J. Mol. Biol. (1997) **266**, 424–440

JMB

Molecular Docking to Ensembles of Protein Structures

Ronald M. A. Knegtel, Irwin D. Kuntz and C. M. Oshiro*

Department of Pharmaceutical Chemistry School of Pharmacy University of California, San Francisco, CA, 94143-0446, USA

**Corresponding author*

Until recently, applications of molecular docking assumed that the macromolecular receptor exists in a single, rigid conformation. However, structural studies involving different ligands bound to the same target biomolecule frequently reveal modest but significant conformational changes in the target. In this paper, two related methods for molecular docking are described that utilize information on conformational variability from ensembles of experimental receptor structures. One method combines the information into an ''energy-weighted average'' of the interaction energy between a ligand and each receptor structure. The other method performs the averaging on a structural level, producing a ''geometry-weighted average'' of the inter-molecular force field score used in DOCK 3.5. Both methods have been applied in docking small molecules to ensembles of crystal and solution structures, and we show that experimentally determined binding orientations and computed energies of known ligands can be reproduced accurately. The use of composite grids, when conformationally different protein structures are available, yields an improvement in computational speed for database searches in proportion to the number of structures.

Keywords: DOCK; receptor flexibility; ensemble; structure-based drug design

Introduction

The discovery of new drugs has evolved from a random process of screening natural products to a suite of sophisticated procedures that include components from computational and structural chemistry. The availability of high-resolution data on enzymes involved in critical metabolic pathways has triggered the development of techniques utilizing such data in the quest for novel compounds of therapeutical relevance. Using this information, computer-based approaches help identify or design ligands that possess good steric and chemical complementarity to various sites in the macromolecular target. This process is often referred to as ''structure-based design'' (Kuntz, 1992).

The many different algorithms for structure-based design can be divided into roughly two classes: *de novo* design, which builds ligands tailored to fit the target, and docking, which searches for existing compounds with good complementarity to the target. In both these paradigms, the enzyme or receptor has traditionally been treated as a rigid body and only one conformation of the enzyme is considered. (Examples of *de novo* design include Lewis (Lewis, 1992) and Miranker (Miranker & Karplus, 1995) and the program LUDI (Böhm, 1992a,b); examples of molecular docking include the works of Kuntz *et al*. (1982), Nussinov *et al*. (Lin, 1994; Norel *et al*., 1994), and Bacon & Moult (1992).)

While increasing attention is being paid to exploring the conformational space of putative ligands in molecular docking (Leach & Kuntz, 1992; Mizutani *et al*., 1994; Clark & Ajay, 1995; Judson *et al*., 1995; Oshiro *et al*., 1995; Welch *et al*., 1996; Rarey *et al*., 1996), relatively little effort has been expended on the conformational state of the receptor (Leach, 1994; Jones *et al*., 1995). Obviously, using a single protein conformation ignores important dynamic aspects of protein–ligand binding. In particular, ''induced fit'' effects (Koshland, 1958; Jorgensen, 1991) are ignored. The general problem, however, of docking or designing fully flexible ligands to match fully flexible receptors remains daunting. Free energy calculations, using flexible enzyme and ligand structures and explicit solvent molecules, can reproduce ligand binding affinities and structures (see, for example, Kollman, 1994). Thermodynamic quantities must, however, be extracted from appropriately weighted ensembles of a system and adequate sampling of such configurations

Abbreviations used: rmsd, root-mean-square deviation; cpu, central processing unit; PCB, polychlorobiphenyl; PDB, Protein Data Bank.

0022–2836/97/070424–17 $25.00/0/mb960776

is computationally demanding (McCammon & Harvey, 1987; Kollman, 1994).

Rather than exploring the receptor conformational space by theoretical means, one could instead make use of the increasingly available experimental data on protein structure and flexibility. This would focus computational efforts on a more limited, but still important, aspect of the problem, where partial information about the receptor conformational status is available from either multiple crystallographic or NMR solution structure determinations. Both types of experiments produce collections of structures, but with different physical interpretations.

A set of related crystal structures are ''snap shots'' of a dominant conformation perturbed by different ligands, different crystallization conditions, point mutations, etc. The structural changes observed in such a collection of receptor structures include regions capable of accommodating differently shaped ligands as well as areas of an induced fit of the ligand. Different crystal forms of the same protein–ligand complex provide conformational changes due to different crystal packing environments.

On the other hand, the result of a structure determination by means of high-resolution NMR spectroscopy is usually not a single structure, but rather an ensemble of structures, all nominally equally in agreement with nuclear Overhauser effect and *J*-coupling data (Wüthrich, 1986). Although it is possible to calculate an energy-minimized average structure, an ensemble or a subset of conformers may provide a more accurate representation of the protein structure (Sutcliffe, 1993; Bonvin & Brünger, 1995). Structural variability in ensembles of solution structures can be due to true conformational flexibility as reflected in decreased nuclear relaxation rates or a lack of sufficient experimental data. Truly flexible regions within a binding site may be relevant to the plasticity of the fit between receptor and ligand. In either case, however, parts of the receptor structure will be ill-defined and therefore cannot be represented accurately by a single structure.

With the availability of multiple crystal or solution structures, a method that allows for the use of an entire ensemble in molecular docking, rather than a single structure, could prove to be advantageous in cases where conformational changes in the receptor are expected. Such a method would be useful for molecular design purposes and, in particular, computational database screening. The general issue for both kinds of ensembles is the same: what procedures best utilize the additional, but still incomplete, information contained in a set of related structures?

Here, methodology is presented for docking ligands to a suite of target structures. The critical constructs are weighting protocols for the intermolecular force field used to measure ligand–target interactions. Force field terms are combined and stored on a single ''scoring grid'' that is used in DOCK 3.5 (Kuntz *et al*., 1982; Meng *et al*., 1992) to evaluate ligand orientations. The weighting of the force fields was developed to meet two basic criteria. First, the correct orientations of the ligand(s) should have a favorable score. Some loss of accuracy can be expected when a ligand orientation is scored against an ensemble of receptor structures, only one of which is adapted to that specific ligand. Nevertheless, a ligand that scores well when docked to its native receptor structure should still do so when scored against the ensemble. Secondly, the evaluation should still be efficient enough to allow for fast screening of large databases for novel lead compounds to be used in the drug design process.

The most straightforward method one might consider is to evaluate the ligand–receptor binding energies with each structure in the ensemble, then use a Boltzmann-weighted energy average at each grid point as the score. With more than a few enzyme conformations, however, evaluating the energy of every ligand orientation with every receptor conformation becomes prohibitive in terms of both computer memory and computational time. Furthermore, this kind of sum cannot be implemented on a single scoring grid, since such a weighted sum depends upon all the ligand atom positions, which cannot be precalculated. One could calculate a score by summing the individual interaction energies, giving each energy equal weight. Such ''mean-value'' energy is dominated by a few bad contacts of the ligand with the receptor. Alternatively, one could calculate an energy-minimized positionally averaged structure from the ensemble of structures. Such an average structure, however, does not adequately take into account any large conformational variability present and known ligands which dock to some members of the ensemble, will not dock to the average structure.

Since the direct method presents difficulties, more tractable approaches were sought. Two general types of averaging appear to be applicable. One approach, referred to as the ''energy-weighted average'', attempts to construct an appropriately averaged ligand–protein interaction energy. Another procedure, referred to as the ''geometry-weighted average'', constructs an average score by considering the variance in atom positions. The first method calculates, for every atom of the macromolecule in every structure, the van der Waals and Coulombic potential energy factors. For each atom, a weighted potential is then calculated, averaging over all structures. The contribution to the potential of each atom is weighted to make the final energy resemble a Boltzmann energy-weighted sum, with the simplification of approximating the free energy with the interaction energy. The actual (unnormalized) weight assigned to each atom from each structure is a sigmoidal function of the distance from the receptor atom to the ligand atom or grid point. This function approaches zero at short distances and increases to unity at longer

distances. As a result, only the attractive potential is considered when there is a mixture of both attractive and repulsive contributions from an atom in different structures; i.e. repulsive potentials have a negligible contribution to the average in a manner analogous to a true Boltzmann energy-weighted sum. Repulsive potentials are only included when the potentials from atoms from all the structures are repulsive. In this manner, a ligand atom can have a reasonable attractive interaction energy, rather than a large repulsive energy, when positioned close to a receptor atom of only one of the several receptor structures. Such an averaging method allows for small variations in local geometry and attempts to model a Boltzmann-like sum. We refer to this method of averaging as the ''energy-weighted average''.

The second method performs the averaging on the structural level and calculates for every atom of the macromolecule, a mean position, averaged over all structures, and its corresponding variance. When averaging the position of a particle, the nature of its motion needs to be considered in order to validate the averaging. For small harmonic motions, the average position is a physically relevant quantity, but for larger anharmonic motions, possibly involving multiple minima, the entire distribution of atom positions offers a more accurate description. As depicted in Figure 1, a ligand binding site can be divided in structurally well-determined regions where the atoms of different structures in the ensemble overlay with small root-mean-square deviations (rmsd values) and less conformationally restricted regions where the same atom can occupy very different positions in space. Atomic motions in these two categories can be considered to represent harmonic and anharmonic oscillators, respectively. For generating position-weighted scoring grids, atoms are assumed to be disordered if their positional standard deviation is above a user-defined threshold. Such atoms are treated as volumeless and copies of all positions in the ensemble are used with fractional occupancy to generate scoring grids. For atoms with a variance less than the threshold, the average position is used. This allows larger ligands to bind by occupying flexible portions of the binding site, without detailed analysis of conformational changes in the protein induced by the ligand. We refer to this method of averaging as geometry-weighted averaging.

In Methods we describe both types of averaging in detail. In Results we evaluate the usefulness of both types of composite grids in regenerating known crystal ligand orientations, using several crystal structures of HIV protease and retinol binding protein as well as an ensemble of NMR structures of ras p21 and uteroglobin. We also evaluate the ability of both types of composite grids to identify known ligands from a small database of compounds. Finally, a comparison of the two methods and a discussion of their limitations is given.

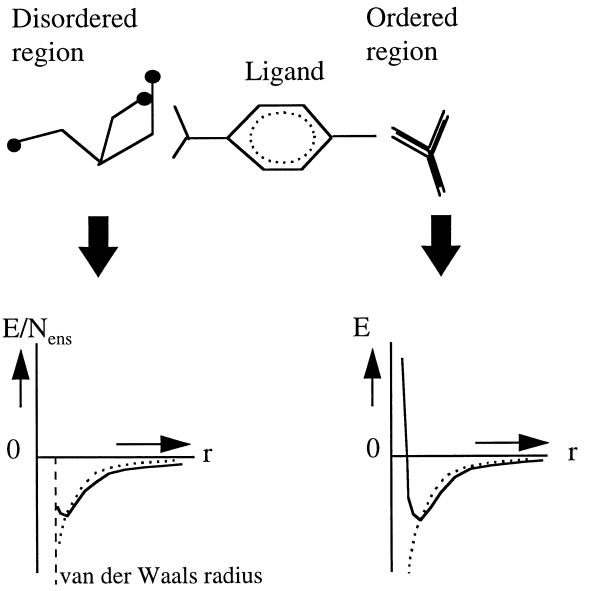


Figure 1. Schematic depiction of how binding site atoms exhibiting different degrees of positional disorder are represented in a single scoring grid. For atoms in static regions of the binding site (on the right), the average position is used to calculate AMBER force field potential E. Atoms that have positional standard deviations above a user-defined threshold (on the left) are represented by all copies of those atoms in N_{ens} protein structures. Their van der Waals volume and AMBER potential are set to zero within a user-defined radius for either polar or non-polar atoms. The remaining attractive portion of the potential E is scaled by the number of copies N_{ens}.

Results

Choice of test cases

Test cases were chosen from a limited set of systems where data on multiple target structures interacting non-covalently with a variety of ligands are publicly available. Two cases, HIV protease and ras p21 protein, were selected for their therapeutic interest. HIV protease is a realistic and well-studied target for structure-based drug design and many crystal structures of complexes with different inhibitors have been reported. The inhibitors used in our studies are generally extended peptide-like molecules, except for the thioketal haloperidol (see Figure 2), and are almost completely enclosed by the fairly hydrophobic active site of the enzyme. The disorder observed in the HIV protease active site is mainly due to the occurrence of different side-chain conformers. The ras p21 protein belongs to a class of oncogene proteins whose mode of action is actively studied. It cycles between an active and an inactive state as a function of ligand bind-

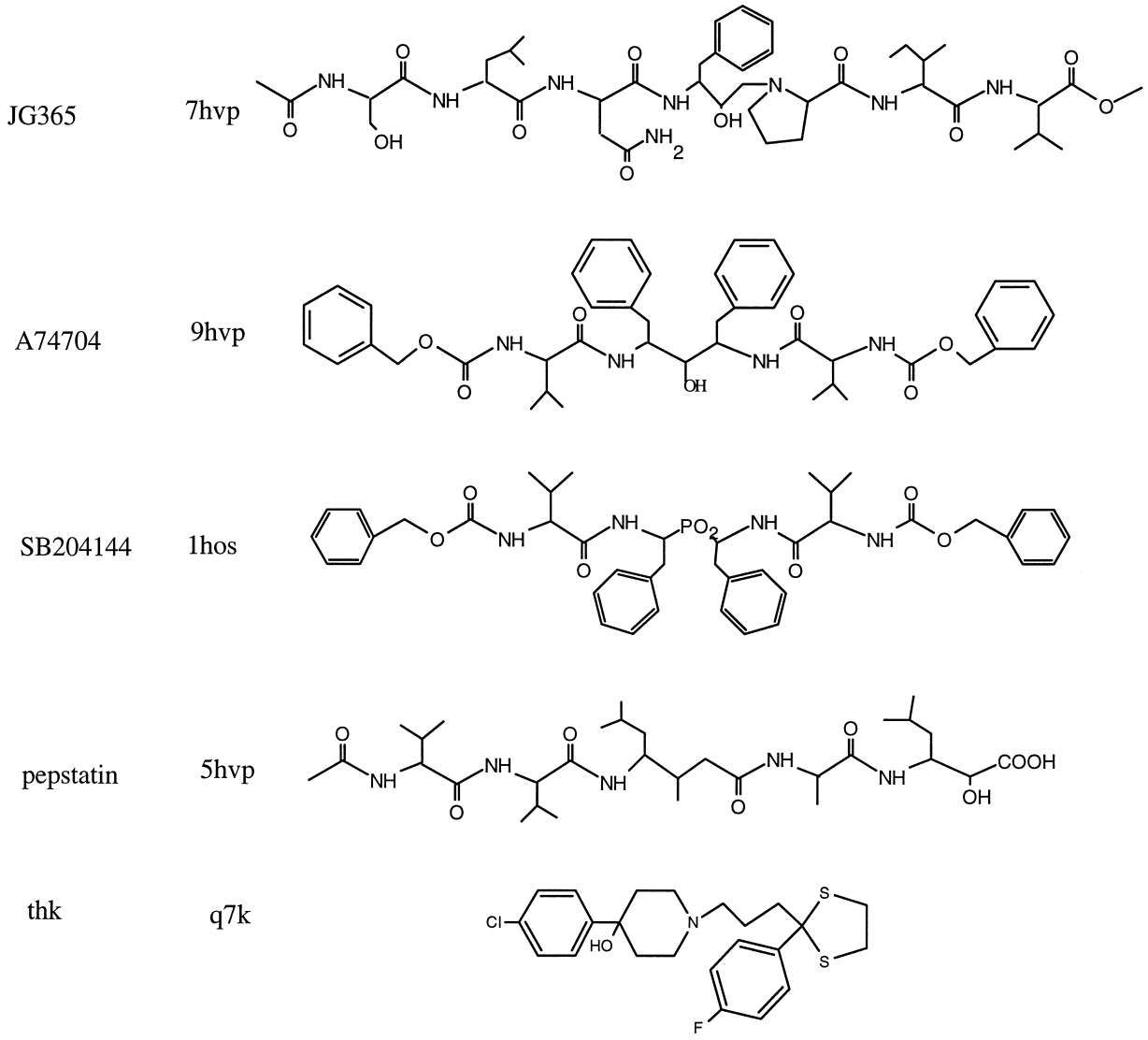


Figure 2. Schematic representation of the chemical structures of HIV protease inhibitors used in docking studies. Ligands and the PDB entry codes of the enzyme–inhibitor complexes listed (see Methods).

ing (Milburn *et al.*, 1992). In contrast to HIV protease, the ras p21 solution structure has a charged and open binding site where the observed disorder involves both side-chains and loop regions. The other two test cases are less well studied targets, but they are also of pharmacological and structural interest. The uteroglobin solution structure has a well determined backbone conformation and a hydrophobic but still solvent-accessible binding site. Random conformational variability is observed for several side-chains. In this case, few ligands are known and our main focus was on a comparison of the behavior of the energy-minimized average structure with the ensemble in molecular docking using structures determined by NMR. Finally, in the ensemble of retinol-binding protein crystal structures the structural variability is mainly due to conformational changes involving two side-chains. The ensemble is, however, dominated by the less permissive conformations. For this test case, the ability of our methods to reflect relatively small conformational differences in the average scoring grid can be evaluated. A more elaborate discussion of the test cases is given in Methods.

Table 1. Force field scores (kcal/mol) of HIV protease inhibitors docked to individual enzyme structures

	7hvp	9hvp	1hos	5hvp	q7k
JG365	**−65**	+27	−52	−53	−3
A74704	−14	**−57**	−59	−11	−25
SB204144	−28	−25	**−49**	+56	+30
Pepstatin	−37	−15	−42	**−30**	−24
thk	−35	−27	−29	−30	**−31**

All values listed are based on single docking calculations. Diagonal elements are values for the ligand docked to its co-crystallized enzyme structure (printed in bold face).

Table 2. Root-mean-square deviations (Å) and force field scores (kcal/mol) of HIV protease inhibitors docked to an ensemble of enzyme structures

	Individual structures		Ensemble G-weighted		Ensemble E-weighted	
	rmsd	Energy	rmsd	Energy	rmsd	Energy
JG365	0.3	−65	0.6	−63	0.8	−44
A74704	0.6	−57	0.4[1]	−50	0.7	−47
SB204144	0.5	−49	0.6	−55	0.8[a]	−43
Pepstatin	0.9	−30	0.4	−57	0.8	−41
thk	0.4	−31	1.1	−35	4.0[b]	−29

The standard deviation threshold above which 65 receptor atoms were treated as flexible was 0.5 Å for the geometry-weighted average grids. The rmsd per non-hydrogen atom located in the box used for the generationof scoring grids was 0.49 Å. rmsd values and energy values are are listed for energy (E)-weighted and geometry(G)-weighted composite grids. Ligands are depicted in Figure 4.

[a]rmsd values after ligand was rotated by 180 degrees along the C2 axis of the HIV protease dimer.

[b]Ligand translated in binding pocket caused by its interaction with a water molecule instead of a chlorine ion. The correct orientation was found with 0.9 Å rmsd and a −27 kcal/mol force field energy.

Use of composite grids in single ligand docking

HIV protease test case

The accuracy of DOCK 3.5 in regenerating the ligand orientation observed in crystal structures of HIV protease using composite scoring grids is compared with that using the scoring grids of the five individual enzyme structures (see Methods) in Tables 1 and 2. Table 1 shows that not every ligand can be docked successfully to every enzyme structure. Several combinations of ligand and enzyme structure yield high interaction energies that would prevent these inhibitors from being identified in docking searches. When either composite ensemble-based grid is used, the crystal ligand orientations and energies can be reproduced reasonably well (Table 2). Both composite grids find orientations within 1 Å of the crystal orientation for almost all ligands within 10 to 100 CPU seconds/ligand. The thk inhibitor was translated by 4 Å in the binding site when using the energy-weighted composite grid. This displacement is caused by the replacement of the chlorine ion, observed in the crystal structure, by a water molecule in order to make this enzyme structure equivalent to the other HIV protease structures (see Methods). When the chlorine ion is reintroduced at its original position, the native ligand orientation is retrieved (Rutenber *et al.*, 1993).

The energies obtained with individual grids compare reasonably well with those obtained with the composite grids although the scores obtained with the energy-weighted grid are not as favorable. An exception is pepstatin, which scores much better when docked using the ensemble-based grids instead of to its own (5hvp) structure (see Table 2). In the experimental structure, steric clashes are present between the inhibitor and the enzyme, yielding a force field score of +140 kcal/mol. Docking and simplex minimization of the ligand to the native enzyme structure yields a more favorable energy value (see Table 2). The ensemble-based grid abolishes the steric hindrance between enzyme and inhibitor, allowing the inhibitor to improve its interaction energy with the enzyme. For the energy-weighted average, a grid spacing of 0.2 to 0.25 Å was found to give results closer to those of the individual grids (data not shown).

Table 3. Root-mean-square deviations (in Å) and force field scores (kcal/mol) of GDP and GTP analogs docked to the ras p21 GDP-bound solution structure

	20 NMR structures (ranges)		Minimized Average		Ensemble G-weighted		Ensemble E-weighted	
	rmsd	Energy	rmsd	Energy	rmsd	Energy	rmsd	Energy
GCP	1.0…10.0	−57… − 23	1.5	−52	1.4	−52	1.2	−39
GDP	0.5…9.2	−55… − 27	0.8	−48	0.6	−49	0.6	−39
GNP	1.6…10.7	−50… − 22	1.9	−47	1.6	−48	1.6	−32
GNQ	1.2…11.1	−51… − 20	**10.4**	−34	0.4	−54	1.2	−41
GNR	0.7…14.2	−61… − 21	**9.9**	−38	0.4	−48	0.5	−36

The standard deviation threshold above which receptor atoms were treated as flexible (181 heavy atoms in total and rmsd/atom of 0.66 Å) was 0.5 Å. rmsd and energy values are are listed for energy (E)-weighted and geometry (G)-weighted composite grids. The ensemble consisted of 20 solution structures. See the text for ligand abbreviations. The rmsd values of ligands that could not be docked successfully are printed in bold face.

ras p21 protein test case

The results of docking GDP and GTP analoges (see Figure 2) to ras p21 solution structures using energy and geometry-weighted averaging are listed in Table 3. No single structure of the ensemble generates orientations closest to the native orientation for all ligands. Also the minimized average structure does not allow for the successful docking of all GTP analogues. Using composite scoring grids, however, the GTP analogs can be docked to yield near-native orientations with deviations of 1.6 Å or less from the crystal structure. These deviations are due to differences in the placement of the charged phosphate group relative to the Mg ion. Relatively large changes in the electrostatic interaction can occur with relatively small changes in position, even though the formal charges have been reduced (see Methods). For the computed energies the same trend is observed as with HIV protease where the energy-weighted grids display a less broad range and somewhat higher energies than are obtained using individual grids.

It is interesting to note that the NMR ensemble of the GDP-bound solution structure apparently contains sufficient information on the conformational flexibility in this system to allow for the larger GTP analogs to be successfully docked. Residues 27 to 32, 58 to 66 and 107 to 109 were identified to be flexible in solution on the basis of reduced ^{15}N relaxation rates while a different orientation was observed for the second helix composed of residues 60 to 75 (Kraulis *et al*., 1994). Additional disorder in the NMR ensemble was observed near residues 30 to 38, 47 to 50 and 120 to 122, although this is not clearly reflected in the reported ^{15}N relaxation rates. These regions overlap reasonably well with residues 32 to 38 and 60 to 75, which are observed to change conformation upon GTP binding in crystal structures (Milburn *et al*., 1990).

To assist in comparing the two composite grid methods, the volumes available for favorable hydrophobic and polar interactions have been evaluated for the above two test cases in Table 4. Carbon and nitrogen probe atoms were used to determine favorable volumes for hydrophobic and polar groups, respectively.

Table 4. Comparison of accessible volumes for carbon and nitrogen atoms in energy-weighted and geometry-weighted ensemble-based grids

	ras p21 (NMR)	
	Volume (Å^3) favorable for C	Volume (Å^3) favorable for N
Energy-weighted		
Ensemble (range)	173–319	291–416
Composite	287	318
Geometry-weighted		
Ensemble (range)	163–299	271–337
Composite (s.d. 1.0 Å)	244	360
Composite (s.d. 0.5 Å)	319	476
	HIV protease (X-ray)	
	Volume (Å^3) favorable for C	Volume (Å^3) favorable for N
Energy-weighted		
Ensemble (range)	454–537	232–328
Composite	523	270
Geometry-weighted		
Ensemble (range)	425–504	211–267
Composite (s.d. 1.0 Å)	540	268
Composite (s.d. 0.5 Å)	740	410

Volumes within a box enclosing the binding site that have an attractive energy for a carbon atom (C, partial charge +0.23) or a nitrogen atom (N, partial charge −0.5) are given in Å^3. The total volume of the box used in the case of HIV protease was 1920 Å^3 while for ras p21 it was 1314 Å^3.

Uteroglobin test case

The solution structure of reduced uteroglobin was solved with a polychlorobiphenyl (PCB) metabolite bound to the protein (Härd *et al*., 1995). The protein is also known to bind progesterone, albeit with a much lower affinity. We were motivated to ask if the progesterone–uteroglobin interaction could be recovered from the solution structure of uteroglobin with an unrelated ligand. Progesterone was docked to both the energy-minimized average structure and an ensemble consisting of 25 solution structures generated with simulated annealing techniques. The average atom rmsd for the binding site is 0.42 Å, making it a fairly well determined structure. When progesterone was docked against the energy-minimized average uteroglobin structure a force field score of only −4.3 kcal/mol was obtained. Using the ensemble grid prepared with a positional standard deviation limit of 0.5 Å yielded a force field score of −25.5 kcal/mol. The energy-weighted composite grid found a similar orientation for progesterone with a force-field score of −22.7 kcal/mol.

The rmsd between the solutions produced with geometry and energy-weighted average grids is 1.3 Å and mostly due to a translation along the long axis of the progesterone molecule. The orientation of the progesterone molecule in both cases is similar to that reported by Dunkel *et al*. (1995), and co-workers based on computer modelling and mutational studies, with the O-3 and O-20 atoms of progesterone positioned between the side-chains of Tyr23 and Thr62. If the orientation of progesterone, as it is docked to the ensemble, were placed in the minimized average structure, severe steric hindrance with Leu15 and Ile65 especially would occur. The conformations of these residues are not very well restrained by the NMR data and were treated as being disordered in the ensemble-based grids. Thus, the minimized average solution structure does not allow for the successful docking of a known ligand, whereas using information from the entire ensemble allows the study of the progesterone–uteroglobin interaction.

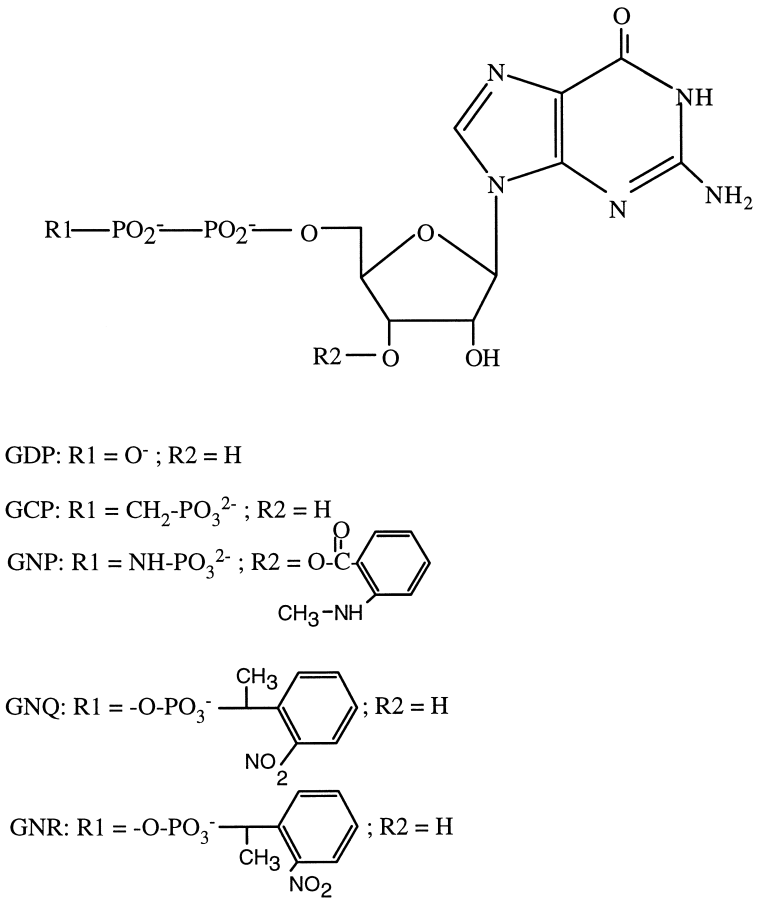


Figure 3. Schematic representation of the chemical structures of GDP and GTP analogs used in docking studies of the ras p21 protein. See Methods for an explanation of the abbreviations used for the different compounds.

Retinol binding protein test case

A total of five crystal complexes of bovine retinol binding with retinol derivatives (see Figure 3) was used for ensemble-based docking. Of the five corresponding ligands, fenretinide especially is observed to induce changes in the side-chain conformations of Leu35 and Leu63 (Zanotti *et al.*, 1994). As is shown in Table 5, fenretidine does not dock correctly to the retinol-bound form of the protein, but all ligands can be docked successfully to the fenretinide-bound form. When composite grids are used, all ligands can be docked with rmsd values and force field energies comparable to those obtained when using the fenretidine-bound crystal structure. The energy-weighted scoring grid again yields a narrower range of energies. The retinol binding protein ensemble is dominated by structures similar to the less permissive retinol-bound form and has the lowest average per atom rmsd of all test cases. Still, both ensemble-based grids allow for successful docking of all ligands.

The force field energy of fenretinide docked against its native protein structure is higher than that obtained when ensemble-based grids

Table 5. Root-mean-square deviations (Å) and force field scores (kcal/mol) of retinol derivatives docked to retinol binding protein

	Retinol bound		Fenretinide-bound		Ensemble G-weighted		Ensemble E-weighted	
	rmsd	Energy	rmsd	Energy	rmsd	Energy	rmsd	Energy
AZE	0.6	−34	0.4	−34	0.7	−34	0.4	−33
ETR	0.6	−29	0.7	−39	0.4	−40	0.4	−36
FEN	**4.7**	−23	0.7	−44	0.8	−38	0.8	−29
REA	1.0	−38	0.8	−39	0.8	−39	0.6	−36
RTL	0.5	−36	0.9	−36	0.8	−36	0.8	−35

The standard deviation threshold above which 35 receptor heavy atoms were treated as flexible (with a receptor rmsd/atom of 0.26 Å) was 0.5 Å. The ensemble consisted of five crystal structures. Results derived with geometry (G)-weighted and energy (E)-weighted composite grids are listed. See the text for ligand abbreviations. The rmsd values of ligands that could not be docked successfully are printed in bold face.

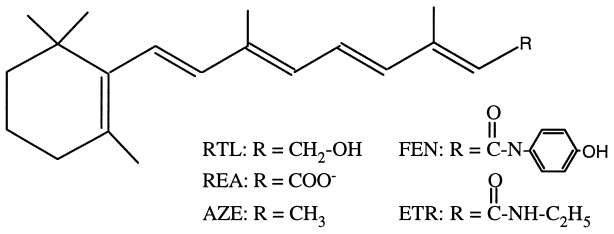


Figure 4. Schematic representation of the chemical structures of retinol analogs used in docking studies of uteroglobin. See Methods for an explanation of the abbreviations used for the different compounds.

are used. This is probably due to the fact that only one out of five protein structures has the side-chains of Leu35 and Leu63 in a conformation adjusted to fenretinide binding, and therefore only one-fifth of the potential function of these residues will interact favorably with this ligand.

Database experiments

To determine the usefulness of composite grids in database searches, a small test database was prepared that included the known HIV and ras p21 binding compounds, as well as about 150 other compounds (Methods). In order to determine whether all known inhibitors or ligands of these systems could be identified, all compounds in this database were docked using the composite grids as well as the standard DOCK grids. In the latter case, a grid was generated for each structure in the ensemble. The relative ranks of known inhibitors within the test database are reported in terms of the portion of the top scoring compounds they constitute, i.e. their fractional rankings are listed in Tables 6 and 7, respectively. The DOCK parameters for these runs were selected to generate a reasonable number of orientations for searching a database (see Methods).

For both the energy-weighted average and geometry-weighted average composite grids, all five HIV inhibitors are ranked within the top 17% of the database. For grids constructed from the separate enzyme structures, at most three inhibitors were identified within the top 25% of the database (Table 6). Here the use of composite grids clearly outperforms the use of a single structure for docking purposes. For the ras p21 test case, all ligands were found within the top 5% of the database using the geometry-weighted average grid and within the top 11% when using the energy-weighted grids. In this case, however, many of the grids generated from individual structures in the ensemble performed better than the composite grids by ranking all inhibitors within the top 5% of the database (Table 7). However, since it is not known *a priori* which of the structures in an ensemble should be used, a composite grid is more likely to identify all ras p21 ligands than a grid constructed from a randomly selected structure in the ensemble.

In order to examine to what extent the composition of the ensemble determines the likelihood with which a composite grid identifies an inhibitor, composite grids were generated using only four out of five HIV protease structures. These grids were used to dock our test database and the resulting fractional rankings for all five permutations of one deleted enzyme structure show that the energy-weigthed average grids find all five inhibitors within 12 to 33% of the database while the geometry-weighted grid allow identification of all inhibitors within 11 to 21% of the database. Deleting the 1hos and q7k systems, especially, from the ensembles causes problems with the successful retrieval

Table 6. Fractional rankings of inhibitors of HIV protease within the test database

No. inhibitors identified	Grid	7hvp	9hvp	1hos	5hvp	q7k	Energy-weighted average	Geometry-weighted average
1		0.01	0.01	0.01	0.01	0.19	0.01	0.01
2		0.06	0.30	0.01	0.42	0.36	0.01	0.01
3		0.12	0.51	0.02	0.59	0.45	02	0.02
4		0.16	0.97	0.03	0.96	0.52	0.03	0.03
5		0.28	1.00	0.35	0.99	0.99	17	0.16

Table 7. Fractional rankings of ligands of ras p21 within the test database

No. ligands identified	Grid	ras p21 1,2,3,4 5,6,7,8, 9,10,11	ras p21 12,13	ras p21 14,15 16,18	ras p21 19	ras p21 20	Energy-weighted average	Geometry-weighted average
1		0.02	0.01	0.02	0.02	0.01	0.01	0.01
2		0.02	0.01	0.02	0.07	0.05	0.02	0.01
3		0.03	0.02–0.04	0.03	0.14	0.17	0.02	0.02
4		0.04	0.03–0.05	0.04–0.08	0.16	0.23	0.04	0.03
5		0.04	0.07–0.10	0.13–0.17	0.30	0.46	0.11	0.05

of all five inhibitors for both grids. However, also in this case the composite grids are still more likely to identify all ligands than a grid based on a randomly selected structure.

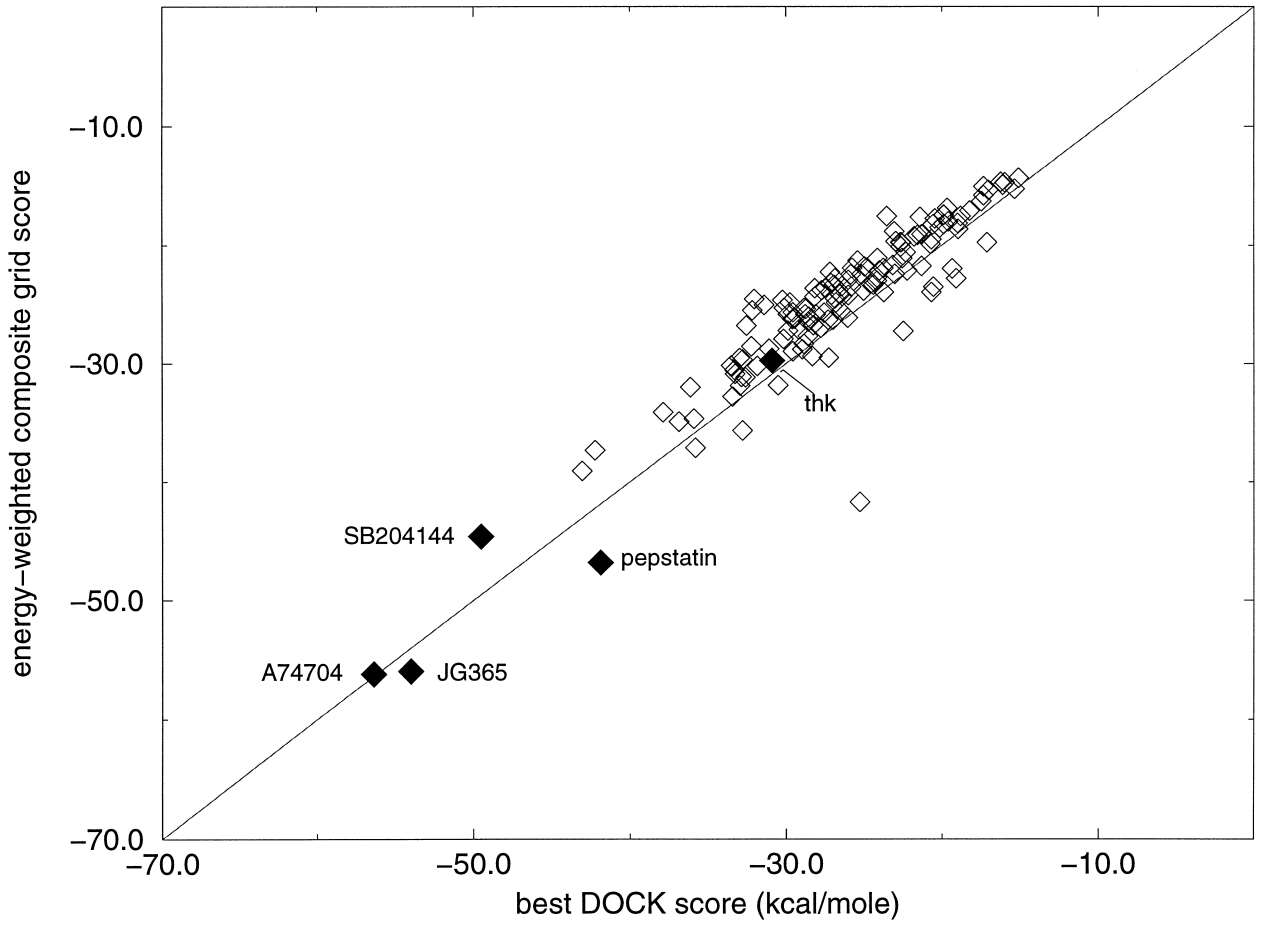
In addition, we compared the score of a particular compound derived from both composite grids with the best score derived from any of the five standard HIV protease and the 20 ras p21 structures. The two sets of scores are plotted against each other in Figure 5 for the HIV protease case and in Figure 6 for the ras p21 test case. Known inhibitors are indicated. A reference line of equal scores is also plotted.

The correlation between the best DOCK score and the composite scores is high: 0.9 for the both composite grids for the HIV protease test case and 0.8 for the ras p21 test case. In addition, the inhibitors and ligands, on average, have a more favorable composite grid score than compounds randomly selected from the database and are identifiable at least as well by the composite grid score as by the DOCK score. The relative rank ordering of the compounds is roughly preserved; that is, the top 10% of compounds scored with composite-grids are also among the top 10% of the best scoring compounds found using all individual structures. Most notably, there are considerable savings in computational time when using composite grids, the order of the number of enzyme systems under consideration.

Discussion

Sensitivity of composite grids to parameters

To be generally applicable, the composite grids must be robust to parameter choices. Our docking experiments showed that the grid spacing for the energy-weighted grids needs to be finer than that of standard or geometry-weighted grids in order to reproduce continuum-calculated values. In particular, the energy-weighted average energies are consistently less favorable when a coarser spacing is used in single ligand docking runs (0.3 Å) compared to the database searches (0.25 Å). The construction of a coarser grid can be visualized as the



(a)

Figure 5(a) *legend opposite*

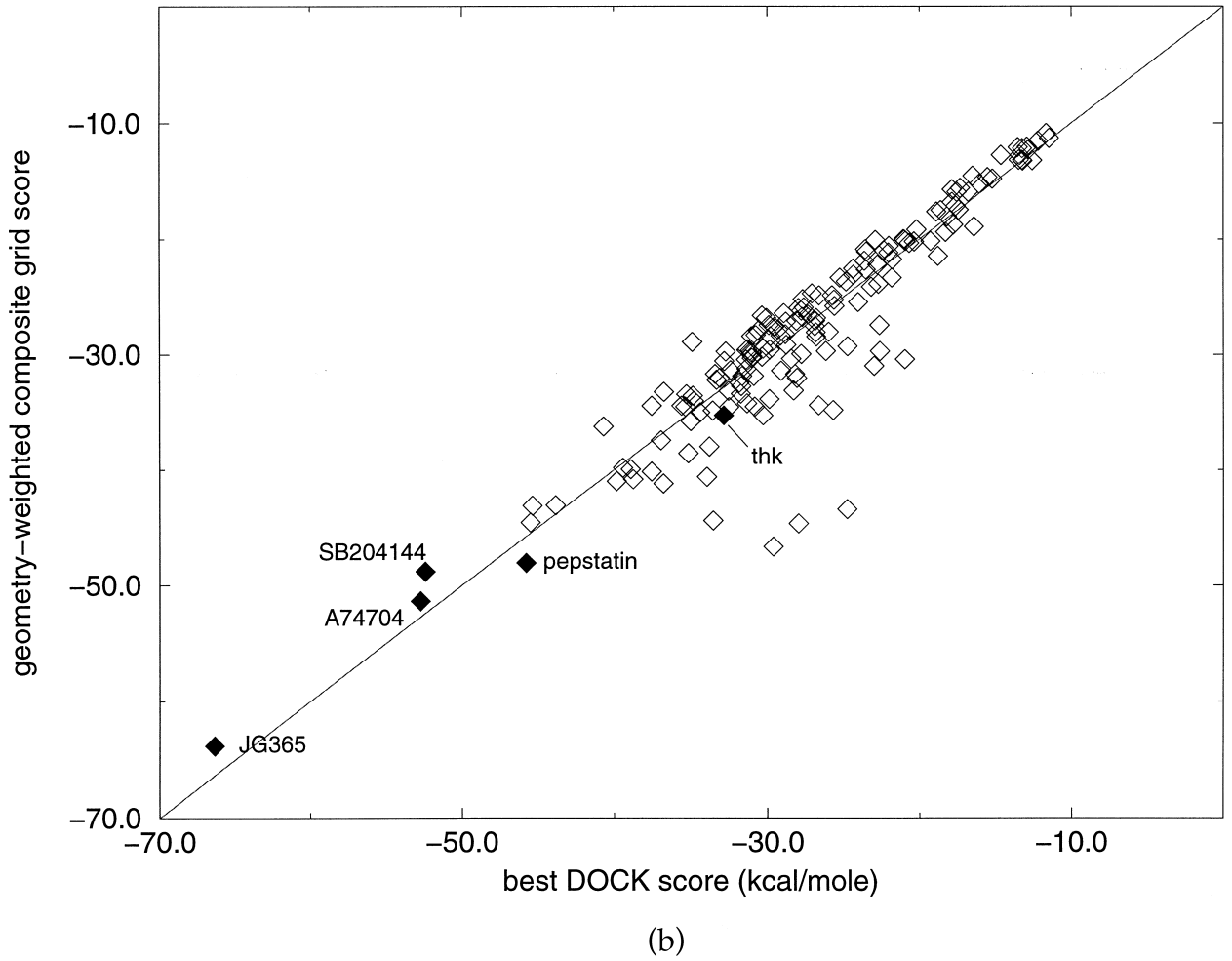


Figure 5. Best DOCK force field energy from five individual HIV protease scoring grids *versus* the energy-weighted (a) and geometry-weighted (b) composite grid scores. Known HIV protease inhibitors are indicated and the line depicting equal scores is plotted for comparison.

averaging of points of a finer grid. Due to the large variation in the repulsive van der Waals potential, this average can be dominated by repulsive interactions, resulting in less favorable scores as observed. For the purpose of reproducing experimentally determined orientations, a coarser spacing was found to be adequate. For studies requiring more accurate energies, however, the finer grid spacing is recommended for this method of averaging. A finer grid spacing demands more random access memory and requires more computer time during grid generation but does not increase the computer time required for docking.

The scores obtained with geometry-weighted average grids are somewhat sensitive to the standard deviation threshold that is chosen. Using a 0.5 Å threshold is recommended, since it assures that small conformational variability is represented in the composite grid. No detrimental effects on the quality of docking solutions were observed when the van der Waals radii, which determine at what point the potential of disordered atoms is set to zero, were changed by 0.1. In addition, the use of spheres generated from a randomly chosen structure of the ras p21 ensemble, instead of the minimized average structure, did not influence the quality of the docked ligand orientations.

Finally, neither method appears to be highly sensitive to the mode of superpositioning. For both the HIV protease and ras p21 test cases, equally good results were obtained when superposition was performed on a different set of residues or different structures.

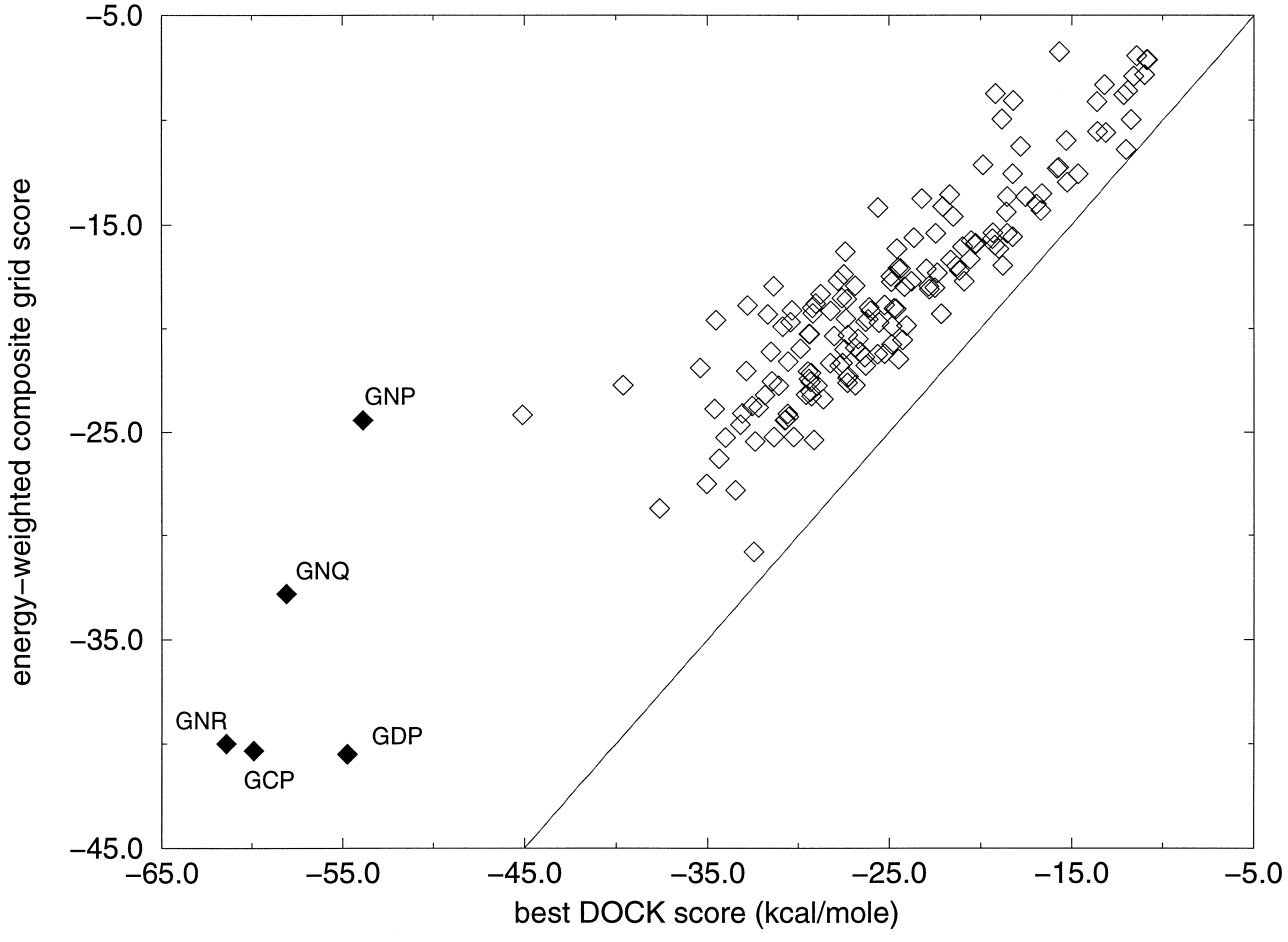
Limitations

The use of ensembles of discrete conformations to mimic receptor flexibility poses certain limitations on the accuracy and scope of the methodology described here. The use of only a limited number of conformations cannot adequately represent the full range of the conformational space available to the receptor and some ligands may therefore not be identified in database searches with composite grids. In this respect, the methodology presented here differs from truly flexible docking in that

different conformational states of the receptor are not generated but taken from available experimental data. Nevertheless, there is more information contained in a set of multiple conformations than in any single conformation. We have addressed the problem of how best to utilize this additional information. In the following paragraphs we will discuss some of the physical issues raised by energy and geometry-weighted averaging methods.

The energy-weighted averaging method attempts to model a Boltzmann-weighted average. Rather than one Boltzmann weight for each structure, each receptor atom has its own separate Boltzmann-like weight. Thus an appropriate portion of each conformation contributes to the energy. Treating atoms in the receptor molecule independently is not physically correct, but attempts to model the accommodation of the ligand by the receptor. There is, of course, some loss of geometric accuracy and the binding energies computed on this composite grid are less favorable than is found for the individual grids. This is to be expected, since an average of energies cannot, by definition, be as favorable as the most attractive enzyme–ligand interaction energy. Nevertheless native ligand orientations can be regenerated and known binding molecules can be identified with this composite grid.

The different treatment of flexible regions by energy and geometry-weighted averaging methods results in differences in the accessibility of disordered regions in the binding site. Energy-weighted averaging smooths repulsive interactions between individual protein structures and a ligand, unless all the structures in the ensemble contribute with a repulsive interaction to the inter-molecular energy at a grid point. The region of overlap between van der Waals radii of receptor atoms from different structures, which decreases with increasing spatial disorder of the atom, can therefore be unaccessible to ligand atoms on the grid. The geometry-weighted average grid, on the other hand, removes the repulsive potentials of flexible atoms entirely, under the assumption that the position of such atoms is uncertain and representing them with their full potentials could interfere with the docking of larger ligands. Therefore, the geometry-weighted average retains less information in the composite grids on which regions of the binding site are inaccessible to ligands. This can yield a



(a)

Figure 6(a) *legend opposite*

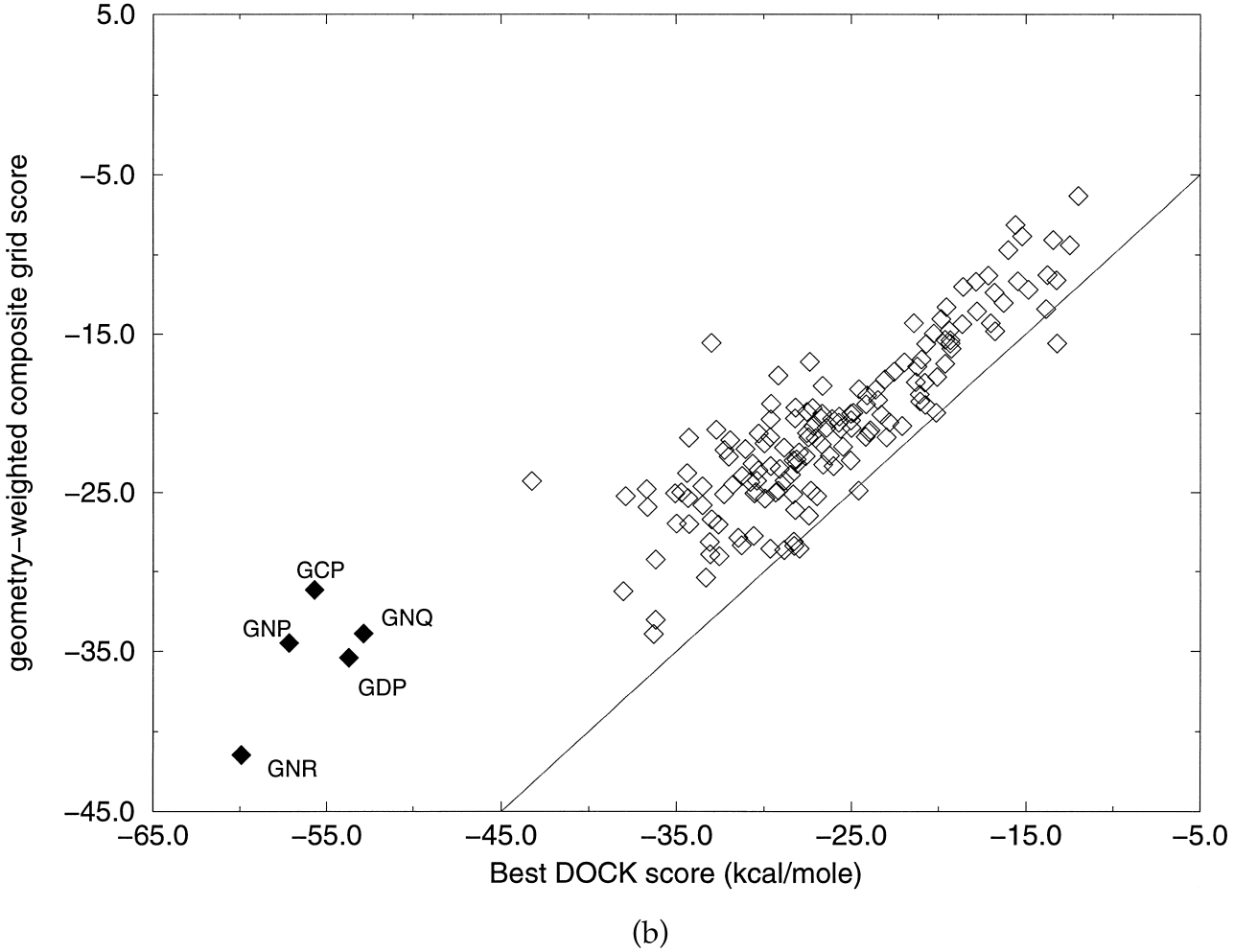


Figure 6. Best DOCK force field energy from 20 individual ras p21 scoring grids *versus* the energy-weighted (a) and geometry-weighted (b) composite grid scores. Known ras p21 ligands are indicated and the line depicting equal scores is plotted for comparison.

geometry-weighted average grid that is more permissive as is illustrated by the results obtained when docking our test database to ensembles consisting of only four out of five HIV protease structures. Using the geometry-weighted average grid, all known ligands dock within the top 21 % of the database while the energy-weighted grids find all inhibitors within the first 33%. The difference is caused by the fact that the geometry-weighted average does not depend on the exact positions of flexible atoms, but only on the variance in the atomic coordinates. Therefore, geometry-weighted averaging is less sensitive to which structure is deleted. This more permissive treatment of disorder in the ensemble may result, however, in protein–ligand complexes being generated that do not correspond to a physical realistic protein conformation.

The differences in how flexible regions are treated are also reflected in the volumes favorable to hydrophobic or polar probe atoms as shown in Table 4. The accessible volumes of energy-weighted composite grids fall within the range displayed by the ensemble. For geometry-weighted composite grids, they increase with the number of disordered atoms. Since disordered atoms are removed from the bump filter, a larger number of ligand orientations will be scored and minimized, resulting in somewhat increased execution times compared to energy-weighted composite grids. This is compensated by the fact that the calculation of energy-weighted average grids for 20 structures takes of the order of ten hours, which could prove to be prohibitive when large numbers of NMR structures are available for docking studies. Geometry-weighted averaging for the same number of structures requires execution times of the order of ten minutes.

Both methods reported here have been pragmatic in nature and do not improve the quality of individual scores *per se*. The development of a new force field, adapted for multiple structures as well as incorporating solvation, hydrophobic and entropic effects, would be highly desirable, but is beyond the scope of this work. Here, we sought to incorporate the information at hand in available

structural data in order to generate a feasible method of evaluating ligand orientations in molecular docking. Our test cases have shown that our composite grid can effectively replace all the individual scoring grids and, for database searching in particular, can be a more useful representation than any single snap shot. If a more accurate evaluation of binding modes and energies is required, ensemble-based grids could be used as a first filter in three dimensional database searches. More accurate scores and orientations can be obtained at a later stage by docking the highest scoring ligands of the composite grids to each of the individual protein structures.

Conclusions

It has been shown that the inclusion of receptor conformational flexibility in molecular docking can be achieved by averaging ensembles of NMR or crystal structures on the basis of interaction energy or structural variability to yield a single scoring grid. Both scoring schemes are based upon the DOCK molecular mechanics interaction energy and are highly correlated with the best DOCK score of a ligand with the available macromolecular conformations. By smoothing or removing the repulsive part of the chemical potentials of flexible atoms, differently shaped ligands can be docked to yield orientations close to known binding modes near to flexible regions of protein structures. Under the same sampling conditions, the relative rank ordering of the compounds is grossly preserved. In the case of HIV protease and ras p21, both methods can identify known inhibitors while affording a saving in time of the order of the number of structures. In all test cases presented here, the composite grids outperform many of the grids derived from individual structures of the ensemble, or in the case of NMR-derived structures, the minimized average structure.

Although the energy-weighted and geometry-weighted averaging methods are based on different assumptions regarding the treatment of multiple structures in molecular docking, both performed equally well for the test cases presented here. These methods do not explore receptor flexibility in a dynamic fashion but, rather, they introduce an averaging algorithm that uses information from an ensemble of structures to allow for a broader range of realistic ligands to be identified when searching three-dimensional databases.

Methods

This section is divided into three parts: (1) a description of the computational procedures; (2) a description of the test systems used in this work; and (3) the parameters used.

Computational procedures

The starting point of all our calculations is a set of protein–ligand complexes, determined either by X-ray crystallography (HIV protease, retinol binding protein) or by NMR (ras p21, uteroglobin). The general docking procedure that we have used is implemented in the programs SPHGEN and DOCK 3.5, which have been described (Kuntz *et al.*, 1982; Meng *et al.*, 1992, 1993; Shoichet *et al.*, 1992). The active site of each enzyme is first characterized with a set of overlapping spheres. These packed spheres form a negative image of the enzyme active site and are used to orient the ligand. Ligands oriented within the binding site are evaluated on a scoring grid yielding an inter-molecular interaction energy. In the following we describe the construction of energy-weighted and geometry-weighted composite scoring grids as well as the individual grids for each enzyme structure.

The standard method of scoring in DOCK 3.5 uses the AMBER potential function (Weiner *et al.*, 1984, 1986) to approximate the ligand–enzyme binding energy with a molecular mechanics interaction energy. The van der Waals and Coulombic interactions are summed over all pairs of ligand atoms and enzyme atoms within a distance cutoff. To calculate this interaction energy, the individual terms in the potential function are separated into protein factors and ligand factors. The protein contributions are precalculated and stored on a scoring grid; this grid fills the volume accessible to ligand atoms. Each ligand orientation is then evaluated by combining these factors with ligand factors to generate a score, the intermolecular interaction energy.

Standard DOCK interaction energy

The DOCK grid-based interaction energy, E, of the ligand with a receptor, is given by (Meng *et al.*, 1992):

$$E = \sum_i \left[A_i \left(\sum_j (\boldsymbol{A}_j / r_{ij}^{12}) \right) - B_i \left(\sum_j (\boldsymbol{B}_j / r_{ij}^{6}) \right) + K q_i \left(\sum_j (\boldsymbol{q}_j / \varepsilon r_{ij}) \right) \right]$$

where i indexes the ligand atoms and j indexes the receptor atoms; A and B are the van der Waals attractive and dispersive parameters, respectively; q is the partial charge on the atom; K is the scaling constant which converts electrostatic energy into kcal/mol; ε is the dielectric constant of the medium; r_{ij} is the distance between ligand atom i and receptor atom j.

Terms in boldface are the receptor van der Waals attractive, repulsive and Coulombic terms, which have been precalculated and stored on a grid.

Energy-weighted average interaction energy

The energy-weighted average interaction energy over the N receptor structures is defined as:

$$E = (1/N) \sum_i \left[A_i \left(\sum_j \sum_n \boldsymbol{wt}_{jn} (\boldsymbol{A}_{jn} / r_{ijn}^{12}) \right) - B_i \left(\sum_j \sum_n \boldsymbol{wt}_{jn} (\boldsymbol{B}_{jn} / r_{ijn}^{6}) \right) + K q_i \left(\sum_j \sum_n \boldsymbol{wt}_{jn} (\boldsymbol{q}_{jn} / \varepsilon r_{ij}) \right) \right]$$

where i and j are indexes over ligand atoms and receptor atoms, respectively, as above; A,B, q, K ε and r_{ij} are are defined above; n, indexes over receptor structures; $wt_{j\ n}$

is a (normalized) sigmoidal-like function of the distance from the receptor atom j of structure n to ligand atom i (at a grid point).

Terms in bold face are precalculated and stored on a grid.

For every receptor atom j, the potential (of the atom) from each structure is calculated, then combined with a normalized weighted sum. (The sum is over the jth atom from all N structures.) The unnormalized weight, wt, for receptor atom j in each structure n, increases from near zero at short distances to unity at large distances *via* a truncated cosine term as shown below:

$$wt = \begin{cases} \approx 0 \text{ if } d_{ij} \leqslant \min R_j \\ 0{:}5(1 - cos(\pi(d_{ij} - \min R_j)); \\ \quad \text{if } \min R_j < d_{ij} < \min R_j + 1 \\ 1; \text{ otherwise} \end{cases}$$

where d_{ij} is the distance between receptor atom j (of structure n) to ligand atom i (at a grid point);

$\min R_j =$ (well-depth of atom $j/\beta)^{1/12}$ and β is an adjustable parameter, set to 50 in our systems and $\min R_j$ can be interpreted as the van der Waals radius of the atom. At distances (from the ligand atom at a grid point) where the van der Waals repulsive potential is greater than or equal to β, the unnormalized weight of an enzyme atom potential is approximately equal to zero and set to δ. Note that if the (unnormalized) weight for the enzyme atom potential from each structure is the same and equal to δ, then the normalized weight becomes unity. Roughly speaking, if, at one point, the repulsive term from an atom from one structure is large, it does not contribute to the potential function at that point. However, if the potential of the atom from all structures is repulsive, then all contribute equally to the potential.

For each orientation, ligand terms, stored separately, are combined with receptor terms to calculate a score. Ligand atom positions falling between grid points use values calculated by trilinear interpolation between grid points (Meng *et al.*, 1992).

In order to limit the number of ligand orientations that are actually evaluated in DOCK, each orientation is first tested to see if ligand atoms intersect the enzyme; if too many ligand atoms ''bump'' into the enzyme, the orientation is not scored. The energy-weighted average bump grid is constructed in a manner similar to the standard DOCK bump grid: for every grid point k, the number of enzyme atoms from every structure that is close to the grid point k is counted. Grid points with counts equal to or greater than the number of structures are considered to be a bump.

Geometry-weighted average interaction energy

The geometry-weighted average method distinguishes between structurally well-defined and disorderd atoms in the ensemble in order to arrive at a composite grid for molecular docking. Ensembles of protein structures are superimposed on residues near the binding site and averages and standard deviations are calculated for all individual atom positions. A single structure is then derived in which, for atoms with standard deviations above a user-defined threshold, all copies of the original atom positions are kept, while for less-flexible atoms only the average position is used. To reduce the number of atoms in the resulting structure, this procedure is only followed for atoms within the boundaries of the box that is used to define the volume and position of the scoring grid. All atoms located outside of this box are averaged, regardless of their positional standard deviation.

In order to allow larger ligands to occupy space within flexible regions of the binding site, the repulsive portion of the force field potentials of disordered atoms needs to be adjusted. In DOCK 3.5 the scoring grids consist of a bump grid and chemical grids (Meng *et al.*, 1992). The bump grid serves as a crude initial filter for orientations generated by the matching algorithm. A grid point of the bump grid is considered to be unaccessible to ligand atoms when the van der Waals volume of a receptor atom overlaps with it. Orientations that place ligand atoms at such grid points are rejected without the need for computationally more demanding scoring. The volume of receptor atoms is determined by two distinct van der Waals radii for either polar or non-polar atoms. For the purpose of creating bump grids, disordered atoms were considered to be volumeless. For the chemical grids, both the van der Waals and electrostatic potentials were set to zero for grid points within the van der Waals radius of a disordered atom. Beyond the van der Waals radius the attractive part of the potential remains intact but is scaled by the number of copies, as is depicted in Figure 1. Thus, although the disordered atom is no longer present as a van der Waals barrier for potential ligands, its chemical characteristics are still encoded in the scoring grid. Due to the fact that the force field grid is no longer continuous, the trilinear grid interpolation implemented in DOCK 3.5 (Meng *et al.*, 1992) was sometimes observed to produce spurious results during simplex minimization of docked ligands (Meng *et al.*, 1993). Therefore we conducted the docking simulations that made use of ensemble-based grids without the use of grid interpolation.

Test cases

Suitable test cases are critical to evaluating methodology for docking to ensembles of receptor structures. For crystal structures to be useful as test cases, several structures of complexes with different ligands at high resolution (preferably less than 2 Å) are required. The ligands should not be covalently linked to the protein and should induce large enough conformational changes in the protein, such that not all ligands fit all receptor structures. For NMR structures, a conclusive definition of resolution has not yet been proposed, but more than ten restraints per residue and a rmsd smaller than 1 Å for the backbone would be preferred. Unfortunately, there are few ensembles of crystal sctructures available through the Protein Data Bank (PDB) that fulfil all the requirements listed above and even less NMR structures of proteins complexed with small molecules.

As test cases we used ensembles of solution structures of the ras p21 protein and uteroglobin as well as crystal structures of retinol binding protein and HIV protease. The NMR solution structure of HIV protease bound to a cyclic urea has been reported recently (Yamazaki *et al.*, 1995) but the coordinates for these structures are not yet available for comparative dockings.

Five complexes of the HIV-1 aspartyl protease with bound inhibitors were used for docking studies. Four complexes were taken form the PDB, namely entries 7hvp (the complex with JG365, a substrate-like hydroxyethylamine inhibitor at 2.4 Å resolution; Swain *et al.*, 1990), 9hvp (the complex with A74704 at 2.8 Å resolution; Erickson *et al.*, 1990), 1hos (the complex with SB204144 at 2.3 Å resolution; Abdel-Meguid *et al.*, 1993),

and 5hvp (complex with acetyl pepstatin at 2.0 Å resolution; Fitzgerald *et al.*, 1990). The structure of the the Q7K mutant of the homodimeric protein complexed with thioketal haloperidol (UCSF8) (Rutenber *et al.*, 1993) was obtained from R. Keenan (Biochemistry Department, UCSF). These ligands will be referred to as JG365, A74704, SB204144, pepstatin and thk, respectively. Their chemical structures are depicted in Figure 2.

The ras p21 protein is a product of the human *ras* oncogene and can exist in a GTP-bound activated state or a GDP-bound inactive state. In both conformations the guanine binding pocket is structurally conserved but large structural changes of the order of 3 to 7 Å are observed in the region where the β- and γ-phosphate groups of GTP and a magnesium ion are bound (Milburn *et al.*, 1990; Privé *et al.*, 1992). The crystal structures of five complexes of ras p21 with GDP and GTP analogs were retrieved from the PDB (Bernstein *et al.*, 1977). These were PDB entries 4q21 (the complex with GDP at 2 Å resolution; (Milburn *et al.*, 1990), 6q21 (the complex with a GTP-like inhibitor at 1.95 Å resolution; Milburn *et al.*, 1990; Privé *et al.*, 1992), 1gnp (the complex with 3^0-*O*-(*N*-methyl-anthraniloyl-2^0-deoxyguanosine-5^0-(β,γ-imido)-triphosphate, at 2.7 Å resolution), 1gnq (the complex with P^3-1[*S*]-(2-nitrophenyl) ethyl-guanosine-5^0-triphosphate at 2.5 Å resolution) and 1gnr (the complex with P^3-1[*R*]-(2-nitrophenyl) ethyl-guanosine-5^0-triphosphate at 1.85 Å resolution; Scheidig *et al.*, 1995). These ligands will be referred to as GDP, GCP, GNP, GNQ and GNR, respectively, and their chemical structures are shown in Figure 3. The latter three GTP analogs have large aromatic substituents at either the ribose (GNP) or the γ-phosphate (GNQ and GNR) and are expected not to dock correctly into the GDP-bound form, since their larger volume requires structural adaptation by the protein. In order to obtain ligand coordinates that were not dependent on the mode of superposition of the different protein structures, all ligands were superimposed with their guanine moieties on that of the GDP molecule of the minimized average NMR structure. This could give rise to somewhat larger root-mean-square deviations due to slightly different placing of the ligands in the different ras p21 structures.

The magnesium ion bound by the ras p21 protein was kept in place but its formal charge and those of the ligand phosphate groups were divided by a factor of 5. This was empirically shown to reduce the dominance of the electrostatic interactions between these highly charged ions on the total energy, without affecting other electrostatic interactions in the binding site.

The NMR solution structure of the GDP-bound form of ras p21 has recently been solved (Kraulis *et al.*, 1994) and the minimized average structure (PDB entry 1crq) together with 20 structures generated using a simulated annealing protocol (PDB entry 1crr) were retrieved from the PDB for use in docking studies. For residues 1 to 166 of the protein coordinates were given in all crystal and NMR structures and therefore only these residues were used.

The structure of the reduced rat uteroglobin protein complexed with $4,4^0$-bismefylsulphonyl-$2,2^0,5,5^0$-tetrachlorobiphenyl ((MeSO2)2-TCB) was recently solved using NMR spectroscopy (Härd *et al.*, 1995). This homodimeric protein has a high affinity for polychlorinated biphenyls (PCB)s ($K_d \approx 10$ nM) (Härd *et al.*, 1995) and is also known to weakly bind progesterone ($K_a \approx 15\ mM^{-1}$) (Dunkel *et al.*, 1995). Its endogenous ligand is still unknown. The binding site of this protein is extremely hydrophobic, with only residues Tyr23 and Thr62 supposedly being capable of forming hydrogen bonding interactions with ligands. Upon binding of the ligand the binding site is thought to close due to oxidation of cysteine residues. The minimized avaraged structure of the complex was obtained from the PDB (entry 1utr) while an ensemble of 25 simulated annealing structures was kindly provided by Dr T. Härd (Karolinska Institute, Sweden).

In the structures of complexes between bovine plasma retinol binding protein and retinol analogs (Zanotti *et al.*, 1993, 1994), changes in side-chain conformations of residues near the retinol binding site have been observed. In particular, fenretidine, which has a hydroxyphenyl amide group replacing the retinol hydroxyl group, altered the side-chain conformations of Leu35 and Leu63 (Zanotti *et al.*, 1994). Five complexes were retrieved from the PDB, namely entries 1erb (the complex with *N*-ethyl retinamide at 1.9 Å resolution; Zanotti *et al.*, 1993), 1fel (the complex with fenretinide at 1.8 Å resolution), 1fem (the complex with reitnoic acid at 1.9 Å resolution), 1fen (the complex with axerophtene at 1.9 Å resolution) and 1hbp (the complex with retinol at 1.9 Å resolution; Zanotti *et al.*, 1994). For convenience, these ligands will be referred to as ETR, FEN, REA, AZE and RTL, respectively, and their structures are shown in Figure 4. Only the ordered residues 3 to 174 of the protein were used for docking studies.

All protein structures were stripped of water molecules except in the case of HIV protease where the water molecule placed between the flaps and the inhibitors was kept in place. In the thioketal haloperidol structure a chloride ion was observed to replace this water (Rutenber *et al.*, 1993). In order to make the thioketal haloperidol complex homologous to the other four complexes the chloride ion was replaced by water. In addition, residues 7, 14, 37, 41, 43, 61, 63, 64, 67, 72 and 95 in both momomers were replaced by alanine, since the amino acids at these positions varied among the five complexes. These residues are remote from the active site and are not expected to influence the docking process. Protons were added to all protein structures and in the case of the HIV protease structures a proton was placed between the catalytic aspartates.

Parameters used for grid generation and docking

All ligand structures were taken from the PDB files, after which proton positions and Gasteiger-Marsili charges (Gasteiger & Marsili, 1980) were calculated using SYBYL V6.02 (Tripos Associates, St Louis, MO).

Spheres were generated with the program SPHGEN as described (Kuntz *et al.*, 1982) and edited in order to remove spheres remote from the binding site. For ensembles of solution structures the minimized average structure was used for sphere generation. A box enclosing the spheres was generated with the box faces placed at 3 Å from the nearest sphere. Protein structures were superimposed on the C^α atoms of residues within 5 Å from the ligand. For ras p21 these were residues 11 to 18, 28 to 32, 34, 57, 116 to 120 and 144 to 147, for uteroglobin residues 8, 11, 15, 23, 40, 43, 58, 61, 62 and 65 in both monomers, for retinol binding protein residues no superposition was performed, since the structures taken from the PDB already superimposed well. The HIV protease structures were superimposed on residues 8, 23, 25, 27 to 32, 45 to 53, 76 and 80 to 86 of both monomers in the dimer.

In the case of geometry-weighted averaging, force field scoring grids were generated with the CHEMGRID program (Meng *et al.*, 1992), which was modified to remove the force field potentials of disordered atoms within the same van der Waals radii as are used for creating the bump grids. For polar atoms the radius was set to 2.3 Å and for non-polar atoms to 2.8 Å while for polar protons a radius of 1 Å was used. A distance-dependent dielectric constant ($\varepsilon = 1/4r$) was assumed for atoms within 10 Å of each other and the grid spacing was set to either 0.25 Å or 0.3 Å. For the protein the AMBER united atom representation was used (Weiner *et al.*, 1984), while for the ligands all atom parameters were applied (Weiner *et al.*, 1986).

In the case of the energy-weighted averaging, grids were constructed as described in the previous section. The AMBER united atom parameter set was used (Weiner *et al.*, 1984). A distance-dependent dielectric constant ($\varepsilon = 1/4r$) was assumed. All receptor atoms within 10 Å of a grid point were used to calculate the receptor terms. Grid spacing was either 0.25 Å or 0.3 Å, although the former spacing was found to be more appropriate and is suggested for future work.

Favorable volumes of chemical scoring grids were calculated using either a sp^3 carbon atom from the AMBER all atom force field with a 0.235 charge or a sp^2/sp nitrogen atom with charge −0.5 as a probe. For database runs the grid spacing was 0.2 Å while for single ligand docking runs 0.3 Å was used.

Single ligand and ras p21 database docking runs were done with a matching tolerance of 0.7 Å and bin sizes of 0.5 Å with 0.2 Å overlap (Shoichet *et al.*, 1992). For HIV protease database searches matching tolerances were set to 1.2 Å with ligand bin sizes of 0.2 Å with 0.05 Å overlap (0.5 and 0.2 Å, respectively for receptor binsizes). Simplex minimization of the generated orientations (Meng *et al.*, 1993) was done using either 500 or 100 as the maximum number of iterations in the case of single ligand or database docking runs, while the convergence criterion was set to 0.2 or 5 kcal/mol, respectively.

Acknowledgements

We gratefully acknowledge support from NIH grant GM 31497. R.M.A.K. acknowledges financial support from the Netherlands Organization for Scientific Research (NWO). We thank Dr T. Härd for making the uteroglobin ensemble available to us, MDL Information Systems Inc. for the CMC-3D database and the UCSF Computer Graphics Laboratory for use of the Midas molecular graphics software.

References

Abdel-Meguid, S., Zhao, B., Murthy, K. H. M., Winborne, E., Choi, J.-K., Desjarlais, R. L., Minnich, M. D., Culp, J. S., Debouck, C., Tomaszek, T. A., Jr, Meek, T. D. & Dreyer, G. B. (1993). Inhibition of human immunodeficiency virus-1 protease by a C2-symmetric phosphinate. *Biochemistry,* **32**, 7972–7980.

Bacon, D. J. & Moult, J. (1992). Docking by least-squares fitting of molecular surface patterns. *J. Mol. Biol.* **225**, 849–858.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. T. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

Böhm, H. J. (1992a). The computer program LUDI: a new method for the *de novo* design of enzyme inhibitors. *J. Comput.-aided Mol. Design,* **6**, 61–78.

Böhm, H. J. (1992b). LUDI: Rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput-aided-Mol. Design,* **6**, 593–606.

Bonvin, A. M. J. J. & Brünger, A. T. (1995). Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol.* **250**, 80–93.

Clark, K. P. & Ajay, (1995). Flexible ligand docking without parameter adjustment across four ligand receptor complexes. *J. Comput. Chem.* **16**, 1210–1226.

Dunkel, R., Vriend, G., Beato, M. & Suske, G. (1995). Progesterone binding to uteroglobin: two alternative orientations of the ligand. *Protein Eng.* **8**, 71–79.

Erickson, J., Neidhart, D. J., van Drie, J., Kempf, D. J., Wang, X. C., Norbeck, D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Paul, D. A. & Knigge, M. (1990). Design, activity and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science,* **249**, 527–533.

Fitzgerald, P. M. D., McKeever, B. M., van Middlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C.-T., Herber, W. K., Dixon, R. A. F. & Darke, P. L. (1990). Crystallographic analysis of a complex between human imunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0 Å resolution. *J. Biol. Chem.* **265**, 14209–14219.

Gasteiger, J. & Marsili, M. (1980). Iterative partial equalization of orbital electronegativity: rapid access to atomic charges. *Tetrahedron Letters,* **36**, 3219–3222.

Härd, T., Barnes, H. J., Larsson, C., Gustafson, J.-Å. & Lund, J. (1995). Solution structure of a mammalian PCB-binding protein in complex with a PCB. *Nat. Struct. Biol.* **2**, 983–989.

Jones, G., Willett, P. & Glen, R. C. (1995). Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **245**, 43–53.

Jorgensen, W. L. (1991). Rusting of the lock and key model for protein–ligand binding. *Science,* **254**, 954–955.

Judson, R. S., Tan, Y. T., Mori, E., Melius, C., Jaeger, E. P., Treasurywala, A. M. & Mathiowetz, A. (1995). Docking flexible molecules: a case study of three proteins. *J. Comput. Chem.* **16**, 1405–1419.

Kollman, P. (1994). Theory of macromolecule-ligand interactions. *Curr. Opin. Struct. Biol.* **4**, 240–245.

Koshland, D. E. (1958). Application of a thoery of enzyme specificity to protein synthesis. *Proc. Natl Acad. Sci. USA,* **44**, 98–104.

Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., van Aken, T. & Laue, E. D. (1994). Solution structure and dynamics of ras p21-GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry,* **33**, 3515–3531.

Kuntz, I. D. (1992). Structure-based strategies for drug design and discovery. *Science,* **257**, 1078–1082.

Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**, 269–288.

Leach, A. R. (1994). Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.* **235**, 345–356.

Leach, A. R. & Kuntz, I. D. (1992). Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comput. Chem.* **13**, 730–748.

Lewis, R. A. (1992). Automated site-directed drug design using molecular lattices. *J. Mol. Graph.* **10**, 66–78.

Lin, S. L. (1994). Molecular surface representation by sparse critical points. *Proteins: Struct. Funct. Genet.* **18**, 94–101.

McCammon, J. A. & Harvey, S. C. (1987). In *Dynamics of Proteins and Nucleic Acids.* Cambridge University Press, Cambridge.

Meng, E. C., Schoichet, B. K. & Kuntz, I. D. (1992). Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **13**, 505–524.

Meng, E. C., Gschwend, D. A., Blaney, J. M. & Kuntz, I. D. (1993). Orientational sampling and rigid-body minimization in molecular docking. *Proteins: Struct. Funct. Genet.* **17**, 266–278.

Milburn, M. V., Tong, L., de Vos, A. M., Brünger, A., Yamaizumi, Z., Nishimura, S. & Kim, S.-H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic *ras* protein. *Science,* **247**, 939–945.

Mirankar, A. & Karplus, M. (1995). An automated method for dynamic ligand design. *Proteins: Struct. Funct. Genet.* **23**, 472–490.

Mizutani, M. Y., Tomioka, N. & Itai, A. (1994). Rational automatic search method for stable docking models of protein and ligand. *J. Mol. Biol.* **243**, 310–326.

Norel, R., Fischer, D., Wolfson, H. J. & Nussinov, R. (1994). Molecular surface recognition by a computer vision-based techniquw. *Protein Eng.* **7**, 39–46.

Oshiro, C. M., Kuntz, I. D. & Dixon, J. S. (1995). Flexible ligand docking using a genetic algorithm. *J. Comput.-aided Mol. Design,* **9**, 113–130.

Privé, G. G., Milburn, M. V., Tong, L., de Vos, A. M., Yamaizumi, Z., Nishimura, S. & Kim, S.-H. (1992). X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization for GTP hydrolysis. *Proc. Natl Acad. Sci. USA,* **89**, 3649–3653.

Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489.

Rutenber, E., Fauman, E. B., Keenan, R. J., Fong, S., Furth, P. S., Ortiz de Montellano, P. R., Meng, E., Kuntz, I. D., De Camp, D. L., Salto, R., Rose, J. R., Craik, C. S. & Stroud, R. M. (1993). Structure of a non-peptide inhibitor complexed with HIV-1 protease. *J. Biol. Chem.* **268**, 15343–15346.

Scheidig, A. J., Franken, S. M., Corrie, J. E. T., Reid, G. P., Wittinghofer, A., Pai, E. F. & Goody, R. S. (1995). X-ray crystal structure analysis of the catalytic domain of the oncogene product $p21^{H\text{-}ras}$ complexed with caged GTP and Mant dGppNHp. *J. Mol. Biol.* **253**, 132–150.

Shoichet, B. K., Bodian, D. L. & Kuntz, I. D. (1992). Molecular docking using shape descriptors. *J. Comput. Chem.* **13**, 380–397.

Sutcliffe, M. J. (1993). Representing an ensemble of NMR-derived protein structures by a single structure. *Protein Sci.* **2**, 936–944.

Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H. & Wlodawer, A. (1990). X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate based hydroxyethylamine inhibitor. *Proc. Natl Acad. Sci. USA,* **87**, 8805–8809.

Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr & Weiner, P. A. (1984). A new force field for molecular mechanics simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.

Weiner, S. J., Kollman, P. A., Hguyen, D. T. & Case, D. A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230–252.

Welch, W., Ruppert, J. & Jain, A. N. (1996). Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chem. Biol.* **3**, 449–462.

Wüthrich, K. (1986). In *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York.

Yamazaki, T., Hinck, A. P., Wang, Y. X., Nicholson, L. K., Torchia, D. A., Wingfield, P., Stahl, S. J., Kaufman, J. D., Chang, C. H., Domaille, P. J. & Lam, P. Y. S. (1995). Three dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy. *Protein Sci.* **5**, 495–506.

Zanotti, G., Malpeli, G. & Berni, R. (1993). The interaction of N-ethyl retinamide with plasma retinol-binding protein (RBP) and the crystal structure of the retinoid-RBP complex at 1.9 Å resolution. *J. Biol. Chem.* **268**, 24873–24879.

Zanotti, G., Marcello, M., Malpeli, G., Folli, C., Sartori, G. & Berni, R. (1994). Crystallographic studies on complexes between retinoids and plasma retinol-binding protein. *J. Biol. Chem.* **26**, 29613–29620.

Edited by B. Honig

(Received 6 October 1996; received in revised form 31 October 1996; accepted 5 November 1996)