing the primary literature or consulting the PDB file header. The Protein Data Bank also supplies biological units for download.

Crystal-Induced Structural Artifacts (Nonbiological Symmetry). Contacts between atoms belonging to different, neighboring asymmetric units are quite common in protein X-ray structures. These "crystal contacts" do not exist when the protein is in solution, and they are known to locally change side chain conformations in X-ray structures.[34] A change in side chain conformation or ligand conformation can have profound effects on the protein−ligand binding energies predicted from such a structure. It is therefore important to minimize the number of structures with crystal contacts at or near the ligand binding site.

As mentioned above, it is also important to carefully consider the validity of the cocrystallized water molecules. Some water molecules are likely to be quite ordered even in the solution form of the protein−ligand complex, and some of these might form water bridges between the protein and the ligand, thereby stabilizing the protein−ligand complex. Other water molecules found in X-ray structures are purely a result of the crystallization process and should therefore be ignored by the scoring function.

The question is now how frequently biological and nonbiological symmetry influences the ligand in protein−ligand X-ray structures. In the following we present statistics on the prevalence of crystal artifacts in protein−ligand X-ray structures contained in the PDBBind database[28,29] and furthermore release a Web server for detecting these in any PDB file.

Materials and Methods

Data Sets and Protein−Ligand Complex Structures. We used the Protein Data Bank (PDB)[35] entries listed in the refined set of the 2007 version of the PDBBind database. This data set consists of 1300 protein structures all containing noncovalently bound ligands. For the statistical analysis of contacts between ligands and symmetry related atoms, all heterogeneous atoms not covalently bound to the protein are treated as ligand atoms unless they fall into the category of common cocrystallized ions that are ignored: Co^{2+}, Na^{+}, Mn^{2+}, Mg^{2+}, Fe^{2+}, Cu^{+}, Ni^{2+}, Zn^{2+}, Ca^{2+}, Cl^{-}, and SO_4^{2-}. For the X-Score analysis we used the unaltered protein and ligand structure files supplied in the PDBBind 2007 refined set. In this set the protein receptors and ligand molecules have been defined by the authors of the PDBBind database, and cocrystallized water molecules and metal ions are included as part of the protein receptor.

Detection of Interactions Influenced by Crystal Contacts. We generate all symmetry related protein and ligand residues using WHAT IF.[36] Symmetry related water molecules are not considered. We calculate interatomic center−center distances between the ligand atoms and the atoms in symmetry related residues and record all interatomic distances shorter than a defined cutoff (5 or 3.5 Å).

Construction of Biological Units. The biological units of all proteins were obtained from the PDB. We compare the asymmetric unit to the biological unit to identify protein atoms in the biological unit that are not found in the asymmetric unit. All interatomic distances between atoms present in the biological unit but not in the asymmetric unit and the ligand atoms were calculated, and all interatomic distances shorter than the defined

cutoff of 5 Å were recorded. If we find any of these close contacts, we classify the biological unit as being of importance for the binding of the ligand (i.e., ligand binding occurs near the biological symmetry interfaces). The protein−ligand complexes where the biological unit is important for ligand binding present a subset of the protein−ligand complexes where symmetry related atoms are found close to the ligand.

Nonbiological Crystal Contacts. Crystal contacts with atoms that are not part of the biological unit are labeled "nonbiological crystal contacts". The number of nonbiological unit related crystal contacts that are close to the ligand is found by evaluating the interatomic distances between the relevant atoms and the ligand atoms using the 5 or 3.5 Å cutoffs as described above.

Comparison of PDB Biological Units and Biological Units Generated by PISA. The Web server facilitates automatic detection of differences between the biological units from PISA[37] and the PDB by comparison of coordinates. Specifically we compare the coordinates of the biological unit(s) from the PDB to the coordinates generated by application of the transformation matrices from PISA. PISA gives multiple probable biological units, and they are all used for comparison. The details of the stepwise comparison are shown in a flow diagram (Supporting Information). Initially the type and number of peptide chains in each biological unit are compared. If the multimer sizes are identical, then the positions of the peptide chains are compared. If the multimer interfaces are identical, then the positions of non-peptide ligands are compared. The biological units are identical if the peptide chains and the ligands are in the same positions.

Water Contacts. Interactions between the ligand and any cocrystallized water molecules are detected by calculating distances between ligand atoms and water oxygen atoms found in the asymmetric unit. We define the ligand to be influenced by water contacts if a water oxygen atom is found to be closer to a ligand atom than 5 Å.

Crystal Contact Ratio. We define the crystal contact ratio, r_{cc}, for a ligand as $r_{cc} = N_{cc}/N_{bu}$, where N_{cc} is the number of nonbiological crystal contact atoms within a defined cutoff of either 5 or 3.5 Å (see text) of the ligand and N_{bu} is the number of biological unit atoms within the cutoff distance of the ligand.

X-Score Predictions. We used X-Score, version 1.2.1, as a representative protein−ligand scoring function. X-Score has three built-in empirical scoring functions that model van der Waals interactions, hydrogen bonds, restriction of ligand rotatable bonds during protein−ligand binding, and the hydrophobic effect. The three built-in scoring functions model the hydrophobic effect in three different ways, and the resulting predictions of X-Score are the consensus scores between the three built-in scoring functions. X-Score was trained using a set of 200 protein−ligand complexes comprising more than 70 different types of proteins. The X-score training set was selected using criteria based on ligand type, size, ligand binding mode, and resolution of the X-ray structure.[27] The X-score training set was thus constructed without considering the presence of crystal contacts close to the ligand binding site.

Test of Statistical Significance of Differences in X-Score Performance. The performance of X-score on subsets of protein−ligand complex structures was evaluated as follows: a subset (B) containing N_B structures (where $N_B \ll$ size of PDBBind 2007 refined set) in which the ligand is influenced by crystal contacts is identified, and the performance (correlation coefficient) of X-score on this data set is found to be CC_B. To establish if X-score performs significantly worse on set B than on all structures without crystal contacts (subset A of size N_A with X-score performance CC_A), we do the following: (1) From the entire PDBBind data set we select 1000 sets of size N_A and calculate the average (AVG_A) and standard deviation (SD_A) of X-score performance. (2) From the entire PDBBind data set we select 1000 subsets of size N_B and calculate the average and standard deviation of X-score performance. (3) We then compare the average correlation coefficient and standard deviations for these two tests. We define X-score to perform worse on B than on A if the correlation coefficient of set B (CC_B) is worse than that of A (CC_A and AVG_A) by more than 1 standard deviation (SD_A).

Results

The PDBBind 2007 refined set contains 1300 protein−ligand complex structures that have been tied to experimentally determined K_d or K_i values. The PDBBind database is thus an immensely useful resource for training and testing protein−ligand scoring functions. In the following we examine how well the structures in the PDBBind database represent the protein−ligand complexes in solution. We determine which structures need biological symmetry corrections applied, which structures contain nonbiological symmetry artifacts, and which structures contain significant water-mediated protein−ligand contacts.

We have three goals: first, to produce a list of PDBBind records that are the least perturbed by crystallization conditions; second, to investigate the prevalence of water-mediated protein−ligand contacts; and third, to illustrate the impact of crystal artifacts on the performance of scoring functions. By achieving these goals, we aim to encourage the development of more accurate scoring functions that are not influenced by symmetry-related and water-related artifacts. We use the 1300 PDBBind records from the 2007 version refined set of the database as the source of protein−ligand binding energies and PDB identifiers.

Biological Units. For 177 structures of the 1300 structures in the data set we find the biological unit to contain more protein chains than the asymmetric unit. In 57% of these structures the ligand is closer than 5.0 Å to atoms in the biological unit that are not contained in the asymmetric unit (Table 2). It is an interesting observation that ligands often bind between monomers, and the significance of this could be investigated in evolutionary terms.

Figure 2 in the Supporting Information shows the distribution of the distances of the closest interaction found for each structure between the ligand and the biological unit atoms. A significant fraction of interactions occur with distances between 3.0 and 4.0 Å, suggesting that hydrogen bonds between the ligand and symmetry-related atoms in the biological unit play a role in complex formation. By analyzing element types of contact atoms and interatomic distances, we find 10% of the contacts within 5.0 Å to be potential hydrogen bonds.

Crystal Contacts. For 470 (36.2%) of the 1300 structures we find contacts shorter than 5 Å between the ligand and symmetry related atoms that are not part of the biological unit. The interactions between the ligand and nonbiological unit-related crystal contacts in the 470 structures presents a challenge for scoring functions, since these crystal contacts are likely to influence the binding of the ligand. Specifically, the crystal contacts are likely to produce small distortions in the geometry of the ligand and protein alike compared to the solution structures of these molecules. These distortions will influence the energies evaluated in scoring functions and thus produce an incorrect binding energy (if we are testing the scoring function) or cause the developer to adjust the evaluation of the energy term during the calibration of the scoring function. In either case, the crystal contact will produce an undesired effect.

Table 2. Number of Protein−Ligand Complexes for Which Interactions Relating to Crystal Contacts, Biological Units, and Water Contacts Were Found

	number of structures for which interactions were found	number of structures for which interactions closer than the threshold of 5 Å were found	reasons why no interactions were found in some structures
crystal contacts	1300	547 (42.1%)	
crystal contacts not related to the biological unit	1300	470 (36.2%)	
biological units	177	101 (57.1%)	The asymmetric unit and the biological unit are identical for 1123 structures.
water contacts	1273	1252 (98.4%)	No water molecules were found in 27 structures.

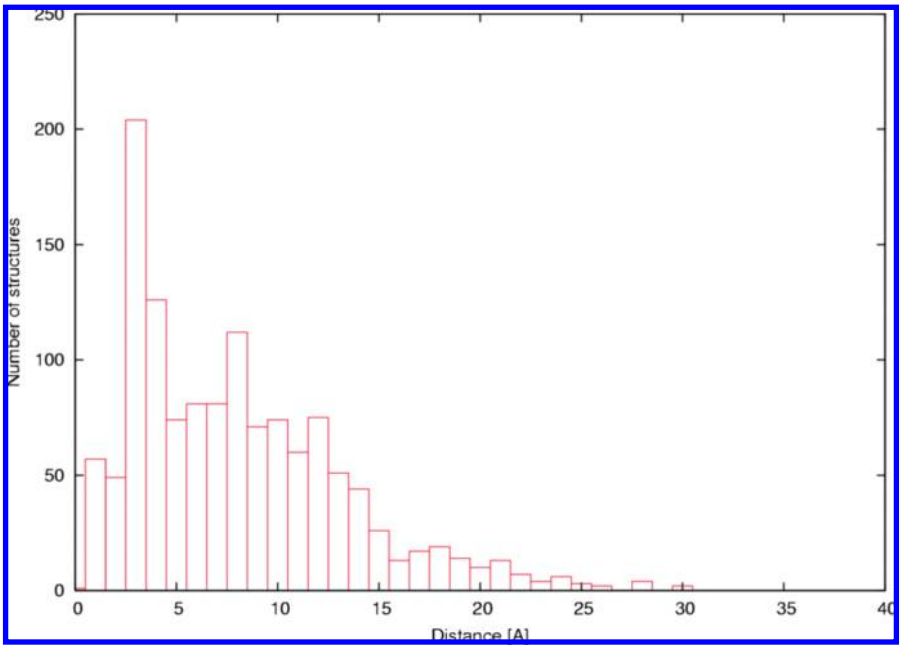

Figure 1. Histogram showing the distribution of closest interactions between any ligand atom and nonbiological unit related crystal contact.

Figure 1 shows the distribution of distances of the closest interaction found in each of the 470 structures. Figure 3 (top panel) in the Supporting Information shows the distribution of all nonbiological crystal contacts in all structures. By analyzing element types and distances of contacts, we find that 6.4% of the crystal contacts form potential hydrogen bonds with the ligand and that most of the closest contacts therefore are likely to be van der Waals interactions.

Water Contacts. Cocrystallized water atoms were found in 1273 out of 1300 protein−ligand complexes, and in 1252 (98.4%) of the complexes we find cocrystallized water atoms closer to the ligand than 5 Å. We find that 31% of the cocrystallized water contacts within 5 Å of the ligand form a hydrogen bond with the ligand.

Figure 2 shows the distribution of the closest ligand−water atom interaction found in each structure. Figure 3 (bottom panel) in the Supporting Information shows the distribution of all contacts between the ligand and cocrystallized oxygen atoms. It is seen that almost all interactions are very close and that hydrogen bonds between ligand and water atoms must be expected in most structures. While cocrystallized water molecules thus are omnipresent, we still have to investigate how many of these water molecules can be expected to be present in solution and we also need to establish the number of water molecules that form water bridges.

Detailed Analysis of Water Contacts. To investigate whether cocrystallized water atoms are artifacts of the crystallization process and therefore could introduce errors in the protein−ligand binding affinity prediction, we perform a detailed water contact analysis. In the detailed water contact analysis we divide cocrystallized water molecules into three different groups:

Group I consists of water molecules that form bridges between the ligand and nonbiological crystal contacts. These water molecules are likely to be influenced by the crystallization process and thereby introduce errors in the prediction of protein−ligand binding affinity.

Group II consists of water molecules that form bridges between the ligand and the biological unit. These water molecules are likely to be present to a significant degree in solution and are therefore important for the accurate prediction of the protein−ligand binding affinity, since they stabilize the protein−ligand complex.

Group III consists of water molecules forming bridges between nonbiological crystal contacts atoms and the biological unit. These water molecules are most likely a result of the crystallization process but will probably not affect the predicted protein−ligand binding affinity, since they do not contact the ligand directly.

Figure 3 shows the detailed water contact analysis for 2′GMP bound to ribonuclease T1 E46Q as described by the PDB entry 1rgl. Statistics of detailed water analysis of all

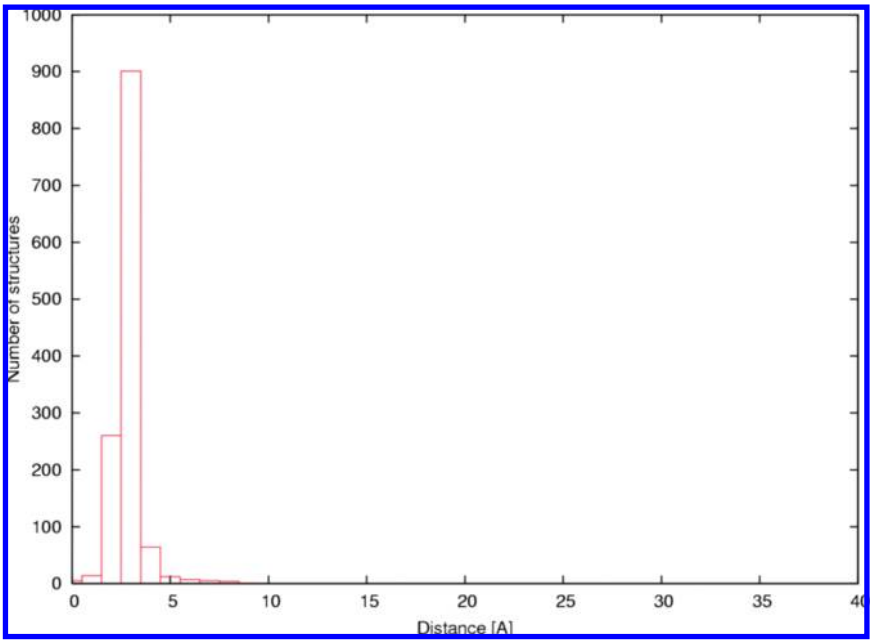

Figure 2. Distribution of distances between "closest interacting" ligand atoms and crystallographic water oxygen atoms.

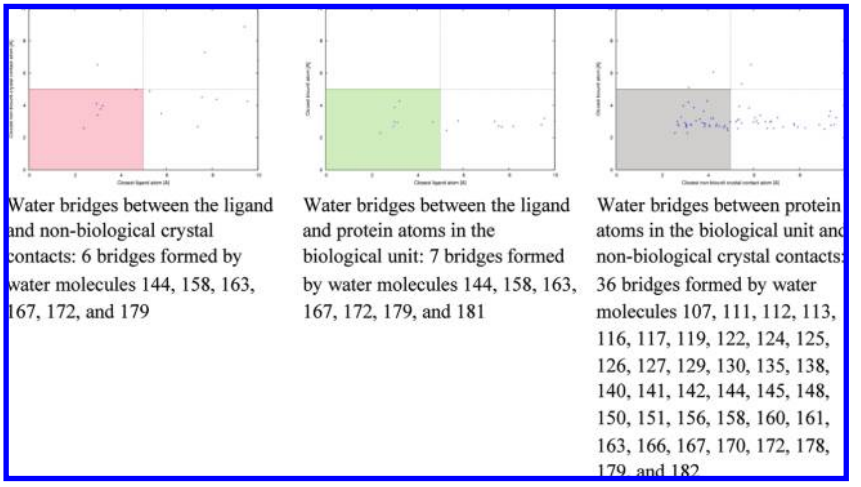

Figure 3. Detailed water contact analysis of 2′GMP bound to ribonuclease T1 E46Q (PDB entry 1rgl). Distances between water oxygen atoms and protein atoms in the biological unit, the ligand, and nonbiological crystal contacts are plotted in three different ways. Left plot: Distance to closest ligand atom (*x*-axis) plotted against the distance to the closest nonbiological crystal contact (*y*-axis). Water oxygen atoms in the red-shaded area are closer than 5 Å to both and thus form a bridge between the nonbiological crystal contacts. Center plot: Distance to closest ligand atom (*x*-axis) versus distance to closest biological unit atom (*y*-axis). These water atoms form water bridges between the ligand and the biological unit, making them important for accurate prediction of protein−ligand binding affinity. Right plot: Distance to closest nonbiological crystal contact (*x*-axis) plotted against distance the closest biological unit protein atom (*y*-axis). These water atoms form contacts between the biological unit and crystal contacts and could therefore be an artifact of the crystallization process; however, unless they are close to the ligand (i.e., they should appear in the left or middle plot), they will most likely not influence protein−ligand binding affinity prediction.The residue numbers of water atoms in the shaded area of each graph are listed. Note that several of the water oxygen atoms appear in more than one category.

1300 structures in the PDBBind refined set are presented in Figure 4. The statistics show that when a 5 Å cutoff is used, 481 structures contain water bridges between the ligand and a nonbiological crystal contact. If the water bridge cutoff is lowered to 3.5 Å, only 187 structures contain water bridges between the ligand and nonbiological crystal contacts. However, almost all structures (1245 of 1300 using a cutoff at 5 Å and 1146 of 1300 with a cutoff at 3.5 Å) contain water bridges between the ligand and the biological unit, thus emphasizing the importance of modeling water bridges explicitly in scoring functions.

Influence of Crystal Contacts on X-Score Performance. Using a 5 Å cutoff, we divided the PDBBind 2007 structures into a subset where the ligands are not influenced by crystal contacts (set A) and into a subset (set B) where ligand atoms form crystal contacts with a protein or ligand atom (for a list of PDB identifiers for which ligands are influenced by crystal contacts, see Table 1 in the Supporting Information). Under these conditions set A contains 1037 structures and set B contains 263 structures. We predict the protein−ligand binding energies for both sets using X-Score[27] to investigate how the presence of crystal contacts influences the performance of

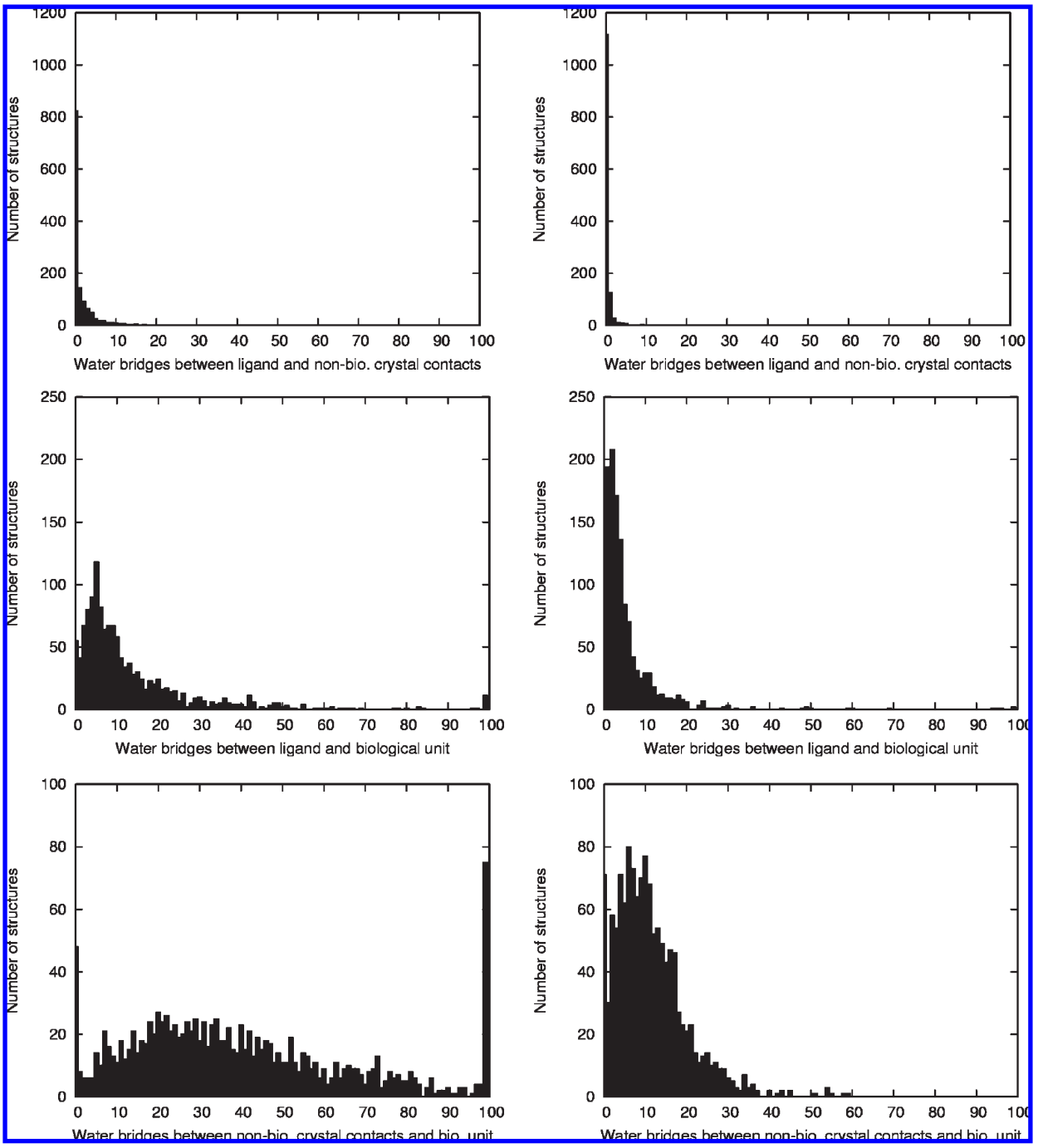

Figure 4. Detailed water bridge analysis for all structures in the PDBBind 2007 refined set. Top: Distribution of the number of water bridges between the ligand and nonbiological crystal contacts. Middle: Distribution of the number of water bridges between the ligand and the biological unit. Bottom: Distribution of the number of water bridges between the biological unit and nonbiological crystal contacts. In the left panel the distance cutoff for water bridge interactions is 5.0 Å, and in the right panel it is 3.5 Å.

X-Score. The performance of X-Score for set A (correlation coefficient of 0.58) is similar to that of set B (correlation coefficient of 0.61), and there is thus no evidence that "long-range" crystal contacts influence X-score performance.

We now lower the distance threshold for crystal contacts to 3.5 Å (the maximum distance for a hydrogen bond). This results in set A containing 1190 structures and in set B containing 110 structures. Under these conditions we obtain a greater difference in performance for set A (correlation coefficient of 0.59) and set B (correlation coefficient of 0.47), thus pointing to a potential detrimental effect of hydrogen bonds with crystal contacts on X-score performance.

To test the significance of this result, we generate 1000 random structure sets that each contain 1190 structures and use these to obtain statistics on the performance of X-score. X-score runs on the 1000 random structure sets yield a correlation coefficient of 0.58 and a standard deviation 0.019, which is similar to the performance of X-score on set A. Setting the sample size to 110 structures (corresponding to the size of set B) and regenerating 1000 random structure sets produce an X-score performance with a correlation coefficient of 0.58 and a standard deviation of 0.064.

The X-score performance on set B (correlation coefficient 0.47) is thus significantly worse than the performance on the

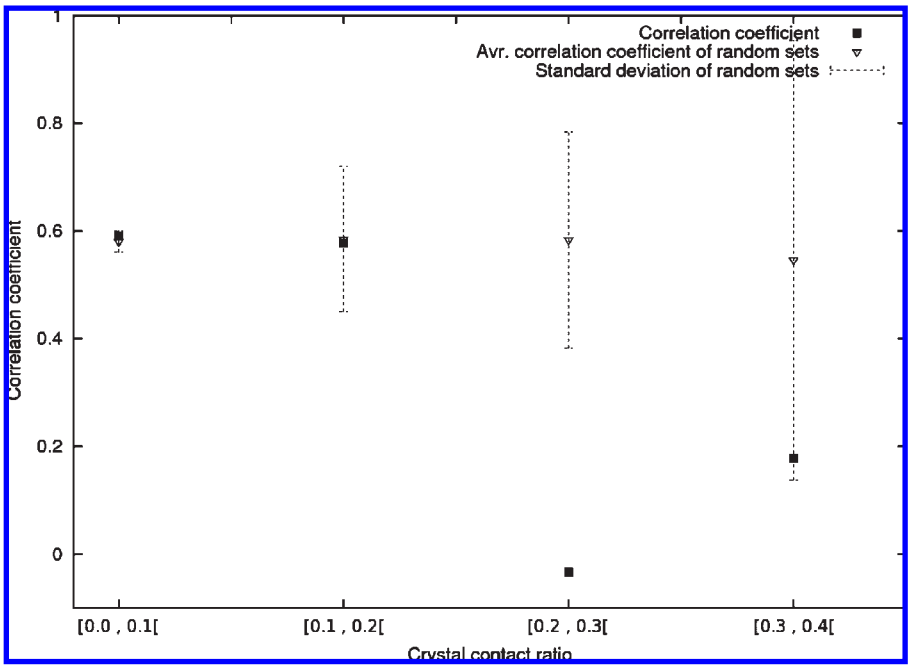

Figure 5. Performance of X-Score on subsets containing structures with crystal contact ratios in the intervals [0.0, 0.1] (1232 structures), [0.1, 0.2] (28 structures), [0.2, 0.3] (13 structures), and [0.3, 0.4] (5 structures) compared to the performance of X-Score on 1000 randomly generated subsets of equal sizes.

randomly generated structure sets, with the set B correlation coefficient being 1.9 standard deviations lower than that of set A when using a sample size of 110 structures and being 6.3 standard deviations lower when using a sample size of 1190 structures. This clearly indicates the adverse influence of crystal contacts on the performance of X-Score.

Figure 4 in the Supporting Information shows the correlation between predicted and measured binding free energies for these four subsets of the PDBBind database.

We furthermore calculate the distribution of the average crystal contact ratio using a 3.5 Å cutoff (see Materials and Methods) and plot it against the performance of X-Score (Figure 5) and find that the performance of X-Score is worse for structures with higher crystal contact ratios. For structures where the ligand forms none or few crystal contacts (crystal contact ratios in the intervals [0.0, 0.1] and [0.1, 0.2]), the correlation coefficient is close to the average of the correlation coefficients of 1000 randomly generated structure sets of equal sizes. For structures where the ligand forms more crystal contacts (crystal contact ratios in the intervals [0.2, 0.3] and [0.3, 0.4]), however, the correlation coefficient is much lower than the average correlation coefficients of 1000 randomly generated structure sets of equal sizes (3.1 standard deviations for structures with crystal contact ratios in the interval [0.2, 0.3] and 0.9 standard deviations for structures with crystal contact ratios in the interval [0.3, 0.4]). Collectively, these results on the performance of X-score show that presence of either hydrogen-bond-mediated crystal contacts or a large number of crystal contacts significantly affects the performance of scoring functions, as exemplified by X-score.

Four examples of the influence of crystal contacts on X-Score performance are discussed in the following. We expect these examples to be of general validity to the performance of scoring functions, and we use X-score only to exemplify the problems that can be expected with any scoring function.

The experimental binding energy for spinach Rubisco in complex with the inhibitor 2-carboxyarabinitol-1,5-diphosphate (PDB code 1rbo, Figure 6a)[38] is −73.1 kJ/mol; however, X-Score predicts a binding energy of −31.4 kJ/mol (error of −41.7 kJ/mol). Upon inspection of the crystal environment we find numerous crystal contacts surrounding the ligand binding site (the crystal contact ratio is 0.17). It is therefore not unexpected that X-score has problems reproducing the solution binding free energy, since the crystal environment of the ligand binding site is significantly different from the solution environment. It is unknown how the solution ligand conformation differs from the crystal conformation, but with the number of crystal contacts it is hard to imagine that the two structures are identical.

For HIV-1 protease complexed with the substrate analogue P2-NC (PDB code 2aoc)[39] there is also a large discrepancy between the experimental binding energy (−28.1 kJ/mol) and the X-Score prediction (−48.4 kJ/mol). Some nonbiological crystal contacts are present close to the substrate analogue, and the complex has a crystal contact ratio of 0.22.

However, not all errors in scoring functions originate from crystal contacts as exemplified by the binding of peptidyl-prolyl cis−trans isomerase A to (3*R*)-1-acetyl-3-methylpiperidine (PDB code 1w8l, Figure 6c).[40] The binding site for this ligand does not affect crystal contacts (the crystal contact ratio is 0.0); however, there is still a large discrepancy between the experimental binding energy (−2.8 kJ/mol) and the X-Score prediction (−28.2 kJ/mol). In this case there are therefore other factors that cause the discrepancy between the predicted and measured binding energy.

Similarly it cannot be concluded that the binding energies for all ligands influenced by crystal contacts are hard to predict. The complex between retropepsin and the TL-3 inhibitor (PDB code 5fiv, Figure 6d)[41] has the highest crystal contact ratio in the entire data set (the crystal contact ratio is 0.59). However, with an experimental binding energy

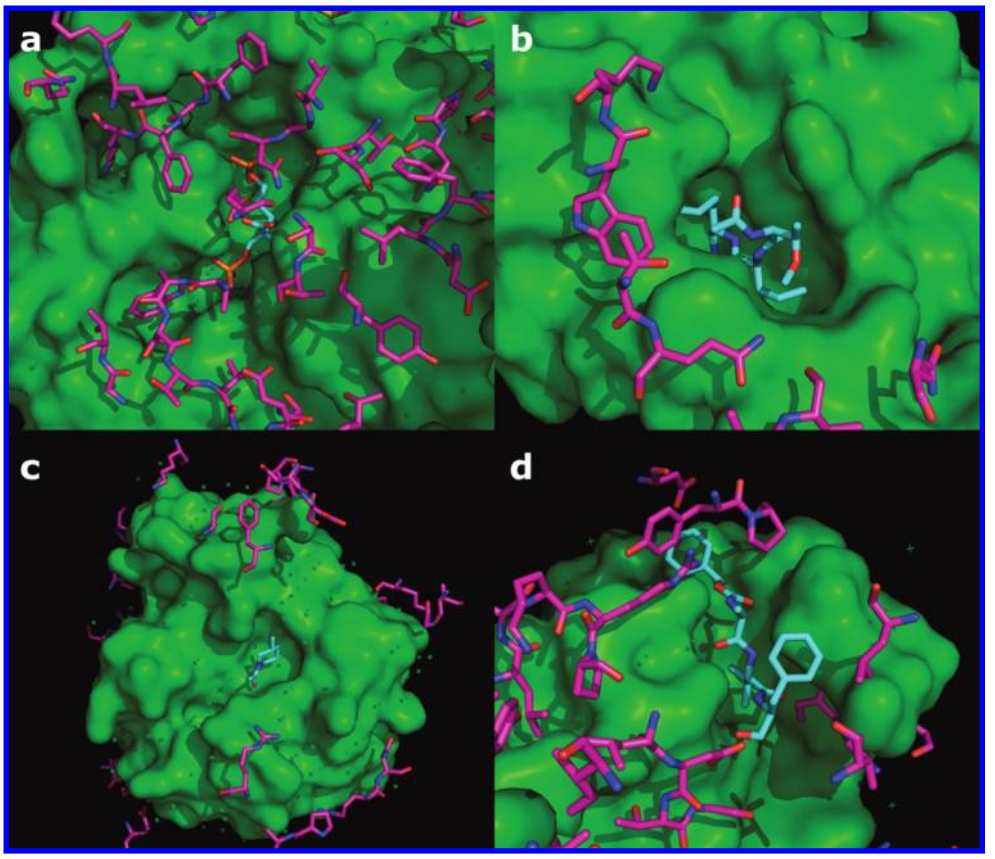

Figure 6. Some example structures for which X-Score performance was poor. The biological units are shown in surface representation and colored green. Nonbiological crystal contacts are colored magenta, red, and blue, and ligands are colored cyan, red, and blue. The PDB codes are (a) 1rbo, (b) 2aoc, (c) 1w8l, (d) 5fiv. Figure 6 was prepared using Pymol.[46].

of −44.0 kJ/mol and a X-Score prediction of −32.7 kJ/mol, the performance of X-Score is relatively good, thus showing that crystal contacts in some cases do not influence ligand conformation or, alternatively, that X-score achieves a correct prediction for a wrong reason.

It is tempting to blame the poor performance of scoring functions on crystal contacts, but although a high crystal contact ratio is associated with poor performance of scoring functions, it is obvious that other factors affect the performance of X-score.

We have illustrated that scoring functions are likely to get binding energies wrong for protein−ligand complexes with high crystal contact ratios. However, we do not know if this also affects the performance of scoring functions when it comes to discriminating decoy ligand poses from the true, experimentally verified, binding pose. To investigate this point, we examined the performance of X-score using a standard decoy data set.[15] This data set contains 100 protein structures that each are associated with 101 ligand binding poses (100 computer generated decoy binding modes of a ligand and the experimental binding mode of the same ligand). As before, we divide the data set into two subsets: set A, which is not influenced by nonbiological crystal contacts, and set B, which is influenced by crystal contacts. Each of the 101 binding poses for each protein are ranked according to their X-Score binding energy, and Figure 7 shows ROC plots describing the fidelity with which X-score identifies the experimental binding pose.

Using a cutoff value of 5.0 Å for constructing sets A and B (Figure 7 left panel), we find 29 of the 100 ligands in the experimental binding mode to be influenced by crystal contacts. X-Score performs better for the 71 experimental binding poses that are not influenced by crystal contacts (the area under curve, AUC, is 0.95, where 1.00 is a perfect score) than it does for the 29 structures influenced by crystal contacts (the AUC value is 0.88 in this case). Calculating AUC values for 1000 randomly selected structure sets each containing 29 structures gives an average AUC value of 0.93 and a standard deviation of 0.020. Hence, the X-Score performance on the 29 structures for which the ligand is influenced by crystal contacts is significantly worse (2.6 standard deviations lower) than the performance on the randomly selected structure sets.

Likewise, using a 3.5 Å cutoff (Figure 7, right panel), X-Score performs better for the 86 structures that are not influenced by crystal contacts (AUC value of 0.95) than it is for the 14 structures that are influenced by crystal contacts (AUC value of 0.81). Calculating AUC values for 1000 randomly generated structure sets each containing 14 structures gives an average AUC value of 0.93 and a standard deviation of 0.028. The X-Score performance on the 14 structures for the ligand is influenced by crystal contacts and is therefore significantly worse (4.2 standard deviations lower) than the performance on randomly generated structure sets. Table 2 in the Supporting Information lists the PDB entries for which the ligand forms crystal contacts.

We continue to test the influence of crystal contacts on the performance of X-Score by rerunning X-Score on the 100 protein−ligand structures but this time including nonbiological crystal contacts as part of the protein receptor. With the

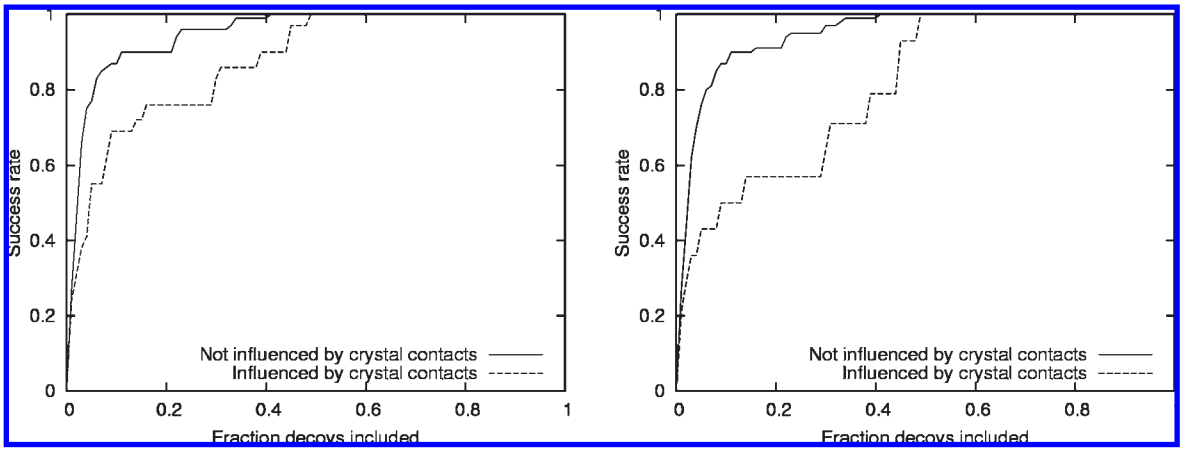

Figure 7. Accumulated success rates for X-Score prediction of binding mode for structures influenced and not influenced by crystal contacts. For each of the 100 protein structures the 101 ligand binding modes are ranked according to the binding affinity predicted by X-Score. The franction of ranked ligand binding modes included in the evaluation set is plotted on the *x*-axis, and the fraction of proteins for which the experimental binding mode was found in the evaluation set is plotted on the *y*-axis. In the left panel the maximal crystal contact cutoff value for defining crystal contacts is 5.0 Å, whereas the cutoff distance is 3.5 Å in the right panel.

nonbiological crystal contacts included, the X-Score performance is now worse for structures not influenced by crystal contacts (AUC value of 0.92 at both the 5.0 and the 3.5 Å thresholds); however, it is better for structures influenced by crystal contacts (AUC values of 0.89 and 0.84 at the 5.0 and 3.5 Å threshold, respectively). This indicates that while the inclusion of nonbiological crystal contacts introduces new errors (e.g., in the form of false contacts outside the true binding site that sometimes provides favorable contacts to decoys), the performance of X-Score is better for structures where the ligand is influenced by crystal contacts.

The presence of crystal contacts close to the ligand thus significantly influences the ability of a scoring function to identify the experimental binding pose. This result has significant implications for the use of virtual drug screening, since results from a crystal-influenced crystal structure should be scrutinized more closely than those obtained with a "crystal-free" protein structure.

Binding Site Validator Web Server. As part of this project we implemented the binding site validator Web server available at http://enzyme.ucd.ie/LIGCRYST. The Web server allows researchers to perform the analysis described in this paper on single protein−ligand complexes deposited in the PDB. After supplying a PDB identifier, the researcher can chose which components of the protein−ligand complex should be treated as protein or ligand or should be ignored. This is useful if the protein−ligand complex contains molecules that should not be included in the analysis or if a peptide chain is to be treated as a ligand. The Web server will supply lists of interactions between the ligand and nonbiological crystal contacts, biological crystal contacts, and cocrystallized water molecules. Furthermore, the Web server can perform detailed water analysis as illustrated in Figure 3.

Discussion

It is well-known that a large fraction of surface residues in protein X-ray structures interact with symmetry related atoms. For small proteins such as hen egg white lysozyme (HEWL) 37% of the total number of residues and 67% of the surface residues are within 5 Å of symmetry-related protein atoms. These crystal contacts rarely perturb the overall structure of a protein significantly[42] as measured by Cα atom displacement. However, it is also well established that crystal contact can significantly change the structures of individual side chains and in particular can "lock" surface protein side chains into conformations that rarely would be observed in solution. While such minor perturbations have no implications for the overall protein fold, they can influence the energetics of specific interactions significantly. Indeed it has been found that crystal contacts significantly influence the result of protein pK_a calculations[43] and that care should be exercised in interpreting calculated pK_a values of surface residues.

Surface residues thus frequently make crystal contacts, and perhaps it is therefore not surprising that ligands bound at a protein surface occasionally should be influenced by crystal contacts. The current study shows that 36% of the ligands in the PDBBind database are influenced by crystal contacts, thus making the presence of crystal contacts more than occasional.

While we have demonstrated that a large fraction of ligands are influenced by crystal contacts, we have not demonstrated the extent to which these crystal contacts perturb the conformation of the protein and the bound ligand. It is entirely possible that crystal contacts in some cases do not perturb the geometry of a protein−ligand complex whatsoever (see the example in Figure 6d) but simply perfectly emulate the transient contacts that water forms with the ligand in the protein−ligand complex.

Regrettably, we have not been able to find an example of the same protein−ligand complex crystallized in two different space groups (one with crystal contacts in the ligand binding pocket and one without), since such a pair of structures clearly would have demonstrated the differences in structure that the crystal contacts induce. Instead we have to satisfy ourselves with the statistics on the prevalence of crystal contacts, and the knowledge that these interactions most likely perturb the structure in some way. Future experimental studies addressing this problem will be very interesting to follow and could lead to a better understanding of protein−ligand binding energetics.

From the analysis of the X-score performance it is clear that while ligand crystal contacts play a role, they are not solely responsible for the relatively poor performance of most protein docking scoring functions. Other deficiencies such as our inability to accurately model changes in ligand and water

entropy upon ligand binding and the modeling of the protein−ligand complex as a single static structure are also likely to introduce major errors in the prediction. However, in order to improve protein docking scoring functions, it is essential that the effect of ligand crystal contacts are addressed in a meaningful way, either through representation of the protein−ligand complex as a solution-like ensemble or by the construction of protein−ligand complex data sets that exclude structures with crystal-contact-influenced ligand binding sites.

In addition to ligand crystal contacts we also find that water contacts occur quite frequently. The detailed water bridge analysis performed here grouped water contacts into different types that should be treated differently. Water bridges between the ligand and the biological unit of the protein occur very frequently (1245 of 1300 structures, 96%) and are expected to stabilize the ligand in the protein binding site. This type of water bridge should therefore be included in scoring functions, and it therefore becomes imperative for docking algorithms to create poses of the protein−ligand complex that include water molecules.

Water bridges between the ligand and nonbiological crystal contacts occur in 481 out of 1300 (37%) cases. These water bridges present another potential source of ligand binding site distortions, since hydrogen bonds to water molecules fixed in the crystal lattice could perturb the conformations of amino acid side chains in the crystal compared to the conformations observed in solution.

Correcting for Crystal Artifacts. Crystal contacts have a large effect on structure-based energy calculations, and it is therefore prudent to consider constructing methods for making X-ray structures more "solution-like". We are not aware of any specific efforts in this direction other than the use of molecular dynamics derived ensembles for various purposes. Although MD ensembles are improving, there are still significant problems with the sampling of conformational space and MD force fields, and it is therefore perhaps worth considering if one can construct algorithms that perform a local "solubilization" of an X-ray structure to alleviate the effects of crystal contacts rather than subjecting the entire protein structure to an MD simulation. In this way one could hope to minimize the artifacts introduced by the solubilization procedure while removing the most unrealistic crystal contact effects.

Conclusion

Protein docking scoring functions occupy a central position in the development and application of virtual screening techniques in modern biology because they ultimately determine the validity of the identified lead compounds. In the present article we have shown that 36% of ligands in protein−ligand X-ray structures are influenced by crystal contacts. We find that a higher crystal contact ratio for ligands correlate with a poor X-Score performance, thus supporting the conclusion that crystal contacts influence the conformation of ligands bound to proteins. We furthermore find that X-Score is more successful at identifying the experimental binding pose in a set of decoy ligand poses, when the experimental binding pose is not influenced by crystal contacts. This finding has potentially large implications for the analysis of the results of small-molecule docking algorithms.

Finally we find a high prevalence of water-mediated protein−ligand contacts in the structures examined here. Almost all of the ligands (96%) in the data set are influenced by a water bridge, and from this analysis it is clear that the effect of water bridges between the protein and the ligand must be included in scoring functions.

A Web server for identifying crystal contacts in protein−ligand X-ray structures can be accessed at http://enzyme.ucd.ie/LIGCRYST.

Acknowledgment. The authors thank Gert Vriend for access to WHAT IF. This work was supported by a Science Foundation Ireland Research Frontiers Programme Grant 05/RFP/CMS0029 to G.P., a CSCB grant, and a Science Foundation Ireland President of Ireland Young Researcher Award 04/YI1/M537 to J.E.N.

Supporting Information Available: Flow diagram, histogram of distances, contour map, performance of X-Score, and list of PDB identifiers. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Schueler-Furman, O.; Wang, C.; Bradley, P.; Misura, K.; Baker, D. Progress in modeling of protein structures and interactions. *Science* **2005**, *310*, 638–642.
(2) Bolon, D. N.; Voigt, C. A.; Mayo, S. L. De novo design of biocatalysts. *Curr. Opin. Chem. Biol.* **2002**, *6*, 125–129.
(3) Lazar, G. A.; Marshall, S. A.; Plecs, J. J.; Mayo, S. L.; Desjarlais, J. R. Designing proteins for therapeutic applications. *Curr. Opin. Chem. Biol.* **2003**, *13*, 513–518.
(4) Mayo, K. H. Recent advances in the design and construction of synthetic peptides: for the love of basics or just for the technology of it. *Trends Biotechnol.* **2000**, *18*, 212–217.
(5) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190–195.
(6) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387–1391.
(7) Kirchmair, J.; Distinto, S.; Schuster, D.; Spitzer, G.; Langer, T.; Wolber, G. Enhancing drug discovery through in silico screening: strategies to increase true positives retrieval rates. *Curr. Med. Chem.* **2008**, *15*, 2040–2053.
(8) Rester, U. From virtuality to reality. Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 559–568.
(9) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(10) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
(11) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **1992**, *13*, 505–524.
(12) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
(13) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(14) Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, B. S.; Johnson, A. P. eHiTS: an innovative approach to the docking and scoring function problems. *Curr. Protein Pept. Sci.* **2006**, *7*, 421–435.
(15) Wang, R.; Lu, Y.; Wang, S. Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* **2003**, *46*, 2287–2303.
(16) Nicholls, A.; Honig, B. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson−Boltzmann equation. *J. Comput. Chem.* **1991**, *12*, 435–445.

(17) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
(18) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein−ligand interactions. *J. Mol. Biol*. **2000**, *295*, 337–356.
(19) Huang, S.-Y.; Zou, X. An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials. *J. Comput. Chem*. **2006**, *27*, 1866–1875.
(20) Zhang, C.; Liu, S.; Zhu, Q.; Zhou, Y. A knowledge-based energy function for protein−ligand, protein−protein, and protein−DNA complexes. *J. Med. Chem*. **2005**, *48*, 2325–2335.
(21) 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.
(22) 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.
(23) Mitchell, J. B. O.; Laskowski, R. A.; Alex, A.; Forster, M. J.; Thornton, J. M. BLEEP, potential of mean force describing protein−ligand interactions: II. Calculation of binding energies and comparison with experimental data. *J. Comput. Chem*. **1999**, *20*, 1177–1185.
(24) Böhm, H. The development of a simple empirical scoring function to estimate the binding constant for a protein−ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des*. **1994**, *8*, 243–256.
(25) Krammer, A.; Kirchhoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. LigScore: a novel scoring function for predicting binding affinities. *J. Mol. Graphics Modell*. **2005**, *23*, 395–407.
(26) 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.
(27) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des*. **2002**, *16*, 11–26.
(28) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind database: collection of binding affinities for protein−ligand complexes with known three-dimensional structures. *J. Med. Chem*. **2004**, *47*, 2977–2980.
(29) Wang, R.; Fang, X.; Lu, Y.; Yang, C. Y.; Wang, S. The PDBbind database: methodologies and updates. *J. Med. Chem*. **2005**, *48*, 4111–4119.
(30) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a Web-accessible database of experimentally determined protein−ligand binding affinities. *Nucleic Acids Res*. **2007**, *35*, D198–D201.
(31) Block, P.; Sotriffer, C. A.; Dramburg, I.; Klebe, G. AffinDB: a freely accessible database of affinities for protein−ligand complexes from the PDB. *Nucleic Acids Res*. **2006**, *34*, D522–D526.
(32) Hartshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T. M.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein−ligand docking performance. *J. Med. Chem*. **2007**, *50*, 726–741.
(33) Wu, S.-y.; Dornan, J.; Kontopidis, G.; Taylor, P.; Walkinshaw, M. D. The first direct determination of a ligand binding constant in protein crystals. *Angew. Chem., Int. Ed*. **2001**, *40*, 582–586.
(34) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol*. **2002**, *320*, 597–608.
(35) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res*. **2000**, *28*, 235–242.
(36) Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graphics* **1990**, *8*, 52–56.
(37) Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol*. **2007**, *372*, 774–797.
(38) Taylor, T. C.; Fothergill, M. D.; Andersson, I. A common structural basis for the inhibition of ribulose 1,5-bisphosphate carboxylase by 4-carboxyarabinitol 1,5-bisphosphate and xylulose 1,5-bisphosphate. *J. Biol. Chem*. **1996**, *271*, 32894–32899.
(39) Tie, Y.; Boross, P. I.; Wang, Y.-F.; Gaddis, L.; Liu, F.; Chen, X.; Tozser, J.; Harrison, R. W.; Weber, I. T. Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 Å resolution crystal structures of HIV-1 protease mutants with substrate analogs. *FEBS J*. **2005**, *272*, 5265–5277.
(40) Kontopidis, G.; Taylor, P.; Walkinshaw, M. D. Enzymatic and structural characterization of non-peptide ligand−cyclophilin complexes. *Acta Crystallogr., Sect. D* **2004**, *60*, 479–485.
(41) Li, M.; Morris, G. M.; Lee, T.; Laco, G. S.; Wong, C.-H.; J., O. A.; Elder, J. H.; Wlodawer, A.; Gustchina, A. Structural studies of FIV and HIV-1 proteases complexed with an efficient inhibitor of FIV protease. *Proteins: Struct., Funct., Genet*. **2000**, *38*, 29–40.
(42) Eyal, E.; Gerzon, S.; Potapov, V.; Edelman, M.; Sobolev, V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J. Mol. Biol*. **2005**, *351*, 431–442.
(43) Nielsen, J. E.; McCammon, J. A. On the evaluation and optimization of protein X-ray structures for pK_a calculations. *Protein Sci*. **2003**, *12*, 313–326.
(44) Zhang, J.; Aizawa, M.; Amari, S.; Iwasawa, Y.; Nakano, T.; Nakata, K. Development of KiBank, a database supporting structure-based drug design. *Comput. Biol. Chem*. **2004**, *28*, 401–407.
(45) Hu, L.; Benson, M. L.; Smith, R. D.; Lerner, M. G.; Carlson, H. A. Binding MOAD (mother of all databases). *Proteins* **2005**, *60*, 333–340.
(46) DeLano, W. L. The PyMOL Molecular Graphics System. http://www.pymol.org.