JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Application of the Docking Program SOL for CSAR Benchmark

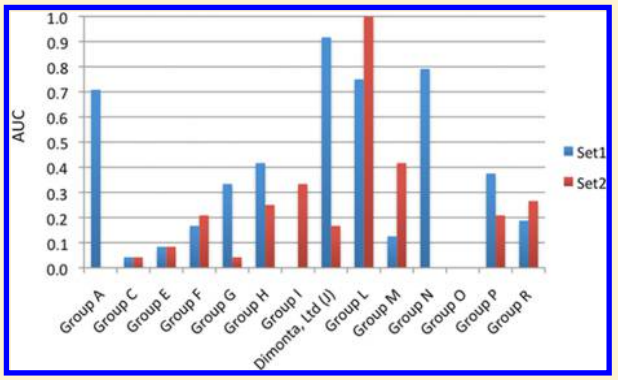
Alexey V. Sulimov,[†,‡] Danil C. Kutov,[†] Igor V. Oferkin,[†,‡] Ekaterina V. Katkova,[†,‡] and Vladimir B. Sulimov[†,‡,*]

[†] Research Computer Center, Moscow State University, Leninskie Gory 1, bldg 4, Moscow 119992, Russia

[‡] Dimonta, Ltd., Nagornaya Street 15, bldg 8, Moscow 117186, Russia

Supporting Information

ABSTRACT: This paper is devoted to results obtained by the docking program SOL and the post-processing program DISCORE at the CSAR benchmark. SOL and DISCORE programs are described. SOL is the original docking program developed on the basis of the genetic algorithm, MMFF94 force field, rigid protein, precalculated energy grid including desolvation in the frame of simplified GB model, vdW, and electrostatic interactions and taking into account the ligand internal strain energy. An important SOL feature is the single- or multi-processor performance for up to hundreds of CPUs. DISCORE improves the binding energy scoring by the local energy optimization of the ligand docked pose and a simple linear regression on the base of available experimental data. The docking program SOL has demonstrated a good ability for correct ligand positioning in the active sites of the tested proteins in most cases of CSAR exercises. SOL and DISCORE have not demonstrated very exciting results on the protein−ligand binding free energy estimation. Nevertheless, for some target proteins, SOL and DISCORE were among the first in prediction of inhibition activity. Ways to improve SOL and DISCORE are discussed.



INTRODUCTION

Many pathologies are secondary to impaired regulation of activity (or production) of one or another enzyme and as a result are also secondary to the imbalance between the activity of that enzyme and its natural inhibitors in the body. Overproduction of a particular enzyme may be counterbalanced by administration of drugs specifically inhibiting that particular enzyme. This simplified conception allows drug development using the purposeful design of new organic compounds—inhibitors for the given target protein. Selection of the effective ligands for inhibition of the target enzyme is usually a very laborious, long, and expensive process. Contemporary molecular modeling tools can accelerate this process and make it much less expensive. Virtual screening by means of ligand docking is widely recognized as a helpful approach in modern drug design.[1,2] We present here a short description of the original docking program SOL, the post-processing program DISCORE, peculiarities of their application for CSAR benchmark, and finally CSAR benchmark results.

Docking Program SOL. Many existing docking programs use either simplified descriptions of protein−ligand interactions, often neglecting solvation effects to be champions in the docking rate (e.g., FlexX[3] with a docking rate of several seconds per flexible ligand at one CPU), or very expensive software (e.g., Sybyl of Tripos and Discovery Studio of Accelrys) to be used by a small research group. Each docking program differs from the others by its unique combination of scoring function type (from rigorous functions based on different force fields to simplified empirical and knowledge-based scores), algorithm of global energy minimum search (from systematic investigation of the whole available conformational space to heuristic algorithms), level of molecule flexibility, etc. All of these features implemented in different docking programs enhanced by a user-friendly interface make each docking program a unique tool for new inhibitor design. The genetic algorithms are admitted to be the most popular docking algorithms. Two programs implementing these docking algorithms, GOLD[4] and AutoDock,[5,6] are the most commonly cited docking programs,[7] but these programs employ too simplified force field scoring functions, either neglecting electrostatic interaction in GoldScore or too simplified treatment of desolvation terms in AutoDock.

We developed our original docking program SOL[8] with the goal to accumulate the best approaches that have been demonstrated to be effective by other docking programs. Our leading idea was to describe with maximal possible accuracy the protein−ligand interactions using a docking procedure based on contemporary molecular mechanics, but we did not worry about the docking rate. However, keeping in mind that practical applications need to perform virtual screening of several

Special Issue: 2012 CSAR Benchmark Exercise

Received: February 7, 2013
Published: July 5, 2013

   dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

millions of ligands in a reasonable time, we have developed facilities for multi-processor parallelized calculations including distributed meta-computing.

The main features of SOL are (1) a rigid target protein with the active site represented by a set of grids for different type potentials, describing protein interactions with all possible types of ligand probe atoms (electrostatic, van der Waals (vdW) forces) in the frame of a Merck molecular force field (MMFF94);[9] (2) quite rigorous descriptions of solvation−desolvation effects upon the ligand binding process based on the generalized Born approximation[10,11] and included in the set of potential grids; and (3) the genetic algorithm for global protein−ligand interaction energy optimization and calculation of the ligand internal strain energy in terms of MMFF94. The grid of potentials representing protein−ligand interactions are calculated separately by the original SOLGRID[8] program before initiation of the docking procedure, and it needs several hours at one processor for a given target protein. The grid of potentials and, in particular, time of its calculation depend only on the target protein. Throughout the docking studies, all ligands are considered as fully flexible, i.e., all topologically available torsion degrees of freedom were unfrozen and allowed to rotate freely, directed only by ligand internal energy preferences in the frame of MMFF94; bond lengths and valence angles have been frozen in the course of the docking procedure. All possible ligand poses within a cube around the center of the target protein active site are considered in the docking procedure with the cube of the optimal size of 22 Å. Electrostatic, vdW, and solvation−desolvation potentials were calculated on the 101 × 101 × 101 grid inside this cube. For any ligand pose inside the cube, its energy in the protein field is calculated as a sum of contributions from all the ligand atoms, and each of these contributions is calculated by linear interpolation of the respective probe atom (the same MMFF94 type as the respective ligand atom) energies using eight closest grid nodes.

Parameters of the genetic algorithm (population size 30 000, mating pool size 70, number of generation 1000, number of runs 50, and parameters for definition of elitism, mutation, crossover, and niching)[8] are usually chosen to get the best docking results for the native ligands of different target proteins. The careful validation[12] of the SOL program was carried out by the two validation protocols, and it was demonstrated that the SOL docking quality was high and comparable with the docking quality of AutoDock. For validation of positioning accuracy, we selected 80 protein−ligand complexes with the experimental structures taken directly from PDB.[13] The docking quality criterion for this type of validation protocol was the value of the root mean square deviation (RMSD) between the docked pose and the experimental pose of the native ligand taken from the respective PDB complex. It is possible to distinguish four scales of docking quality: RMSD < 1 Å, excellent docking qualit; 1 Å < RMSD < 2 Å, good quality; 2 Å < RMSD < 3 Å, satisfactory quality; and 3 Å < RMSD, unsatisfactory quality. The validation has shown that the program SOL is able to dock with excellent and good quality (respective RMSD did not exceed 2 Å) 50 native ligands from 80. Meanwhile, the program AutoDock 3.05 has demonstrated such docking quality for 48 native ligands. The results of the docking quality comparison have also demonstrated that the number of native ligands docked by SOL with RMSD ≤ 1 Å is almost two times larger than the respective number of ligands docked by AutoDock 3.05. The situation for 1 Å < RMSD < 2 Å is quite opposite. The relative number of complexes with excellent, good, and satisfactory quality (RMSD < 3 Å) was 68.7% for SOL and 67.5% for AutoDock.

After docking, the protein−ligand binding energy can be refined by performing a local energy optimization in respect with positions of the ligand and protein atoms in the frame of the MMFF94 model and the surface GB implicit solvent model.[11] SOL has been adapted to perform virtual screening at cluster-type supercomputers of Moscow State University Chebyshev (60 Tflops, 5000 cores) and Lomonosov (400 Tflops, 35 776 cores).[14] Recently, near 800 000 ligands from the ZINC lead compounds database have been docked at Lomonosov during several days.

SOL has demonstrated good performance at general type validation tests[12] and has been employed for the thrombin inhibitors development when docking was the main driving force of the new inhibitors discovery.[15] The first step of SOL validation in this development was virtual screening of the whole NCI diversity set (about 2000 compounds) and selection of 114 virtual hits for the subsequent experimental verification. Criteria for this selection were sufficiently negative SOL scores (inhibitors must have scores less than −5.5 kcal/mol) and reasonable ligand poses in the thrombin active site, verified by visual inspection. Among these selected 114 compounds, 15 were proved to be real thrombin inhibitors in different experimental tests, and five of them had sufficiently high activities to be interesting for subsequent drug design.[15] These exercises have shown SOL's ability to reveal unknown thrombin inhibitors in the course of virtual screening. Moreover, during this stage, the isothiuronium group has been found as a P1 moiety of thrombin inhibitors for the first time. Then, different virtual libraries of ligands as possible thrombin inhibitors were generated, taking into account all discovered patterns. The overall number of compounds computed by SOL in virtual screening experiments was near 6000. Some new compounds with the best scores were synthesized and tested experimentally, supporting inhibitor selection on the basis of the SOL score. As a result of this process of new compound design, SOL docking scoring, and best candidate selection, synthesis and experimental testing new effective direct thrombin inhibitors have been discovered[15] after a couple of dozen new compound syntheses. The best new inhibitors have demonstrated K_I = 0.2 nM in the hydrolysis assay and strong anticoagulant activity in plasma − IC_{50} was approximately 100 nM in the thrombin generation assay.

Parallel Implementations of Docking Programs SOLGRID and SOL. Because of large computational demands for virtual screening of molecular databases, the implementation of parallel multi-processor facilities into docking tools is required.

We have made MPI parallel implementations for one ligand docking by the program SOL[16] as well as for the program SOLGRID,[16] generating grids of protein−ligand interaction potentials. Our calculations were performed at the cluster supercomputer Chebyshev, where each node had two 4-core processor Intel Xeon E5472 3.0 GHz and 8 GB of RAM and InfiniBand DDR (Mellanox ConnectX) with latency of 1.3−1.95 microseconds, and a bandwidth of 1540 MB per second was used as the MPI network. Efficiency of the MPI program SOLGRID on 16 nodes (128 cores) is 70%, thus the grid calculation takes several minutes instead of several hours for a single CPU version of the program SOLGRID. Efficiency of the

MPI program SOL on 128 nodes (1024 cores) is 30%. So, MPI versions of programs SOLGRID and SOL allow for reduction of calculation time significantly without loss of accuracy. More details on efficiency of parallel implementations of SOLGRID and SOL can be found in the Supporting Information.

Post-Processing Program DISCORE. DISCORE is a post-processing program that calculates the protein−ligand binding free energy after a docking procedure. DISCORE input is the 3D structure of a protein−ligand complex obtained after docking the ligand into the active site of the protein. DISCORE performs local energy optimization in the frame of the MMFF94 force field using an implicit (continuum) solvent model, estimates protein−ligand binding free energy (score), and presents it as a sum of different energy contributions such as direct electrostatic interaction between the ligand and protein atoms, van der Waals interactions, desolvation energy term, etc., with coefficients fitted to reproduce experimental values.

So, DISCORE calculates the ligand binding position more accurately than the docking program SOL and refines the binding free energy estimation given by SOL. Inaccuracies of the SOL protein−ligand binding energy estimation are connected with some simplifications introduced when the protein−ligand interactions in the frame of the MMFF94 force field in solvent are represented by the energy grid generated by the SOLGRID program. These inaccuracies are eliminated in the DISCORE program.

The local energy minimization is based on the limited memory BFGS[17] method. All the ligand atoms and neighboring protein atoms can move during energy minimization, while other protein atoms are fixed. The target function to be optimized is the sum of E_{MMFF94} and $E_{Solvent}$, where E_{MMFF94} is calculated by the MMFF94 force field methodology and $E_{Solvent}$ is the solvation energy.

DISCORE includes the MMFF94 force field and three implicit solvent models: SGB (surface generalized Born), COSMO, and PCM realized in the program module DISOLV.[18,19] During DISOLV development, a lot of attention was paid to construction of the smooth solvent-excluded surface (SES). SES is built by the original program TAGSS[18,20] performing two main stages of calculations: (i) executing primary and secondary rolling[21,22] and (ii) generation of the triangulating grid using parameters obtained in the first stage. The primary rolling procedure was performed with a hard sphere of radius 1.4 Å (imitating a water molecule) rolled over the solute molecule represented by a set of hard spheres centered at the molecule atoms with radii being parameters of the model. The secondary rolling removes singularities: self-intersections of tores and intersections of some spherical elements of the primary rolling. Smoothness of the surface has a heavy impact on the precision of the implicit solvent models. For nonelectrostatic solute−solvent interactions, we used the usual energy term proportional to the solvent accessible surface (SAS) area[23]

$$\Delta G_{np} + \Delta G_{cav} = \sigma \times S_{SAS} + b \quad (1)$$

where ΔG_{np} is the nonpolar (van der Waals) interactions with solvent term; ΔG_{cav} is the cavitational term; S_{SAS} is the solvent accessible surface area, where SAS is constructed from SES by similarity transformation; and σ = 0.00387 kcal/(mol Å^2) and b = 0.698 kcal/mol for water.

Binding energy components are calculated as E_i(Complex) − E_i(Ligand) − E_i(Protein), where E_i(Complex) is i-th energy component of the optimized protein−ligand complex, E_i(Ligand) is i-th energy component of the optimized ligand, and E_i(Protein) is i-th energy component of the optimized protein. There are seven components (i = 1, 2, ... 7) in the binding energy: atom−atom electrostatic interactions term obtained from the MMFF94 force field, van der Waals atom−atom interactions term from the MMFF94 force field, valence interactions term from the MMFF94 force field (ligand internal strain energy), electrostatic interactions with solvent, van der Waals interactions with solvent, loss of ligand torsion degrees of freedom, and loss of ligand rotational−translational degrees of freedom. Final protein−ligand binding free energy is a weighted sum of these components.

Difference between Scoring Functions in SOL and DISCORE. SOL is the docking program finding the ligand pose in the active site of the target protein and estimating the protein−ligand binding energy (SOL scoring function) for this ligand pose. DISCORE is the post-processing program estimating the protein−ligand binding energy (DISCORE scoring function) for the ligand pose found by SOL program. The method of the scoring functions calculations implemented in both programs goes back to the Bohm publication.[24] The scoring function is formed as a linear combination of rigorously calculated protein−ligand interaction terms obtained by molecular mechanics in the frame of the MMFF94 force field for either docked (SOL) or docked (SOL) and locally optimized ligand poses (DISCORE); coefficients of this linear combination have been fitted to obtain the best agreement with experimentally measured protein−ligand binding energies. This method differs from other established approaches like CoMFA, GRID, and 4D-QSAR in two main features. It does not use descriptors of any type, and it does not use molecular dynamics conformational sampling.

The SOL scoring function is the sum of interactions between the ligand and the protein and loss of ligand torsional degrees of freedom

$$G^{SOL} = k_1 \times (E_{grid}{}^{vdW} + E_{grid}{}^{Coulomb} + E_{grid}{}^{Solvent}) + k_2 \times N^{Torsions} \quad (2)$$

where k_1 is the coefficient equal to 0.1, k_2 is the coefficient equals to 0.33 kcal/mol, $E_{grid}{}^{vdW}$ is the van der Waals protein−ligand interaction energy, $E_{grid}{}^{Coulomb}$ is the electrostatic protein−ligand interaction energy, $E_{grid}{}^{Solvent}$ is the protein−ligand interaction energy provided by solvent polarization, and $N^{Torsions}$ is the number of bonds in the ligand with torsional mobility.

The ligand internal energy calculated in the frame of MMFF94 is also taken into account during global energy optimization where the target function is the sum of the ligand grid energy and ligand internal energy. However, when the global energy minimum is found and the ligand best pose is determined, the internal ligand energy is not taken into account in the SOL scoring function.

Grid energies ($E_{grid}{}^{vdW}$, $E_{grid}{}^{Coulomb}$, and $E_{grid}{}^{Solvent}$) are calculated for the ligand pose corresponding to the global energy minimum found by the docking procedure. These energies are calculated as sums of respective contributions from all the ligand atoms calculated from the grid of potentials generated by the SOLGRID program.

DISCORE scoring function is calculated as the difference between three scores

$$G^{\mathrm{DISCORE}} = G_{\mathrm{complex}}{}^{\mathrm{DISCORE}} - G_{\mathrm{free\ protein}}{}^{\mathrm{DISCORE}} - G_{\mathrm{free\ ligand}}{}^{\mathrm{DISCORE}} \quad (3)$$

where $G_{\mathrm{complex}}{}^{\mathrm{DISCORE}}$ is the score of the protein−ligand complex, $G_{\mathrm{free\ protein}}{}^{\mathrm{DISCORE}}$ is the score of the free protein, and $G_{\mathrm{free\ ligand}}{}^{\mathrm{DISCORE}}$ is the score of the free ligand.

Thus DISCORE takes into account the internal energy of the protein and ligand (unlike SOL).

The score of the protein−ligand complex, free protein, or free ligand is calculated by the following expression

$$G^{\mathrm{DISCORE}} = k_1 \times E^{\mathrm{Coulomb}} + k_2 \times E^{\mathrm{vdW}} + k_3 \times E^{\mathrm{Valence}} + k_4 \times E^{\mathrm{Solvent\ Coulomb}} + k_5 \times E^{\mathrm{Solvent\ vdW}} + k_6 \times N^{\mathrm{Torsions}} + k_7 \times E^{\mathrm{Rot-Trans}} + B \quad (4)$$

where k_1, k_2, k_3, k_4, k_5, k_6, k_7, and B are fitting coefficients; E^{Coulomb} is the energy of electrostatic interactions; E^{vdW} is the energy of van der Waals interactions; E^{Valence} is the energy of valence interactions, $E^{\mathrm{Solvent\ Coulomb}}$ is the energy of solvent polarization, $E^{\mathrm{Solvent\ vdW}}$ is the nonelectrostatic interaction energy with the solvent, N^{Torsions} is the number of ligand bonds with torsional flexibility (calculated only for the ligand), and $E^{\mathrm{Rot-Trans}}$ is the rotational−translational free energy of the molecule as a whole[25] (calculated only for the ligand).

E^{Coulomb}, E^{vdW}, and E^{Valence} are directly calculated in the frame of the MMFF94 force field without a precalculated grid of potentials. $E^{\mathrm{Solvent\ Coulomb}}$ is obtained from the surface GB, PCM, or COSMO models. The solvent excluded surface (SES) is generated around the molecule in the form of a triangular grid. Partial charges of atoms comes from the MMFF94 force field. $E^{\mathrm{Solvent\ vdW}}$ is calculated through the solvent accessible surface (SAS) area as presented in the previous section.

DISCORE performs only the local energy optimization by the L-BFGS (limited memory Broyden−Fletcher−Goldfarb−Shanno) method of protein−ligand complex and free ligand, while SOL performs global energy optimization of the ligand position by the genetic algorithm.

Program GFIT for Scoring Function Fitting. GFIT is a program for selection of the empirical (fitting) coefficients of the DISCORE scoring function. Scoring function (calculated binding free energy) is computed with respect to the next expression

$$G^{\mathrm{theor}} = k_1 \times E_1 + k_2 \times E_2 + k_3 \times E_3 + k_4 \times E_4 + k_5 \times E_5 + k_6 \times E_6 + k_7 \times E_7 + B \quad (5)$$

where B and k_i are the fitting coefficients, E_i is the following energy components: E_1 is the atom−atom electrostatic interactions term, the first energy component from DISCORE.out file; E_2 is the van der Waals atom−atom interactions term, the second energy component from DISCORE.out file; E_3 is the valence interactions term (internal strain energy), the third energy component from DISCORE.out file; E_4 is the electrostatic interactions with solvent term, the fourth energy component from DISCORE.out file; E_5 is the nonpolar (van der Waals) interactions with solvent term, the fifth energy component from DISCORE.out file; E_6 is the loss of ligand torsion degrees of freedom when binding (number of single noncyclic bonds that rotate the heavy atoms multiplied by 0.33 kcal/(mol torsion)), the sixth energy component from DISCORE.out file; E_7 is the loss of ligand rotational−translational degrees of freedom,[25] the seventh energy component from DISCORE.out file.

Coefficients B and k_i are fitted in such a way that the difference between theoretical (calculated) and experimental protein−ligand binding free energy goes to minimum.

Program GFIT calculates coefficients k_1, ..., k_7, and B to minimize the difference function D.

The difference function D represents the difference between the experimental and theoretical binding free energies and is evaluated by the following equation

$$D(k_1, ..., k_7, B) = \sqrt{\frac{\sum_i (G^{\mathrm{exp}}{}_i - G^{\mathrm{theor}}{}_i)^2}{N}} + KRL \times \sum_{n,m} \frac{k_n{}^2}{k_m{}^2} \quad (6)$$

where the coefficient KRL is equal to 10^{-4} kcal/mol for any protein from the benchmark set, N is the number of ligands in the training set, $G^{\mathrm{theor}}{}_i$ is calculated by the equation (4), and $G^{\mathrm{exp}}{}_i$ is the experimental binding free energy. Summation in the first term is performed over the all ligands from the training set (i = 1, 2,..., N) and n, m = 1, 2, ...,7 in the second term.

The second term in the difference function D increases the difference between theoretical and experimental free energies only slightly, but fitted coefficients k_1, ..., k_7 become much closer to each other. In other words, the second term keeps coefficients k_1, ..., k_7 physically realistic. If a ligand is inactive and its $\Delta G^{\mathrm{exp}}{}_i$ is not available, then parameter "dG_bad_bottom" is used as a $\Delta G^{\mathrm{exp}}{}_i$, and $\Delta G^{\mathrm{theor}}{}_i$ after calculation by the equation is also set equal to "dG_bad_bottom", if $\Delta G^{\mathrm{theor}}{}_i$ > "dG_bad_bottom".

To minimize the difference function D, we use L-BFGS, which is the local optimization algorithm of a differentiable function. To prevent negative values of the coefficients k_1, ..., k_7, local optimization is performed in variables x_1, ..., x_7, which are the square roots of variables k_1, ..., k_7.

CSAR 2011 Benchmark Exercise. We have taken part in CSAR 2011 Benchmark Exercise to get the independent quality assessment of our docking and postprocessing programs. The molecular database of four target proteins (urokinase, Chk1, ERK2, and LpxC (pseudomonas)) and ligand sets for each of the target proteins have been placed at the CSAR Web site http://www.CSARdock.org.

In the course of the research, we used the SOL docking program (5.1.4 version), SOLGRIG program for generating grids of protein−ligand interaction potentials (2.01 version), DISCORE program for calculating protein−ligand binding free energy using the implicit (continuum) solvent model (1.14 version), as well as the GFIT (1.07) program for the coefficient fitting. All the calculations were performed on the supercomputer clusters SKIF MSU "Chebyshev" and "Lomonosov". For more information please visit http://parallel.ru/cluster.

Data Preparation. The ligand set for each target protein has been presented in the SMILES format (*.smi). Program MarvinSketch[26] ver. 5.11 was used for isomer and conformer building. The ligands in the SMILES format (.SMI) were transformed into the 3D structures with the option "tetrahedral stereo isomers and double bond stereo isomers" to create stereoisomers. Then the obtained stereoisomers were saved into the .MOL format, and the conformers were created from each stereoisomer. The final 3D structures were saved into the .SDF format. Because Marvin did not add hydrogens to the

ligands, we added hydrogens with Open Babel[27] ver. 2.3.1. Then we checked the correctness of the results manually by visual analysis and comparison with the initial structures in the .SMILES format (Table 1).

Table 1. Number of Ligands for Different Target Proteins

name of target protein	initial number of ligands	number of ligands after 3D transformation
Chk1	47	356
Erk2	39	104
LpxC	16	66
urokinase	20	108

Thereafter, the database in the SDF format was split into the separate molecules by our APLITE program (version 0.46), and the geometries of all these 3D structures have been optimized in vacuum by APLITE (version 0.47 and above) using the limited memory BFGS[17] method.

In the frame of CSAR benchmark, it was suggested to download the target protein structures from the Protein Data Bank.[13] The structure of urokinase corresponds to the 1owe PDB ID. The structure of the Chk1 protein corresponds to 2e9n, ERK2 corresponds to 3i5z, and LpxC corresponds to 3p3e. We added hydrogen atoms to these structures by the APLITE program (version 0.46). Further, we determined the coordinates of the active site of each target protein as a geometrical center of its native ligand and built the grid of protein−ligand interaction potentials corresponding to four target proteins by the SOLGRID program using 0.4 Å broadening for vdW potentials. Then, we docked all the ligands into their target proteins by the SOL program and performed the post-processing procedure by the DISCORE program.

Calculations Review. In the frame of the benchmark, the results can be presented in the form of several data sets. The first data set should contain the results of the main prediction method. This set included the ranking of the ligands for each protein by the SOL score function in our case. Some additional methods, algorithms, and programs, which improve the prediction quality, could be used in the second set. In our case, we used the DISCORE program to perform local optimization of the structures and to calculate protein−ligand binding free energy more carefully taking into account the implicit (continuum) solvent model. All the results should be presented as a table with ligands ranked by the score function, and all the docked ligands in the complex with its proteins should be sent with indication of the best poses.

In the course of docking with the SOL program, almost all conformers and isomers of all the ligands have been docked successfully. A small part of the structures failed to be docked due to different reasons, but they could not affect the results because we should send only the best three poses of each ligand under the contest conditions. In this case, the best three poses of each ligand were selected by the SOL scoring function because we did not know the native poses of the ligands, and we could not calculate the RMSD value between the native and docked poses.

Hence, three molecular structures of each ligand in the PDB format, four structures of the target proteins also in the PDB format, and four ranking files for each target protein containing the list of the ligands and the best score values corresponding to one of the three poses were sent according to the first set of contest.

The post-processing procedure was carried out to refine the SOL score. The DISCORE program performs local energy optimization with respect to positions of all ligand atoms with protein atoms kept fixed. As a result, DISCORE calculates the protein−ligand binding free energy using the implicit (continuum) solvent surface generalized Born model.[11] For these purposes, the initial ligand poses were taken from the SOL results. For each ligand structure, the best pose was considered as the pose with the most negative grid energy among the different poses obtained in 50 independent runs of the SOL genetic algorithm. We took all docked poses of all conformers and isomers of each ligand in our case. As in the case of SOL, almost all the ligand structures were computed successfully with the DISCORE program.

The DISCORE program also provides all contributions to the protein−ligand binding free energy, and the respective scoring function is presented as the linear combination of all these contributions with respective coefficients (see Program GFIT for Scoring Function Fitting section). The coefficients have been obtained in advance (Table 2). Coefficients for the urokinase protein were prepared on the basis of the training set of ligands with experimentally known urokinase−ligand binding energies derived from respective inhibition constants; coefficients for other proteins (Chk1, Erk2, and LpxC) were prepared using a training set of the ligands with known binding constants for different proteins. Overall about a hundred of different proteins were taken from the literature.

Data for the second set of the benchmark were formed on the basis of the DISCORE scoring function. Three molecular structures in the PDB format, which have the best values of the scoring function among all the conformers and isomers, were selected for each ligand. The ranking files also were prepared for each four target proteins. These files include a list of the ligands with the best score values corresponding to one of the three poses.

RESULTS

The AUC (area under the curve) value was used by CSAR organizers to estimate the quality of predictions of the different research groups' software. The AUC value is the area under the ROC curve. For the best prediction, AUC is equal to 1, and for the worst prediction AUC is equal to 0.5. The ROC curve is a plot, which can be used to estimate the credibility of the predicted inhibitor affinity. There is a simple method to plot the ROC curve. First, inhibitors are sorted by theoretical scores from low to high. Then, we go to the beginning of the sorted inhibitor list and start drawing a curve from point (0, 0). If the current inhibitor from the list is experimentally good, then we

Table 2. GFIT Coefficients for Urokinase Protein and for Other Proteins (Chk1, Erk2, and LpxC)

terms	k_1	k_2	k_3	k_4	k_5	k_6	k_7
urokinase coefficients	0.056	0.054	−0.035	0.051	2.34	0.31	0.209
other proteins coefficients	0.023	0.074	−0.008	0.023	1.655	0.432	−0.072

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

draw a unit vector up; otherwise, we draw a unit vector to the right. Then, we go to the next inhibitor on the list. After processing all inhibitors from the list, a final plot is scaled to the size 1 × 1. In the case of a good prediction, if all experimentally good inhibitors are on the top of the list, the curve will dramatically go up at the beginning and go right at the end of drawing. In case of bad predictions, if experimentally good inhibitors are uniformly mixed on the list, then the curve will be close to the line from (0, 0) to (1, 1), and AUC will be near 0.5. So, the area under the ROC curve can be used to estimate credibility of the prediction.

For ROC curve plotting, it is necessary to take into account the ligands with good affinities as well as the ligands with bad affinities. All Erk2 ligands had good affinities in the CSAR 2011 data set, and the AUC value cannot be calculated for this target protein. So, we examined the quality of the docking into this protein only using full the CSAR 2012 data set.

On the basis of the CSAR 2011 results obtained by all groups, the organizers have chosen the comparative ligand data and have calculated the AUC value for each two sets. There was a unique data set for each protein. We have drawn the diagrams to compare the AUC values obtained by different research groups. The descriptions of the comparative ligand sets and the AUC values are presented in Table 3. The diagrams are shown in Figures 1a, 2a, and 3a.

Table 3. Descriptions of Ligand Sets Used for Building of Diagrams to Estimate Prediction Quality

protein	Chk1		LpxC		urokinase
total number	39		11		19
number of active ligands	30		3		15
number of inactive ligands	9		8		4
ligand series	series 1	series 2	series 1	series 2	series 1
AUC, Dimonta	0.44	0.68	0.92	0.17	0.97

We have also built ROC curves and calculated AUC values for each protein for the first and second ligand sets docked and post-processed by our programs SOL and DISCORE, respectively. The number of the Chk1 and urokinase ligands presented in the second set was less than the first set because some complexes failed to be computed by the DISCORE program. The descriptions of the ligand sets used for ROC curve building and respective AUC are presented in Table 4, where the ligand numbers are different from those presented in Table 3 because our docking was successful for larger ligand sets than comparative ones that have been chosen by the organizers to build diagrams in Figures 1a, 2a, and 3a.

To build the ROC curves, the list of ligands should be sorted by scoring function first. If the first ligand is active, the unit vector (0,1) must be drawn, and if not, the unit vector (1,0) must be drawn. Then, we take the next ligand and repeat the same procedure using the terminus of the first vector as the origin of the new vector. Thereby, we draw the ROC curve for the ranking list of the ligands and normalize it to unity. When using normalized units, the area under the curve is equal to the AUC value. ROC curves are presented in Figures 1b, 2b, and 3b.

Comparison of the binding free energies estimated by the docking program SOL and post-processing program DISCORE with experimental results shows that for some target proteins these programs have a good ability of ligand ranking with respect to binding energy (scoring) separating active compounds from inactive ones. This result confirms success of the previously made SOL validation for several target proteins in the search of inhibitors among large amounts (about 2000) of inactive compounds (to be more precise among supposedly inactive compounds).[12]

At the same time, for the Chk1 protein (SOL and DISCORE) and LpxC (for DISCORE), AUC values turned out to be too small to separate active and inactive compounds on the basis of their binding free energy estimation (scoring).

CSAR 2012 Benchmark Exercise. After finishing the CSAR 2011 contest, the organizers laid out the CSAR 2012 database containing two new target proteins, CDK2 and CDK2−CYCLINA, in addition to four previous target proteins. Ligand sets for the previous target proteins have been enlarged in the CSAR 2012 database in comparison with the CSAR 2011 database. Numbers of active and inactive ligands for each of the target proteins are presented in Table 5.

All molecules of CSAR 2012 have been obtained in MOL2 (*.mol2) format. Ligands were given in 3D structures with all their hydrogens. For native ligands, structures of complexes and

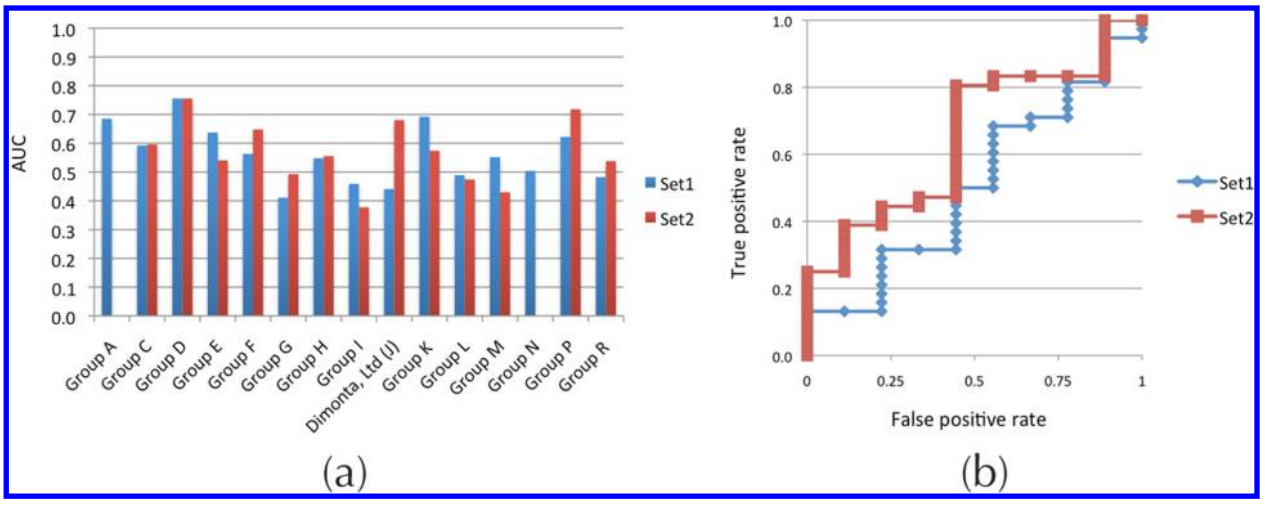


Figure 1. Estimation of prediction quality for the Chk1 protein. (a) AUC values comparative diagram for different research groups. (b) ROC curves for the Dimonta group.

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

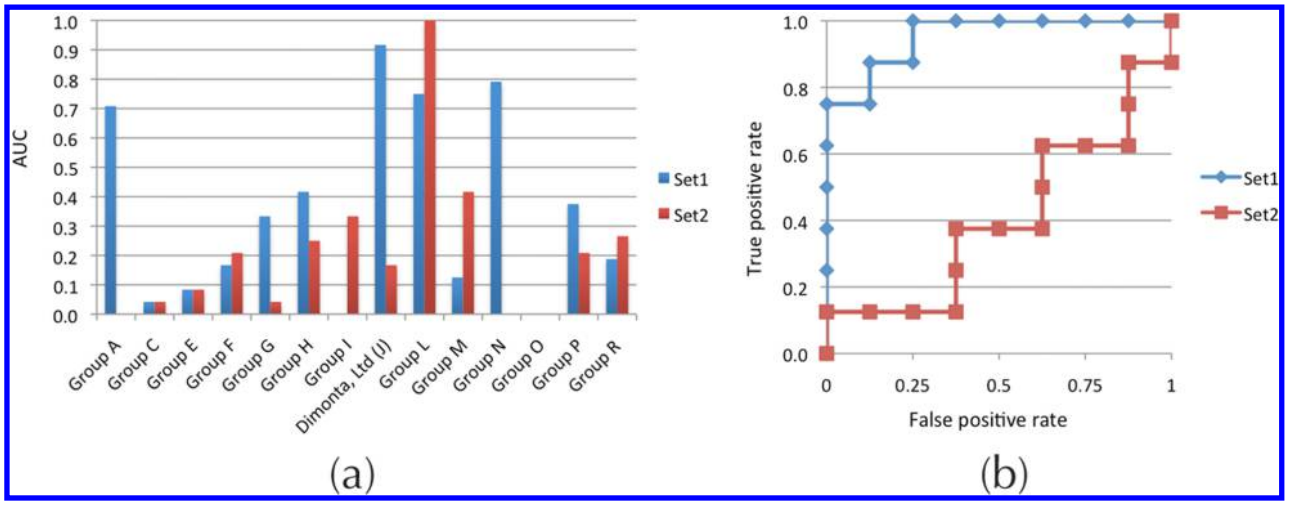


Figure 2. Estimation of prediction quality for the LpxC protein. (a) AUC values comparative diagram for different research groups. (b) ROC curves for the Dimonta group.

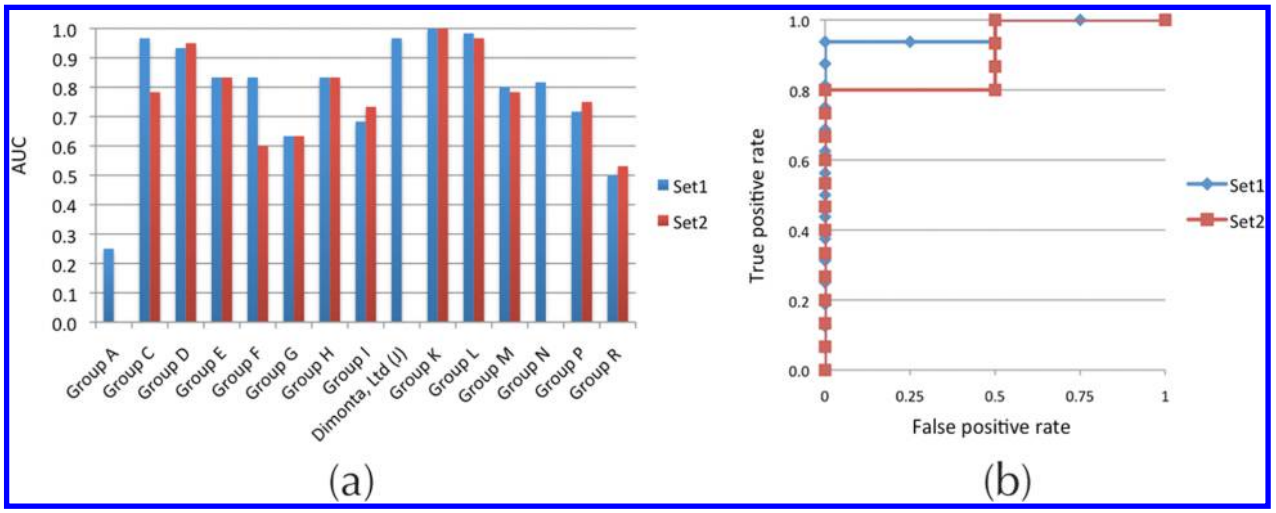


Figure 3. Estimation of prediction quality for the urokinase protein. (a) AUC values comparative diagram for different research groups. (b) ROC curves for the Dimonta group.

Table 4. Descriptions of Ligand Sets Used for ROC Curve Building by the Dimonta Research Group

protein	Chk1		LpxC		urokinase	
ligand set	set 1	set 2	set 1	set 2	set 1	set 2
total number	47	45	16	16	20	17
number of active ligands	38	36	8	8	16	15
number of inactive ligands	9	9	8	8	4	2
AUC	0.51	0.65	0.95	0.40	0.97	0.9

Table 5. Ligand Number in CSAR 2012 for Six Different Target Proteins

target protein	Chk1	Erk2	LpxC	urokinase	CDK2	CDK2–CYCLINA
ligand total number	137	52	32	64	118	115
active ligand number	107	52	20	53	26	23
inactive ligand number	30	0	12	11	92	92

separate molecules (proteins and ligands) were given. Using these data, we could analyze quality of ligand positioning for different target proteins. The same versions of all programs have been used as in CSAR 2011. We converted all ligands into the MOL (*.mol) format by our APLITE program for docking procedures, and added hydrogen atoms to protein structures by the APLITE program.

Ligand Positioning. Accuracy of the native ligand positioning by the docking program SOL is presented in Table 6. Here, the native ligand set 2011 consists of ligand structures used in the CSAR 2011 contest. These ligands have been transformed into 3D structures as described above in the Data Preparation section, and they have been docked into the protein structures taken from the Protein Data Bank.[13] Docked ligand poses have been sent to the CSAR 2011 organizers who sent to us back calculated RMSD values between docked and native poses. The CSAR 2012 database contains crystallized ligands and respective proteins structures, which are ideally suited for the ligand positioning validation goal.

As shown, the SOL docking accuracy is good (RMSD < 2 Å) for LpxC, urokinase, CDK2, and CDK2-CICLINA proteins with no bad results (RMSD > 3 Å) at all. SOL docking native ligand positioning was noticeably worse for Chk1 and Erk2 proteins. Approximately one-third of the native ligands could

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

Table 6. Ligand Positioning Quality by Docking Program SOL

	Chk1		Erk2		LpxC		urokinase		CDK2	CDK2− CYCLINA
protein structures	2e9n	2012	3i5z	2012	3p3e	2012	1owe	2012	2012	2012
native ligand structures	2011	2012	2011	2012	2011	2012	2011	2012	2012	2012
total ligand number	14	17	12	12	4	5	4	7	15	1
RMSD < 2 Å	6	10	3	7	4	5	3	3	14	1
RMSD = 2−3 Å	0	2	0	1	0	0	1	4	1	0
RMSD > 3 Å	8	5	9	4	0	0	0	0	0	0

Table 7. GFIT Coefficients for Chk1, LpxC, Urokinase Proteins

terms	k_1	k_2	k_3	k_4	k_5	k_6	k_7	B
Chk1 coefficients	0.037	0.017	0.012	0.02	0.029	0.015	0.014	−8.487
LpxC coefficients	0.155	0.107	0.372	0.071	0.1	0.061	0.077	−5.413
urokinase coefficients	0.039	0.025	0.008	0.006	0.016	0.013	0.013	−3.813

not be docked correctly (RMSD > 3 Å). However, the other two-thirds of the native ligands have been docked with accuracy better than 2 Å. Docking accuracy of the native ligands from CSAR 2012 was found to be better than in the case of CSAR 2011. In the latter case, docking was performed into protein structures taken from the Protein Data Bank,[13] but the native ligands have been taken from other protein−ligand structures—the structures used by CSAR organizers. On the other hand, in CSAR 2012 testing, each native ligand has been docked into the respective co-crystallized protein, and this gave definite improvement in docking accuracy.

Comparison Scoring Function with Experimental Data. In CSAR 2011, we performed docking for four proteins, and each of these proteins had its own ligands with known native positions. As shown in Table 6, we have good docking quality only for the proteins Chk1, LpxC, and urokinase; therefore, we used the program GFIT only for these three proteins.

All the experimental data were presented by organizers in terms of pIC50, pK_i, and pK_d. It is necessary to transform these experimental values into ΔG for using the program GFIT to fit the DISCORE scoring function. For this purpose, the next equation was used

$$\Delta G = R \times T \times \ln(E) \times \left(\frac{1}{4.18 \times 10^3}\right) \quad (7)$$

where E is the binding free energy in the terms of IC50, K_i, and K_d; R is the gas constant (8.31 J/(K mol)); T is the temperature (accepted as room temperature 300 K); and ΔG is the binding free energy (kcal/mol).

Also for program GFIT, we need to determine the boundary parameter "dG_bad_bottom" to separate active and inactive ligands. Parameter "dG_bad_bottom" has it own value for each protein.

As a result of GFIT performance, the fitting coefficients k_i (i = 1, ..., 7) and B have been obtained, and they have been used in the post-processing program DISCORE for prediction of CSAR 2012 ligand binding (Table 7).

With respect to the CSAR 2012 benchmark, we also added the data from CSAR 2011 to the GFIT training set and calculated independent coefficients for urokinase, Chk1, and LpxC proteins (hereinafter DISCORE GFIT). Because there were no ligands with bad affinities in CSAR 2011 for the ERK2 protein, we did not calculate new coefficients for this protein and used coefficients "for other proteins" (Chk1, Erk2, and LpxC), which we used in CSAR 2011 (hereinafter DISCORE INIT).

Comparison of experimental results with calculated SOL, DISCORE INIT, and DISCORE GFIT scores is presented in Figure 4, where the DISCORE INIT score is the score obtained

Figure 4. Comparison of experimental results with calculated SOL, DISCORE INIT, and DISCORE GFIT scores for LpxC and urokinase.

with universal coefficients for all proteins (Table 2); the DISCORE GFIT score is the score obtained for Chk1, LpxC, and urokinase proteins individually (Table 7). Ligands were sorted by affinity in ascending order for each protein (LpxC and urokinase). Horizontal green lines mark the ligands with good experimental affinity; horizontal red lines mark the ligands with bad experimental affinities. Vertical green lines mark the correspondence between theoretical prediction and the experiment; vertical red lines mark the contradiction between theoretical prediction and the experiments.

Examination of the difference between scores in SOL and DISCORE shows that all three scores (SOL, DISCORE INIT, DISCORE GFIT) have high correlation with each other for urokinase and have low correlation with each other for Chk1 (Table 8). For LpxC SOL and DISCORE GFIT, scores correlate with each other and differ from the DISCORE INIT score (Table 8). The results of the examination are presented in Figures 5−7.

Quality of Estimated Binding Free Energy: Quality of Scoring. Quality of ligand rankings in respect with the scores by the SOL and DISCORE programs for the CSAR 2012 data set are similar to those from CSAR 2011. For some target proteins, AUC values are sufficiently large (AUC > 0.9), but for

Table 8. Correlation Values between Scores in SOL and DISCORE (SOL, DISCORE INIT, DISCORE GFIT)

correlations values	Chk1	LpxC	urokinase
SOL_vs_DISCORE_INIT	0.33	0.11	0.91
SOL_vs_DISCORE_GFIT	0.24	0.72	0.92
DISCORE_INIT_vs_DISCORE_GFIT	0.37	0.07	0.96

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

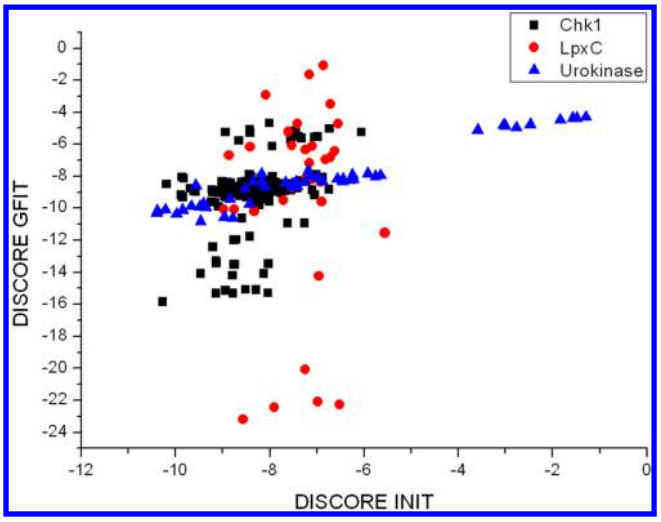


Figure 5. Correspondence between DISCORE INIT score and DISCORE GFIT score. *X* and *Y* directions measured in kcal/mol.

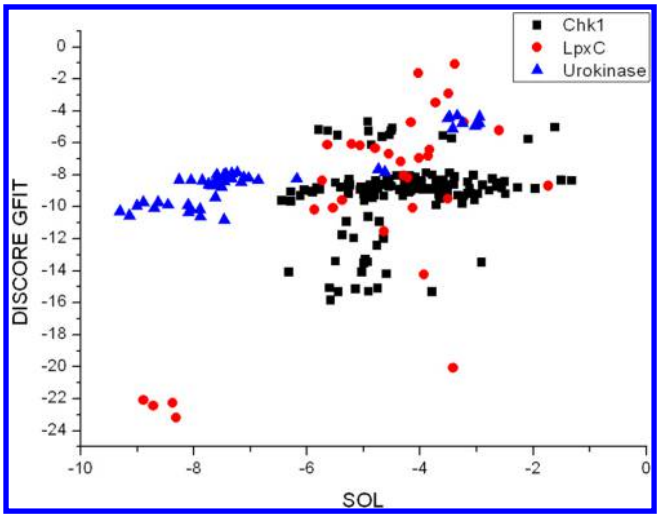


Figure 6. Correspondence between SOL score and DISCORE GFIT score. *X* and *Y* directions measured in kcal/mol.

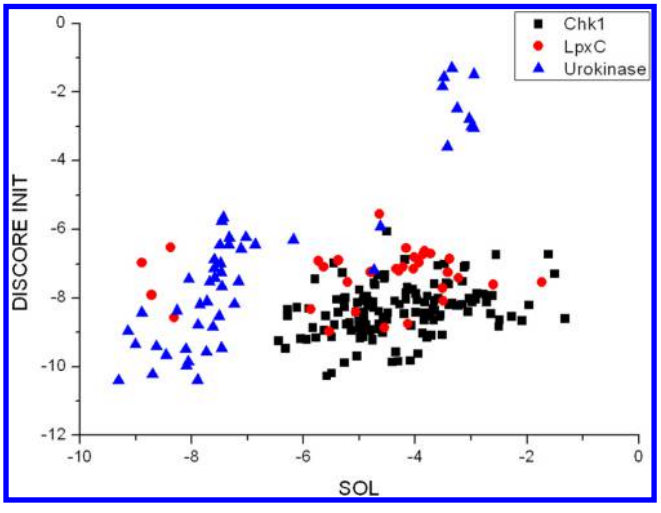


Figure 7. Correspondence between SOL score and DISCORE INIT score. *X* and *Y* directions measured in kcal/mol.

some target proteins AUC values are low (AUC ≈ 0.5−0.6). The AUC value is an average characteristic of the prediction quality, and it is possible to interrogate the prediction quality by analysis of true and erroneous predictions. In particular, this can be done with the next procedure.

All ligands for a given target protein are ranked in respect with their calculated score (estimated binding free energy). Then, a boundary scoring value separating predicted active compounds from predicted inactive ones is defined in such a way that ligands with scores larger than the boundary value are considered as inactive compounds for the given target protein, and ligands with scores lower than the boundary value are considered as active compounds. Results of such predictions are compared with experimental information about the ligands under consideration, and for experimental binding energies (recalculated usually from binding constants), the boundary binding energy separating experimentally active compounds from experimentally inactive ones is also defined. The latter can be different from the scoring boundary value.

If a given ligand calculated score and experimentally defined binding energy are below the respective boundary values, we designate such a case as TP (true positive), i.e., the prediction that this ligand is active for the given target protein is confirmed by experiment. If the calculated and experimental values are above the respective boundary values, we designate this case as TN (true negative). These two cases indicate predictive correctness of the program (true). The other two cases indicate erroneous prediction (false). The first case, when the ligand score is above the boundary value but the experimental energy is below the respective boundary value, is designated as FN (false negative), and it means that the calculating program predicts inactivity of the ligand that is really active. The second case, when the ligand score is below the boundary value but the experimental energy is above the respective boundary value, is designated as FP (false positive), and it means that the calculating program predicts activity of the ligand that is really inactive. Such comparison predictions with experiments in the binary scale allows for an estimation of prediction quality closer to the needs of real inhibitors design.

The respective results for all six target proteins are summarized in Tables 9−13, where the source of the protein structures is designated.

Table 9. Estimation of Predictive Quality by Docking Program SOL and Post-Processing Program DISCORE for Chk1 Protein

protein structure		2e9n			Chk1_4_pro
scoring program		SOL	DISCORE INIT	DISCORE GFIT	SOL
total ligand number		137			
true	positive	24	104	94	48
	negative	22	1	2	18
false	positive	8	30	29	12
	negative	83	2	12	59

It is shown in Table 9 that SOL gives twice as many false predictions as true ones for the Chk1 protein structure taken from the Protein Data Bank.[13] In the case of usage of the Chk1 structure obtained from CSAR organizers, numbers of correct and erroneous predictions become almost equal.

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

Table 10. Estimation of Predictive Quality by Docking Program SOL for Erk2 Protein

protein structure		3i5z	ERK2_000501_pro
scoring program		SOL	
total ligand number		52	
true	positive	30	16
	negative	0	0
false	positive	0	0
	negative	22	36

Table 11. Estimation of Predictive Quality by Docking Program SOL and Post-Processing Program DISCORE for LpxC Protein

protein structure		3p3e			LpxC_CS263_Bmonomer_pro
scoring program		SOL	DISCORE INIT	DISCORE GFIT	SOL
total ligand number		32			
true	positive	10	20	18	0
	negative	11	0	5	12
false	positive	1	12	7	0
	negative	10	0	2	20

Table 12. Estimation of Predictive Quality by Docking Program SOL and Post-Processing Program DISCORE for Urokinase[a]

protein structure		1owe			urokinase_8_pro
scoring program		SOL	DISCORE E-Ur	DISCORE GFIT	SOL
total ligand number		49			
true	positive	38	36	38	38
	negative	9	10	9	9
false	positive	2	1	2	2
	negative	0	2	0	0

[a]The DISCORE E-Ur score is the score obtained with old coefficients for urokinase (Table 2).

Table 13. Estimation of Predictive Quality by Docking Program SOL for CDK2 and CDK2−CYCLINA Proteins

protein structure		CDK2_CS20_pro	CDK2CyclinA_CS260_pro
scoring program		SOL	SOL
total ligand number		118	115
true	positive	19	15
	negative	69	74
false	positive	23	18
	negative	7	8

In spite of these non-exciting results, FP (false positive) numbers are rather low in these two cases of SOL predictions. But just such errors (FP) result in considerable time and expensive losses in the course of real inhibitors development. Application of post-processing DISCORE results in a sizable increase in TP predictions but also in an increase in FP cases.

For the Erk2 protein (Table 10) structure taken from PDB,[13] SOL gives approximately 1.5 times more correct predictions (true) than erroneous ones. Usage of the protein structure obtained from the organizers increases the FN number considerably. Nevertheless, there are no FPs at all in both cases.

For the LpxC protein (Table 11) structure taken from PDB[13], SOL gives approximately twice as many correct predictions (true) than erroneous ones, and application of post-processing DISCORE increases twice the number of TP predictions at the expense of a considerable decrease in the FN number. Unfortunately, there were no TP predictions when SOL worked with the protein structure obtained from the organizers.

Both programs, SOL and DISCORE, give correct predictions for the overwhelming majority of ligands for urokinase, being erroneous only in 2−3 cases out of 49 (Table 12).

SOL gives three times as many correct predictions than wrong ones for CDK2 and CDK2−CYCLINA proteins (Table 13), but unfortunately it demonstrates too many FP cases.

CONCLUSIONS

The execution of CSAR exercises has shown[28] that the docking program SOL has demonstrated a good ability for correct ligand positioning in the active sites of the tested proteins in most cases. Concerning the estimation of the protein−ligand binding free energy, the docking program SOL and the respective post-processing program DISCORE have not demonstrated very exciting results. Nevertheless, for some cases, e.g., LpxC and urokinase (Figures 2a and 3a), SOL won one of the first places in respect with AUC calculation; DISCORE demonstrated one of the best AUC values for the Chk1 protein (Figure 1a).

Obvious improvement in SOL predictions could be made by rejection of the grid approximation for calculation of protein−ligand interactions. Such a direct docking approach gives us an opportunity to include to some extent target−protein flexibility and to calculate more accurately the desolvation contribution of the binding energy. However, these improvements demand a considerable increase in the performance speed of the docking algorithm and energy evaluation.

The post-processing program DISCORE did not demonstrate considerable improvement in binding energy predictions compared with SOL ones for all cases. However, in many cases, the improvements were noticeable (Figure 1a, Tables 3 and 4, Chk1 column, Tables 9 and 11). Apart from the cases when SOL docking gave bad accuracies of ligand positioning, and it was impossible to wait for good DISCORE estimations of the binding energy, there are two main reasons for the ineffective performance of DISCORE.

First is the insufficient parametrization of the employed implicit water surface generalized Born model.[11] Respective parameters of this model, atomic radii for solvent excluded surface construction, have been fitted to the experimental data on hydration energies of small molecules[23] and a fairly small number of molecular ions. Taking into account that the desolvation energy in protein−ligand binding is a noticeable value (dozens of kcal/mol) and compensation of usually part of the direct protein−ligand Coulomb interactions, further improvement of the implicit water model parametrization must result in sizable improvement of DISCORE post-processing.

Second, the inadequate model used for the configurational entropy loss upon binding used in SOL and DISCORE programs should be a serious source of errors in binding energy calculations. This loss is proportional to the number of ligand internal torsion degrees of freedom in SOL as well as in DISCORE programs. However, such an approach can be quite wrong as shown by Chang et al.[29]

dx.doi.org/10.1021/ci400094h | J. Chem. Inf. Model. 2013, 53, 1946−1956

So, modification of DISCORE by improving the surface GB parametrization and refining the configurational entropy loss upon ligand binding should result in better post-processing performance in the binding energy prediction and approaching all fitting coefficients in DISCORE to unity thereby making DISCORE scoring function more universal and objective.

It is too early to speak now about a good correlation between calculations with SOL and DISCORE and experimentally measured binding free energies. Further improvement of the binding free energy calculations accuracy is needed.

■ ASSOCIATED CONTENT

Supporting Information

Parallel implementations of docking programs SOLGRID and SOL. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: +7 (495) 9394004. E-mail: vladimir.sulimov@gmail.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the organizers of CSAR benchmark, all people who submited their experimental results for it, the organizers of the "CSAR 2011−2012 Benchmark Exercise" session of Division of Computers in Chemistry at the 244th ACS National Meeting (Philadelphia, Pennsylvania, August 19−23, 2012), and especially Professor Heather A. Carlson.

■ ABBREVIATIONS

TP, true positive; TN, true negative; FP, false positive; FN, false negative; AUC, area under curve; ROC, receiver operating characteristic; vdW, van der Waals; MMFF, Merck molecular force field; SGB, surface generalized Born model

■ REFERENCES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935−949.

(2) Zoete, V.; Grosdidier, A.; Michielin, O. Docking, virtual high throughput screening and in silico fragment-based drug design. *J. Cell. Mol. Med.* **2009**, *13*, 238−248.

(3) Gastreich, M.; Lemmen, C.; Briem, H.; Rarey, M. In *Virtual Screening in Drug Discovery*; Alvarez, J., Shoichet, B., Eds.; CRC Press: Boca Raton, FL, 2005; pp 25−46.

(4) Cole, J. C.; Nissink, J. W. M.; Taylor, R. In *Virtual Screening in Drug Discovery*; Alvarez, J., Shoichet, B., Eds.; CRC Press: Boca Raton, FL, 2005; pp 379−415.

(5) 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 empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(6) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−2791.

(7) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein−ligand docking: Current status and future challenges. *Proteins* **2006**, *65*, 15−26.

(8) Romanov, A. N.; Kondakova, O. A.; Grigoriev, F. V.; Sulimov, A. V.; Luschekina, S. V.; Martynov, Ya.B.; Sulimov, V. B. The SOL docking package for computer-aided drug design. *Numer. Methods Program.* **2008**, *9* (2), 64−84 , (in Russian).

(9) Halgren, T. A. Merck molecular force field. *J. Comput. Chem.* **1996**, *17* (5&6), 490−641.

(10) Ghosh, A.; Rapp, C. S.; Friesner, R. A. Generalized Born model based on a surface integral formulation. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.

(11) Romanov, A. N.; Jabin, S. N.; Martynov, Y. B.; Sulimov, A. V.; Grigoriev, F. V.; Sulimov, V. B. surface generalized Born method: A simple, fast and precise implicit solvent model beyond the Coulomb approximation. *J. Phys. Chem. A* **2004**, *108*, 9323−9327.

(12) The validation of SOL docking program. http://www.keenbase.ru/en/index.htm (accessed June 21, 2013).

(13) Protein Data Bank. http://www.rcsb.org/ (accessed June 21, 2013).

(14) Moscow University Supercomputing Center. http://hpc.msu.ru/ (accessed June 21, 2013).

(15) Sinauridze, E. I.; Romanov, A. N.; Gribkova, I. V.; Kondakova, O. A.; Surov, S. S.; Gorbatenko, A. S.; Butylin, A. A.; Monakov, M. Y.; Bogolubov, A. A.; Kuznetsov, Y. V.; Sulimov, V. B.; Ataullakhanov, F. I. New synthetic thrombin inhibitors: Molecular design and experimental verification. *PLoS One* **2011**, *6* (5), e19969 DOI: 10.1371/journal.pone.0019969.

(16) Oferkin, I. V.; Sulimov, A. V.; Kondakova, O. A.; Sulimov, V. B. Implementation of parallel computing for docking programs SOLGRID and SOL. *Numer. Methods Program.* **2011**, *12* (1), 205−219 , (in Russian).

(17) Nocedal, J.; Wright, S. J. *Numerical Optimization*, 2nd ed.; Springer Science+Business Media, LLC: New York, 2006; section 7.2.

(18) Kupervasser, O. Y.; Zhabin, S. N.; Martynov, Y. B.; Fedulov, K. M.; Oferkin, I. V.; Sulimov, A. V.; Sulimov, V. B. Continual model of solvent: The DISOLV software—Algorithms, implementation, and validation. *Numer. Methods Program.* **2011**, *12*, 246−261 , (in Russian).

(19) Sulimov, V. B.; Katkova, E. V.; Oferkin, I. V.; Kochugaeva, M. P. PCM, COSMO, and Surface GB Implementation in Program DISOLV: Development and Validation, 244th ACS National Meeting, Philadelphia, Pennsylvania, August 19−23, 2012, oral presentation COMP 441.

(20) Zhabin, S. N.; Sulimov, V. B. Construction of the adaptively triangulated smooth molecular surface: Program TAGSS. *Sci. Visualization* **2011**, *3* (2), 27−53.

(21) Connolly, M. L. Solvent-accessible surfaces of proteins and nucleic acids. *Science* **1983**, *221* (4612), 709−713.

(22) Vorobjev, Y. N.; Hermans, J. SIMS: computation of a smooth invariant molecular surface. *Biophys. J.* **1997**, *73*, 722−732.

(23) Bordner, A. J.; Cavasotto, C. N.; Abagyan, R. A. Accurate transferable model for water, n-octanol, and n-hexadecane solvation free energies. *J. Phys. Chem. B.* **2002**, *106*, 11009−11015.

(24) Bohm, H.-J. 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. Design* **1994**, *8*, 243−256.

(25) Schwartzl, S. M.; Tschopp, T. B.; Smith, J. C.; Fischer, S. Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction? *J. Comput. Chem.* **2002**, *23*, 1143−1149.

(26) ChemAxon. http://www.chemaxon.com/ (accessed June 21, 2013).

(27) Open Babel: The Open Source Chemistry Toolbox. http://openbabel.org/ (accessed June 21, 2013).

(28) Sulimov, A. V.; Kutov, D. K.; Katkova, E. V.; Sulimov, V. B. In *Application of the Docking Program SOL for CSAR Benchmark*, 244th ACS National Meeting, Philadelphia, Pennsylvania, August 19−23, 2012, oral presentation COMP 48.

(29) Chang, C. A.; Chen, W.; Gilson, M. K. Ligand configurational entropy and protein binding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1534−1539.