J Comput Aided Mol Des (2006) 20:131–144
DOI 10.1007/s10822-006-9041-7

ORIGINAL PAPER

Protein Alpha Shape (PAS) Dock: A new gaussian-based score function suitable for docking in homology modelled protein structures

Kristin Tøndel · Endre Anderssen · Finn Drabløs

Received: 17 December 2003 / Revised: 8 August 2005 / Accepted: 27 February 2006 /
Published online: 2 May 2006

Abstract Protein Alpha Shape (PAS) Dock is a new empirical score function suitable for virtual library screening using homology modelled protein structures. Here, the score function is used in combination with the geometry search method Tabu search. A description of the protein binding site is generated using gaussian property fields like in Protein Alpha Shape Similarity Analysis (PASSA). Gaussian property fields are also used to describe the ligand properties. The overlap between the receptor and ligand hydrophilicity and lipophilicity fields is maximised, while minimising steric clashes. Gaussian functions introduce a smoothing of the property fields. This makes the score function robust against small structural variations, and therefore suitable for use with homology models. This also makes it less critical to include protein flexibility in the docking calculations. We use a fast and simplified version of the score function in the geometry search, while a more detailed version is used for the final prediction of the binding free energies. This use of a two-level scoring makes PAS-Dock computationally efficient, and well suited for virtual screening. The PAS-Dock score function is trained on 218 X-ray structures of protein–ligand complexes with experimental binding affinities. The performance of PAS-Dock is compared to two other docking methods, AutoDock and MOE-Dock, with respect to both accuracy and computational efficiency. According to this study, PAS-Dock is more computationally efficient than both AutoDock and MOE-Dock, and gives a better prediction of the free energies of binding. PAS-Dock is also more robust against structural variations than AutoDock.

Keywords Computational docking · Empirical score function · Gaussian property fields · Protein Alpha Shape Similarity Analysis (PASSA) · Virtual screening

K. Tøndel (✉) · F. Drabløs
Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Laboratory Center, Erling Skjalgsons St. 1, N-7006 Trondheim, Norway
E-mail: kristin.tondel@ntnu.no
Tel.: +47 72 57 33 23
Fax: +47 72 57 64 00

E. Anderssen · K. Tøndel
Department of Chemistry, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

Introduction

The knowledge about genes and proteins associated with pathological states is increasing, especially following the human genome project. This has highly increased the potential of computer-aided drug design and virtual screening. A large variety of methods is available for small-molecular docking and virtual library screening. However, most methods are highly time consuming. Many methods also have limitations such as neglection of receptor flexibility, inaccuracies in determination of partial charges and underestimation of hydrophobic effects. Docking methods typically use a search method to explore the conformational space of the ligand bound to the protein, and a score function to guide the geometry search and to estimate the binding affinity for the different conformations. Search methods range from rigorous search methods such as simulated annealing to faster methods such as Tabu search [1] and genetic algorithms [2]. Since the number of available experimentally determined protein structures is not increasing at the same speed as the number of available protein sequences, homology modelling has

great potential in structure-based drug design. However, homology models are not accurate enough to be used with most existing docking methods. It is therefore a need for new docking methods that are robust against small structural errors.

Many score functions exist for ranking drug candidates and prediction of binding affinities between a receptor and a ligand. Existing score functions can be divided into three main classes: force field-based methods, empirical score functions and knowledge-based methods. The use of score functions in drug design has recently been reviewed by Böhm and Stahl [3].

Force field-based scoring methods use nonbonded energies of molecular mechanics (MM) force fields to estimate the binding affinity. Force field-based methods are generally time consuming and sensitive to errors in the protein structure models, estimated partial charges and protonation states. Examples of force field-based methods showing some success include the score function implemented in the AutoDock program [4] which utilises parameters from the AMBER force field [5], MM PB/SA [6] which complement the electrostatic interactions by a solvation term calculated by the Poisson-Boltzmann equation [7] and the OWFEG (one window free energy grid) method [8].

Empirical score functions are generally faster than force field-based methods. The underlying idea is that the binding free energy can be interpreted as a weighted sum of localised interaction terms. The interaction terms typically represent hydrogen bonding terms, ionic interactions, hydrophobic interactions, entropy change associated with binding, etc. In addition, penalty functions for e.g. steric clashes can be added. The interaction terms are usually calculated using experimental 3D structures of receptor-ligand complexes, and the weights are estimated by multiple linear regression of experimental binding affinities. One disadvantage of empirical score functions is the dependency on the set of experimental structures used to train the functions. Usually, between 50 and 100 complexes are used to train the score functions, but recently it was shown that more than 100 complexes are needed for convergence [9]. Examples of empirical score functions showing some promise include PLP [10, 11], ChemScore [12] and X-Score [13]. PLP uses a sum of pairwise linear potentials between ligand and protein heavy atoms with parameters dependent on interaction type. Each pair of interacting atoms is assigned one of three interaction types: donor and acceptor hydrogen bonding, repulsive donor–donor and acceptor–acceptor interactions and generic dispersion of other contacts. The ChemScore function is a weighted sum of hydrogen-bonding terms, terms accounting for metal–ligand interactions, hydrophobic effects and the number of rotors. The X-Score regression equation contains a van der Waals interaction term, a hydrogen bonding term, a term representing the hydrophobic effect and a torsional entropy penalty.

Knowledge-based score functions are derived by statistical analysis of structural data alone, without reference to experimentally determined binding affinities. They are based on the inverse formulation of the Boltzmann law. The frequency of occurrence of individual contacts is used as a measure of their energetic contribution to binding. The frequencies are compared to frequencies from a random or average distribution. A high frequency indicates an attractive interaction, while a low frequency indicates a repulsive interaction. Knowledge-based score functions include the Potential of Mean Force (PMF) score function [14–16] and DrugScore [17]. The PMF score function is a sum of distance-dependent interaction potentials for atom pairs, where both enthalpic and entropic effects are assumed to be included implicitly. In the DrugScore equation also solvent-accessible surface dependent singlet potentials for protein and ligand atoms are included.

Recently, a comparison of eleven score functions using the same set of experimental structures was published [18]. This study indicates that X-Score and DrugScore are the score functions most suited for use with conformational sampling, since they produce a funnel-shaped energy surface for protein–ligand complexation, and therefore will most likely lead to a faster convergence to the global minimum. The study also indicates that a combination of several different score functions might be advantageous. X-Score, DrugScore and PLP were the score functions showing most promise in the study. However, these score functions give only moderate correlation between the predicted and the experimental binding affinities for these 100 structures, using the experimentally determined conformations. Hence, the study indicates a need for improved methods for ranking a large number of ligands according to success of binding to a receptor.

In this work we have developed a new empirical score function, Protein Alpha Shape (PAS) Dock, for estimation of binding affinities to a receptor using gaussian property distributions for both the protein and the ligands. The score function evaluates only the match between the lipophilicity and hydrophilicity of the receptor and the ligand, in addition to describing van der Waals effects. This makes it easy to interpret and robust against overfitting. The fact that the score function is based on gaussian density estimates makes it more robust against the errors typically found in homology models, since gaussian functions give a less detailed representation than force field models, and they have neither steep derivatives nor singularities [19]. Hence, this score function is well suited for virtual screening using protein structure models built by homology modelling. Including protein flexibility is also less important

compared to many other docking methods, since a probability distribution is used to represent the molecular properties instead of one single value. Because of the simplified description of the protein–ligand interactions, the accuracy of our new score function can not be compared to score functions that take e.g. partial charges and electrostatics into account. However, the speed of our calculations makes this method an effective tool for exploration of the ligand conformational space and generation of starting conformations for more accurate docking methods.

Two different score functions are used in PAS-Dock, a rough evaluation of the match between the protein and ligand structures for the geometry search, and a more accurate scoring for the final estimation of the binding affinities. In this way, we get a good prediction of the free energy of binding, in addition to a fast search for the best ligand conformation. Since our calculations are independent of placement of hydrogen atoms, the fact that a hydrogen atom can form hydrogen bonds only with atoms pointing towards it is not accounted for. This is one of the major limitations of our method. We plan to account for this in a future version.

Methods

Definition of score functions

In PAS-Dock, the protein binding site properties are mapped using the newly developed method Protein Alpha Shape Similarity Analysis (PASSA) [20], while the ligand properties are described using gaussian property distributions similar to those used in Comparative Molecular Similarity Index Analysis (CoMSIA) [21]. Both PASSA and CoMSIA work by assigning property distributions to each atom, so that the molecules are described by the spatial distribution of their interactions. At each point of a 3D grid, the values of the molecular similarity fields are computed. A molecular similarity index is based on atomic parameters such as lipophilicity, hydrogen bond donor and acceptor properties, etc. A gaussian function with the intensity of the atomic parameter, and a standard deviation (σ) proportional to the atomic radius is centred at each atom. For each physicochemical property, the value in a grid point is computed as the sum of the contributions from all gaussian functions representing that property (see Eq. 1).

$$F(q,j) = \sum_{i=1}^{n} \frac{\omega_{ik}}{(\sigma_i\sqrt{2\pi})^3} \cdot e^{\frac{-r_{iq}^2}{2\sigma_i^2}} \quad (1)$$

F is the value of the similarity field in grid point q of molecule j, ω_{ik} is the value of the physicochemical property k of atom i, r_{iq} is the distance between grid point q and atom i and σ_i is proportional to the atomic radius of atom i.

PASSA [20] converts the discreet information contained in the placement of geometrical objects known as alpha spheres and the positions of protein atoms to a continuous field using gaussian density estimates. An alpha sphere is a sphere that contacts four protein atoms on its surface and has no atoms contained internally. Centres of alpha spheres have been found to correspond well with the placement of atoms in bound ligands [22]. Alpha spheres are determined geometrically, using only the positions and radii of the heavy atoms. This eliminates the need for placing hydrogens and determining protonation states and partial charges. The alpha spheres are classified as hydrophobic or hydrophilic depending on the protein atoms that they contact. In PASSA, gaussian functions (as shown in Eq. 1) are centred at dummy atoms placed at each alpha sphere centre, and at all protein atoms. A 3D grid is placed around the binding site of the protein, and in each grid point the sum of the contributions from all gaussian functions is computed. The use of gaussian functions with a very simple partitioning according to the hydrophilic or hydrophobic nature of the alpha spheres reduces some of the problems associated with force field models [19].

The dummy atoms placed in each alpha sphere centre has the properties of either an oxygen atom or a carbon atom, depending on whether the alpha spheres were classified as hydrophilic or hydrophobic. The gaussian functions centred at the dummy atoms are given unit weight for either the hydrophilic or the hydrophobic field, according to the properties of the alpha spheres. To include steric effects, the gaussian functions centred in protein atoms are given the weight −1 for both fields, since it is unfavourable to have a ligand atom in the position of a protein atom. In PAS-Dock, a separate van der Waals field based on gaussian functions is generated for the protein, in addition to the hydrophilic and hydrophobic field. Hence, van der Waals effects are taken into account both implicitly in the calculation of the hydrophilic and hydrophobic fields, and explicitly as a separate property field. This makes us able to tune the parameters in the score function better, to avoid collision between ligand and protein atoms. Gaussian functions are also centred in each ligand atom, and classified according to the atom types. The property fields are generated by summing the contributions from all gaussian functions of a class in each grid point. The following fields are used to describe the ligands: hydrophilicity, lipophilicity, van der Waals, hydrogen acceptor and hydrogen donor. The same 3D grid is used to calculate all gaussian property fields.

In each grid point, the products of the ligand and protein gaussian fields are computed. These product values are

then summed over all grid points, giving one variable describing the match between the given ligand and protein fields. The variable "protein lipophilicity * ligand lipophilicity (lip_lip)", for example, describes the match between the protein and ligand lipophilicity fields, summed over all grid points. The PAS-Dock score function is given in Eq. 2.

$$S = x_1\ \text{hyd\_hyd} + x_2\ \text{lip\_lip} + x_3\ \text{hyd\_lip} + x_4\ \text{lip\_hyd} + x_5\ \text{lip\_hacc} + x_6\ \text{vdw\_vdw} + x_7\ \text{f(vdw\_vdw)} \quad (2)$$

For the protein, only lipophilicity and hydrophilicity are considered, since the generation of the gaussian property fields for the protein is based on calculation of alpha spheres, which are classified as either hydrophilic or lipophilic. No hydrogens or partial charges are taken into account in the calculations. Hence, our calculations do not separate hydrogen donors from acceptors on the protein, and no directions are considered when estimating hydrophilic interactions.

In the score function used to predict the free energy of binding ("$\Delta G_{binding}$"), all grid points are used to calculate the values of the variables (Table 1), and gaussian functions are used to describe both the protein and the ligand properties, as described above. In the geometry search score function ("Score") only the values of the protein fields in the positions of the ligand atoms are used (found by multi-linear interpolation), instead of all grid point values. This speeds up the computation of the score values, since the number of calculations that have to be done is reduced. In addition, the ligand property fields are substituted by vectors with zero or unit entries according to the atomic properties. Hence, no gaussian functions are used for the ligand. This speeds up the calculations further. The van der Waals radii of the ligand atoms are used instead of the ligand van der Waals field. To further penalise steric clashes, a sigmoid function of the van der Waals term (Eq. 3) was added to this score function. The receptor atoms are kept in fixed positions during the geometry search. Hence, protein flexibility is not explicitly taken into account in the docking calculations. However, the use of gaussian functions to describe the protein and ligand properties will partly compensate for this, since this makes the score function robust against small structural variations. Hence, protein flexibility is less critical than for many other docking methods.

$$f(\text{vdw\_vdw}) = 100/\left(1 + e^{(-50^* \text{vdw\_vdw} + 8)}\right) \quad (3)$$

Gaussian property fields

The properties of the protein binding site were mapped in the same way as in PASSA [20]. A 3D grid centred at the ligand and extended to 3 Å outside the ligand was used to compute the gaussian property fields. A grid spacing of 0.5 Å was used. The standard deviation of the gaussian functions (σ in Eq. 1) was equal to half the van der Waals radius for both dummy atoms and protein atoms. The standard deviation of the gaussian functions used to compute the protein van der Waals field was also 0.5 times the van der Waals radius for the protein atoms, and the gaussian functions were all given unit weight (ω).

The ligand properties were described using gaussian property fields similar to those used in CoMSIA [21]. In the same way as for the protein, gaussian functions with unit weight and standard deviation of 0.5 times the atomic van der Waals radius were centred in each ligand atom. The atomic properties were determined using the pharmacophore functions [23] in Molecular Operating Environment (MOE) [24]. These functions return either the value zero or one, depending on whether the atom is of the specified pharmacophoric type or not. The van der Waals field for the ligand was computed using gaussian functions with standard deviation of 0.5 times the van der Waals radius of the ligand atoms and weight equal to the van der Waals radius.

Training of the empirical score functions

Five different sets of experimental structures of protein–ligand complexes for which the binding affinity is known were used to train both versions of the score function. One was the set of 50 complexes used by Baxter et al. [1], to validate a flexible docking method using Tabu search. In

Table 1 Variables used in the score functions

Variable	Description	Weight	Weight in "Score"	Weight in "$\Delta G_{binding}$"
hyd_hyd	protein hydrophilicity * ligand hydrophilicity	x_1	−2.976	−2.154
lip_lip	protein lipophilicity * ligand lipophilicity	x_2	−9.187	−8.719
hyd_lip	protein hydrophilicity * ligand lipophilicity	x_3	2.775	3.199
lip_hyd	protein lipophilicity * ligand hydrophilicity	x_4	−4.1	−2.93
lip_hacc	protein lipophilicity * ligand hydrogen acceptor	x_5	5.028	4.035
vdw_vdw	protein van der Waals * ligand van der Waals	x_6	1	−3.464
f(vdw_vdw)	Sigmoid van der Waals term	x_7	100	0

addition to these 50 structures, we used the 100 complexes reported in [18]. These 100 structures have been used by Wang *et al.* to evaluate the performance of eleven score functions for molecular docking [18]. The five peptide structures reported in [25], the 170 protein–ligand complexes used to train the empirical score function SCORE [9] and the 19 complexes reported in [26] were also added to our training set. Redundancies were removed from the training set, and the designed samples from the set made by Baxter *et al.* [1] (DFR4, TSC2 and TMT1) were excluded since they were not real X-ray structures. PDB [27, 28] entries 1FKB and 2XIS were also removed from the training set due to problems with the interpretation of the connection of ligand atoms in the X-ray structures. All together, the training set contained 218 different protein–ligand complexes.

Ions and other co-factors were treated as a part of the receptor when not directly bound to the ligand. If more than one ligand molecule were present in the structure, only one of them was kept. Water molecules (and hydrogen atoms) are not considered in the calculations.

We used Partial Least Squares (PLS) regression in Unscrambler [29] with full cross-validation to fit the parameters for the different terms of the score function. The variables describing the match between the protein and ligand fields were centred and standardised (divided by the standard deviation for each variable) prior to the regression analysis.

The variables used to score the match between the protein and ligand properties are shown in Table 1, together with the parameters used in the two versions of the score function.

Because van der Waals effects are also included implicitly in the variables in Table 1 (described in the Introduction), ''lip_hyd'' has a negative weight, even though one would expect it to have a positive weight based on its physical meaning. Furthermore, the weight for ''vdw_vdw'' does not have the same sign for the two score functions. The reason is that the parameters in ''$\Delta G_{binding}$'' are trained using X-ray structures without steric clashes between the protein and ligand atoms. Therefore only stabilising van der Waals interactions are observed. In ''Score'', the van der Waals term was given the positive weight of 1, since this score function is used to identify a ligand conformation with a minimum of steric clashes, possibly in a non-optimal binding site. Hence, we need a positive steric clash penalty.

The parameters of the sigmoid function added to the van der Waals term of the geometry search score function were chosen based on observed values for the X-ray structures in the training set. Using this function leads to a steep increase in the steric clash penalty at values of vdw_vdw above 0.1, since this is the highest value of vdw_vdw observed in the training set (Fig. 1a). The steric clash penalty reaches a constant value when vdw_vdw passes 0.2. The parameter for this term (having the value 100) was chosen to give this term high weight compared to the other terms. This sigmoid function has no effect for low values of vdw_vdw. Structures for which this sigmoid function has a higher value than 90 are considered so wrongly placed that they are given a high, positive value of 1000 for the score. The steric clash penalty (Eq. 3) is shown as a function of vdw_vdw in Fig. 1b.

To avoid intramolecular steric clashes between ligand functional groups, the Lennard-Jones potential of the ligand, representing the ligand intramolecular energy, is calculated. A threshold value of 500 kcal/mol for this potential is used in the geometry search. All conformations having a higher Lennard-Jones potential than this threshold value were discarded. Several different threshold values were tested (100, 200, 500 and 1000 kcal/mol), but changing this parameter did not have a significant effect on the results.

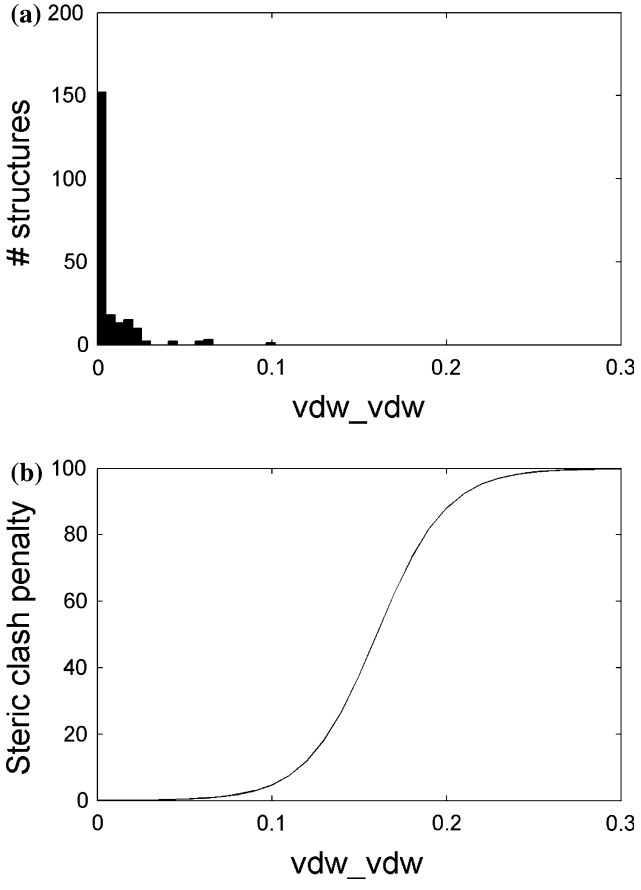

Fig. 1 (**a**) Histogram over the values of the variable vdw_vdw (protein van der Waals field * ligand van der Waals field) for the X-ray structures in the training set. (**b**) The steric clash penalty as a function of vdw_vdw

Docking using Tabu search

The geometry search routines used are the same as those used for the Tabu searching in MOE-Dock [1, 24]. Tabu search is a stochastic searching algorithm that maintains a list of previously visited conformations. These conformations are forbidden (tabu) to future moves. A new conformation is compared to the conformations in the list by calculating the root mean square deviation (RMSD) between the Cartesian coordinates of the new conformation and those of every entry in the list. If the RMSD value is below a specified value, the conformations are considered to be the same, and the move is tabu.

The structures resulting from each Tabu search run (the results from several iterations) are ranked using the same score function as used in the geometry search, and the best ranked conformation is used for the final prediction of the binding affinity.

All scripts are written in Scientific Vector Language (SVL) [24] and can be obtained from the authors upon request.

Testing of the docking performance

The method accuracy was evaluated by calculation of the RMSDs between the coordinates of the docked ligand structures and the X-ray ligand structures, and the PAS-Dock performance was compared to MOE-Dock [1, 24] and AutoDock [4] for all protein–ligand complexes in the training set. The accuracy of the predictions of the free energy of binding for PAS-Dock was also compared to that obtained with MOE-Dock and AutoDock, by comparison to the known, experimental data. The computational efficiency for the three methods was also compared.

With PAS-Dock, ten Tabu search runs of 1000 iterations each were performed, and binding affinities were predicted using the score function that utilises the entire grid to calculate the variables. The docking calculations were started from the X-ray structures.

Also with MOE-Dock [1, 24] ten Tabu search runs of 1000 iterations each were performed, using the molecular mechanics force field MMFF94 [30] and a smooth non-bonded cut-off of 10–12 Å. Prior to the docking analysis with MOE-Dock, hydrogen atoms were added to the X-ray structures, and optimised to an RMS gradient of 1 with the same force field. MOE-Dock calculates the potential energy grids only once, at the beginning of the docking procedure. Hence, protein flexibility is not taken into account in the MOE-Dock calculations.

With AutoDock [4], ten runs of Lamarckian genetic docking were performed, using the largest ring in the ligand structure as a rigid scaffold structure in the calculations. All other torsion angles were allowed to rotate. For ligands without ring structures the first carbon atom in the list was used as the root atom. Default settings were used for the AutoDock calculations, except that the maximum number of torsions allowed was set to 128.

In order to evaluate the suitability of the method for homology modelled protein structures, the robustness against small structural errors was tested. This was done by running molecular dynamics (MD) simulations on the X-ray structures (the protein–ligand complexes in the training set for PAS-Dock) using the AMBER force field [5], and selecting ensembles of five different structures from the MD results for each protein–ligand complex. The RMSDs between the protein structures resulting from the MD simulations and the X-ray structures were approximately 1.0 Å, 1.5 Å, 2.0 Å, 2.5 Å and 3.0 Å, respectively. Docking calculations were started from each of the MD generated structures, and the results were compared for the different structures of each protein. A high degree of stability of the results in the ensemble of structures for each protein is an indication that the docking method is robust against small structural variations, and therefore suitable for use with homology modelled protein structures. The results from PAS-Dock were compared to those obtained with AutoDock. This test was not performed for MOE-Dock, since that method is much more computationally expensive than PAS-Dock and AutoDock.

Results and discussion

Empirical score function

Training of the fast version of the score function resulted in a correlation coefficient from the PLS regression of 0.62 between the score value (from the cross-validation, keeping the van der Waals terms out) and the experimental binding affinity. The van der Waals terms were kept out of the regression analysis because the training set consists only of structures without severe steric clashes. The van der Waals regression coefficient would therefore not be useful in the geometry search, where steric clashes have to be accounted for. We tested the performance of the docking method using several different values for the weight of the van der Waals parameter, and for our training set, the docking method performed best when the van der Waals parameter was given unit weight. The fast score function can not be used for binding affinity prediction, due to the form of the van der Waals terms. Hence, this function is only suitable for finding the best ligand conformations in a geometry search.

The score function "$\Delta G_{binding}$" uses gaussian functions for both protein and ligand property fields and summation over all grid points. This function is more accurate, and

was used to estimate the free energy of binding for the best conformations from each docking run. The correlation coefficient between the predicted binding affinity (from the PLS cross-validation) and the experimental binding affinity for this score function is 0.64. Since we assume that the best conformation resulting from a docking run can be compared to an X-ray structure in the sense that they contain no severe steric clashes with the protein structure, we use the regression coefficient from the PLS regression as the weight for the van der Waals parameter in this score function. Hence, this score function should give realistic estimates for the free energy of binding.

Testing of the docking performance

To test the performance of PAS-Dock, a docking analysis was performed with all protein–ligand complexes in the training set. The same set of structures was used both for training of the score function and for testing of the docking method. However, the wide variety of structures present in the data set combined with extensive cross-validation and a relatively small number of parameters ensure that the score function has not been overfitted. The test will therefore still be valid. The RMSD between the X-ray ligand structures and the ligand structures resulting from the docking analysis was calculated, and the results were compared to the results obtained with MOE-Dock and AutoDock for the same set of structures. The results are given in Table 2 and Fig. 2. The predicted binding affinities are correlated to the experimental binding affinities in Fig. 3. For PAS-Dock, the score function "$\Delta G_{binding}$" was used to predict the binding free energies. The results may be biased, since the test set and the training set is the same set of structures for PAS-Dock. However, except for a couple of complexes, the structures used to calibrate AutoDock and MOE-Dock are also included in the test set.

As shown in Table 2 and Fig. 2, the differences between the three compared docking methods are quite small with respect to the ability to reproduce the ligand X-ray structures. MOE-Dock is slightly more accurate than the other two methods, but it is also the most time consuming of the three methods. With AutoDock, a slightly higher number of structures are modelled with very high accuracy (RMSD < 2 Å between the docked ligand conformations and the ligand X-ray structures) than with PAS-Dock, but with this method we also get a higher number of severe mismatches (RMSD > 3 Å). This indicates that the method works very well for some types of compounds, but also sometimes gives large errors. The shape of the curve in the histogram for PAS-Dock corresponds to what one might expect from a score function that is robust against small structural errors. PAS-Dock gives a reasonable result for most ligands, and relatively few severe mismatches. For initial screening purposes methods that give reasonable, although not perfect, results for all compounds are most suitable, since the optimal binding conformations can be found in later stages of the modelling. The main purpose of virtual screening is to get a fast separation between potentially active and inactive compounds. It is very important not to overlook good drug candidates, and a severe mismatch can sometimes lead to such false negatives. The histograms in Fig. 2 indicate that PAS-Dock performs much better than both AutoDock and MOE-Dock for ligands having high affinity for the protein (shown in red). This is a large advantage for virtual screening applications. PAS-Dock is also more computationally efficient than AutoDock and MOE-Dock. This is important in initial screening of large compound libraries. The level of accuracy of our score function might not be sufficient for a stand-alone docking procedure, but since our docking method is fast, it is well suited for generation of starting conformations for more accurate docking.

Figure 3 indicates that the predictions of the free energy of binding are more accurate for PAS-Dock than for AutoDock and MOE-Dock. The correlation coefficient between the predicted and the experimental binding free energy is 0.59 for PAS-Dock, as compared to 0.26 for AutoDock and 0.43 for MOE-Dock. The regression line for MOE-Dock also has a negative slope, which is the opposite of what is expected in a docking study. For MOE-Dock,

Table 2 Results from comparison of the three docking methods PAS-Dock, AutoDock and MOE-Dock*. The RMSD values refer to the RMSD between the coordinates of the docked ligand conformations and the ligand X-ray structures

	PAS-Dock	AutoDock	MOE-Dock
RMSD < 2 Å (% of docked structures)	47	51	55
2 Å≤RMSD≤3 Å (% of docked structures)	24	14	24
RMSD>3 Å (% of docked structures)	29	35	21
Average calculation time (minutes pr. structure)	3.7	5.15	37

*These docking calculations were done with version 2002.03 of MOE. This version has been shown earlier to use 10 times as much computation time as PAS-Dock, on a different computer than the one used for this study. Since we were unable to run both AutoDock and MOE 2002.03 on the same computer, the value for the average calculation time for MOE-Dock is derived from the result for PAS-Dock. Version 2004.03 of MOE has a lower performance than both PAS-Dock and AutoDock, but uses 6.5 min.pr. structure

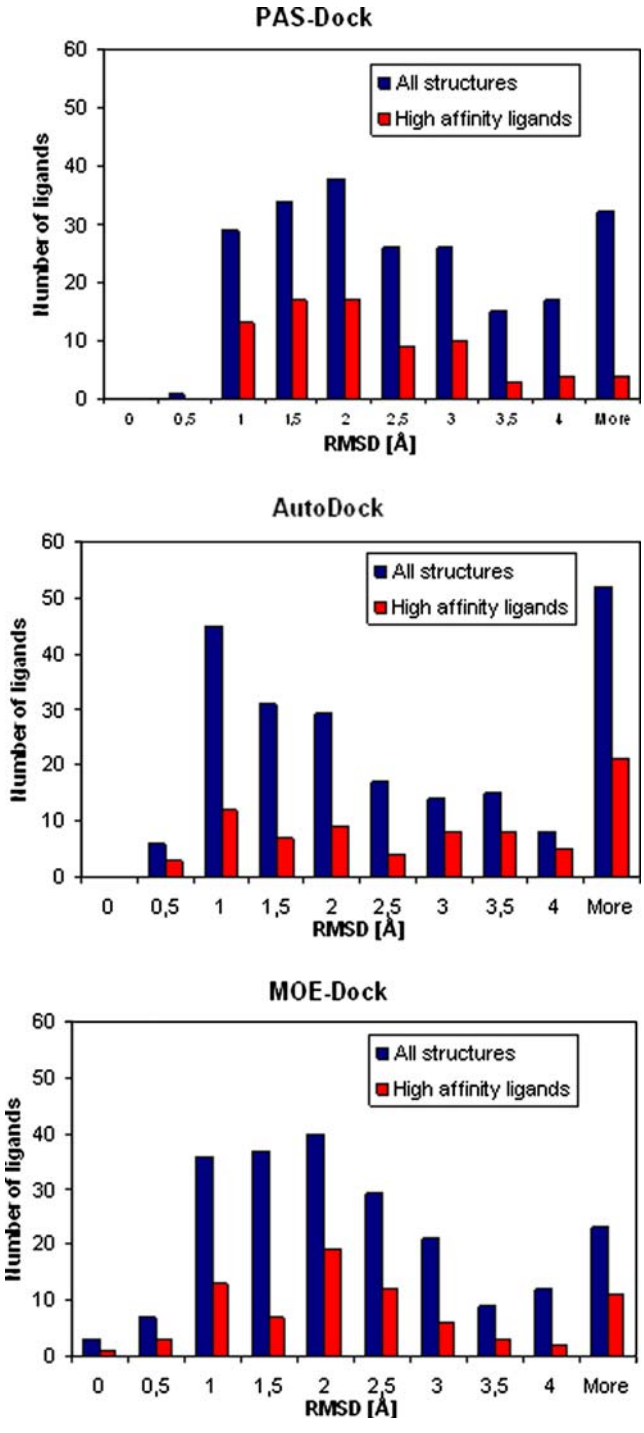

Fig. 2 Histograms over RMSD values between the X-ray ligand structures and the ligand structures resulting from the docking calculations with PAS-Dock, AutoDock and MOE-Dock. The fraction of the complexes having experimental binding affinities below −40 kJ/mol is shown in red

the docking energy equals the sum of the electrostatic and the dispersive protein–ligand interaction energy and the intramolecular energy of the ligand.

The following X-ray structures (from the PDB [27, 28]) contain bond lengths and angles that are not frequently observed, caused by e.g. an ion-containing ring structure: 1L83, 2TMN, 5TMN, 6TMN, 2SNS, 9RUB and 2CTC. This leads to a very high value for the estimated internal energy. They were therefore severe outliers both in the initial analysis and the docking results from PAS-Dock and MOE-Dock, and were kept out of the plots in Fig. 3. These structures were not outliers in the results from AutoDock, but removing them from the data did not affect the correlation coefficient. They are therefore also kept out of Fig. 3b, to make the comparison of the three methods more valid. Fig. 3a shows that 1STP, 6CPA, 7CPA and 1EBG are false negatives for PAS-Dock (shown in the lower, right part of Fig. 3a). 1EBG contains two ion bonds, and therefore binds much stronger than predicted. 1STP binds to the protein through several hydrogen bonds, and since PAS-Dock does not fully represent hydrogen bonds, the binding affinity is underestimated. 6CPA and 7CPA contain hydrophobic groups that protrude towards the solvent. Alpha spheres can only represent ligand atoms bound in a protein cavity. Hence, interactions on the outer surface of the protein are ignored. The contribution of the protruding groups to binding is therefore not included, and the binding affinity is underestimated. As shown in Fig. 3b, 1STP, 6CPA and 7CPA are also outliers in the AutoDock results. In addition, 4HVP and 2RNT were kept out of the correlation plot for AutoDock. The structures 6CPA and 7CPA are severe false negatives in the AutoDock results. In the plot for MOE-Dock, 1ULB was kept out, in addition to the seven structures that were also kept out for the other two methods. 1STP and 7CPA are also outliers in the MOE-Dock results. In addition, 3CPA is an outlier. This might have been caused by the Zn-atom in the ligand in PDB-entry 3CPA.

Robustness against small structural variations

To verify the robustness of the methods towards the small structural errors found in homology models, a test of the stability of the docking results towards structural variations generated by MD simulations was carried out. By using MD generated structures instead of using real homology models, we can control the degree of structural variation in the test set, and compare the results in a more systematic way. The accuracy of homology models depends to a large degree upon the choice of template structures, and it is therefore more difficult to generate homology models with varying deviations from the X-ray structures in a systematic way. The MD generated structures were selected according to their RMSD from the X-ray structure from which the MD simulation was started, representing RMSDs of approximately 1.0 Å, 1.5 Å, 2.0 Å, 2.5 Å and 3.0 Å.

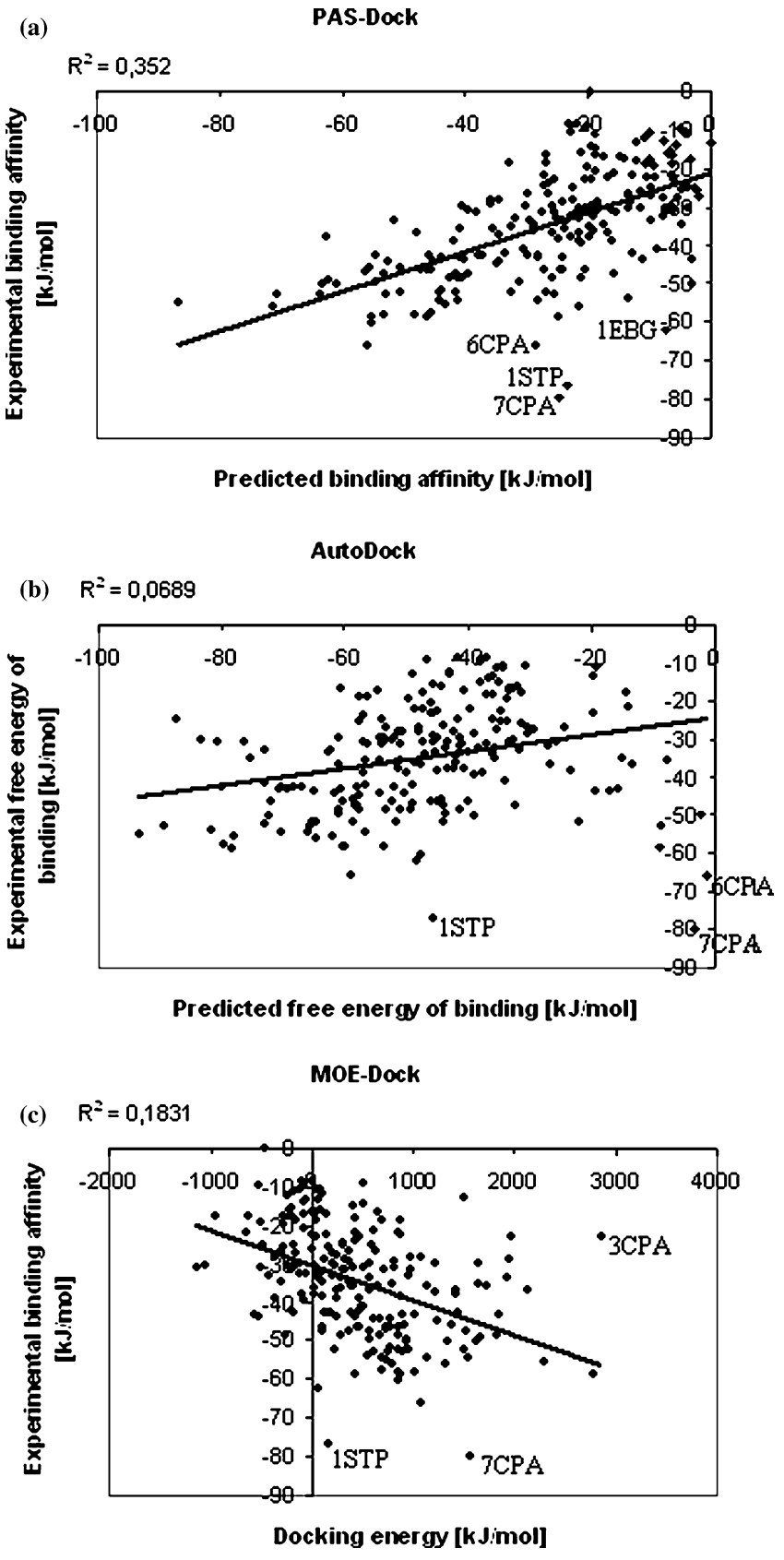

Fig. 3 Predicted binding free energies for the ligand structures resulting from the docking analysis plotted against the experimental binding affinities. (**a**) Results from PAS-Dock (calculated using Eq. 2). (**b**) Results from AutoDock. (**c**) Results from MOE-Dock

The results were averaged over all structures in each category, and the results are shown in Fig. 4. The standard deviations for the docking results over each ensemble of structures of each protein (the five MD snapshot structures) were computed, and the average standard deviations over all ensembles were calculated (Table 3). The results for PAS-Dock and AutoDock are compared in Fig. 5.

The results in Fig. 4a and Table 3 indicate that the predictions of the ligand binding conformations are quite robust against structural variations, both for PAS-Dock and for AutoDock. The difference between the average RMSD values for the two methods is also small, and on average, the two compared docking methods are equally stable. However, as shown in Fig. 5a, AutoDock results in a larger number of binding mode predictions having very low and very high standard deviations, while PAS-Dock gives a higher number of medium standard deviations for the RMSD values than AutoDock. Hence, AutoDock gives more results at the extremes, and PAS-Dock is less exact but more robust. The results from the predictions of the free energy of binding, however, are much more robust against structural variations for PAS-Dock than for AutoDock, since AutoDock gives larger variations in Fig. 4b, and a larger number of structure ensembles have low standard deviations for PAS-Dock than for AutoDock (Fig. 5b and Table 3). This indicates that PAS-Dock is more suitable for use with homology modelled protein structures than AutoDock.

Examples from the docking analysis with PAS-Dock

To illustrate in what cases our gaussian-based docking method is likely to succeed in predicting the binding affinity and the structure of a ligand–receptor complex, we show some examples from the docking analysis performed on our test set.

Three examples where our method succeeds in reproducing the ligand X-ray structure are PDB [27, 28] entries

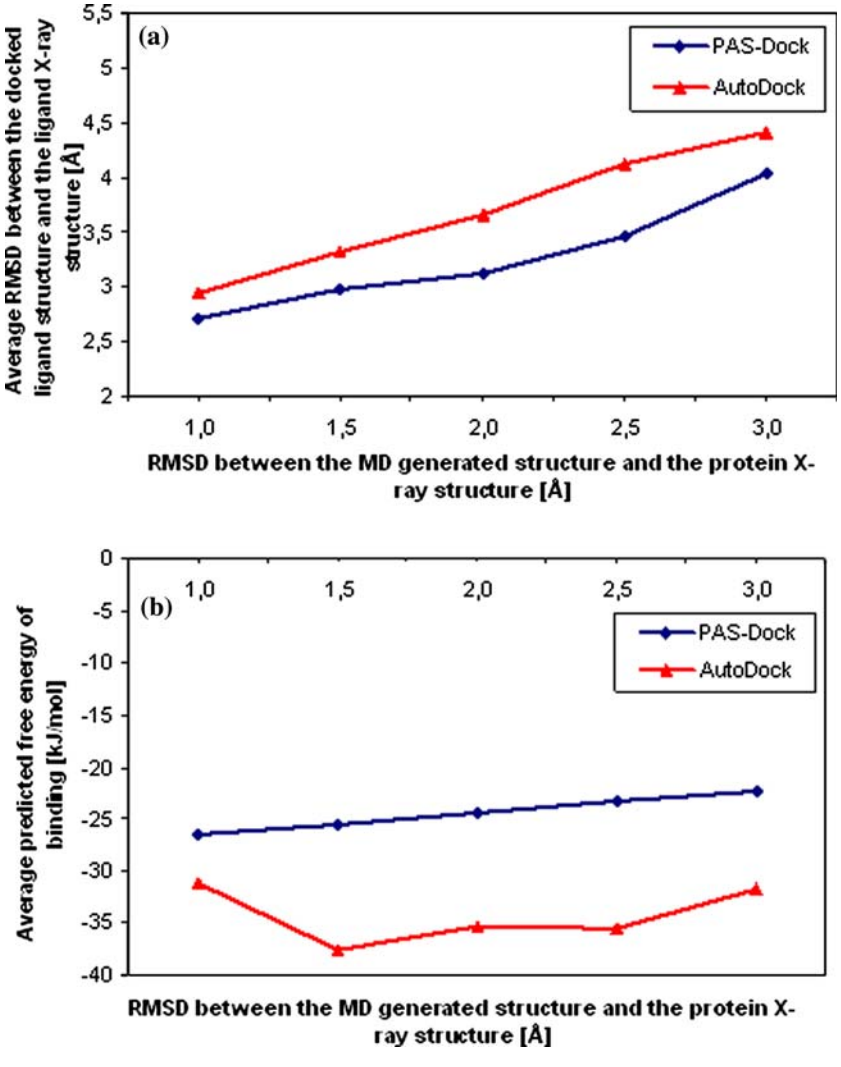

Fig. 4 Average docking results for all structures in each category of MD generated structures. The structures in the five categories deviate from the X-ray structures by RMSDs of 1.0 Å, 1.5 Å, 2.0 Å, 2.5 Å and 3.0 Å, respectively. (**a**) Average RMSDs between docked ligand structures and the ligand X-ray structures. (**b**) Average predicted free energies of binding

Table 3 Standard deviations for the results for the five different structural models for each protein, averaged over all ensembles of MD structures

	PAS-Dock	AutoDock
RMSD [Å]	1.3	1.2
Predicted binding free energy [kJ/mol]	4.8	8.8

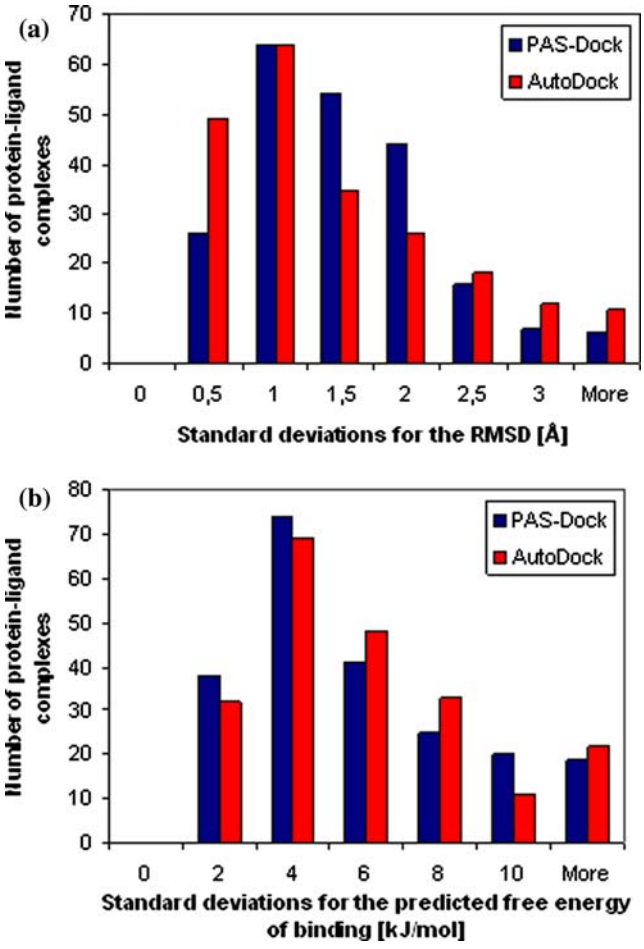

Fig. 5 Histograms over the standard deviations for the docking results in the ensembles of MD generated structures for PAS-Dock and AutoDock. (**a**) Standard deviations for the RMSD values between the docked ligand conformations and the ligand X-ray structures. (**b**) Standard deviations for the predictions of the free energy of binding

1E96, 1ETS and 1HVI. The RMSD values between the docked conformations and the ligand X-ray structures are 0.46 Å, 0.57 Å and 0.93 Å, respectively. Figs. 6–8 show the docking results for these three PDB entries.

One example where our docking calculations resulted in a high RMSD value between the X-ray structure and the docked structure is PDB entry 1TET. Figure 9 shows the docked ligand conformation together with the X-ray structure of the complex.

Figure 9 shows that for 1TET we have very good match between the docked conformation and the X-ray ligand conformation in the region binding to the protein, while the part of the ligand pointing towards the solvent is flipped outwards. The reason is that our algorithm considers only cavities in the protein structure, since alpha spheres are placed in protein cavities. This leads to a very high RMSD value, even though our docking calculations succeeded for the part of the ligand that is relevant for binding to the protein.

Our docking calculations also failed to reproduce the ligand X-ray structure in PDB entry 1CPS. As shown in Fig. 10 (left), the structure of the ligand in PDB entry 1CPS is flipped 180° in the docked conformation. One possible reason is that our algorithm does not represent hydrogen bonds fully. As seen from Fig. 10 (right), the ligand in PDB entry 1CPS is stabilised in the bound conformation by several hydrogen bonds.

The result from the docking analysis of PDB entry 1DBK (Fig. 11) demonstrates that our algorithm predicts hydrophobic interactions quite well. The RMSD value between the docked ligand conformation and the ligand X-ray structure is very high for 1DBK. In the same way as for 1CPS, the ligand structure is flipped 180°. However, Fig. 11 shows that the ligand structure is almost symmetric, and our algorithm succeeds in placing the hydrophobic ring structures. Since the ligand is flipped, an almost correct ligand placement leads to a very high RMSD value (5.86 Å).

The results from our docking analysis show that our method predicts hydrophobic interactions better than hydrophilic interactions. One reason is that hydrophilic

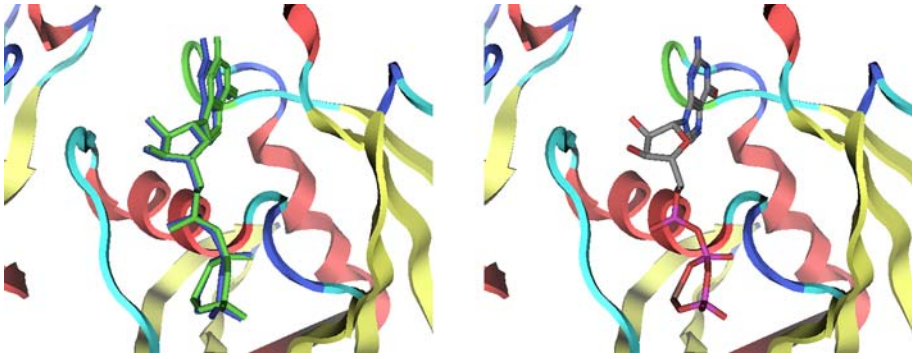

Fig. 6 Left: Result from docking of PDB entry 1E96. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 0.46 Å. Right: The X-ray structure from PDB entry 1E96

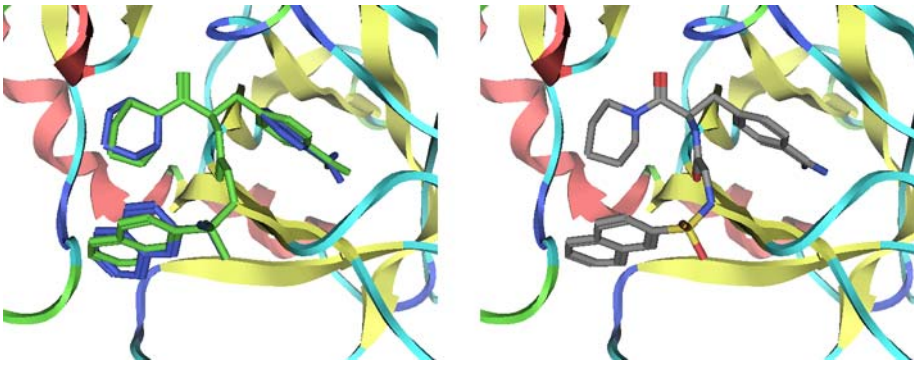

Fig. 7 Left: Result from docking of PDB entry 1ETS. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 0.57 Å. Right: The X-ray structure from PDB entry 1ETS

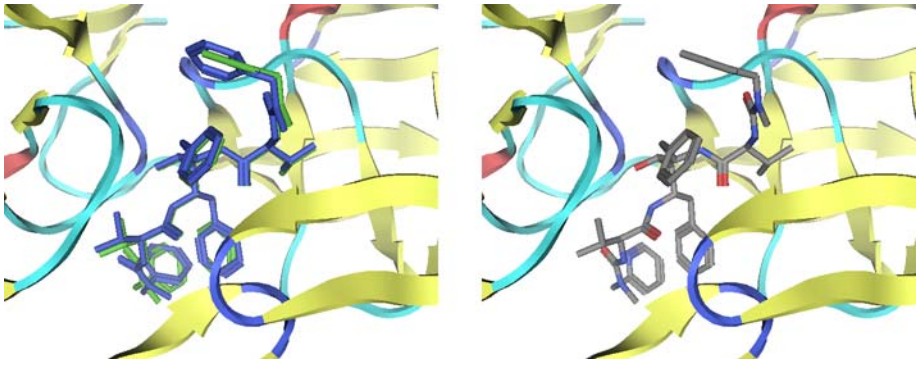

Fig. 8 Left: Result from docking of PDB entry 1HVI. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 0.93 Å. Right: The X-ray structure from PDB entry 1HVI

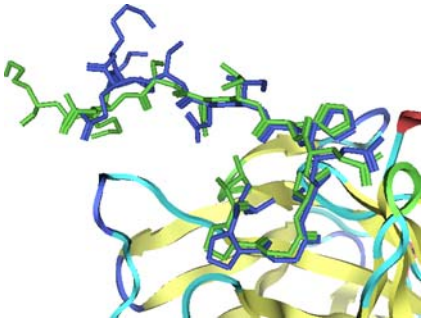

Fig. 9 Result from docking of PDB entry 1TET. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 3.70 Å

interactions are more direction-specific than hydrophobic interactions. The ligands in PDB [27, 28] entries 1ETS and 1HVI both contain several hydrophobic groups (Figs. 7 and 8). Since our method is independent of placement of hydrogen atoms and estimation of partial charges, we do not fully account for the formation of hydrogen bonds. Hydrogen bond formation is very dependent on the direction in which the hydrogen atom points. We are planning to account for possible hydrogen bond formation and to include variables representing hydrogen donors and acceptors on the protein in a future version of the score function. Our method is not suitable in cases where the ligand makes an ion bond to the protein. However, our method succeeds to a high degree in placing the hydrophobic parts of the ligands. The results also show that our method works best in cases where the ligand is bound in a well-defined cavity of the protein. This is not surprising since alpha spheres only describe protein cavities, not the outer surface of the protein. Hence, our gaussian-based docking method is most likely to succeed in cases where the protein has a well-defined binding pocket and the ligand is not bound to the protein by an ion bond. The docking method presented here gives a reasonable reproduction of the ligand conformations found in X-ray structures in most cases, and the speed of the calculations makes it a useful tool for initial drug candidate screening. The use of gaussian functions to describe the molecular properties makes this docking method suitable for use with homology modelled protein structures.

Other docking methods have also been reported that utilise gaussian functions to describe protein–ligand interactions. The method reported by McGann *et al.* [31] only accounts for shape, and is trained on a much smaller set of protein–ligand complexes than our docking method. By including hydrophilicity and hydrophobicity in addition to van der Waals effects, and by using a larger training set, we aim to describe protein–ligand interactions better. Another gaussian-based method that also accounts for hydrophilicity and hydrophobicity in addition to shape has been reported by Schafferhans and Klebe [32]. As for our method, gaussian functions are used to represent the physicochemical properties of the receptor and the ligand, and the

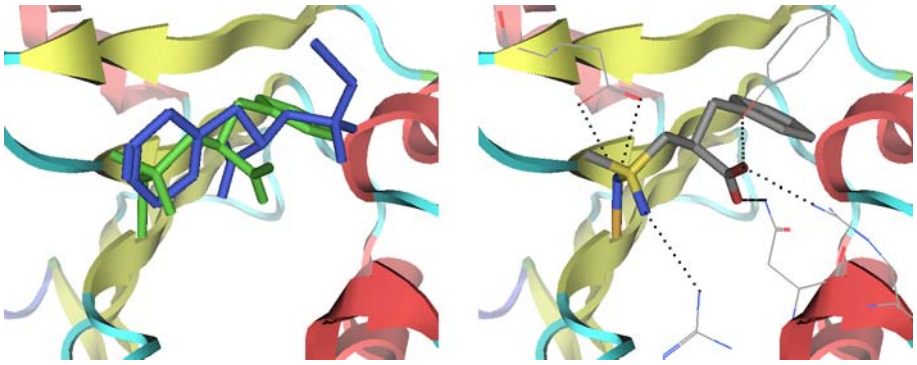

Fig. 10 Left: Result from docking of PDB entry 1CPS. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 6.18 Å. Right: The X-ray structure from PDB entry 1CPS. Possible hydrogen bonds to the protein are shown (drawn using MOE [24])

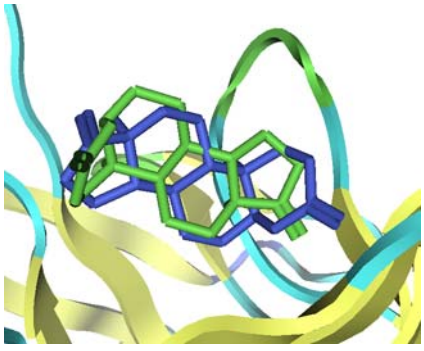

Fig. 11 Result from docking of PDB entry 1DBK. The ligand conformation from the X-ray structure is rendered in green, while the docked conformation is rendered in blue. The RMSD value between the docked conformation and the ligand X-ray structure is 5.86 Å

overlap between the functional descriptions of the receptor binding site and the ligand is optimised. This method uses interaction sites identified by the *de novo* ligand design program LUDI [33, 34] to generate a description of the protein binding site. As the method developed by McGann et al. [31], this docking method is also trained on a relatively small set of complexes. Hence, this method is sensitive to deviations between the target system and the training set. It is also indicated that this method predicts both hydrophobic interactions and electrostatics insufficiently [32]. The results indicate that gaussian functions are too soft to model electrostatics sufficiently. In contrast to our docking method, this method is dependent on assignment of charges and protonation states. This information is not trivial to generate, especially not for protein models with potential inaccuracies, such as homology models. One advantage with this method is that it can take several different homology models into account simultaneously by an averaging of the property densities of the models. This increases the robustness of this method.

Because of the simplified description of the protein–ligand interactions, the accuracy of our new score function can not be compared to that of score functions that take e.g. electrostatics into account. However, the speed of our calculations makes this method an effective tool for prescreening for virtual drug design. We use a fast and less detailed version of the score function in the geometry search, while a more detailed score function is used for the final prediction of the binding free energies. The use of a two-level scoring makes PAS-Dock computationally efficient. In virtual screening, the purpose is to identify a set of promising drug candidates from a large collection of ligand structures. Hence, the binding affinity has to be estimated for a large number of structures. This makes computational efficiency an important factor to consider. In virtual screening, it is not always necessary to predict the absolutely correct binding mode for all ligands, or predict the binding affinity with a high level of accuracy. It is most important to effectively separate the active compounds from the non-active ones. The correct ranking of the promising drug candidates and the correct binding conformations can be found with more accurate and time consuming methods, once the number of structures to consider has been reduced.

Conclusion

A new score function for virtual library screening is introduced, that use gaussian functions to describe protein–ligand interactions. This score function accounts for hydrophilicity and lipophilicity of the protein, and hydrophilicity, lipophilicity, hydrogen donor and acceptor potential for the ligand. In addition, van der Waals effects are taken into account. Neither hydrogens nor partial charges are considered. The use of a two-level scoring makes PAS-Dock computationally efficient, and the use of gaussian functions makes the score function relatively robust against small structural errors. The gaussian functions also partly compensate for the lack of protein flexibility in the docking calculations. A comparison of PAS-Dock to two other dock-

ing methods shows that PAS-Dock is more computationally efficient than both AutoDock and MOE-Dock, and gives a better prediction of the free energies of binding. PAS-Dock is also more robust against structural variations than the other two docking methods. The accuracy of our score function can not be compared to that of more complex score functions that account for e.g. hydrogens, partial charges and electrostatics, but the combination of speed, reasonable accuracy and robustness makes our method more suitable for use in pre-screening.

Acknowledgement We thank The Norwegian Research Council for financial support.

References

1. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD (1998) Prot Struct Func Gen 33:367
2. Halperin I, Ma B, Wolfson H, Nussinov R (2002) Prot Struct Func Gen 47:409
3. Böhm H-J, Stahl M, Lipkowitz KB, Boyd DB (eds) (2002) *Reviews in computational chemistry*, The use of scoring functions in drug discovery applications. Wiley-VCH, New York, 18, pp 41–87
4. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639
5. Weiner SJ, Kollman PA, Case DA (1984) J Am Chem Soc 106:765
6. Massova I, Kollman PA (1999) J Am Chem Soc 121:8133
7. Nicholls A, Honig B (1995) Science 268:1144
8. Pearlman DA, Charifson PA (2001) J Med Chem 44:502
9. Wang R, Liu L, Lai L, Tang Y (1998) J Mol Model 4:379
10. Gelhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Fogel LJ, Freer ST (1995) Chem Biol 2:317
11. Gelhaar DK, Bouzida, D, Rejto PA, Parrill, L, Reddy, MR (eds) (1999) In *Rational drug design: novel methodology and practical applications*. American Chemical Society, Washington, DC, pp 292–311
12. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) J Comput Aided Mol Des 11:425
13. Wang R, Lai L, Wang S (2002) J Comput Aided Mol Des 16:11
14. Muegge I, Martin YC (1999) J Med Chem 42:791
15. Muegge I (2000) Perspect Drug Discovery Des 20:99
16. Muegge I (2001) J Comput Chem 22:418
17. Gohlke H, Hendlich M, Klebe G (2000) J Mol Biol 295:337
18. Wang R, Lu Y, Wang S (2003) J Med Chem 46:2287
19. Wieman H, Tøndel K, Anderssen E, Drabløs F (2004) Mini-Rev Med Chem 4:793
20. Tøndel K, Anderssen E, Drabløs F (2002) J Comput Aided Mol Des 16:831
21. Klebe G, Abraham U, Mietzner T (1994) J Med Chem 37:4130
22. Liang J, Edelsbrunner H, Woodward C (1998) Protein Sci 7:1884
23. Sheridan RP, Bush BL (1993) J Chem Info Comp Sci 33:756
24. Molecular Operating Environment™ (2002) Version 2002.03, Chemical Computing Group, Inc.
25. Rognan D, Lauemøller SL, Holm A, Buus S, Tschinke V (1999) J Med Chem 42:4650
26. Rarey M, Kramer B, Lengauer T, Klebe G (1996) J Mol Biol 261:470
27. The RCSB Protein Data Bank, http://www.rcsb.org/pdb/
28. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28:235
29. The Unscrambler® (2000) Version 7.6 SR-1, CAMO ASA
30. Halgren TA (1996) J Comp Chem 17:490
31. McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK (2003) Biopolymers 68:76
32. Schafferhans A, Klebe G (2001) J Mol Biol 307:407
33. Böhm H-J (1992) J Comput Aided Mol Des 6:593
34. Böhm H-J (1992) J Comput Aided Mol Des 6:61