JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

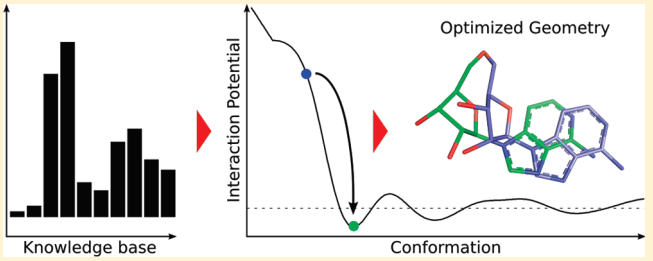
MiniMuDS: A New Optimizer using Knowledge-Based Potentials Improves Scoring of Docking Solutions

Andreas Spitzmüller,[†] Hans F. G. Velec,[†,‡] and Gerhard Klebe*[,†]

[†]Institute of Pharmaceutical Chemistry, Philipps-Universität Marburg, Marbacher Weg 6, D-35032 Marburg, Germany

Supporting Information

ABSTRACT: In small molecule docking, the scoring and ranking of generated conformations is an important, though still not a completely resolved problem. Rescoring schemes often improve the quality of the obtained rankings. It is known that a local optimization is essential before a valid rescore value can be calculated. Here, we present a method that improves rescoring results obtained with the DrugScore function due to a new optimization technique. The method implements a more sophisticated search algorithm compared to the classic local optimization procedures used in this context. We validated the proposed method on a set of 192 protein−ligand complexes. Results show substantial improvements compared to original docking results with success rates increased by up to 10% for top scored solutions below 2 Å root-mean-square deviation to the native state and up to 18% increase below 1 Å, respectively.



INTRODUCTION

In structure-based drug design, small molecule docking is a well established and widely used technique to predict the geometry of protein−ligand complexes. It is applied in lead identification as well as lead optimization to predict binding modes and affinities of compounds in complex with a target protein. From a computational point of view docking consists of two subproblems. The first one is the generation of reasonable ligand geometries inside the binding pocket. This is often referred to as the docking problem. Usually multiple solutions are suggested. The second one is the detection of those poses corresponding to the biologically relevant state out of the whole set of generated solutions. This is called the scoring problem.

A community-wide accepted criterion for reasonable ligand conformations is a root-mean-square deviation (rmsd) value below 2 Å compared to an experimentally determined structure. In the case of redocking a compound into its native receptor conformation, modern docking programs are generally able to explore conformational space sufficiently to generate such ligand geometries. This was shown in various studies evaluating popular docking algorithms on independently compiled test sets.[1−3]

In contrast, the problem of identifying the best pose among all generated solutions still remains as insufficiently solved. Dozens of scoring functions have been reported in literature in the past. Research in this field is still going on as reflected in regular benchmark studies on scoring functions.[4−6] Rescoring and consensus scoring are based on the idea of making the maximal use of the available scoring functions through their combination within one experiment.[7−9]

One such scoring function is DrugScore,[10,11] a knowledge-based scoring function that uses distance-dependent statistical potentials to evaluate pairwise atomic contacts between ligand and protein. In all the above-mentioned benchmark studies, DrugScore is listed among the best performing scoring functions concerning both the recognition of near-native binding modes and the ranking of different ligands with respect to their predicted binding affinity.

Originally, the statistical potentials were derived from noncovalent atomic contacts in protein−ligand complexes as stored in the Protein Data Bank (PDB).[12] In 2005, Velec et al. reported a new version, DrugScore^{CSD},[10] where the potentials were retrieved from the spacial distribution of nonbonding interactions in small molecule crystals as stored in the Cambridge Structural Database (CSD).[13] This was motivated by the assumption that atom−atom contacts in small molecule crystals are subject to the same physical principles as those in protein crystals. Considering the CSD, the statistical basis for individual contact types is much broader as compared to the PDB. This led to a highly increased statistical significance of many pair potentials along with an improved performance.

In addition, the substantially higher resolution and experimental accuracy of small molecule crystal structures is related to lower positional uncertainties for individual atoms compared to the PDB. This, however, introduces a remarkable side effect: The CSD-derived potentials show steeper energy wells and are more sharply bounded. Figure 1 illustrates these effects for a charged interaction as it occurs between a deprotonated carboxylate oxygen and a protonated amidine nitrogen. This in turn makes DrugScore^{CSD} more sensitive to small conformational and positional variations. Bearing in mind that DrugScore is typically

Received: February 24, 2011
Published: April 29, 2011

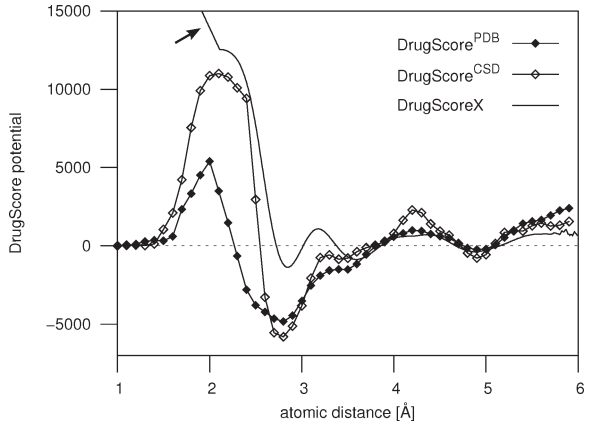


Figure 1. Pair potentials for an O.co2−N.pl3 contact in DrugScore^{PDB} (filled squares, ■) and DrugScore^{CSD} (empty squares, □) and an O.co2−N.mih contact in DrugScoreX (no squares).

used for rescoring, this effect becomes even more relevant. The poses subjected to rescoring by DrugScore have originally been generated and optimized under the influence of a different scoring function. However, local and global optima of the latter function and DrugScore will not be exactly the same. Thereby, a given ligand geometry positioned in a local optimum of the original scoring function can obtain a quite unfavorable DrugScore value even though a local DrugScore optimum can be in very close vicinity. For this reason a local optimization of the poses according to the function applied for rescoring is highly advisible.[14] O'Boyle et al. recently emphasized this point again,[7] noting however, that a strictly local optimization cannot escape the energy well on the potential surface the pose resides in. Thus, a conformationally or spatially very close but better scoring pose from a neighboring well might not be considered.

Here we introduce an approach to optimize ligand conformations with respect to DrugScore potentials that accounts for and tries to efficiently avoid the above-mentioned problem of strictly local optimizations. Therefore, a global minimization strategy based on local optima smoothing[15] was implemented into the program MiniMuDS (**Mini**mizing **M**olecules **u**sing **D**rug**S**core potentials). We want to emphasize that it is not our intention to generate new binding modes during the optimization. This would be equivalent to the development of an alternative docking program. We want to postoptimize given binding modes to agree to the local optima of the DrugScore function. Therefore, our implementation does not perform a global search in the entire conformational and configurational space but rather a global search within a tightly restricted area of the search space. However, this restricted space is still significantly larger than the area covered by a strictly local search.

We validated our method on a data set comprising 195 high-quality protein−ligand complex structures.[4] First, we optimized the considered crystal structures to evaluate by how much the optima in our objective function deviate from the crystallographically determined native states. Second, we analyzed by how much the given geometries obtained from docking are modified through optimization. Third, we compared the rankings obtained from minimized and non-minimized docking solutions.

■ MATERIALS AND METHODS

The Energy Model. In the applied energy model DrugScore potentials account for all intermolecular interactions between ligand and protein. However, to avoid a collapse of the ligand in the course of the optimization, the intramolecular interactions have to be modeled too. Bond lengths and angles are kept fixed, assuming that the pose generating program already adjusted these parameters to reasonable values. Thus, only torsion energies and van der Waals interactions are modeled during optimization. The variables to be optimized comprise the ligand's torsion angles and its free translation and rotation in space, whereas the protein is kept rigid. The resulting objective function is given as

$$E = \alpha E_{DS} + \beta E_{tors} + \gamma E_{vdW}$$

where the factors α, β, and γ allow for different weightings of the three interaction types.

The van der Waals interactions are described using a standard 12-6 Lennard-Jones potential as used in the Sybyl-X[16] implementation of the Tripos force field.[17] Similar to the DrugScore interactions, the torsional degrees of freedom are modeled through knowledge-based torsion angle potentials also derived from the CSD, as described by Klebe and Mietzner.[18]

Applied Pair Potentials. The current implementation of MiniMuDS uses a new version of DrugScore potentials developed since 2005 by Gerd Neudert (to be published). This section briefly describes the major differences between DrugScore^{CSD} and this new version, called "DrugScoreX" (see Figure 1 for the corresponding new potential for the carboxylate-amidine interaction).

DrugScoreX potentials are also derived from CSD structures. However, a new set of atom types has been applied. DrugScore^{CSD} was based on atomic contacts according to the Sybyl atom-type conventions. In contrast, DrugScoreX is based on the new fconv atom types[19] which are much more differentiated. The bin width for the sampling of contact frequencies was originally chosen to be 0.1 Å. As potential differences in neighboring bins can still be very large (cf. Figure 1), an interpolation was necessary for intermediate atomic distances. In DrugScoreX this interpolation is already included in the derivation procedure of the potentials. Discrete potential values are then recorded using a step size of 0.01 Å. This makes an ad hoc interpolation unnecessary, as instead a much more efficient look-up table of the discrete values can be used. In addition, the new potentials are extended by a linear term for very short atomic distances (arrow in Figure 1). Since obviously no observations can be made in the knowledge base at these distances, potentials are artificially extended in this range to account for the strong repulsion that would occur between two nonbonded atoms.

However, the presented method is not restricted to the fconv atom-type model. In general, any consistent set of knowledge-based potentials can be used for the minimization.

The Optimization Algorithm. Generally, an overall funnel shape can be expected for the energy surface of a small molecule ligand binding to a target protein.[20] However, this funnel can be locally strongly perturbed. We assume a similar behavior for the objective function in our optimization problem, since the interactions between a small molecule ligand and its protein receptor are to be optimized.

The optimization procedure implemented in MiniMuDS is a modification of the "algorithm based on local smoothing for optimization" (ALSO) described by Addis et al.[15] A technically very detailed description of this algorithm is given as Supporting Information. There, also all differences between the original ALSO and the MiniMuDS implementation are explained, and

the actual parametrization is documented. Here, only the general principals will be described.

ALSO tries to reveal the overall funnel shape of the objective function $f(x)$ which is free of any local perturbations and therefore a much easier target for optimization. To this end, the piecewise constant step function $L(x) = f(\mathrm{LS}_f(x))$ is determined, where $\mathrm{LS}_f(x)$ is the result of a local search on f starting at position x. $L(x)$ already shows the underlying funnel structure. However, as a step function exhibits a constant gradient of zero (if defined at all), it is not suitable for optimization. Thus, $L(x)$ is smoothed using a Gaussian kernel:

$$g(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$$

The smoothed function is then given by

$$\langle L\rangle_g(x) = \frac{\int g(\| x - y \|)L(y)\mathrm{d}y}{\int g(\| x - y \|)\mathrm{d}y}$$

Due to the nature of the structure optimization problem, it is neither possible to get an analytical expression for $\langle L\rangle_g(x)$ nor to obtain a numerical estimate. That is why an approximation within a local environment is used. Based on a current point x_c, K random samples are drawn from L within a neighborhood $B(x_c)$ around x_c. Using this information $\langle L\rangle_g(x)$ can be approximated within B:

$$\hat{L}_g^B(x) = \frac{\sum_{i=1}^{K} g(\| x - y_i \|)L(y_i)}{\sum_{i=1}^{K} g(\| x - y_i \|)}$$

and subjected to an optimization to get the next optimum x_{c+1}. In an iterative way x_{c+1} is used as the new center for sampling in $B(x_{c+1})$ in the next optimization cycle. To speed up this procedure, a new iteration is already started whenever a x_i better than the current x_c is sampled from $B(x_c)$.

Due to the discrete pair potentials, a gradient-free method is necessary to determine $L(x)$ during the sampling phase. This is done using a Nelder–Mead–Simplex algorithm.[21] A variant of this algorithm is applied to optimize on the approximated function $\hat{L}_g^B(x)$. It is modified to ensure that the search stays within the boundaries defined by B based on the ideas of the so-called "complex" method described by Box.[22]

Parametrization. The number of samples K drawn from $B(x_c)$ in each iteration is set to $2 \times n$, i.e., twice the dimensionality of the optimization problem, where n is the number of rotatable bonds + 3 (translation) + 3 (rotation). The neighborhood B is limited to a maximal rotation of $\pm 5°$ around each rotatable bond, a maximal translation of 0.1 Å in each spacial direction, and a maximal total rotation of about 10°. Random points are drawn from within a n-dimensional hypersphere of radius 1. Then every dimension is scaled according to the bounds of B. Random rotation matrices are generated according to the algorithm proposed by Arvo,[23] which guarantees uniformly distributed rotation matrices (which would not be the case for randomly sampled Euler angles).

A newly found optimum is accepted at the end of each iteration whenever it improves the score. In addition, the user can apply a further acceptance criterion, namely a maximal rmsd with respect to the input structure. This reduces the available search space and thereby the computational complexity of the optimization problem. By default this parameter is set to 2 Å. This also ensures that the overall binding mode of the optimized geometry is conserved.

To determine convergence of the iterative procedure, the number of successive iterations that do not improve the current optimum with respect to its score is limited to 10. If a newly found optimum differs by less than 0.1 Å from the previously best point five times in a row, then the optimization is also terminated.

To get suitable values for the weighting factors of the energy model, α was set to 1 while a systematic search was performed on β and γ. The Astex Diverse Set[24] consisting of 85 high-quality protein structures was chosen as a training set. It is freely available at http://www.ccdc.cam.ac.uk/. Since both E_{tors} and E_{vdW} model internal ligand interactions, they are only affected by modifications of the torsion angles. Thus, four complexes without any freely rotatable bond were excluded from the training set. For each of the remaining 81 complexes, 10 randomly deflected conformations were generated such that the torsion angle of every rotatable bond deviated by 10°–60° from the native state. The resulting 810 conformations were systematically optimized assigning all combinations of the values 0 and 10^i for $i = -3, ..., 3$ and in a second, more fine-grained run all combinations of 0 and 2^i to β and γ. This range of values was selected to simulate situations from completely neglected to strongly dominating individual terms. For each combination the mean rmsd with respect to the atomic coordinates and with respect to the torsion angles between optimized and native states were calculated. As best performing values $\beta = 0.5$ and $\gamma = 8$ were selected.

Workflow. MiniMuDS expects a protein file in pdb or mol2 and a ligand file in mol2 format. Prior to the actual optimization, a few preprocessing steps are performed. A binding pocket around the given ligand is extracted, freely rotatable bonds are detected, and fconv atom types are assigned. All these tasks are performed using the external tool fconv.[19]

Validation Data. The introduced method was validated on a data set compiled by Cheng et al.[4] for a comparative assessment of scoring functions. It contains 65 diverse proteins, each represented by three structures in complex with different ligands (195 complexes in total). This set is freely available at http://www.pdbbind.org.cn/. The structures are readily prepared for docking experiments and were used without any modification. Three complexes were in common with the above-mentioned training data set and were thus excluded from any evaluation.

For each complex, two sets of 50 docking solutions each were generated using two different docking engines: (i) FlexX[25] in the LeadIT release 1.2 with default settings and (ii) Gold[26,27] version 4.1 using default settings with 50 GA runs and early termination not allowed. To get a broad coverage of the available space in the pocket, the generation of diverse solutions was activated with an rmsd cutoff of 1 Å and a cluster size of 1. Both FlexX and Gold are widely used programs, and they are built on two completely different types of docking algorithms. In both cases an area of 5 Å around any ligand atom in the crystal structure was used to define the binding region. To prevent a bias in the direction of the crystal structure, relaxed input conformations of the ligands as contained in the data set were used for docking.

Principally, also any other docking engine, such as DOCK,[28] Glide,[29] or ICM[30] could be used to generate a set of reasonable docking solutions. As a stand-alone tool, MiniMuDS generally

Table 1. Impact of MiniMuDS on Crystal Structures

	all	lead-like	nonlead
no. of complexes[a]	192	107	85
average rmsd [Å][b]	0.53 (±0.34)	0.50 (±0.30)	0.57 (±0.37)
rmsd >1 Å [%][c]	9.9 (19)	8.4 (9)	11.8 (10)
rmsd >1.9 Å [%][c]	0 (0)	0 (0)	0 (0)

[a] Total number of optimized complexes in the optimized data set. [b] Mean rmsd between input and output conformation (standard deviation in parentheses). [c] Fraction of optimized structures with a rmsd above the given threshold (number of cases in parentheses).

operates independently of the initially applied docking program. All rmsd values were calculated with fconv[19] regarding non-hydrogen atoms only. Results from MiniMuDS are also compared to a strictly local optimization according to the DrugScoreX potentials (named "Local" in the Results section). This was done using a standard local optimization[21] like it is also used by MiniMuDS to determine the step function.

RESULTS AND DISCUSSION

We evaluated the performance of the described method in three steps. First, we applied it to the original crystal structures to see how well the optima of our objective function are in agreement with crystallographically determined native states of the complexes. Second, we optimized all generated docking solutions. To assess the influence of the optimization on the structural quality, the rmsd values to the native state before and after minimization are compared. Finally, the influence of the optimization on the ranking of the solutions was examined.

The employed test set contains a large range of compounds from small, fragment-like ones up to oligo-peptides. The latter molecules are usually highly flexible due to their large number of rotatable bonds. In addition, they often bind to the surface of a protein receptor in rather flat binding regions. This makes them a difficult task for docking algorithms and also provides a challenge for the presented method. However, such big ligands are usually not considered as drug- or lead-like. Therefore, we divided the data set into two subsets: (i) the "lead-like" subset containing 107 complexes satisfying the definition of Oprea[31] and (ii) the "non-lead" subset containing the remaining 85 structures. Descriptors to determine lead-likeness were calculated with the program MOE.[32]

Optimization of Crystal Structures. We started with the optimization of the given crystal structures. Table 1 shows the impact of the minimization in terms of the average rmsd between native and minimized ligand geometries. Only small modifications are recorded (0.53 Å on average). We also observed that only very few structures show movements of more than 1 Å (9.9%), although modifications up to 2 Å had been allowed during optimization.

Whereas lead-like structures show movements around 0.50 Å, the generally larger nonlead compounds seem to be slightly stronger modified (around 0.57 Å). This tendency is also reflected in the fraction of structures modified by more than 1 Å (8.4% of the lead-like versus 11.8% of the non-lead ligands). However, these differences are statistically not significant.

No complex reaches the upper limit of the allowed search space with a final rmsd between 1.9 and 2.0 Å. This indicates that the examined space of 2 Å around the input structure is large enough to find a suitable optimum with respect to the DrugScoreX potentials, at least in the case crystal structures are used as input.

The question remains whether the rather small shifts in the range of 0.5 Å indeed reflect a good agreement between the implemented energy model and the experimentally determined native states or whether this reflects the general behavior of the optimization algorithm on any input data. To clarify this we subsequently optimized large sets of generated docking solutions.

Table 2. Impact of MiniMuDS on Docking Poses[a]

	all	lead-like	nonlead
no. of poses	18.736	10.341	8.395
average rmsd [Å]	0.99 (±0.41)	0.99 (±0.43)	0.98 (±0.40)
rmsd >1 Å [%]	41.0 (7.683)	41.8 (4.327)	40.0 (3.356)
rmsd >1.9 Å [%]	4.3 (811)	4.9 (503)	3.7 (308)

[a] Rows correspond to Table 1.

Optimization of Docking Poses. We applied the two docking engines Gold and FlexX to generate up to 50 poses per complex (see Methods Section for details). Table 2 summarizes the results of the optimization of these docking poses. First result to note is that the poses are modified twice as much as in the case of the crystal structures (about 1 Å average rmsd). This time, more than 40% of the optimized structures show movements by more than 1 Å. This confirms that the small modifications observed for the crystal structures are not the general behavior of the optimization algorithm but indeed reflect a high agreement between the optima of the implemented objective function and the experimentally determined native states of the structures. Yet, still only a small fraction of the optimization runs fully exploit the allowed search space (below 5%). This supports the reasonable setting to restrict the search space to 2 Å rmsd with respect to the input geometry. In contrast to the crystal structures, the extent of the modifications seems to be independent from the compound type. There is no observable difference between lead-like and non-lead subsets.

Next, we analyzed the influence of the optimization on the structural quality of the docked conformations. Here, the quality is expressed in terms of the rmsd between a docking pose and the corresponding native state. Taking the difference between the rmsd before and after the optimization, we get positive values if the optimized geometry is closer to the crystal than the non-optimized state. Negative ones are received if the optimized geometry departs further from the crystal structure than the docking pose.

Figure 2 shows the results for all generated docking solutions. On average, a pose is improved by about 0.12 Å, as depicted by the solid black line. This means that in general the optimization moves a docking pose toward the native state. The bars give the improvement for all docking poses that fall within the same range of rmsd to the crystal structure, averaged across bins of 0.1 Å. This allows for a more differentiated view. As the individual bins below 0.5 Å each contain less than 100 poses, they are summarized in the first bar (light gray). This prevents the calculation of statistically meaningless average values based on only a few poses within one bin. The same has been applied for the last bin, which comprises all poses above 6.8 Å.

We observe that the optimization of geometries between 1 and 2 Å rmsd performs well above average with an at least 2-fold better improvement. Notably, this is the most interesting range in structure-based drug design. Beyond about 3 Å the improvement

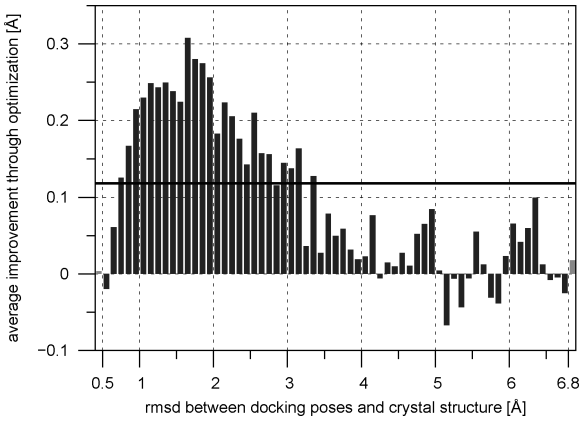


Figure 2. Average improvement for all 18 736 docking poses (solid black line) and for poses within a common range of rmsd to the crystal structure (bars). The light-gray bars pool all poses below 0.5 Å and above 6.8 Å rmsd, respectively.

that can be expected from the optimization rapidly decreases with the increasing deviation of the input structure from the native state. Nevertheless, even in this range the docking solutions are rarely perturbed in a way to deviate further from the crystal structure. Again, only marginal differences between the FlexX and the Gold poses were recorded.

Overall, rather small improvements are observed (generally below 0.3 Å). This reflects once more our intention not to generate new poses but to preserve the binding mode of a given docking solution. However, considering the steepness of the applied potentials, these slight adaptions can be essential for scoring. They subsequently allow DrugScoreX to better discriminate between well-docked and nonrelevant poses. This finally leads to the most important question: How much do these small modifications affect the scoring and finally the ranking of the poses?

Pose Recognition Performance. To evaluate the influence on the ranking of the solutions we examined four different scenarios: (i) the ranking obtained from the docking program; (ii) the one obtained from rescoring the unmodified docking solutions with DrugScoreX; (iii) the one from rescoring solutions applied to a standard local optimization according to the DrugScoreX potentials; and (iv) the one obtained from rescoring poses optimized using the new procedure implemented in MiniMuDS. Furthermore, we wanted to assess if it is necessary to optimize all 50 generated docking solutions. Thus, we also evaluated the rankings when only the 10 top-scored poses by the docking program are considered for optimization (denoted by the superscript "Top10" label added to the method names).

The ranking performance is given in terms of success rates for identifying near native poses on rank 1. The success rate is defined as the fraction of complexes for which the top-ranked pose is found within a certain rmsd cutoff compared to the native state. Figure 3 gives the results for the complete test set. The upper part shows the success for poses generated by FlexX, the lower part for the Gold solutions, respectively.

FlexX yields a rank 1 solution within 2 Å to the crystal structure for 44.3% of the cases (Figure 3, top left). In 21.9% the top-ranked solution is found within 1 Å. When the unmodified poses are rescored with DrugScoreX, these success rates drop by 1.0 and 4.7%, respectively. An increase of 6.3% at a 2 Å cutoff compared to the original docking poses is obtained when a local optimization is performed before the solutions are rescored. At the 1 Å level the

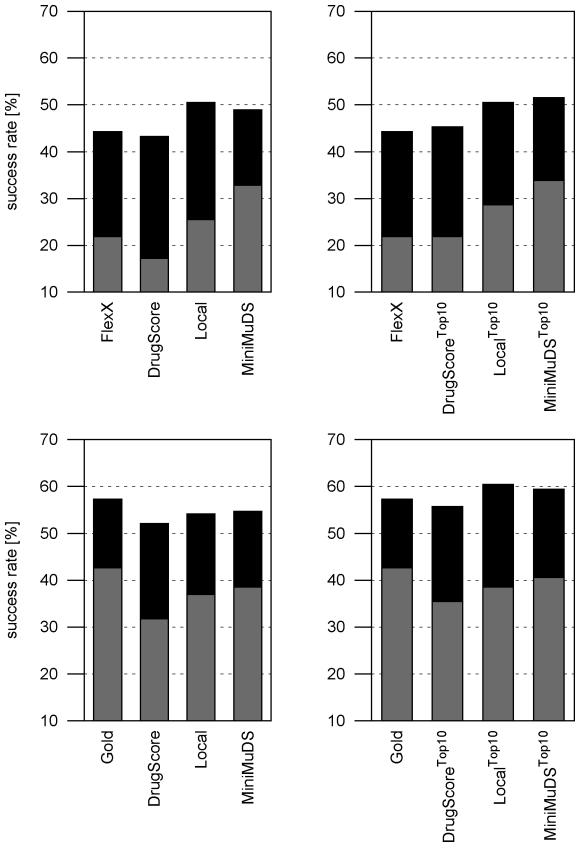


Figure 3. Success rates for the complete test set with top ranked solution below 1 Å rmsd (gray) and below 2 Å rmsd (black). Top: FlexX docking solutions; bottom: Gold docking solutions; left: optimization of all poses; and right: optimization of top 10 docking poses.

improvement reaches 3.6%. An optimization with MiniMuDS increases the success rates by 4.7% at 2 Å and 10.9% at 1 Å.

Slightly better results are obtained when only the 10 top-ranked FlexX solutions are subjected to following rescoring procedures (Figure 3, top right). The success rates for direct rescoring do not drop any more. Local^{Top10} optimization again increases success by 6.3% at 2 Å and 6.8% at 1 Å. MiniMuDS^{Top10} raises success by 7.3 and 12.0% to finally reach 50.8% at 2 Å and 33.9% at 1 Å, respectively. In summary, we observe a steadily increasing improvement of the results, starting from the direct rescoring scheme using DrugScoreX, through the local optimization according to the rescoring function, to the point of minimizing with MiniMuDS. Notably, it seems to be sufficient to consider the 10 top-ranked solutions only for the optimization. This strategy actually yielded the best results.

In case of Gold docking we receive a rank 1 solution within 2 Å for 57.3% of the complexes and one within 1 Å for 42.7% (Figure 3, down left). Here we observe a drop by 5.2 and 10.9%, respectively, when DrugScoreX is directly used for rescoring. Neither a local minimization nor the use of MiniMuDS can fully compensate for this effect. Only if the optimization is limited to the top 10 Gold solutions we finally obtain results comparable to the original docking poses with an increase of 2.1% at 2 Å and a decrease of 2.1% at 1 Å (Figure 3, down right).

The reason for the minor performance especially on the Gold solutions can be found in the composition of the data set. We indicated already that non-lead structures are more strongly

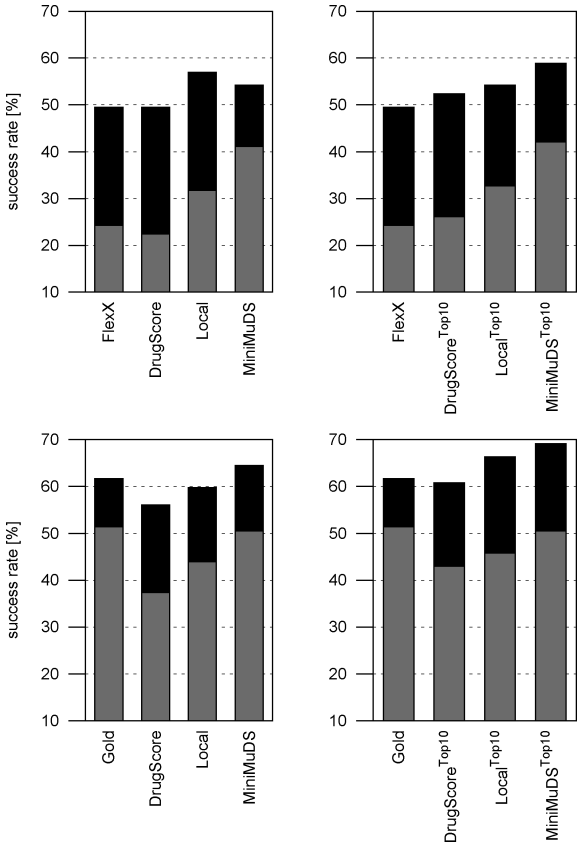


Figure 4. Success rates for the lead-like subset. Considered docking poses in the individual graphs correspond to Figure 3.

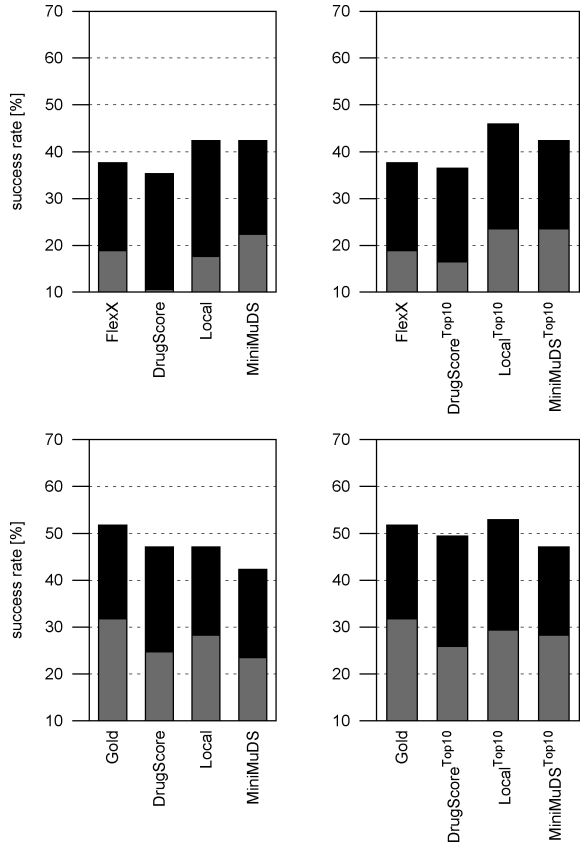


Figure 5. Success rates for the nonlead subset. Considered docking poses in the individual graphs correspond to Figure 3.

modified upon optimization compared to lead-like structures. Considering the obtained rankings, we observe an analog effect. Figure 4 shows the success rates only for the lead-like subset. The absolute level of success rates is for both docking programs about 5−10% higher compared to the complete test set. And although the general tendency across the different methods is very much the same, the improvement from FlexX to MiniMuDS^{Top10} raises from 7.3 to 9.3% at a 2 Å cutoff and from 12.0 to 17.8% at 1 Å. The improvement on the Gold results also raises from 2.1 to 7.5% at 2 Å, whereas almost no deterioration is observed any more at 1 Å. On the contrary, we found a decrease of success rates for Gold solutions of compounds in the nonlead subset applying the different rescoring and optimization methods, as shown in Figure 5.

The observed difference between the optimization of all or only the 10 top-ranked solutions is worth a closer inspection. For the FlexX poses, MiniMuDS^{Top10} achieved 2.6% better results at the 2 Å level compared to MiniMuDS. On the Gold poses the difference is 4.7%. In particular the GoldScore function appears to be already quite successful in identifying the correct pose out of the 50 widely spread decoys. This obvious advantage is deteriorated if all solutions are subjected to MiniMuDS. In consequence, some of the actually less optimally placed solutions are artificially optimized into geometries that even receive a superior DrugScoreX value. In contrast, if we rely on GoldScore to detect the most promising solutions among the top ranked ones, then we avoid these artifacts. Only the most relevant poses are considered for the optimization. This finally leads to the improved success rates.

SUMMARY AND CONCLUSIONS

The aim of this work was to develop a tool for the optimization of in-silico generated protein−ligand complexes according to the DrugScoreX potentials. The scoring function DrugScoreX is typically used to rescore bound ligand geometries that were generated by any docking program. Thus, these ligand poses are optimized according to the internally implemented scoring function used by the docking algorithm. Considering the steepness of the DrugScoreX potentials, this does not necessarily guarantee reliable and relevant scoring even if the docking poses are geometrically very close to a DrugScoreX optimum. In general, this holds true for every rescoring scenario, so a local optimization with respect to the subsequently used scoring scheme is strongly recommended in these cases.

In 2009, O'Boyle et al.[7] stated that a local optimization is always constrained to the energy well on the potential surface in which the original pose resides. So there may be an even deeper well nearby which will not be considered in the local optimization but would be equally valid. Our new tool MiniMuDS should account for this problem. On the other hand, we do not want to perform a global optimization since this would require exhaustive search for the overall best-scoring binding mode and, at the end, would result in a new docking algorithm. Instead, we want to keep close to the pose generated by the original docking engine and simply adapt this solution to the DrugScoreX function.

The developed method was to combine these two tasks by avoiding a strictly local optimization without extending to a full global search at the same time. To this end, we implemented an

optimization strategy that contains elements of a global optimization but is still restricted to a local part of the search space. Simply speaking, the applied algorithm can overcome small hills on the potential surface but only if the following valley is deeper than the current one. Thus, major energetic barriers between different conformations will not be passed.

In the validation of MiniMuDS we showed several important properties: (i) The optima of our applied energy model correspond impressively well to the experimentally determined native states of the evaluated complexes. This was shown by the optimization of the original crystal structures, which resulted in an average rmsd of about 0.5 Å, a value much smaller than the one observed in case of optimized docking solutions. This deviation has also to be seen in light of the positional accuracy estimated for experimental structure determination. The observed deviations virtually fall into the same range. (ii) The aim of keeping the given binding modes was achieved. The presented method allows for modifications up to 2 Å rmsd compared to the input geometry. Remarkably, not even 5% of the optimized docking poses fully exploited this available space. On average a modified geometry shows an rmsd of about 1 Å to the input structure. (iii) MiniMuDS in general improves a given docking solution by about 0.1 Å considering its rmsd to the native state. The best performance was observed for well-docked poses between 1 and 2 Å rmsd, which could be improved by up to 0.3 Å on average. (iv) We could show that an optimization that exceeds the restrictions of a strictly local search can strongly improve the resulting ranking. We received up to 4.7% better success rates at a 2 Å cutoff and up to 9.3% improvement at the 1 Å level when comparing MiniMuDS to a local optimization. (v) Most importantly, considering computational efforts, we could show that it is sufficient to only subject the 10 top-ranked docking solutions to optimization. This consistently yielded slightly better ranking results for all applied protocols compared to an optimization of all generated docking solutions. At 80% less computational effort, we received up to 4.7% higher success rates at 2 Å and 2.1% higher once at a 1 Å cutoff.

Especially the last aspect mentioned above confirms that it is advisible to focus only on those docking poses for optimization that were already well ranked by another scoring function. We thus consider only poses that score well with respect to two different scoring functions, thereby taking advantage from some kind of consensus effect. In light of these findings we have to strongly recommend the usage of at least a local optimization before applying DrugScore for rescoring purposes. Beyond that, we also suggest the application of a more sophisticated search strategy, such as the one implemented in MiniMuDS. In particular when dealing with small, lead-like structures, this method can substantially improve the rescoring results.

We are aware of the fact that the presented results do not take protein flexibility and cross-docking effects into account. Although this is a known drawback, it is still the standard way to evaluate docking algorithms and scoring functions, to validate their general applicability. Nevertheless, to address this problem there is ongoing research done in our group attempting to incorporate these effects into MiniMuDS, which will be reported elsewhere. In addition, we plan to make MiniMuDS available online at http://www.agklebe.de, beside the online version of the DrugScore function.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Detailed description of the implemented optimization algorithm and its parametrization. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Klebe@staff.uni-marburg.de.

Present Addresses

‡ Abbott GmbH & Co. KG, Knollstrasse 50, D-67061 Ludwigshafen, Germany

■ ACKNOWLEDGMENT

We gratefully acknowledge Thomas Mietzner (BASF Ludwigshafen, Germany) for his ideas and support, especially at the beginning of this project. The help of Gerd Neudert (University of Marburg, Germany) particularly for making latest versions of DrugScore available for this study is highly appreciated and gratefully acknowledged. We also thank the Cambridge Crystallographic Data Center (CCDC Cambridge, UK) for providing us with the Gold Suite, the BiosolveIT (Siegburg, Germany) for providing us with FlexX and the LeadIT software, and the Chemical Computing Group (Toronto, Canada) for making the MOE software tool available to us.

■ REFERENCES

(1) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *49*, 1455–1474.

(2) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.

(3) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.

(4) Cheng, T.; Li, X.; Li, Y.; Liu, Z.; Wang, R. Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.* **2009**, *49*, 1079–1093.

(5) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032–3047.

(6) Wang, R.; Lu, Y.; Wang, S. *J. Med. Chem.* **2003**, *46*, 2287–2303.

(7) O'Boyle, N. M.; Liebeschuetz, J. W.; Cole, J. C. Testing assumptions and hypotheses for rescoring success in protein-ligand docking. *J. Chem. Inf. Model.* **2009**, *49*, 1871–1878.

(8) Feher, M. Consensus scoring for protein-ligand interactions. *Drug Discovery Today* **2006**, *11*, 421–428.

(9) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100–5109.

(10) Velec, H. F. G.; Gohlke, H.; Klebe, G. DrugScore(CSD)—knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. *J. Med. Chem.* **2005**, *48*, 6296–6303.

(11) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict proteinligand interactions. *J. Mol. Biol.* **2000**, *295*, 337–356.

(12) 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.

(13) Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr.* **2002**, *B58*, 380–388.
(14) Cole, J. C.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Comparing proteinligand docking programs is difficult. *Proteins* **2005**, *60*, 325–332.
(15) Addis, B.; Locatelli, M.; Schoen, F. Local optima smoothing for global optimization. *Optim. Meth. Software* **2005**, *20*, 417–437.
(16) *Sybyl-X*; Tripos International: St. Louis, MO, 2009
(17) Clark, M.; Cramer, R. D.; Van Opdenbosch, N. Validation of the general purpose tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(18) Klebe, G.; Mietzner, T. A fast and efficient method to generate biologically relevant conformations. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 583–606.
(19) Neudert, G.; Klebe, G. fconv: format conversion, manipulation and feature computation of molecular data. *Bioinformatics* **2011**, *27*, 1021–1022.
(20) Tsai, C. J.; Kumar, S.; Ma, B.; Nussinov, R. Folding funnels, binding funnels, and proteinfunction. *Protein Sci.* **1999**, *8*, 1181–1190.
(21) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313.
(22) Box, M. J. A New Method of Constrained Optimization and a ComparisonWith Other Methods. *Comput. J.* **1965**, *8*, 42–52.
(23) Arvo, J. Fast random rotation matrices. *Graphics Gems III*; Academic Press: San Diego, CA, 1992; pp 117−120.
(24) 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.
(25) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
(26) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved proteinligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.
(27) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(28) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269–288.
(29) 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.
(30) Totrov, M.; Abagyan, R. Flexible protein ligand docking by global energy optimization in internal coordinates. *Proteins* **1997**, *29*, 215–22.
(31) Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput-Aided Mol. Des.* **2000**, *14*, 251–264.
(32) *MOE (The Molecular Operating Environment)*, version 2009.10; Chemical Computing Group Inc.: Montreal, Canada, 2009.