Empirical Scoring Functions for Advanced Protein−Ligand Docking with PLANTS

Oliver Korb,[†] Thomas Stützle,[‡] and Thomas E. Exner*,[†]

Theoretische Chemische Dynamik, Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany, and IRIDIA, CoDE, Université Libre de Bruxelles, Brussels, Belgium

Received August 25, 2008

In this paper we present two empirical scoring functions, $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$, designed for our docking algorithm PLANTS (Protein−Ligand ANT System), which is based on *ant colony optimization* (ACO). They are related, regarding their functional form, to parts of already published scoring functions and force fields. The parametrization procedure described here was able to identify several parameter settings showing an excellent performance for the task of pose prediction on two test sets comprising 298 complexes in total. Up to 87% of the complexes of the *Astex diverse* set and 77% of the CCDC/Astex *clean list*^{nc} (noncovalently bound complexes of the *clean list*) could be reproduced with root-mean-square deviations of less than 2 Å with respect to the experimentally determined structures. A comparison with the state-of-the-art docking tool GOLD clearly shows that this is, especially for the druglike *Astex diverse* set, an improvement in pose prediction performance. Additionally, optimized parameter settings for the search algorithm were identified, which can be used to balance pose prediction reliability and search speed.

INTRODUCTION

Computer methodologies have become a crucial part of drug discovery projects, from hit identification to lead optimization, and approaches such as ligand- or structure-based *virtual screening* techniques are widely used in many discovery efforts.[1] One key methodology - *docking* of small flexible molecules (ligands) to protein binding sites (receptors) - remains a highly active area of research.

A large variety of approaches to this problem has been proposed and frequently reviewed.[1−9] They can roughly be classified as *incremental*, as *stochastic optimization* methods for finding a global minimum structure with respect to a given objective function (often also called scoring function), or as *multiconformer docking* methods. Here, we will concentrate on stochastic optimization methods and especially *swarm-intelligence*-based approaches, which were developed in recent years. Examples are *particle swarm optimization* (PSO) in the programs SODOCK[10] as well as ClustMPSO[11] and *ant colony optimization* (ACO) in our own docking approach PLANTS.[12,13] In these methods, the quality of the scoring function used for evaluating the complex structure is of major importance for the success of the method. Recent studies comparing different docking tools on a large test set of complex structures report success rates of 30 to 80%.[6−9] This shows that a universal docking tool (algorithm and scoring function) that outperforms all others on every system is not available at the moment. Besides insufficient conformational sampling, the used scoring functions were identified as the main reason for these failures.[6−9] Therefore, many research efforts are devoted to the improvement of existing and the development of new scoring functions. Thus, besides a large number of stand-alone, general purpose scoring functions, almost every docking approach has its own specially adapted scoring function. These can be grouped into three partly overlapping categories: *force field-based*, *empirical*, or *knowledge-based*. Force field-based scoring functions including (even if there is sometimes a smooth transition to empirical scoring functions) GOLDscore,[14,15] DOCK,[16−19] and AutoDock,[20,21] are primarily derived from force fields like AMBER[22] and CHARMm[23] frequently used in molecular dynamics simulations. Empirical scoring functions, like PLP,[24−27] Chemscore,[28,29] Glide SP/XP,[30] and $PLANTS_{CHEMPLP}$ as well as $PLANTS_{PLP}$ introduced in this paper, are derived to reproduce data of experimentally determined complex structures. The third class of scoring functions, the knowledge-based type, relies on the idea that a sufficiently large data sample can serve to derive rules and general principles inherently stored in this knowledge base. In this sense, crystallographic data of a large number of protein−ligand complexes is used to determine pair distribution functions for different pairs of ligand and protein atom types. From these pair distributions, distance-dependent pairwise potentials of mean force are extracted. Existing implementations are, e.g., PMF,[31,32] DrugScore[33,34] and its derivatives like DrugScore^{CSD},[35] and the Astex statistical potential.[36]

The PLANTS core algorithm along with results for re- and cross-docking experiments considering flexible protein side-chains as well as *virtual screening* studies have already been published in refs 12 and 13. In the work presented here, we focus on the detailed description and parametrization of the available empirical scoring functions which constitute, together with the sampling algorithm, an overall high-performing protein−ligand docking approach.

MATERIALS AND METHODS

Problem Representation and Algorithm. Here, we concisely describe the main aspects on the problem repre-

* Corresponding author phone: +49 7531 882015; fax: +49 7531 883587; e-mail: thomas.exner@uni-konstanz.de.
[†] Universitat Konstanz.
[‡] Université Libre de Bruxelles.

10.1021/ci800298z CCC: $40.75
Published on Web 01/06/2009

sentation we use and the main steps of the PLANTS algorithm. A more in-depth description can be found in refs 12 and 13.

In PLANTS, the *protein−ligand docking problem* (PLDP) is treated as a continuous optimization problem, where each component of the vector of variables $\vec{x}$ corresponds to one of the degrees of freedom of the protein or the ligand. The goal of the optimization algorithm then is to find a global minimum of a given objective function f:

$$\min_{\vec{x}\in\mathbb{R}^n} f(\vec{x}) : \mathbb{R}^n \rightarrow \mathbb{R} \quad (1)$$

The problem dimension n is equal to the number of degrees of freedom considered for the particular PLDP instance that is to be tackled. We consider for the ligand three translational, three rotational, and r_l torsional degrees of freedom and for the protein r_p torsional degrees of freedom for flexible protein side-chains or rotatable hydrogen bond donor groups (*OH*- and NH_3^+-groups). Thus, the problem dimension is $n = 6 + r_l + r_p$. The objective function f to be minimized is given by the *scoring function*.

In previous research efforts, we have developed a hybrid *ant colony optimization* (ACO) algorithm, called PLANTS, for searching an optimum solution for the problem in eq 1. Ant Colony Optimization is a class of algorithmic techniques that is inspired by the foraging behavior of real ants, which are able to find the *shortest paths* between their nest and a food source.[37,38] Crucial for this ability is the real ants' *pheromone* trail laying and following behavior. In fact, ants deposit, while walking, *pheromones* to mark paths they follow. When they have the choice between several paths to follow, paths with a high pheromone concentration are chosen with a higher probability. ACO algorithms use *artificial pheromones* to imitate this behavior. These are represented as numerical values that are associated with each possible solution component. In ACO, ants are randomized procedures that generate a complete candidate solution using a constructive mechanism. For a more detailed explanation of the inspiring source for ACO we refer to ref 39.

ACO algorithms are usually applied to combinatorial optimization problems. Therefore, instead of developing a new ACO algorithm for continuous problems, PLANTS is based on one of the known best performing ACO algorithms for combinatorial optimization problems, *MAX--MIN* Ant System.[40] PLANTS works on a discretization of the continuous domain of each variable. For each of the three translational degrees of freedom, we use a discretization step size of 0.1 Å and for each rotational and torsional degree of freedom a discretization step size of 1°. Hence, the number of solution components for a translational degree of freedom depends on the dimension of the predefined binding site. This results in 120 to 400 discrete points for the complexes of the CCDC/Astex data set[41] considered in this paper, depending on the particular problem instance.

Artificial pheromones τ_{ij} refer to the desirability of assigning the value j to degree of freedom i. An ant then iteratively constructs a solution by assigning a value to each degree of freedom, whereas the probability of choosing value j for degree of freedom i is proportional to the already deposited amount of pheromone. Once each ant has constructed a solution and possibly improved it by a local search algorithm, feedback is given by increasing the pheromone values of the solution components (in our case here, the values assigned to a specific degree of freedom) that were found in the best solutions generated and by decreasing the pheromone levels of the others. [Note that ACO algorithms are population-based algorithms where several solutions are generated at the same time. Typical population sizes in ACO algorithms range from a few to several tens or hundreds of ants.] For details, we refer to ref 13.

PLANTS is a hybrid algorithm, since it combines the randomized solution construction by the ants with a local search algorithm. For the PLDP, we found the *Nelder-Mead simplex* (NMS) algorithm for nonlinear, continuous function optimization[43] to be particularly effective. In PLANTS, all ants improve their solution by applying the NMS algorithm. It is noteworthy that, while the ants work on a discretization of the problem, the NMS algorithm uses the continuous space for its search. The NMS algorithm is terminated either after a predefined number of scoring function evaluations or if the fractional range[44] from the highest to the lowest point in the simplex with respect to the function value is less than a tolerance value nms_{tol}. Hence, a final solution may be further improved by reapplying the NMS algorithm. We apply this refinement local search to the best solution among all the solutions obtained after the first local search application. The tolerance value used for the termination of this refinement local search is called *ref-*nms_{tol}. The NMS algorithm is described in detail in ref 44.

An algorithmic outline of the PLANTS algorithm is given in Algorithm 1. After some parameters and the pheromones are initialized, the main algorithm loop is repeated *max_iterations* times. First, each of the *number_of_ants* ants constructs a solution s_a (procedure ConstructSolution) and improves it by the NMS local search algorithm (procedure LocalSearch). The iteration-best solution, s^{ib}, which is determined by the procedure GetIterationBestSolution gets further refined by the procedure RefinementLocalSearch. After that, the new pheromone limits are calculated by procedure UpdatePheromoneLimits, and the pheromone values are adapted by procedure UpdatePheromones. In contrast to the PLANTS variants presented in refs 12 and 13, here an additional intensification strategy was added to improve the search performance. While in the former versions only the iteration-best solution, s^{ib}, was allowed to update the pheromone distribution, here additionally the best solution found since the last search diversification, s^{db}, is allowed to perform an update. If for a certain number of iterations, $iter_{db-update}$ (usually 5 or 10), no iteration-best solution s^{ib} is better than s^{db}, also s^{db} is used for the update. In this way, divergence from already encountered favorable solutions is likely to be avoided. Finally, a check whether search diversification should be applied is performed. If the conditions for a search diversification are met, procedure ApplySearchDiversification either applies a proportional pheromone smoothing or a random restart. Once the algorithm stops, the set M of all solutions found by any of the applications of the NMS local search algorithm is returned. This set of solutions M is then postprocessed to generate a set of diverse, high-quality conformations. To do so, first the solutions in M are sorted according to increasing values of the scoring function. Next, a specific number of solutions, typically 10, are extracted from that ordered list such that the minimum root-mean-square deviation (rmsd) between any

Algorithm 1 PLANTS

```
InitializeParametersAndPheromones()
M ← ∅
for iter = 1 to max_iterations do
  for a = 1 to number_of_ants do
    s_a ← ConstructSolution(τ)
    s*_a ← LocalSearch(s_a, nms_tol)
    M ← M ∪ s*_a
  end for
  s^ib ← GetIterationBestSolution()
  s^ib ← RefinementLocalSearch(s^ib, ref-nms_tol)
  M ← M ∪ s^ib
  UpdatePheromoneLimits()
  UpdatePheromones()
  if (diversification criteria met) then
    ApplySearchDiversification()
  end if
end for
s^gb ← GetGlobalBestSolution(M)
return M, s^gb
```

pair of solutions is larger than 2 Å. Finally, the global best solution found, s^{gb}, as well as the set of diverse conformations is returned. For further details of the PLANTS algorithm we refer to ref 13.

Scoring Functions. The scoring functions $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$ are based, at least regarding the functional form, on parts of already published scoring functions and force fields. The *piecewise linear potential* (PLP)[24−27] scoring function is used in both cases to model steric complementarity of the protein and the ligand. In the $PLANTS_{CHEMPLP}$ case, terms of GOLD's Chemscore implementation[45] are used to introduce angle-dependent terms for hydrogen bonding and metal binding. The torsional potential from the Tripos force field[46] together with a heavy-atom clash term is employed to account for intraligand interactions.

$$f_{PLANTS} = f_{plp} + f_{clash} + f_{tors} + c_{site} \quad (2)$$

$$f_{PLANTS_{CHEMPLP}} = f_{plp} + f_{hb} + f_{hb-ch} + f_{hb-CHO} + f_{met} + f_{met-coord} + f_{met-ch} + f_{met-coord-ch} + f_{clash} + f_{tors} + c_{site} \quad (3)$$

Table 1. Atom Typing Rules

class	rule
	$PLANTS_{PLP}$ Types
donor	nitrogen with at least one attached hydrogen and no accessible lone pair
acceptor	oxygen or nitrogen without attached hydrogens with one or two heavy-atom neighbors, except -O- in ester groups
don./acc.	oxygen and nitrogen that can act as both (e.g., hydroxyl, water)
nonpolar	non-hydrogen atoms that are not classified according to the other rules
metal	metal-ion (protein only)
	Additional $PLANTS_{CHEMPLP}$ Types
H	any hydrogen attached to oxygen or nitrogen
charged H	donor hydrogen's neighbor has positive formal charge
CH-donor	hydrogen attached to a carbon atom neighbored to an aromatic ring nitrogen acceptor
charged acc.	acceptor has negative formal charge

The meaning and functional form of each individual term will be discussed in the following sections.

Atom Typing. All protein and ligand atoms in PLANTS are typed according to the rules given in Table 1. These atom types are used for the appropriate selection of protein−ligand interaction potentials. For special cases, like nitro-groups, the formal-charge assignments are overridden and the acceptors are treated as neutral.

Table 2. Piecewise Linear Potential (PLP) Interactions

	protein atom type				
ligand atom type	donor	acceptor	don./acc.	nonpolar	metal
donor	repulsive	H-bond	H-bond	buried	repulsive
acceptor	H-bond	repulsive	H-bond	buried	metal
don./acc.	H-bond	H-bond	H-bond	buried	metal
nonpolar	buried	buried	buried	nonpolar	buried

Table 3. Parameters Used for the PLP Interactions

	distance [Å]			score [a.u.]		
interaction	A	B	C	D	E	F
H-bond	2.3	2.6	3.1	3.4	w_{plp-hb}	20.0
metal	1.4	2.2	2.6	2.8	$w_{plp-met}$	20.0
buried	3.4	3.6	4.5	5.5	$w_{plp-bur}$	20.0
nonpolar	3.4	3.6	4.5	5.5	$w_{plp-nonp}$	20.0

	distance [Å]		score [a.u.]	
interaction	A	B	C	D
repulsive	3.2	5.0	$w_{plp-rep}$·0.1	$w_{plp-rep}$·20.0

Piecewise Linear Potential. In both scoring functions, $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$, the *piecewise linear potential* (PLP)[24−27] scoring function (f_{plp}) is used to model the steric complementarity of the ligand molecule and the protein. Two different piecewise functions are defined, one for repulsive/attractive, *plp*, and one for purely repulsive interactions, *rep*:

$$plp(r,A,B,C,D,E,F) = \begin{cases} F\cdot(A-r)/A & \text{if } r<A \\ E\cdot(r-A)/(B-A) & \text{if } A \le r<B \\ E & \text{if } B \le r<\mathrm{C} \\ E(D-r)/(D-C) & \text{if } C \le r \le D \\ 0 & \text{if } r>D \end{cases} \quad (4)$$

$$rep(r,A,B,C,D) = \begin{cases} r\cdot(C-D)/A + D & \text{if } r<A \\ -C\cdot(r-A)/(B-A) + C & \text{if A} \le r \le B \\ 0 & \text{if } r>B \end{cases} \quad (5)$$

$$f_{plp} = \sum_{p\in P_{prot-lig-plp}} plp(p_r,p_A,p_B,p_C,p_D,p_E,p_F) + \sum_{p\in P_{prot-lig-rep}} rep(p_r,p_A,p_B,p_C,p_D) \quad (6)$$

In these equations, $P_{prot-lig-plp}$ and $P_{prot-lig-rep}$ are the sets of protein−ligand atom pairs used for the evaluation of function *plp* and *rep*, respectively. For each pair, the actual distance between a ligand and a protein atom is given by p_r. The parameters p_A to p_F are dependent on the interaction potential chosen and set according to Tables 2 and 3. Both potentials are illustrated in Figure 1. Table 2 shows all types of protein−ligand interactions considered. The protein and ligand atom types correspond to the ones already described in Table 1. For all interactions except donor-donor, acceptor-acceptor, and donor-metal, where the repulsive *rep* potential is used, the *plp* potential is evaluated. This is the case for hydrogen bonding interactions, called *H-bond*, for metal-chelating, called *metal*, and for nonpolar-nonpolar interactions, called *nonpolar*. If polar atoms (donor, acceptor, donor/acceptor, metals) are buried by nonpolar ones, a different potential than in the nonpolar-nonpolar case is used, called

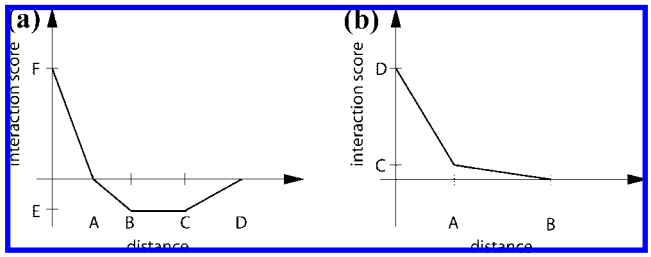

Figure 1. (a) Illustration of the *piecewise linear potential* (PLP) that is used for both scoring functions. Parameters *A* to *F* define the functional form. (b) The repulsive term used for donor-donor, acceptor-acceptor, and donor-metal interactions. The potential is defined by the parameters *A* to *D*.

Table 4. Cutoff Distances Used for the Ligand Clash Potential[a]

interaction	$r_{clash1-4}$ [Å]	$r_{clash>1-4}$ [Å]
class1-class1	2.50	2.50
class2-class3	2.50	2.75
other	2.75	3.00

[a] For a description of the interaction classes see the text.

buried. These interactions turned out to be crucial for the derivation of high-quality parametrizations with respect to pose prediction accuracy. The distance parameters used for all interaction types are given in Table 3 along with the weighting parameters considered for the parameter optimization process.

Ligand Clash Potential. To avoid internal ligand clashes, an empirical heavy-atom potential (f_{clash}) is evaluated for each ligand pose. The functional form of this potential is given by

$$f_{dist}(r, w, r_{clash}) = \begin{cases} 0 & \text{if } r > r_{clash} \\ w \cdot \dfrac{(r_{clash}^2 - r^2)}{r_{clash}^2} & otherwise \end{cases} \quad (7)$$

$$f_{\text{clash}} = w_{\text{clash}} \cdot \sum_{c \in C_{\text{clash}}} f_{\text{dist}}(||\overrightarrow{c_{p_a} c_{p_b}}||, c_w, c_{r_{\text{clash}}}) \quad (8)$$

In these equations, C_{clash} is the set of ligand atom pairs that are at least 3 bonds away from each other and reside in different rigid fragments separated by rotatable bonds. Each ligand atom is classified as at least one of 3 clash types. Clash type *class1* consists of oxygen and nitrogen atoms, which are allowed to have attached hydrogen atoms as well as fluorine, chlorine, and bromine atoms. This class accounts for intramolecular hydrogen bonds. Clash type *class2* consists of the same atoms as *class1* except nitrogens and oxygens, which have hydrogens attached and thus act as donors. Finally, *class3* contains all heavy-atoms that are not in *class1* or *class2*. The potential f_{dist} has no contribution if the distance between the atoms is bigger than the cutoff-distance r_{clash} as predefined for each interacting pair according to Table 4. In this table, $r_{clash1-4}$ is the cutoff-distance for two atoms that are exactly separated by 3 bonds, while $r_{clash>1-4}$ is the cutoff-distance for two atoms separated by more than 3 bonds. All clash weights *w* are treated equally and set to a value of 50. The potential is visualized in Figure 2 for all 3 cutoff-radii used in Table 4. One of the biggest advantages of this potential is that it is fast to evaluate as no hydrogen atoms

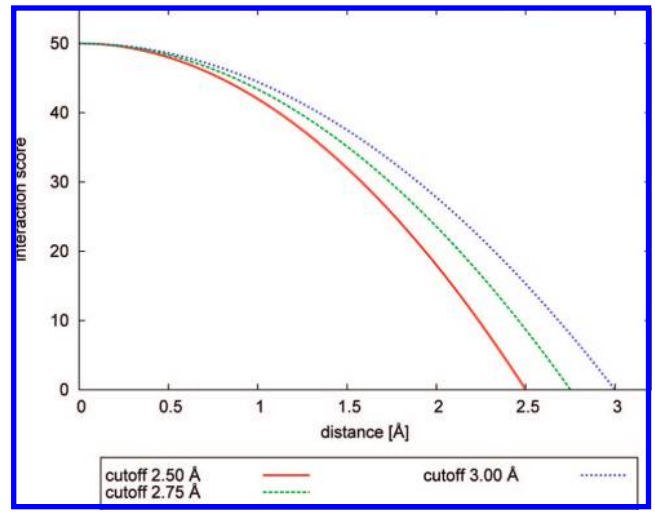

Figure 2. Illustration of the ligand heavy-atom clash potential. The potential is shown for the 3 different cutoff-radii 2.5 Å, 2.75 Å, and 3.0 Å.

are taken into account and only squared distances are needed for the calculation. For all experiments the weighting parameter w_{clash} has been set to 1.0.

Ligand Torsional Potential. The torsional potential (f_{tors}) as implemented in the Tripos force field,[46] is calculated for all rotatable bonds available in the ligand molecule except for terminal hydrogen bond donor groups. The torsional potential is calculated as

$$f_{tors} = w_{tors} \cdot \sum_{p \in P_{tors}} \frac{1}{2} p_V (1 + p_S \cos(|p_n| p_\omega)), \quad (9)$$

where p_V is the torsional barrier, p_S defines whether the minimum is in the staggered ($p_S = 1$) or the eclipsed ($p_S = -1$) conformation, p_n is the periodicity, and p_ω is the actual torsion angle.

Hydrogen Bonding. The distance- and angle-dependent hydrogen bonding terms (f_{hb}, f_{hb-ch}, and f_{hb-CHO}) used in the $\text{PLANTS}_{\text{CHEMPLP}}$ scoring function are based on the Chemscore potentials[28,29,45] with slightly modified parameter settings. Three different types of hydrogen bonds are taken into account. The term f_{hb} with weight w_{hb} is evaluated for all donor−acceptor pairs that are part of the set P_{hb}, where both atoms are uncharged or exactly one of both atoms is charged. Hydrogen bonding pairs that are part of the set P_{hb-ch} consist of a charged donor and a charged acceptor. For these pairs the term f_{hb-ch} is calculated in which the standard hydrogen bond weight w_{hb} is further scaled by a factor w_{hb-ch}. This parameter is usually set to a value greater than 1. If CH-donors as defined in Table 1 are available, the term f_{hb-CHO} is evaluated for all hydrogen bonding pairs containing an oxygen-acceptor and weighted with w_{hb-CHO}. These pairs constitute the set P_{hb-CHO}. All of these potentials rely on a piecewise linear function f as defined in Chemscore[28,29,45] which is given by

$$f(x, x_1, x_2) = \begin{cases} 1 & \text{if } x \le x_1 \\ (x_2 - x)/(x_2 - x_1) & \text{if } x_1 < x \le x_2 \\ 0 & \text{if } x > x_2 \end{cases} \quad (10)$$

All hydrogen bonding equations contributing to the total $\text{PLANTS}_{\text{CHEMPLP}}$ score along with the geometric parameters used are given by

$$f_{hb} = w_{hb} \cdot \sum_{p \in P_{hb}} f(|p_r - 1.85|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 30, 80) \cdot \prod_{q \in P_{acc-nb}} f(|q_\beta - 180|, 80, 100) \quad (11)$$

$$f_{hb-ch} = w_{hb} \cdot w_{hb-ch} \cdot \sum_{p \in P_{hb-ch}} f(|p_r - 1.85|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 30, 80) \cdot \prod_{q \in P_{acc-nb}} f(|q_\beta - 180|, 80, 100) \quad (12)$$

$$f_{hb-CHO} = w_{hb-CHO} \cdot \sum_{p \in P_{hb-CHO}} f(|p_r - 2.35|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 50, 100) \cdot \prod_{q \in P_{acc-nb}} f(|q_\beta - 180|, 80, 100) \quad (13)$$

In these equations, p_r is the actual distance of the hydrogen donor H to the acceptor atom A, and p_α is the angle $\angle DHA$, where D is the donor heavy-atom. On the acceptor side, p_{acc-nb} is the set of atoms directly neighbored to the acceptor atom A. For each of these atoms $q \in p_{acc-nb}$ the angle q_β, $\angle qAH$, is computed, where H is the hydrogen donor atom. Each donor hydrogen H is only allowed to form one interaction, and thus only the best interacting acceptor contributes to the total score.

Metal Interactions. In PLANTS, two types of metal interactions are considered. The following distance- and angle-dependent potential is used for magnesium and calcium atoms, while ideal interaction polyhedra are taken into account for other common metal atoms. In the first case, besides the PLP contributions modeling the steric part, the metal interaction score is computed for all ligand acceptor/metal-ion pairs p that are part of the set $P_{acc-met}$ using the following distance- and angle-dependent potentials:

$$f_{met} = w_{met} \cdot \sum_{p \in P_{acc-met}} f(p_r, 2.6, 3.0) \cdot \prod_{q \in P_{acc-nb}} f(|q_\beta - 180|, 80, 90) \quad (14)$$

In this equation, w_{met} is the weight for a metal interaction, f is the same piecewise linear function as used in the hydrogen bonding section, and p_r is the actual distance between the ligand acceptor and the metal-ion. The set p_{acc-nb} contains all atoms neighbored to the ligand acceptor atom A. For each of these neighboring atoms $q \in p_{acc-nb}$, q_β is the angle $\angle qAM$, where M is the metal-ion.

For the second case, coordination polyhedra are fitted with respect to protein atoms coordinating the metal-ion using the least-squares fitting algorithm of Kabsch.[47,48] For all free vertices of the fitted polyhedron, which can be either a tetrahedron or an octahedron depending on which of both exhibits smaller deviations according to the least-squares fitting procedure, favorable ligand acceptor positions are defined 2.2 Å away from the metal-ion. The result of the fitting of an octahedral interaction geometry to an iron atom is shown in Figure 3 for PDB code 1eoc. In contrast to eq 14, the distance p_{rf} to the fitting point instead of the metal-ion itself is calculated in the potential $f_{\text{met-coord}}$:

$$f_{met-coord} = w_{met} \cdot \sum_{p \in P_{acc-met-coord}} f(p_{rf}, 0.6, 0.8) \cdot \prod_{q \in P_{acc-nb}} f(|q_\beta - 180|, 80, 90) \quad (15)$$

The same angular scaling taking into account the acceptor atom's neighbors is applied as in the previous case. However, it must be noted that for each fitting point only the best interacting acceptor atom contributes to the total score.

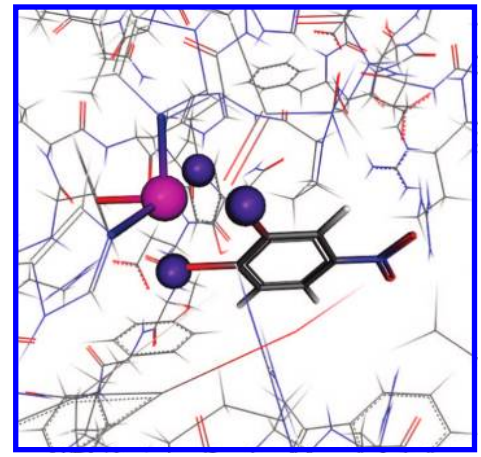

Figure 3. Octahedral interaction geometry matched to the iron atom (shown in magenta) in PDB code 1eoc. The blue spheres define favorable ligand acceptor positions, which are also occupied in two cases by the cocrystallized ligand.

If a charged acceptor atom is involved in a metal interaction, the above-described potentials are additionally scaled with a parameter w_{met-ch}, which is usually, like in the hydrogen bonding case, set to a value greater than 1:

$$f_{met-ch} = w_{met-ch} \cdot f_{met} \quad (16)$$

$$f_{met-coord-ch} = w_{met-ch} \cdot f_{met-coord} \quad (17)$$

Additional Contributions. As outside the binding site definition, which is defined by a sphere in PLANTS, no interaction potentials are calculated, a mechanism is needed to guide the search algorithm toward the interesting regions located at the predefined binding site. A quadratic potential is calculated for the reference point of the ligand (origin of the ligand's coordinate system) if this point lies outside the sphere, that is, the distance r of the reference point to the center of the sphere is greater than the binding site radius r_{site}:

$$c_{site}(r) = \begin{cases} 0 & \text{if } r \le r_{site} \\ r^2 - r_{site}^2 & \text{if } r > r_{site} \end{cases} \quad (18)$$

Additionally, if a ligand heavy-atom lies outside the binding site, a constant value, set to 50, is added to the total score. The latter term was mainly introduced for *virtual screening* purposes to force all ligands to interact with the predefined protein surface and thereby avoiding parts of the ligands to be placed outside the binding site. One may argue that this term reduces the search space. However, experimental results show that setting the penalty value to 0 has nearly no effect on the optimization performance (the success rate decreases by less than 1% on average for the data sets discussed in this paper, which is statistically not significant).

For both scoring functions, a value of −20.0 is added for shifting the scoring values. As the ACO algorithm only deposits pheromone if the scoring function value is lower than zero, in this way, especially at the beginning of the search process, at least some information can be gained even if slightly repulsive protein−ligand conformations are sampled.

Scoring Function Evaluation. The scoring function evaluation is the most time-consuming step of the whole

Table 5. Parameter Values Considered for the Parameter Optimization Process

	parameter values	
parameter	$PLANTS_{CHEMPLP}$	$PLANTS_{PLP}$
w_{plp-hb}	− 0.5, − 1	− 1, − 2, − 3, − 4, − 5, − 6
$w_{plp-met}$	w_{plp-hb}	− 1, − 2, − 3, − 4, − 5, − 6, − 7, − 8, − 9, − 10, − 11, − 12
$w_{plp-bur}$	− 0.05, − 0.1, − 0.15	− 0.05, − 0.1, − 0.15
$w_{plp-nonp}$	− 0.4	− 0.4
$w_{plp-rep}$	0.25, 0.5, 1	0.25, 0.5, 1
w_{tors}	1, 2, 3	1, 2, 3
w_{hb}	− 2, − 4, − 5	-
w_{hb-ch}	1, 1.25, 1.5, 1.75, 2	-
w_{met}	− 3, − 6, − 9	-
w_{met-ch}	1, 1.25, 1.5, 1.75, 2	-
w_{hb-CHO}	− 1, − 2, − 3, − 4, − 5	-

PLANTS approach. Therefore, great care has been taken to implement this step as efficiently as possible. Most of the steps, except the evaluation of the intraligand clash potential, have been reduced to linear time complexity with respect to the number of atoms available in the ligand. For example lookup-tables are used for the torsional potentials, the PLP-potentials are stored in four precomputed grids at a resolution of 0.3 Å using trilinear interpolation at evaluation time, and spatial regular grid data structures are used for fast distance-queries in the case of the angle-dependent hydrogen and metal bonding terms. Clearly, the scoring function evaluation is not exact in that case, especially when the trilinear interpolation is used, but the accuracy is sufficient for docking purposes and can be justified by enormous speedups reached. These techniques were employed in all docking calculations carried out in this paper, that is, for the scoring function parametrization process and the pose prediction experiments.

Scoring Function Parameterization. The main focus of the optimization process was the derivation of the scoring function's weighting parameters, while the geometric parameters were kept fixed. Table 5 shows all parameters considered during the optimization process for $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$. For the sake of reducing the number of parameters to be optimized, parameter $w_{plp-nonp}$ was fixed to −0.4 for both scoring functions, and the parameter $w_{plp-met}$ was set to the same value as w_{plp-hp} for $PLANTS_{CHEMPLP}$. The latter setting is convenient as for $PLANTS_{CHEMPLP}$ the major contributions for hydrogen bonding and metal-interactions can be attributed to the parameters w_{hb} and w_{met}.

A heuristic tree search strategy was applied to identify reasonable parameter settings reducing the immense computer resources needed for a full optimization. For both scoring functions, a training set of 31 protein−ligand complexes from the CCDC/Astex data set[41] was used to remove unfavorable parameter settings from all possible combinations of parameters in Table 5. The 31 complexes were selected to cover a large variety of different protein targets, while using high- to medium-quality crystal structures. Out of these 31 structures 20 have a resolution between 1.6 Å and 2.0 Å, whereas the remaining 11 structures have a resolution between 2.1 Å and 2.8 Å. For the final decision on the best set of values, all complexes of the high-resolution *Astex diverse* set[42] were tested on these preselected combinations as described below.

First, the complexes were divided up into independent subsets accounting for different types of protein−ligand interactions. E.g. the first subset for $PLANTS_{CHEMPLP}$ only contains complexes without charged and weak hydrogen bonds as well as complexes without metal ions. This set was used, as described below in detail, to parametrize all PLP weights w_{plp-X}, the torsional weight w_{tors}, and the hydrogen bonding weight w_{hb}. In the next layer of the search tree a new subset accounting for additional interaction types was added to search for the best parameters describing these new interactions from the possible values given in Table 5. For each scoring function parameter setting considered, the complexes were docked 10 times using the following settings for the search algorithm. Twenty ants, $\rho = 0.15$ and $p_{best} = 0.5$, were used, which corresponds to the standard search setting as given in ref 13, but the number of ACO-iterations was doubled ($\sigma = 2$) and the simplex tolerance for the refinement local search was set to $nms_{tol} = 0.0001$ to minimize the chance of sampling inaccuracies.

Prior to docking the ligand structures were randomized with respect to their translational, rotational, and torsional degrees of freedom to eliminate any potential search bias toward the crystallographic conformation. After docking, the rmsd of the resulting overall best-scoring ligand pose was calculated. If the rmsd was lower than 2 Å with respect to the experimentally determined ligand structure, the docking was counted as a success. Then all settings of the current step were sorted in the first instance according to the number of successes. If two or more parameter settings exhibited the same number of successes, the rmsd values to the crystal structures were also taken into account. For doing so, the rmsd of the best-ranked structure (over ten runs) is calculated for each complex. These are then averaged over all complexes used up to this stage, and this average rmsd is used for sorting in second instance. The best settings according to the sorted list were then taken to the subsequent step where additional parameters were optimized thereby keeping the parameter values from the last step fixed. Clearly, this search strategy is suboptimal and will not necessarily identify the global optimal parameter setting with respect to the given set of parameter values. However, the parametrization process using an explicit docking strategy is extremely time-consuming, and thus heuristic methods are the first choice for keeping the computational demands tractable while still identifying parameter models of reasonable quality. The resulting models for both scoring functions were then tested with respect to pose prediction accuracy on the independent *Astex diverse* set[42] containing 85 protein−ligand complexes, which has no intersection with the training set used. The same search settings as described above for the parameter optimization process were applied, but 25 instead of 10 experiments were performed and the termination criterion for the refinement simplex was set to $nms_{tol} = 0.0001$.

The spherical binding site was defined starting from the ligand's crystallographic conformation considering all protein atoms that are up to 6 Å away from any ligand heavy-atom resulting in binding site diameters ranging from 17.5 Å for PDB code 1w1p to 33.7 Å for PDB code 1t46.

Parameterization and Results for $PLANTS_{CHEMPLP}$. The 31 complex structures have been divided into four independent training sets. Set 1 contained nine PDB codes (1apu, 1c1e, 1ai5, 1fen, 1tyl, 1hsb, 1rt2, 1abf, and 3erd) and was used to parametrize all PLP weights w_{plp-X}, the torsional weight w_{tors}, and the hydrogen bonding weight w_{hb} resulting

Table 6. Parameter Models Derived from the Optimization Procedure for $PLANTS_{CHEMPLP}$

	$PLANTS_{CHEMPLP}$ models						
	M_1	M_2	M_3	M_4	M_5	M_6	M_7
Parameter							
w_{plp-hb}	−0.50	−0.50	−1.00	−0.50	−0.50	−1.00	−1.00
$w_{plp-met}$	−0.50	−0.50	−1.00	−0.50	−0.50	−1.00	−1.00
$w_{plp-bur}$	−0.15	−0.15	−0.10	−0.10	−0.10	−0.10	−0.15
$w_{plp-nonp}$	−0.40	−0.40	−0.40	−0.40	−0.40	−0.40	−0.40
$w_{plp-rep}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00
w_{tors}	2.00	3.00	2.00	2.00	2.00	3.00	2.00
w_{hb}	−4.00	−4.00	−3.00	−4.00	−3.00	−3.00	−4.00
w_{hb-ch}	2.00	1.75	2.00	1.75	2.00	2.00	1.75
w_{met}	−6.00	−6.00	−6.00	−6.00	−6.00	−6.00	−6.00
w_{met-ch}	2.00	2.00	2.00	2.00	2.00	2.00	2.00
w_{hb-CHO}	−3.00	−1.00	−3.00	−2.00	−3.00	−3.00	−3.00
Training Set							
correct (of 31)	30	30	30	30	30	30	30
av rmsd [Å]	0.91	0.92	0.87	0.90	0.91	0.91	0.90
ASTEX Diverse							
correct (of 85)	73	73	73	74	74	73	73
av rmsd [Å]	1.27	1.29	1.27	1.27	1.24	1.29	1.32

in 216 possible parameter settings. 151 of the 216 settings were able to reproduce all nine ligands inside 2 Å with average rmsd values between 0.71 Å and 0.88 Å.

These 151 settings were taken to the second optimization stage where parameter w_{hb-ch} was optimized on set 2 containing seven PDB codes (1a4g, 1tnl, 2ifb, 1epb, 1rob, 1aqw, and 1kel). Testing all five values for w_{hb-ch}, hence, resulted in $5 \times 151 = 755$ models for this stage. None of these models was able to identify the binding modes of all seven ligands correctly, and therefore only the models that reproduced six correct binding modes were further considered. These models were clustered according to the parameters of the first stage, and for each cluster only the w_{hb-ch}-setting with the lowest average rmsd was kept taking into account the rmsd values of all 16 ligands docked up to this stage.

The resulting 51 models were taken to stage 3, where the weight for w_{hb-CHO} accounting for CH-O interactions was determined. In this stage, six PDB codes (1dhf, 1dr1, 2ak3, 1ydt, 6rnt, and 1uvs) were used for the parametrization. Possible values for w_{hb-CHO} have been restricted not to be lower than w_{hb}, that is, a CH-O bond should not have a higher impact than a standard noncharged hydrogen bond. 74 out of 177 models were able to reproduce all six ligand binding modes correctly. Taking into account the obtained rmsd values from the prior stages, 37 models were taken to the last stage for the optimization of the metal bonding parameters, w_{met} and w_{met-ch}.

In this last stage, nine PDB codes (1phd, 2cpp, 5cpp, 1cbx, 1eoc, 1frp, 1lcp, 4aah, and 1mmq) were used to determine favorable parameter settings. For each of the 37 models taken from the last stage, 15 settings resulting from combinations of w_{met} and w_{met-ch} values were tested. Seventeen of these 555 models were able to reproduce the ligand binding modes for all nine metal-containing complexes correctly. These were further reduced to 12 settings by keeping only the best metal parameter for the same settings of the other parameters with respect to the minimum average rmsd.

All 12 models resulting from the optimization procedure were tested with respect to their pose prediction performance on the *Astex diverse* set[42] (with 85 complexes). The performance of these models is quite striking with two models able to dock correctly 74 complexes and five models able to dock correctly 73 complexes. The top-ranking ligand poses have been reproduced with average rmsd values between 1.24 Å and 1.32 Å. Also the five other models were able to identify 71 or 72 complexes correctly with average rmsd values between 1.33 Å and 1.37 Å. However, 12 models were too many to handle for all subsequent tests, and, therefore, only the seven best models listed in Table 6 were kept for the further experiments.

Table 7. Parameter Models Derived from the Optimization Procedure for $PLANTS_{PLP}$

	$PLANTS_{PLP}$ models				
	M_1	M_2	M_3	M_4	M_5
Parameter					
w_{plp-hb}	−4.00	−3.00	−4.00	−4.00	−2.00
$w_{plp-met}$	−7.00	−7.00	−6.00	−8.00	−4.00
$w_{plp-bur}$	−0.05	−0.05	−0.05	−0.05	−0.05
$w_{plp-nonp}$	−0.40	−0.40	−0.40	−0.40	−0.40
$w_{plp-rep}$	0.50	1.00	0.50	1.00	0.50
w_{tors}	3.00	2.00	1.00	3.00	1.00
Training Set					
correct (of 31)	25	25	25	25	25
av rmsd [Å]	1.36	1.55	1.45	1.43	1.58
ASTEX Diverse					
correct (of 85)	67	67	68	67	70
av rmsd [Å]	1.56	1.57	1.67	1.57	1.64

Parameterization and Results for $PLANTS_{PLP}$. For the parametrization of scoring function $PLANTS_{PLP}$, the training sets 1, 2, and 3 used for the parametrization of $PLANTS_{CHEMPLP}$ have been merged, resulting in 22 protein−ligand complexes, because $PLANTS_{PLP}$ does not account for charged and weak CH-O hydrogen bonding in a special way. Therefore, in the first step the PLP parameters w_{plp-X}, except $w_{plp-met}$, and the torsional weight w_{tors} were optimized on the merged set consisting of 22 protein−ligand complexes. 35 of the 162 models were able to reproduce correctly 16 to 18 training set complexes. These 35 models were then transferred to the second optimization stage in which the same nine metal-containing complexes were used as for the parametrization

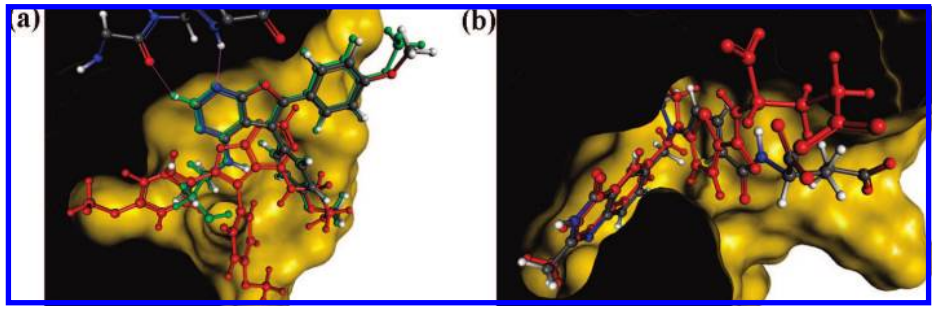

Figure 4. Comparison of the experimentally determined complex structures and the ones predicted by PLANTS: The protein is visualized as a yellow molecular surface, and the experimentally determined ligand conformation is shown in a ball-and-stick representation color-coded according to atom types. To allow for a better insight into the active site, the front part of the surface has been removed. (a) Structure of protein kinase Chk1 with a furanopyrimidine inhibitor (PDB code 2br1): The predicted conformation of $PLANTS_{CHEMPLP}$ model M_3, forming a hydrogen bonding network (magenta) with the hinge region of the kinase, and the wrong prediction of $PLANTS_{PLP}$ model M_1 are shown in green and red, respectively. (b) Structure of human thymidylate synthase with Tomudex (PDB code 1hvy): The ligand conformation predicted by $PLANTS_{CHEMPLP}$ model M_1 is shown in red.

of $PLANTS_{CHEMPLP}$. The values tested for parameter $w_{plp\text{-}met}$ in this stage were restricted to be lower or equal to $w_{plp\text{-}hb}$, that is, an acceptor-metal interaction should be at least as favorable as a hydrogen bond. 32 of the resulting models reproduced correctly 25 to 26 of 31 protein−ligand complexes. Choosing only the best metal-setting of all models with the same $w_{plp\text{-}X}$ (except $w_{plp\text{-}met}$) and w_{tors} settings based on the average rmsd for all 31 complexes reduced the number of candidate models to 13. The average rmsd of these 13 models ranged from 1.35 Å to 1.77 Å. In order to further narrow down the number of scoring function models, all 13 settings were tested on the 85 complexes of the *Astex diverse* set. All models were able to reproduce correctly between 63 and 70 complexes with average rmsd values between 1.56 Å to 1.83 Å, which is quite remarkable considering the simple form and the small number of parameters of the $PLANTS_{PLP}$ scoring function. The best five models, which were able to reproduce correctly between 67 and 70 complexes, are listed in Table 6 and were chosen for subsequent experiments.

Docking Results of Selected Complexes. Not only to demonstrate the predictive power but also to show the deficiencies of the derived scoring function models, several examples from the training set and the *Astex diverse* set will be discussed in the following.

The first example is protein kinase Chk1 with a furanopyrimidine inhibitor (PDB code 2br1). In this complex as well as in many other kinases, a weak CH-O hydrogen bond in the hinge region can be observed (see Figure 4). Scoring function $PLANTS_{CHEMPLP}$ is able to reproduce the correct binding mode, independent of the model used, due to the explicit treatment of this type of interaction. In contrast, none of the $PLANTS_{PLP}$ models is able to make a correct prediction. Besides the missing CH-O interaction term, the radial, angle-independent hydrogen bonding potential could also be a reason for the failure. However, other kinase complexes (e.g., PDB code 1ydt) can be reproduced correctly by several $PLANTS_{PLP}$ models.

For the determination of the parameters describing the metal coordination, especially the cytochrome P450 complexes 2cpp and 5cpp are of interest. Here the acceptor oxygen of camphor is forming a hydrogen bond with the protein, but it does not coordinate to the iron of the heme group. Thus, this complex can be used to limit the strength of an acceptor-metal interaction. $PLANTS_{CHEMPLP}$ model M_4

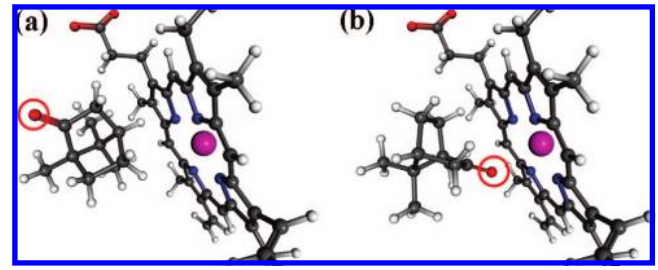

Figure 5. Predicted complex structures of the cytochrome P450 complex (PDB code 2cpp): (a) structure of $PLANTS_{CHEMPLP}$ model M_4, reproducing the correct binding mode ($rmsd$ = 0.71 Å), and (b) structure of a model with the same parameters as model M_4 but with $w_{met} = -9.00$ ($rmsd$ = 3.19 Å). For the sake of clarity, only the camphor molecule and the heme group are shown. In both illustrations red circles mark the acceptor positions of the camphor molecule.

(as well as all other models in Table 6) is able to reproduce correctly all seven complexes of the metal training set. If the weight for the metal interactions is decreased from $w_{met} = -6.00$ to $w_{met} = -9.00$, in all seven training set complexes a ligand acceptor is coordinated to the metal-ion resulting in wrong predictions for PDB codes 2cpp and 5cpp (see Figure 5).

A common problem when using the heavy-atom rmsd between the predicted and the experimentally determined structure as the success criterion will be exemplarily highlighted for human thymidylate synthase in complex with Tomudex (PDB code 1hvy). As it can be seen in Figure 4, the oxoquinazolin group as the main binding motif is correctly predicted. However, the rest of the predicted structure differs from the experimentally determined structure due to an insufficient specificity of the scoring function. First of all, it can be observed that the thiophene ring is rotated by 180. According to our simplified atom typing rules, carbon as well as sulfur are both categorized as *nonpolar*, which makes the distinction between the two orientations of thiophene almost impossible. Second, two carboxylate groups are solvent-exposed in the experimentally determined structure. Because interactions with the solvent are not accounted for in the scoring function, these carboxylate groups are placed at the protein surface to form favorable hydrogen bonding interactions according to the scoring function. Thus, in this as well as in some other cases the binding motif can be identified correctly but the rmsd values greater than 2 Å suffer from unspecific interactions, which cannot be ac-

counted for by the simple form of the scoring functions presented here. Similar observations are made e.g. for the complexes 1gm8, 1jd0, 1oq5, 1owe, and 1ygc.

Search Algorithm Parameter Optimization. Besides the parameters for $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$, the parameter settings of the search algorithm have a major impact on PLANTS's overall performance. In fact, we have tuned the search parameter settings of PLANTS for each setting for the scoring function's parameters because each setting induces its own specific fitness landscape. For this task, we have chosen a training set of seven PDB codes (CCDC/Astex data set), five from the scoring function parametrization (1a4g, 1hsb, 1lcp, 1mmq, and 1rt2) and two additional ones (1fax and 4dfr), and used it to determine reasonable values for parameter σ, scaling the number of iterations of the ACO algorithm, the evaporation factor ρ, the number of nonimproving iterations until an update with solution s^{db} is forced ($iter_{db\text{-}update}$), and the nms_{tol} simplex tolerance for the local search (nms_{tol}) as well as the refinement local search ($ref\text{-}nms_{tol}$). The 5 PDB codes from the scoring function training set have been chosen such that all 12 selected scoring models for $PLANTS_{CHEMPLP}$ (M_1 to M_7) and $PLANTS_{PLP}$ (M_1 to M_5) potentially can reproduce the correct binding mode. PDB codes 1fax (target *factor Xa*) and 4dfr (target *dihydrofolate reductase*) represent relatively large, druglike compounds and are thus preferred to be reproduced correctly with high reliability. The value for the number of ants was fixed to 20, and p_{best} was set to 0.5 following the proposed settings in ref 13.

For the tuning, the following parameter values have been considered: For σ the values (0.25,0.5,1.0,1.25) have been considered for scoring function $PLANTS_{CHEMPLP}$, while for $PLANTS_{PLP}$ additionally a setting of $\sigma = 1.5$ was tested to approximately match the number of scoring function evaluations reached for $PLANTS_{CHEMPLP}$. Values for parameter ρ were chosen from the set (0.1,0.15,0.2,0.25). For $iter_{db\text{-}update}$ the values 5 and 10 were considered. Finally, the pairs (nms_{tol}, $ref\text{-}nms_{tol}$) considered for the tolerance of the local and the refinement local search were (0.02, 0.01), (0.01, 0.01), (0.02, 0.0001), and (0.01, 0.0001). This resulted in 128 parameter settings tested for $PLANTS_{CHEMPLP}$ and 160 for $PLANTS_{PLP}$. An analysis similar to the one presented in ref 13 has been carried out to identify well-performing search parameter settings. All identified search algorithm parameter settings for the 7 $PLANTS_{CHEMPLP}$ and the 5 $PLANTS_{PLP}$ scoring function models can be found in the Supporting Information.

COMPUTATIONAL RESULTS

For the final evaluation of the algorithm performance in pose prediction experiments, that is, docking the ligand back into its experimentally determined protein structure, all complexes of the *clean list* of the *CCDC/Astex* set,[41] restricted to noncovalently bound complexes (*clean list*^{nc}, 213 complexes from which 31 have been used for the scoring function parametrization) as well as the *Astex diverse* test set[42] (85 protein−ligand complexes already used for the scoring function model selection), were used. In the case of the *clean list*^{nc}, the spherical binding site definition as given in the test set was used, while for the *Astex diverse* set the binding site was defined starting from the ligand's crystallographic conformation considering all protein atoms that are up to 6 Å away from any ligand heavy-atom. The problem dimension of the continuous optimization problems treated for these test sets was, on average, 18 taking into account the ligand's and the protein's degrees of freedom. For these data sets also results for the docking program GOLD,[14,15,45] a tool frequently used in the pharmaceutical industry, will be presented in order to allow for a comparison with a state-of-the-art approach. Due to the stochastic nature of PLANTS, each experiment has been repeated 25 times. Thus, all presented success rates and timing results are averaged over 25 experiments. A 64 bit version of PLANTS has been used to perform all experiments on an Intel Xeon X5365 CPU running at 3 GHz. Prior to docking, each ligand structure taken as input for PLANTS has been randomized with respect to its translational, rotational, and torsional degrees of freedom to avoid any bias toward the experimentally determined structure. Tables with the PLANTS results obtained for docking back the ligands into their experimentally determined protein structures are given as Supporting Information (all combinations of scoring function models and speed settings).

We first focus on what we recommend as the standard parameter settings of PLANTS (identified by *speed 1* in the Supporting Information). For the *Astex diverse* set, all scoring function models of $PLANTS_{CHEMPLP}$ perform remarkably well with average success rates for the top-ranked solutions between 84% for model M_1 up to 87% for model M_3. The success rates for solutions up to rank three and ten, as identified by the cluster algorithm, are as high as 97% for some scoring function models, e.g., model M_3. The average search time per ligand varies between approximately 23 and 29 s depending on the scoring function, which definitely makes PLANTS suitable for *virtual screening* tasks.

If very large databases need to be screened within a very short time while still maintaining a certain degree of pose prediction reliability, also search parameter settings *speed 2* and *speed 4* may be considered. For search setting *speed 2*, which approximately results in half the search time of around 12 s per ligand compared to setting *speed 1*, average success rates for the top-ranked solutions between 82% to 84% can be reached. For *speed 4* resulting in average search times of approximately 6 s, still success rates of 75% to 79% can be obtained for the top-ranked solutions.

Considering the top-ranked solutions, $PLANTS_{PLP}$ models reach worse success rates than the $PLANTS_{CHEMPLP}$ models: For example, for the default settings *speed 1*, they reach average success rates between 76% and 83% in around 9 to 13 s per ligand. The higher success rates for $PLANTS_{CHEMPLP}$ can be attributed to the higher specificity due to the inclusion of angle-dependencies in the hydrogen and metal bonding terms. However, the success rates of the $PLANTS_{PLP}$ models up to ranks three and ten are comparable to those of the $PLANTS_{CHEMPLP}$ models. The average success rates decrease slightly if half the search time is used for search parameter settings *speed 2*. Finally, search parameter setting *speed 4* reaches average success rates of 73−76% at average docking times of around 3 s per ligand. Concerning $PLANTS_{PLP}$, model M_5 shows the overall most convincing performance for this druglike test set.

For the 213 complexes of the CCDC/Astex *clean list*^{nc} all 7 $PLANTS_{CHEMPLP}$ models show a similarly good performance as on the *Astex diverse* set. However, the average

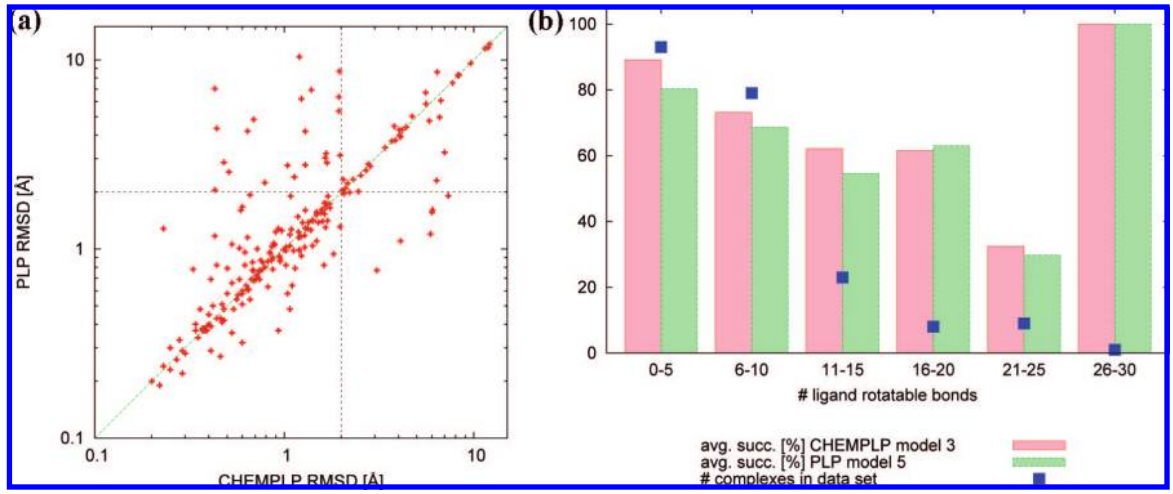

Figure 6. Comparison of $PLANTS_{CHEMPLP}$ M_3 and $PLANTS_{PLP}$ M_5 on the CCDC/Astex *clean list*^{nc}. (a) rmsd values of the top-ranking ligand poses. The dotted black lines mark the 2 Å cutoff imposed for correct predictions (both axes use logarithmic scale). (b) Pose prediction performance with respect to the number of rotatable bonds in the ligand molecule. The blue dots give the number of complexes in the CCDC/Astex *clean list*^{nc} for each interval of rotatable bonds.

success rates are typically lower and the average docking times per ligand are higher because the CCDC/Astex *clean list*^{nc} contains ligands with up to 28 rotatable bonds (e.g., PDB code 1ppi, pancreatic amylase with a carbohydrate inhibitor). Such protein−ligand complexes lead to high average search times due to the automatic scaling of the number of ACO-iterations with respect to the ligand size. When only considering ligands with up to 10 rotatable bonds, similar search times as for the *Astex diverse* set can be observed. The performance of PLANTS with respect to different subsets will be discussed later.

Concerning the particular results, the same trends as for the *Astex diverse* set can be observed. $PLANTS_{CHEMPLP}$ reaches average success rates between 75% and 77% at average search times between 53 and 67 s with parameter setting *speed 1*. Considering solutions up to rank 10, success rates up to 92% can be obtained. For settings that approximately halve the search time (setting *speed 2*) the average success rate diminishes by around 3−4% independent of the PLANTS model. With the shortest computation times (setting *speed 4*, average search time about 14 s), still success rates of around 68% can be observed.

As for the *Astex diverse* set, the $PLANTS_{PLP}$ models perform slightly worse than the $PLANTS_{CHEMPLP}$ models when looking at the top-ranked solutions only. However, the performance for solutions up to rank 10 is similar. For details we refer to the Supporting Information.

Comparison of $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$. Due to the particularly good performance of $PLANTS_{CHEMPLP}$ model M_3 and $PLANTS_{PLP}$ model M_5 across all test sets, these were chosen as the standard settings. Next, we compare in more detail their results on the 213 protein−ligand complexes part of the CCDC/Astex *clean list*^{nc}. Figure 6a shows the rmsd distribution of the top-ranked protein−ligand conformations, averaged across the 25 experiments, for $PLANTS_{CHEMPLP}$ M_3 versus $PLANTS_{PLP}$ M_5; clearly, a correlation of both scoring functions with respect to their rmsd can be observed. Although $PLANTS_{CHEMPLP}$ can identify more complexes within an rmsd of 2 Å, there are a few complexes in the lower right rectangle that are only predicted correctly by $PLANTS_{PLP}$. Of course, for more complexes the contrary is true (see upper left rectangle). The upper right square of this plot contains the complexes that are not identified correctly by both scoring functions.

Figure 6b shows the pose prediction performance of both scoring functions with respect to the number of rotatable bonds in the ligand molecule. The number of rotatable bonds corresponds to the classification by $PLANTS_{CHEMPLP}$, which, different from $PLANTS_{PLP}$, treats *OH*- and NH_3^+-groups rotatable due to the consideration of explicit polar hydrogen atoms in the computation of the distance- and angle-dependent hydrogen bonding terms. In general, both scoring functions show a similar pose prediction performance, but in most cases $PLANTS_{CHEMPLP}$ performs slightly better than $PLANTS_{PLP}$. The pose prediction performance decreases for both scoring functions with increasing numbers of rotatable bonds, which can be attributed to not only sampling but also scoring function errors. Note that most complexes of the data set contain ligands with up to 15 rotatable bonds. Thus, results for ligands with more than 15 rotatable are based on only very few ligands and are, therefore, statistically not significant; in fact, in the interval 26−30 only one protein−ligand complex is included. For this single complex, PLANTS performed extremely well, while for other highly flexible ligands much lower success rates are expected and are actually seen (data not shown). Due to the fact that these large and highly flexible ligands are of minor importance when searching for lead-structures, no more efforts were devoted toward improving the docking performance for these kinds of structures.

Table 8 gives results for subsets of the CCDC/Astex *clean list*^{nc} grouped by specific features of the complexes. $PLANTS_{CHEMPLP}$ reaches similar success rates of about 76% for protein−ligand complexes with metal-ions (subset *metal*) and without metal-ions (subset *no metal*). Differently, $PLANTS_{PLP}$ shows significant differences between the success rates on these two subsets: 59% success rate for the metal-containing subset versus 73% for the subset without metal-ions; this difference can be attributed to the advanced treatment of metal-interactions in $PLANTS_{CHEMPLP}$. For the subset restricted to ligands with up to 10 rotatable bonds (RB 0−10), the success rate of PLANTS is almost 82% at

Table 8. Pose Prediction Results Averaged over 25 Experiments for Different Subsets of the CCDC/Astex *clean list*[nc a]

subset	complexes	success rate [%] up to rank 1	3	10	time [s]	eval. [in 10^6]
			$PLANTS_{CHEMPLP}$ M_3			
metal	40	77 (3)	91 (4)	94 (2)	27	3.19
no metal	173	77 (1)	87 (2)	91 (2)	68	5.35
RB 0−10	172	82 (1)	92 (1)	94 (1)	20	2.68
			$PLANTS_{PLP}$ M_5			
metal	40	59 (2)	84 (2)	92 (2)	18	3.19
no metal	173	73 (1)	83 (1)	89 (2)	42	4.97
RB 0−10	172	75 (1)	87 (1)	93 (1)	11	2.47

[a] The three subsets consist of protein-ligand complexes containing metal-ions (subset metal), containing no metal-ions (subset no metal) and containing ligands with up to 10 rotatable bonds (subset RB 0−10). The average success rate up to rank 1, 3, and 10 (standard deviations given in brackets), the average search time and the average number of scoring function evaluations are presented for $PLANTS_{CHEMPLP}$ M_3 and $PLANTS_{PLP}$ M_5 with respect to the specific subset.

Table 9. Pose Prediction Results Considering Top-Ranked Solutions only for PLANTS and GOLD on the 213 Complexes of the CCDC/Astex *clean list*[nc] and the 85 Complexes of the Astex Diverse Set[a]

	CCDC/Astex *clean list*[nc]					
	PLANTS speed setting					
	1		2		4	
scoring function	succ. [%]	time [s]	succ. [%]	time [s]	succ. [%]	time [s]
$PLANTS_{CHEMPLP}$	77 (1)	61	73 (2)	30	69 (2)	13
$PLANTS_{PLP}$	71 (1)	37	69 (2)	19	65 (2)	8
	GOLD autoscale setting					
	1		0.5		0.25	
scoring function	succ. [%]	time [s]	succ. [%]	time [s]	succ. [%]	time [s]
$GOLD_{GOLDscore}$	72 (1)	102	71 (1)	59	69 (1)	31
$GOLD_{Chemscore}$	67 (2)	42	65 (2)	23	63 (2)	11
	Astex diverse					
	PLANTS speed setting					
	1		2		4	
scoring function	succ. [%]	time[s]	succ. [%]	time [s]	succ. [%]	time [s]
$PLANTS_{CHEMPLP}$	87 (2)	26	84 (2)	13	79 (3)	5
$PLANTS_{PLP}$	84 (2)	13	83 (2)	7	78 (3)	3
	GOLD autoscale setting					
	1		0.5		0.25	
scoring function	succ. [%]	time[s]	succ. [%]	time [s]	succ. [%]	time [s]
$GOLD_{GOLDscore}$	79 (2)	53	78 (2)	29	77 (3)	16
$GOLD_{Chemscore}$	78 (1)	22	75 (2)	11	71 (3)	6

[a] In the case of PLANTS, scoring function model M_3 was used for $PLANTS_{CHEMPLP}$ and model M_5 for $PLANTS_{PLP}$. Due to the stochastic nature of both docking approaches, all presented values are averaged over 25 experiments. Standard deviations for the success rates are given in brackets.

average search times of 20 s per ligand. This is comparable to the results obtained for the *Astex diverse* set. Similarly, $PLANTS_{PLP}$ performs with a success rate of nearly 75% better on this druglike subset than on the whole test set. For this subset, the average search time per ligand is only around 10 s.

Comparison to GOLD. In order to be able to assess the performance of PLANTS more objectively, the same experiments have been carried out with GOLD version 3.2.[14,45,15] For the GOLD docking experiments, three different settings of parameter *autoscale*, which automatically chooses appropriate GA search settings in dependence of the given protein binding site and the ligand molecule, have been tested. The maximum number of GA runs per ligand was set to 10, *early termination* and *cavity detection* were activated, and *GOLDscore* ($GOLD_{GOLDscore}$) as well as *Chemscore* ($GOLD_{Chemscore}$) were used as the scoring functions. The GOLD runs were performed on the same computer architecture as the PLANTS experiments. The results obtained for different settings of parameter *autoscale* can be found in Table 9. As PLANTS, GOLD uses a nondeterministic search algorithm, and thus all experiments were performed 25 times. The presented average docking times exclude the time needed for protein setup and ligand preparation to allow for a direct comparison with the results

obtained for PLANTS. These are shown for the standard settings $PLANTS_{CHEMPLP}$ M_3 and $PLANTS_{PLP}$ M_5.

On both test sets, $GOLD_{GOLDscore}$ reaches higher average success rates than $GOLD_{Chemscore}$, regardless of the parameter value of *autoscale*, which determines the search time.

For the 213 complexes of the *clean list*^{nc}, $GOLD_{GOLDscore}$ reaches an average success rate of around 72% at average docking times of 102 s per ligand for a setting of *autoscale* = 1. Only $PLANTS_{CHEMPLP}$ reaches higher success rates, and it does so at shorter search times for search parameter settings *speed 1* and *speed 2*: average success rates of around 77% and 73% are obtained at average docking times of 61 s and around 30 s, respectively. The latter time is comparable to the fastest settings we examined for $GOLD_{GOLDscore}$ (*autoscale* = 0.25): on a time equalized basis, $PLANTS_{CHEMPLP}$ reaches a success rate that is 5% higher than that of $GOLD_{GOLDscore}$. $GOLD_{Chemscore}$ only reaches an average success rate between 67% and 63% at average docking times between 42 and 11 s per ligand for search settings *autoscale* = 1 and *autoscale* = 0.25, respectively. While the average search times for $GOLD_{Chemscore}$ and $PLANTS_{PLP}$ are in a similar range for each of the different speed/autoscale settings, $PLANTS_{PLP}$ reaches for comparable search times higher success rates than $GOLD_{Chemscore}$.

On the *Astex diverse* test set, the advantage of PLANTS over GOLD is even more pronounced when considering the success rate differences at comparable search times. Apart from this fact, the most noteworthy observations are the following. If one increases the settings of autoscale from 0.25 (search time approximately 5 s) to 1 (search time approximately 53 s) for $GOLD_{GOLDscore}$, the pose prediction performance only improves slightly from a success rate of 77% to approximately 79%. Differently, $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$ profit much stronger from increases of the computation time; for example $PLANTS_{CHEMPLP}$ goes from a success rate of 79% (search time approximately 5 s) to about 87% (search time approximately 26 s). Finally, on this test set, $PLANTS_{PLP}$ shows a better performance than $GOLD_{GOLDscore}$. In fact, the $PLANTS_{PLP}$ setting *speed 2* reaches higher success rates than the highest reached by $GOLD_{GOLDscore}$ (83% versus 79%) at computation times that are more than seven times smaller.

In summary, for both data sets, one can find for each setting of GOLD a setting for PLANTS that reaches (in part much) higher success rates at (slightly) lower average search times. For a sincere comparison of success rates, sampling errors, arising from the fact that the test sets only contain a limited number of complexes, have to be considered.[41] The smaller differences seen between PLANTS and GOLD are very probable within these sampling errors. Nevertheless, the performance of PLANTS is expected to be at least comparable to the performance of state-of-the-art approaches also for other targets.

CONCLUSIONS

In this paper we presented the parametrization procedure used to identify meaningful parameter models for the two scoring functions employed in our docking algorithm PLANTS, $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$. The procedure identified several parametrizations that reach very high performance for the task of pose prediction on two test sets comprising 298 protein−ligand complexes in total. Up to 87% of the complexes of the *Astex diverse* set and 77% of the CCDC/Astex *clean list*^{nc} could be reproduced with root-mean-square deviations of less than 2 Å with respect to their experimentally determined structures. For druglike ligands, the average search time per complex was around 26 s which definitely makes PLANTS suitable even for large-scale *virtual screening* tasks. Compared to the state-of-the-art docking program GOLD this is, especially for the druglike *Astex diverse* set, an improvement in pose prediction performance. Additionally, we identified for each scoring function different parameter settings for the search algorithm that can be used to balance between pose prediction reliability and search speed. Although the scoring functions still exhibit systematic failures on some complexes, the obtained results show that our approach is at least comparable to other state-of-the-art docking approaches with respect to its pose-prediction capabilities.

Some systematic failures can be attributed to missing interactions not accounted for by the scoring functions. While in ref 13 already experiments considering protein side-chain flexibility have been presented, meanwhile also additional degrees of freedom like displaceable explicit water molecules can be handled by PLANTS; this is the topic of a subsequent publication. Additionally, a modified parallel version of the PLANTS algorithm capable of exploiting the enormous floating point performance of today's *graphics processing units* (GPUs) will be presented soon. These computational resources can be used to either speed up a single docking-run or to use more complex scoring functions.

An academic version of PLANTS is available free of charge from the authors.

ACKNOWLEDGMENT

The authors thank Tim ten Brink for his help regarding the preparation of the data sets and Jens Gimmler for helpful discussions and careful reading of the manuscript. Additionally, the CUSS cluster in Ulm and the HPC cluster of the University of Konstanz is acknowledged for allocating computational resources. This work was supported by a scholarship of the Landesgraduiertenförderung Baden-Württemberg awarded to Oliver Korb. Thomas Stützle acknowledges support of the Belgian F.R.S.-FNRS, of which he is a research associate.

Supporting Information Available: Optimized search algorithm parameters as well as the PLANTS results, obtained for docking back the ligands of the *CCDC/Astex* and *Astex diverse* sets into their experimentally determined protein structures for all scoring function models of $PLANTS_{CHEMPLP}$ and $PLANTS_{PLP}$. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Drug Discovery* **2004**, *3*, 935–949.

(2) Lang, P. T.; Aynechi, T.; Moustakas, D.; Shoichet, B.; Kuntz, I. D.; Brooijmans, N.; Oshiro, C. Molecular docking and structure-based design. In *Drug Discovery Research*; Huang, Z., Ed.; Wiley: Hoboken, 2007; pp 3−23.

(3) Krumrine, J.; Raubacher, F.; Brooijmans, N.; Kuntz, I. Principles and methods of docking and ligand design. *Meth. Biochem. Anal.* **2003**, *44*, 443–476.
(4) Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335–373.
(5) Gohlke, H.; Klebe, G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644–2676.
(6) 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.
(7) Krovat, E. M.; Steindl, T.; Langer, T. Recent Advances in Docking and Scoring. *Curr. Comput.-Aided Drug Des.* **2005**, *1*, 93–102.
(8) Perola, E.; Walters, W. P.; Charifson, P. S. A Detailed Comparison of Current Docking and Scoring Methods on Systems of Pharmaceutical Relevance. *Proteins* **2004**, *56*, 235–249.
(9) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A Review of Protein-Small Molecule Docking Methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151–166.
(10) Chen, H. M.; Liu, B. F.; Huang, H. L.; Hwang, S. F.; Ho, S. Y. SODOCK: Swarm Optimization for Highly Flexible Protein-Ligand Docking. *J. Comput. Chem.* **2007**, *7*, 308–313.
(11) Janson, S.; Merkle, D.; Middendorf, M. Molecular Docking with Multi-Objective Particle Swarm Optimization. *Appl. Soft Comp.* **2007**, *8* (1), 666–675.
(12) Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. *Lect. Notes Comput. Sci.* **2006**, *4150*, 247–258.
(13) Korb, O.; Stützle, T.; Exner, T. E. An Ant Colony Optimization Approach to Flexible Protein-Ligand Docking. *Swarm Intel.* **2007**, *1* (2), 115–134.
(14) Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
(15) 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.
(16) 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.
(17) DesJarlais, R. L.; Sheridan, R. P.; Seibel, G. L.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor Binding Site of Known Three-Dimensional Structure. *J. Med. Chem.* **1988**, *31*, 722–729.
(18) Gschwend, D. A.; Kuntz, I. D. Orientational Sampling and Rigid-Body Minimization in Molecular Docking Revisited: On-the-Fly Optimization and Degeneracy Removal. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 123–132.
(19) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search Strategies for Automated Molecular Docking of Flexible Molecule Databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
(20) Goodsell, D. S.; Olson, A. J. Automated Docking of Substrates to Proteins by Simulated Annealing. *Proteins* **1990**, *8*, 195–202.
(21) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(22) Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J. Comput. Chem.* **1981**, *2*, 287–303.
(23) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiner, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3585–3616.
(24) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular Recognition of the Inhibitor AG-1243 by HIV-1 Protease: Conformationally Flexible Docking by Evolutionary Programming. *Chem. Biol.* **1995**, *2*, 317–324.
(25) Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Freer, S. T.; Rose, P. W. Monte Carlo Simulations of the Peptide Recognition at the Consensus Binding Site of the Constant Fragment of Human Immunoglobulin G: The Energy Landscape Analysis of a Hot Spot at the Intermolecular Interface. *Proteins* **2002**, *48*, 539–557.
(26) Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Freer, S. T.; Rose, P. W. Computational Detection of the Binding-Site Hot Spot at the Remodeled Human Growth Hormone-Receptor Interface. *Proteins* **2003**, *53*, 201–219.
(27) Verkhivker, G. M. Computational Analysis of Ligand Binding Dynamics at the Intermolecular Hot Spots with the Aid of Simulated Tempering and Binding Free Energy Calculations. *J. Mol. Graphics Modell.* **2004**, *22*, 335–348.
(28) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical Scoring Functions: I. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
(29) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical Scoring Functions: II. The Testing of an Empirical Scoring Function for the Prediction of Ligand-Receptor Binding Affinities and the Use of Bayesian Regression to Improve the Quality of the Model. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 503–519.
(30) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.
(31) Muegge, I.; Martin, Y. C. A General and Fast Scoring Function for Protein-Ligand Interactions: A Simplified Potential Approach. *J. Med. Chem.* **1999**, *42*, 791–804.
(32) Muegge, I. A Knowledge-Based Scoring Function for Protein-Ligand Interactions: Probing the Reference State. *Perspect. Drug Discovery Des.* **2000**, *20*, 99–114.
(33) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions. *J. Mol. Biol.* **2000**, *295*, 337–356.
(34) Gohlke, H.; Hendlich, M.; Klebe, G. Predicting Binding Modes, Binding Affinities and Hot Spots for Protein-Ligand Complexes Using a Knowledge-Based Scoring Function. *Perspect. Drug Discovery Des.* **2000**, *20*, 115–144.
(35) Velec, H. F. G.; Gohlke, H.; Klebe, G. DrugScoreCSD-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.
(36) Mooij, W. T. M.; Verdonk, M. L. General and Targeted Statistical Potentials for Protein-Ligand Interactions. *Proteins* **2005**, *61*, 272–287.
(37) Goss, S.; Aron, S.; Deneubourg, J. L.; Pasteels, J. M. Self-Organized Shortcuts in the Argentine Ant. *Naturwissenschaften* **1989**, *76*, 579–581.
(38) Deneubourg, J. L.; Aron, S.; Goss, S.; Pasteels, J. M. The Self-Organizing Exploratory Pattern of the Argentine Ant. *J. Insect Behav.* **1990**, *3*, 159–168.
(39) Dorigo, M.; Stützle, T. *Ant Colony Optimization*; MIT Press: Cambridge, 2004.
(40) Stützle, T.; Hoos, H. H. MAX-MIN Ant System. *Fut. Gen. Comput. Syst.* **2000**, *16* (8), 889–914.
(41) Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A New Test Set for Validating Predictions of Protein-Ligand Interaction. *Proteins* **2002**, *49* (4), 457–471.
(42) 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.
(43) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *28* (2), 612–623.
(44) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C: The Art of Scientific Computing*; Cambridge University Press: Cambridge, 1992.
(45) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins* **2003**, *52*, 609–623.
(46) Clark, M.; Cramer III, R. D.; van Opdenhosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(47) Kabsch, W. A Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1976**, *A32*, 922–923.
(48) Kabsch, W. A Discussion of the Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1978**, *A34*, 827–828.

CI800298Z