Automated Docking with Grid-Based Energy Evaluation

Elaine C. Meng, Brian K. Shoichet, and Irwin D. Kuntz*

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

Received 24 September 1991; accepted 4 December 1991

The ability to generate feasible binding orientations of a small molecule within a site of known structure is important for ligand design. We present a method that combines a rapid, geometric docking algorithm with the evaluation of molecular mechanics interaction energies. The computational costs of evaluation are minimal because we precalculate the receptor-dependent terms in the potential function at points on a three-dimensional grid. In four test cases where the components of crystallographically determined complexes are redocked, the "force field" score correctly identifies the family of orientations closest to the experimental binding geometry. Scoring functions that consider only steric factors or only electrostatic factors are less successful. The force field function will play an important role in our efforts to search databases for potential lead compounds.

INTRODUCTION

The opportunities for *de novo* drug design are greater now than ever before. This can be attributed to the discovery of the molecular bases of many diseases and to progress in macromolecular structure determination. A wealth of mechanistic information exists in the atomic coordinates of a macromolecule; in addition, the detailed structure may suggest a means for altering function. Numerous drugs work by specifically binding to a receptor molecule and modulating its biological activity.

The ability to propose reasonable ways of binding a putative ligand molecule to a known receptor site is crucial to the success of structure-based design. One approach is to position or "dock" ligand and receptor molecules together in many different ways and then "score" each orientation according to an evaluation function of some kind.

Docking methods can be subdivided into manual and automatic approaches. In manual docking, the user is responsible for positioning the molecules; this process may be interactive, with continuous feedback on the energy of the system,[1-3] or each energy determination may require a significant amount of computer time. As with any energy calculation, the time demands increase with increasing complexity of the evaluation function, increasing number of atoms, and increasing number of degrees of freedom within the system. The same considerations apply to automated docking,[4-11] in which the molecules are positioned according to algorithms that vary from exhaustive to stochastic to deterministic. Compared to manual docking, automatic methods are less dependent upon, though certainly not independent of, the preconceptions of the user regarding which areas of the receptor are most important for binding.

The complexity of configuration space for systems involving biomacromolecules leads to high computational costs, especially when realistic potential functions are used. A balance between thermodynamic accuracy and computational tractability is desirable. A significant reduction in scoring time can be achieved by precomputing terms in the potential function that are sums over receptor atoms. As in the work of Goodford[12] and several of the docking methods,[1-3,10] receptor terms are calculated for each point on a three-dimensional grid. While this approach requires more preparation, it decreases the time needed to evaluate ligand orientations. Any expression in which the ligand and receptor terms are separable can be treated in this manner. We have chosen a set of functions and parameters that approximate the AMBER force field.[13,14] Using this approach, ligand orientations generated with the rigid-body docking algorithm of Kuntz and coworkers[5,15-17] can be ranked according to molecular mechanics interaction energy. The method is rapid and robust; test runs on known complexes suggest that approximate interaction energies are useful for discerning likely binding orientations.

COMPUTATIONAL METHODS

The major steps of the procedure are: characterization of the receptor site, calculation of grids for evaluating docked structures, docking, and evaluation of

*Author to whom all correspondence should be addressed.

CCC 0192-8651/92/040505-20$04.00

the resulting ligand orientations. Programs involved in the overall process are shown in Figure 1. The computer programs MS[18,19] and DelPhi[20,21] are distributed independently.

Site Characterization

We characterize the site as described previously.[5,15-17] The Connolly MS algorithm[18,19] is used to generate a molecular surface as defined by Richards.[22] Spheres that fill surface indentations are then calculated with the program SPHGEN.[5] Each sphere touches the surface at two points and is centered along the surface normal at one of the points. Only one sphere per surface atom, the largest that does not intersect the surface, is generally retained; groups of overlapping spheres are referred to as clusters. The cluster containing the greatest number of spheres tends to occupy the largest indentation of the surface, typically the active site of an enzyme. The user selects one or more clusters for docking.

Calculation of Grids

We use the following means of evaluating molecular complexes: contact score, electrostatic interaction energy, and molecular mechanics interaction energy.

While each option makes use of a cubic lattice, there are differences in the details of implementation. The contact grid is automatically constructed to enclose the input atoms, which may form part or all of the receptor. The electrostatic grid encloses a cubic volume, which, due to the nature of the calculation, should include the entire receptor molecule.

The volume enclosed by the force field grid may have different x, y, and z extents. All receptor atoms are included in the calculation whether or not they fall within the grid volume. The force field grid may be positioned either by direct specification of its coordinates or by centering within a sphere cluster; in this manner, one can define a box that efficiently encloses the space that docked molecules are likely to occupy.

We next consider how each set of grid values is calculated.

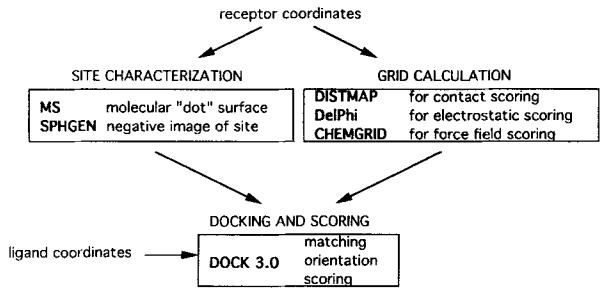


Figure 1. Programs involved in the use of DOCK. MS[18,19] and DelPhi[20,21] are distributed independently.

The program DISTMAP[16,17] produces the grid for contact scoring. The user specifies the grid resoltuion, two "close contact" limits (for receptor polar and nonpolar atoms, respectively), and a cutoff defining the range of pairwise contacts. For every receptor atom within the contact range, the sum at a grid point is incremented by one, unless a close contact limit is violated, in which case a negative number is added. Hydrogens are not included in the calculation. We note that this is a different contact score than used in earlier versions of DOCK.[15]

The electrostatic score is an interaction energy based on potentials calculated with the DelPhi program.[20,21] DelPhi uses a finite-difference algorithm to solve the Poisson–Boltzmann equation. The resulting electrostatic potential is thought to be more realistic than those of standard force fields[23]; internal and external dielectrics of different magnitudes, nonzero ionic strength, and ion exclusion effects can be modeled.[20,21] We assume that a suitable potential can be calculated using the receptor alone (see the discussion); i.e., the potential is not recalculated in the presence of the ligand.

The program CHEMGRID produces the values for computing force field scores. These scores, or molecular mechanics interaction energies, are calculated as a sum of van der Waals and electrostatic components:

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + 332.0 \frac{q_i q_j}{D r_{ij}} \right], \quad (1)$$

where each term is a double sum over ligand atoms i and receptor atoms j, A_{ij} and B_{ij} are van der Waals repulsion and attraction parameters, r_{ij} is the distance between atoms i and j, q_i and q_j are the point charges on atoms i and j, D is the dielectric function, and 332.0 is a factor that converts the electrostatic energy into kilocalories per mole. Equation (1) contains the *intermolecular* terms present in the AMBER[13] molecular mechanics function, except for an explicit hydrogen-bonding term. We assume that hydrogen bond energies can largely be accounted for in the electrostatic term.[24]

Grid-based scoring can be accomplished efficiently when the ligand and receptor terms in the evaluation function are separable. This is generally true for the electrostatic part of a potential function. For the VDW terms, it is necessary to use a geometric mean approximation[2,24]:

$$A_{ij} = \sqrt{A_{ii}}\sqrt{A_{jj}} \quad \text{and} \quad B_{ij} = \sqrt{B_{ii}}\sqrt{B_{jj}}, \quad (2)$$

where the single-atom-type parameters are calculated from van der Waals radius, R, and well depth, ϵ, according to:

$$A = \epsilon(2R)^{12} \quad \text{and} \quad B = 2\epsilon(2R)^{6}. \quad (3)$$

Using this approximation, eq. (1) can be rewritten as:

$$E = \sum_{i=1}^{lig} \left[\sqrt{A_{ii}} \sum_{j=1}^{rec} \frac{\sqrt{A_{jj}}}{r_{ij}^{12}} - \sqrt{B_{ii}} \sum_{j=1}^{rec} \frac{\sqrt{B_{jj}}}{r_{ij}^{6}} + 332.0 q_i \sum_{j=1}^{rec} \frac{q_j}{Dr_{ij}} \right]. \quad (4)$$

Three values are stored for every grid point k, each a sum over receptor atoms that are within a user-defined cutoff distance of the point:

$$aval = \sum_{j=1}^{rec} \frac{\sqrt{A_{jj}}}{r_{jk}^{12}} \quad bval = \sum_{j=1}^{rec} \frac{\sqrt{B_{jj}}}{r_{jk}^{6}}$$

$$esval = 332.0 \sum_{j=1}^{rec} \frac{q_j}{Dr_{jk}}. \quad (5)$$

These values, with or without interpolation, may subsequently be multiplied by the appropriate ligand values to give the interaction energy.

Input to CHEMGRID includes the grid resolution, location, and dimensions, the form of the dielectric function (constant or distance-dependent), a scaling factor for the dielectric function, a nonbonded cutoff distance, and names of parameter files. Two parameter files are read during a run: a table listing charges and VDW types for atoms in each of the 20 standard amino acids, and a table containing $A^{1/2}$ and $B^{1/2}$ for each VDW type. The receptor parameterization step employs hashing and is very rapid, typically taking less than 1% of the total grid calculation time. The present work uses AMBER united-atom parameters[13] for the receptor, with the exception that all hydrogens bonded to noncarbon atoms are considered volumeless. We would like to emphasize, however, that it is possible to use other parameter sets without changing the code.

Docking

Orientations are generated by finding sets of ligand atoms that match sets of sphere centers, then performing a least-squares superimposition.[25] Sets are considered to match if their pairwise internal distances correspond, within some tolerance. Beginning with DOCK 2.0, a modification involving presorting the distances into "bins" has been employed.[16,17] This allows for more systematic searches of orientation space and greater user control over the thoroughness of the searches.

Scoring

For contact scoring, each ligand atom is assigned the score of the nearest point on the grid. The total score is the sum of the atomic scores.

The DelPhi-calculated potential at each ligand atom is obtained by trilinear interpolation of the values at the eight surrounding grid points. The potential is multiplied by the ligand atom point charge to give the electrostatic interaction energy, and the total energy is the sum of the atomic energies.

Force field scoring requires the retrieval of three grid values. These may be the sums corresponding to the nearest point or the results of trilinearly interpolating the values for the eight surrounding points. Substituting eq. (5) into eq. (4), the interaction energy is:

$$E = \sum_{i=1}^{lig} \left[\sqrt{A_{ii}}(aval) - \sqrt{B_{ii}}(bval) + q_i(esval) \right]. \quad (6)$$

Atoms that fall outside the grid, if any, are given interaction energies of zero. Ligand atoms are associated with parameters at read-in time; the present work uses AMBER all-atom VDW parameters[14] except hydrogens bonded to noncarbon atoms are again considered volumeless.

Test Systems and Run Parameters

Four well-determined crystallographic complexes were chosen from the Brookhaven Protein Data Bank[26,27] (Fig. 2 and Table I): dihydrofolate reductase/methotrexate (4dfr),[28] ribonuclease A/uridine vanadate (6rsa),[29] periplasmic binding protein/glucose (2gbp),[30] and carboxypeptidase A/glycyltyrosine (3cpa).[31] Different aspects of complementarity are evident in these systems, including salt bridge formation, hydrogen bonding, and hydrophobic interactions. In each case, crystallographic waters and ions were removed; the ligand and receptor were separated and hydrogens were added as necessary, in standard geometries. A partial molecular surface was calculated for the receptor, excluding roughly the half of the molecule farthest from the site of interest. This surface was used in SPHGEN, and the largest of the resulting sphere clusters was selected for docking. The DISTMAP (contact grid) calculation included atoms contributing to the molecular surface, as well as additional atoms within 5.0 Å of any surface atom. Polar and nonpolar close contact limits were 2.3 and 2.8 Å, respectively, the maximum distance for a "good" contact was 4.5 Å, and the contact grid spacing was 3 points per Å. DelPhi runs included the entire receptor, with AMBER united-atom partial charges.[13] Three-step focusing,[21] in which the protein occupied 20, 60, and then 90% of the electrostatic potential grid, was used to reduce any errors associated with boundary conditions. Internal and external dielectric constants were 4 and 80, respectively, the ionic strength was 0.145 M, the ion exclusion radius was 2.0 Å, and the probe radius was 1.4 Å. Force field grids were calculated in CHEMGRID, using the entire receptor, 0.3-Å spacing, a 10.0-Å cutoff, and $D = 4r$.[32] Grid boxes, as well as

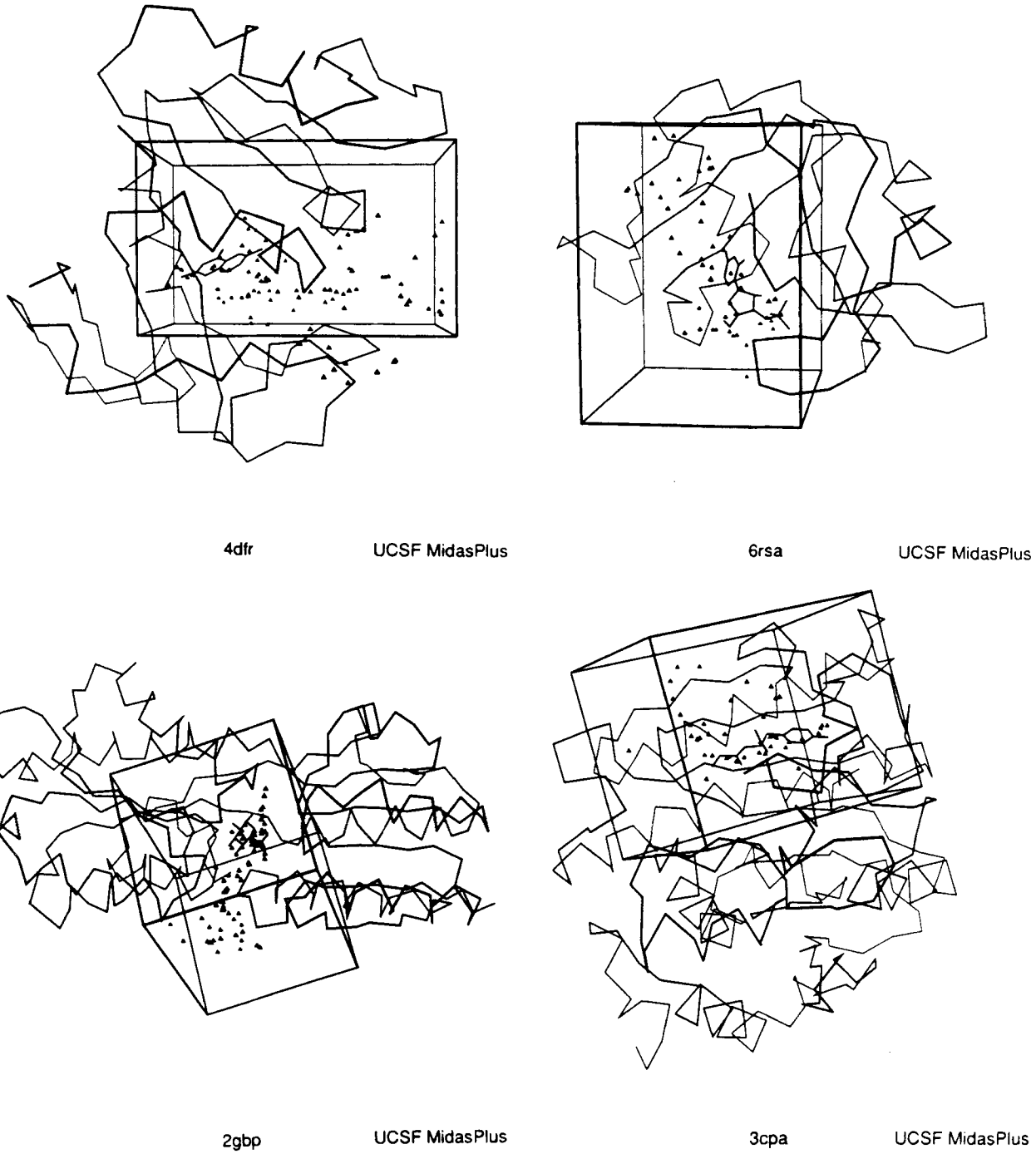


Figure 2. Test systems: Cα representations of the proteins, shown with ligands, sphere centers used for docking (triangles), and boxes outlining the force field grids. Pictures generated with UCSF MidasPlus: Molecular Interactive Display and Simulation, Computer Graphics Laboratory, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA.

the sphere centers used for docking, are shown in Figure 2. Test system ligands are shown in Figure 3.

In DOCK, distance matching parameters (Table II) were chosen such that several thousand sterically allowed orientations were found. The close contact limits set in DISTMAP determine which orientations are sterically acceptable. Molecular mechanics energies were obtained with trilinear interpolation of grid values. Calculations of the root-mean-square deviation (rmsd) from the crystallographic orientation did not include hydrogens.

Time requirements are given in Table III for the predocking calculations and in Table IV for the docking runs. All calculations were performed on Silicon Graphics IRIS 4D/25 workstations with 16 Mb of main memory.

Table I. Test systems.

Brookhaven[a] file	Resolution[b]	Receptor	Complexed ligand	Docked ligand, formal charge
4dfr[c]	1.7	Dihydrofolate reductase	Methotrexate	2,4-Diamino-6-methylpteridine,[d] +1
6rsa[e]	2.0	Ribonuclease A	Uridine vanadate	Uridine 3′-phosphate,[f] −2
2gbp[g]	1.9	Periplasmic binding protein	β-D-Glucose	β-D-Glucose, 0
3cpa[h]	2.0	Carboxypeptidase A	Glycyl-L-tyrosine	Glycyl-L-tyrosine, 0 (zwitterion)

[a]Refs. 26 and 27.
[b]Å.
[c]Ref. 28.
[d]The inflexible part of methotrexate; see text and Fig. 3.
[e]Ref. 29.
[f]Built from uridine vanadate; see text and Fig. 3.
[g]Ref. 30.
[h]Ref. 31.

RESULTS

Results of the docking runs are given in Figures 4–7. The rmsd of the ligand from the crystallographically determined position is plotted vs. score.

In viewing the plots, it should be noted that there is no reason to expect a simple correlation between rmsd and score since the most favorable alternative sites are not necessarily those closest to the crystallographically determined binding site. In addition, since the rmsd represents the collapse of three-dimensional information into a one-dimensional descriptor, there may be more than one orientational family having a particular approximate rmsd value. Whether or not this occurs depends on the symmetry and steric restrictiveness of the site.

Below, we consider each complex and compare the abilities of the scoring functions to identify the orientations closest to the crystallographic geometry.

Dihydrofolate Reductase

N1-protonated 2,4-diamino-6-methylpteridine (Fig. 3) was chosen as the ligand for docking to dihydrofolate reductase. This is the inflexible portion of methotrexate (Fig. 3), in the protonation state believed to be important for binding.[28] Use of the entire methotrexate molecule proved too restrictive for testing scoring methods; relatively few orientations were generated.

STO-3G partial atomic charges[33] were calculated for the ligand and used in docking; 86 spheres were in the cluster of interest and 2617 orientations were written out. The best (highest) contact score (Fig. 4A) corresponds to a low rmsd, although not the lowest. Other families of orientations that receive high contact scores are the 2.8-Å structures, which are barrel-rolled and angled slightly relative to the crystallographic orientation; the 4.0-Å structures, which are angled approximately 90°; and the 4.8-Å structures, which are flipped end-to-end (so that the 6-methyl group is pointing in the opposite direction) and angled slightly. While members of these clusters are also reasonable according to the other scoring functions, members of the lowest-rmsd family receive the best scores (Fig. 4, B–D). The 9.0- and 13.0-Å structures with high contact scores occupy different regions of the site than the experimental orientation.

The use of a somewhat coarser grid, two points per Å, produced similar results, the main effect being an increase in VDW energy for several orientations due to the interpolation approximation. Likewise, using an "infinite" cutoff distance for interactions did not change the output appreciably (data not shown).

Two additional charge sets were generated for the ligand using the Gasteiger–Marsili[34–36] and Gasteiger–Hückel[34–38] options in SYBYL 5.4.[39] These methods are connectivity-based (independent of conformation) and much faster than molecular orbital calculations. For this test system, the Gasteiger–Marsili results are similar to the STO-3G results; the Gasteiger–Hückel charges are less useful for singling out low-rmsd orientations, although members of the lowest-rmsd family again have the best molecular mechanics scores (data not shown).

Ribonuclease A

Uridine 3′-phosphate (Fig. 3) was chosen as the ligand for docking to ribonuclease A so that AMBER all-atom charges[14] could be used. This molecule was constructed from the crystallographic ligand, uridine vanadate (Fig. 3), by changing atom types as necessary and optimizing the phosphate geometry with the Tripos force field.[39]

The cluster for docking contained 47 spheres and 3738 orientations were written out. The rmsd values are somewhat more diffusely distributed than in the dihydrofolate reductase test case (compare Figs. 4 and 5). Of the eight highest contact scores, six correspond to the lowest-rmsd family of orientations (Fig. 5A).

The force field scores (Fig. 5B) and DelPhi scores (not shown) are also able to distinguish the lowest-rmsd dockings from other orientations. The highest-

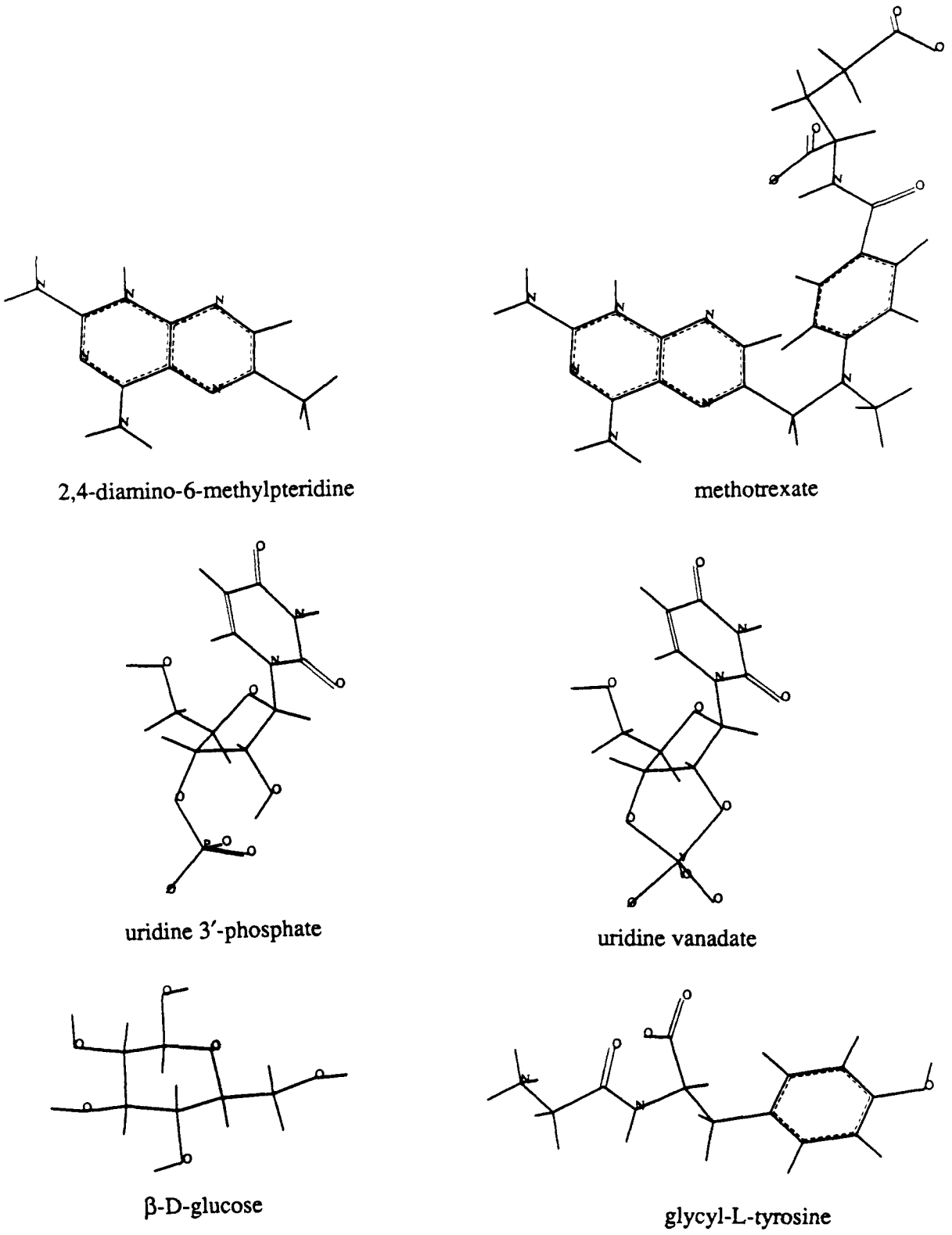


Figure 3. Test system ligands. Pictures generated with SYBYL 5.4: Molecular Modeling System SYBYL, Version 5.4, TRIPOS Associates, Inc., St. Louis, MO.

ranking alternative structures have rmsds of 4.0–6.0 Å and place the phosphate essentially correctly, with the rest of the molecule angled 60–90° relative to the crystallographic orientation. The 10.0- to 13.0-Å structures with favorable force field scores are approximately related to the known orientation by a plane of reflection; the true and image phosphates face each other through the nitrogen of a nearby lysine side chain. Such results are evidence of the weight placed upon charge–charge interactions by the scoring function, especially in this test case where the ligand bears a net charge of −2. However,

Table II. Distance matching parameters (Å).[a]

Test system	Receptor bin width; overlap	Ligand bin width; overlap	Matching tolerance
4dfr	1.0; 0.2	1.0; 0.2	1.5
6rsa	1.0; 0.5	1.0; 0.5	1.5
2gbp	1.0; 0.4	1.0; 0.4	1.5
3cpa	1.5; 0.5	1.5; 0.5	2.0

[a]Receptor site sphere–sphere distances and ligand atom–atom distances are sorted into bins before matching is done. Increasing bin width and overlap increases the number of receptor–ligand internal distance comparisons and thus the number of orientations found. The matching tolerance defines how much distances may differ while still being paired with one another. See ref. 17 for a discussion of these variables.

the total force field score is more helpful in discerning the correct binding mode than the electrostatic component alone.

Virtually indistinguishable results were obtained using Gasteiger–Marsili and Gasteiger–Hückel charges for uridine-3′-phosphate.

Periplasmic Binding Protein

The complex of periplasmic binding protein and glucose was expected to be a relatively difficult test case. Glucose (Fig. 3) bears no net charge and is roughly an oblate ellipsoid. Thus, neither charge nor shape will strongly differentiate among various orientations possible in the context of the site. In addition, periplasmic binding protein has a high affinity for both the α- and β-anomers of D-glucose and D-galactose. The structure of the site suggests that any one of these isomers can participate in 13 hydrogen bonds.[30]

Gasteiger–Marsili charges were calculated for β-D-glucose and used in docking; 75 spheres were in the cluster of interest and 2265 orientations were written out. There are three obvious clusters of rmsd values, corresponding to a family of dockings very similar to the crystallographic orientation, a group of structures with rmsds of 3.0–5.0 Å, and a group with rmsds greater than 10.0 Å (Fig. 6). The intermediate rmsds correspond to orientations that overlay the crystal structure ligand but are flipped or rotated in several different ways. The high rmsds correspond to structures located in either end of the tunnel that traverses the protein (Fig. 2); apparently, there are constrictions that prevent sterically acceptable dockings from being distributed evenly throughout the tunnel. We note that this may pose a problem for methods in which the protein conformation is held constant and the ligand is moved through a representation of real space; a large energy barrier must be surmounted to reproduce the known geometry of the complex.

The simple contact score (Fig. 6A) favors orientations in the correct region of space, but the highest

Table III. Computational time requirements[a] for predocking steps.

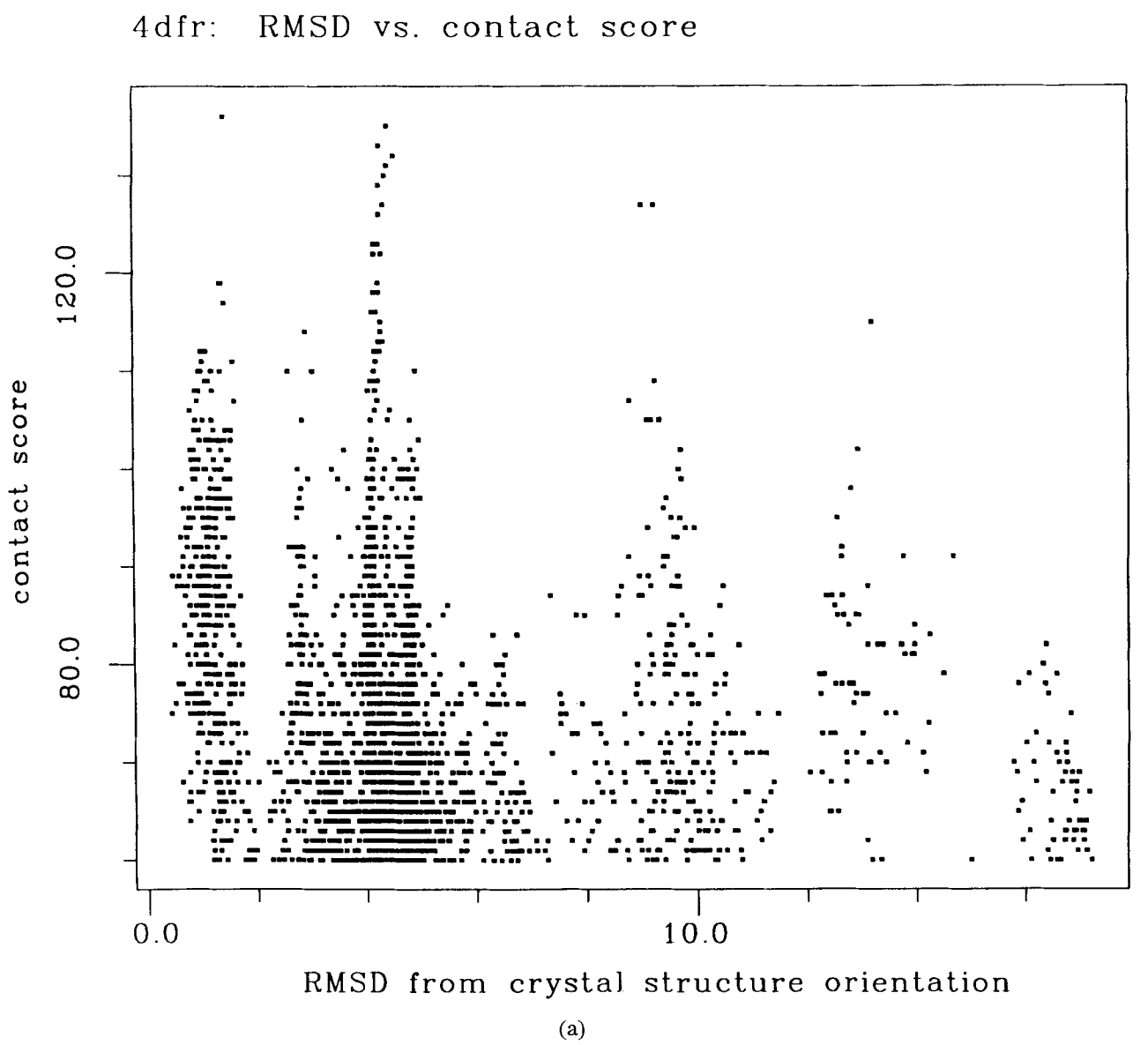
Test system	MS[b,c]	SPHGEN	DISTMAP	CHEMGRID[d,e]	DelPhi[f]
4dfr	2:25	72:11	1:21	16:46	5:07
6rsa	1:02	16:32	1:23	18:56	4:33
2gbp	2:13	27:13	1:47	23:31	4:12
3cpa	0:56	3:37	1:02	16:42	4:40

[a]Min:s on a Silicon Graphics IRIS 4D/25 with 16 Mb of main memory.
[b]Refs. 18 and 19.
[c]Surface area, square Å: 3509 for 4dfr, 1641 for 6rsa, 2108 for 2gbp, and 724 for 3cpa.
[d]0.30-Å spacing, 10.0-Å cutoff.
[e]Number of grid points: 368,475 for 4dfr, 475,190 for 6rsa, 439,280 for 2gbp, and 279,075 for 3cpa.
[f]Refs. 20 and 21; three-step focusing.

Test IV. Computational parameters and time requirements for docking.

Test system	Spheres	Atoms[b]	Found[c]	Written[d]	Times[a] Cw[e]	C[f]	C,DE[g]	C,FF[h]
4dfr	86	13	26,121	2617	3:06	2:13	2:06	2:34
6rsa	47	21	6252	3738	4:43	2:43	2:59	3:29
2gbp	75	12	10,926	2265	5:44	4:54	6:13	5:51
3cpa	44	17	76,684	4327	22:52	20:47	22:12	22:37

[a]Min:s on a Silicon Graphics IRIS 4D/25 with 16 Mb of main memory.
[b]Nonhydrogen atoms in ligand, those used for matching to spheres.
[c]Sterically allowed orientations found; each of these is scored.
[d]Sterically allowed orientations with contact scores greater than a user-specified cutoff.
[e]Contact scoring only, scores and coordinates written out.
[f]Contact scoring only, scores but not coordinates written out.
[g]Contact and DelPhi scoring, scores but not coordinates written out.
[h]Contact and force field scoring, scores but not coordinates written out.



(a)

Figure 4. 4dfr test case using STO-3G charges: rmsd vs. score. (a), contact score, all 2617 orientations with scores of 60 or greater; (b), force field score, the 2404 orientations with energies below 100.0 kcal/mol; (c), electrostatic component of the force field score, all 2617 orientations; (d), DelPhi score, all 2617 orientations.

rankings go to 3.0-Å structures. As expected, electrostatic scores alone are not helpful in identifying the lowest-rmsd dockings. Only the total force field score is successful in identifying structures with rmsds below 1.0 Å (Fig. 6B).

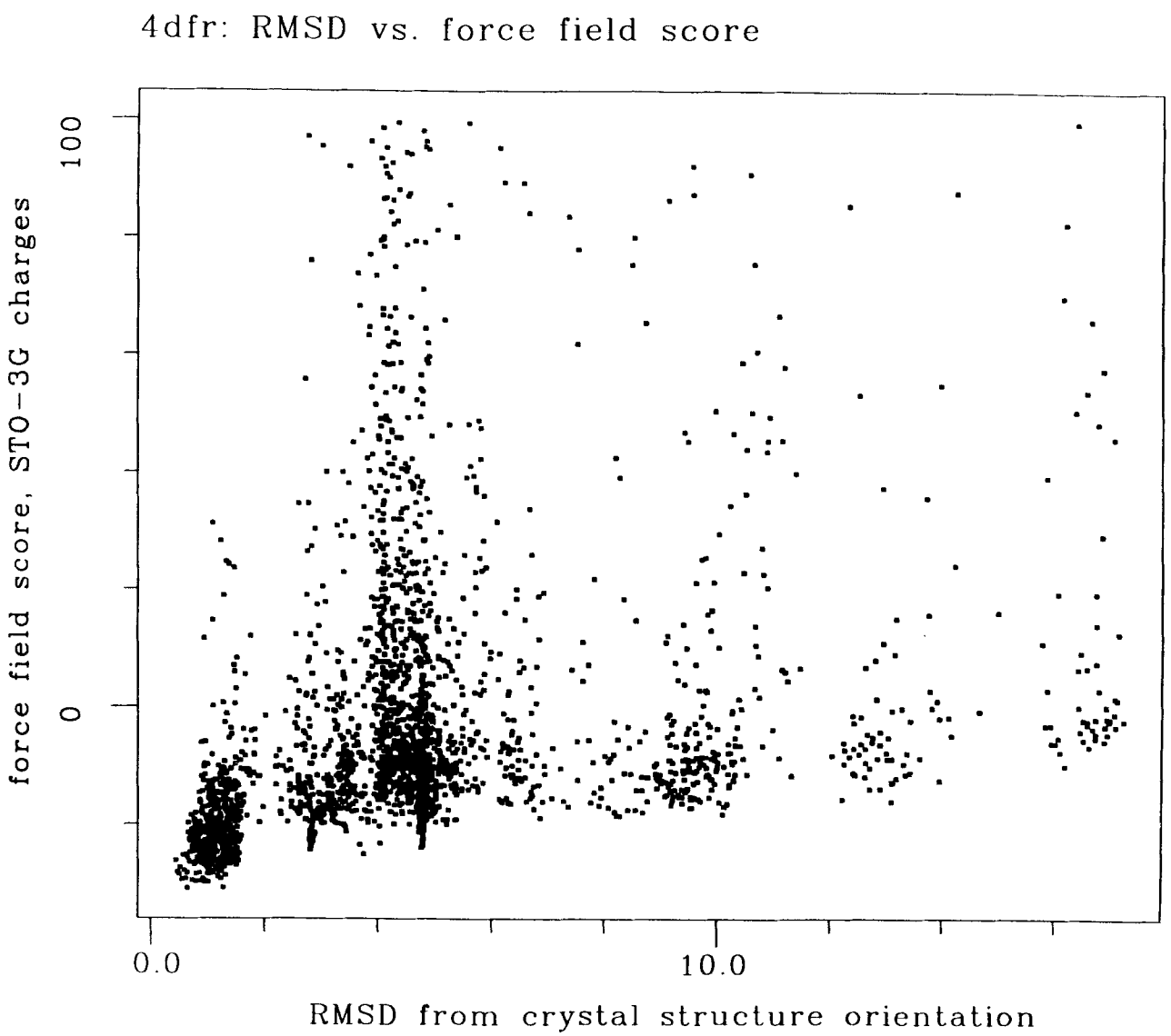
Carboxypeptidase A

AMBER all-atom charges[14] were used for glycyl-L-tyrosine. There were 47 spheres in the cluster of interest and 4327 orientations were written out. Structures with rmsds below 2.0 Å describe essentially the experimental binding mode. Relative to the crystallographic orientation, dockings with rmsds just above 2.0 Å are angled slightly, 3.0- to 4.0-Å structures are barrel-rolled and translated approximately a bond length along the long axis of the molecule, and structures with rmsds greater than 6.0 Å are flipped end-to-end. The best contact score corresponds to a member of the lowest-rmsd family, but good scores are also given to several of the end-to-end flipped structures (Fig. 7A). The same could be said of the DelPhi scores (not shown). For these scoring methods, the best values are outliers. The force field score, however, clearly selects a low-RMSD cluster (Fig. 7B). The lowest rmsd, 0.4 Å, corresponds to the second-best force field score.

Force field score results were essentially the same using Gasteiger–Marsili and Gasteiger–Hückel charge sets for glycyl-L-tyrosine.

Summary

In Table V, we summarize the results of the docking calculations. For each evaluation function, the score of the experimental complexation geometry is com-



(b)

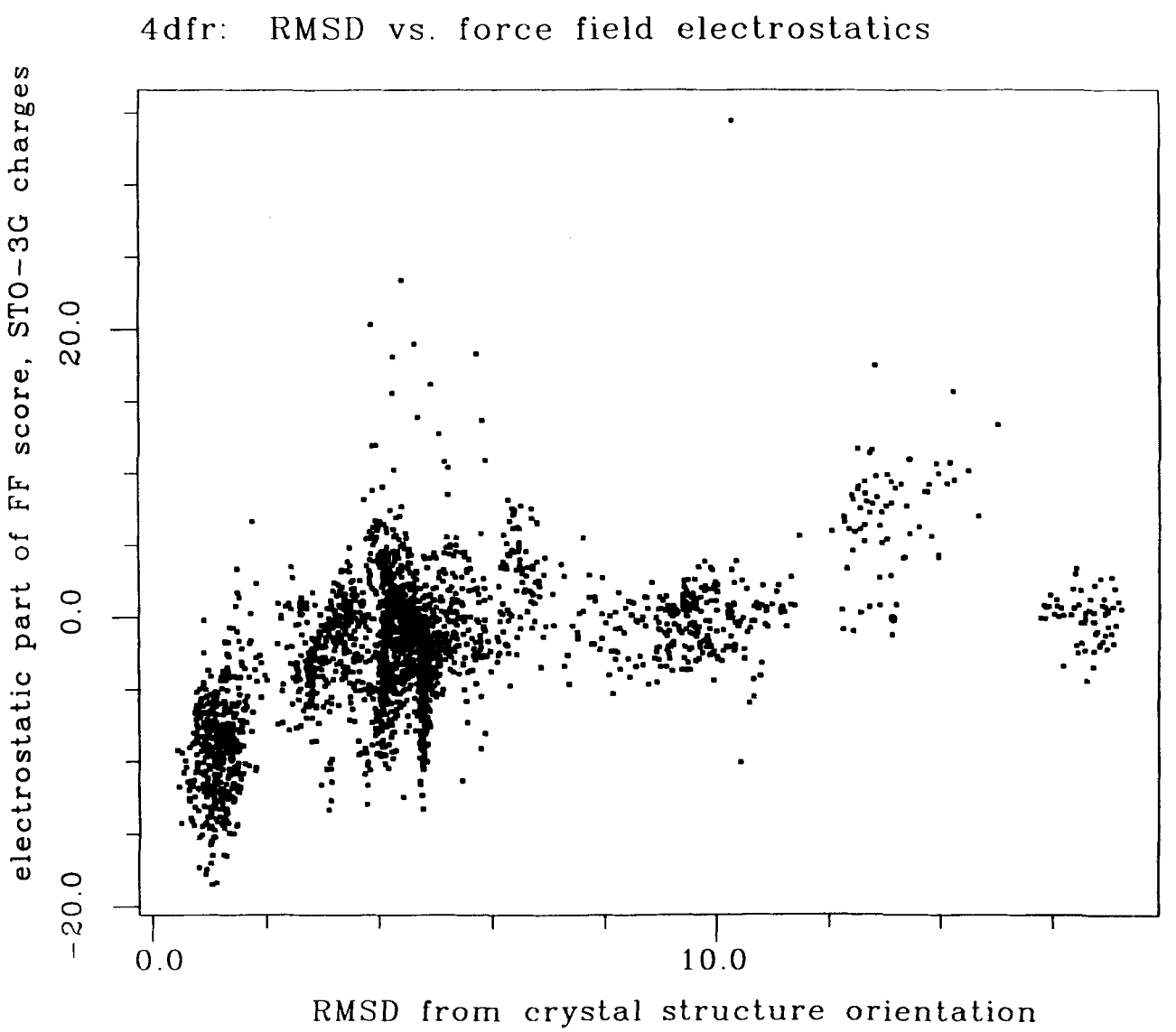
Figure 4. (continued).

pared with the best value found, and the rmsd of the best-scoring orientation is given.

Several features are worth noting. First, only the molecular mechanics score is successful in identifying orientations close to the crystallographic result in all four test cases. The other functions favor alternative modes in one or more of the test cases and, in general, favor dockings with larger rmsds. Second, the force field identification of a family of orientations near the experimental structure is quite robust. This family is well represented in the set of best scores; depending on the test case, the top 8–112 force field scores correspond to members of the lowest-rmsd family (Table VI). Table VII lists the average rmsds for the 10 best orientations according to contact score and force field score. Third, it is apparent that the experimental orientation does not necessarily receive the best score. This is to be expected for the contact score, which is simple and does not include electrostatics, and for the solely electrostatic scores, which do not include VDW energies. For example, a docked ligand that interacts optimally with charges on the receptor often approaches receptor atoms too closely. The force field score is also simplistic and involves numerous assumptions and approximations (see the discussion). Fourth, alternative binding modes with relatively good force field scores are found in each case. These are inherently plausible and may be considered in any design effort.

DISCUSSION

The generation of feasible complexation geometries is important at more than one level to the process of structure-based drug design. Above, we addressed the identification of the preferred mode of binding of a specific conformation of a ligand. Only with this



(c)

Figure 4. (continued).

information can one suggest structural modifications intended to form, enhance, or disrupt specific interactions with the receptor. A further application of automated docking is the discovery of ligands by searching through databases of molecules. To be useful in either task, a docking program should: (1) adequately sample the possible configurations of a ligand–receptor system, (2) score each configuration accurately, and (3) operate in a reasonable amount of time.

We will discuss each of these issues as they pertain to the present work. We will then consider the limitations of our method and compare it to other approaches to the docking problem.

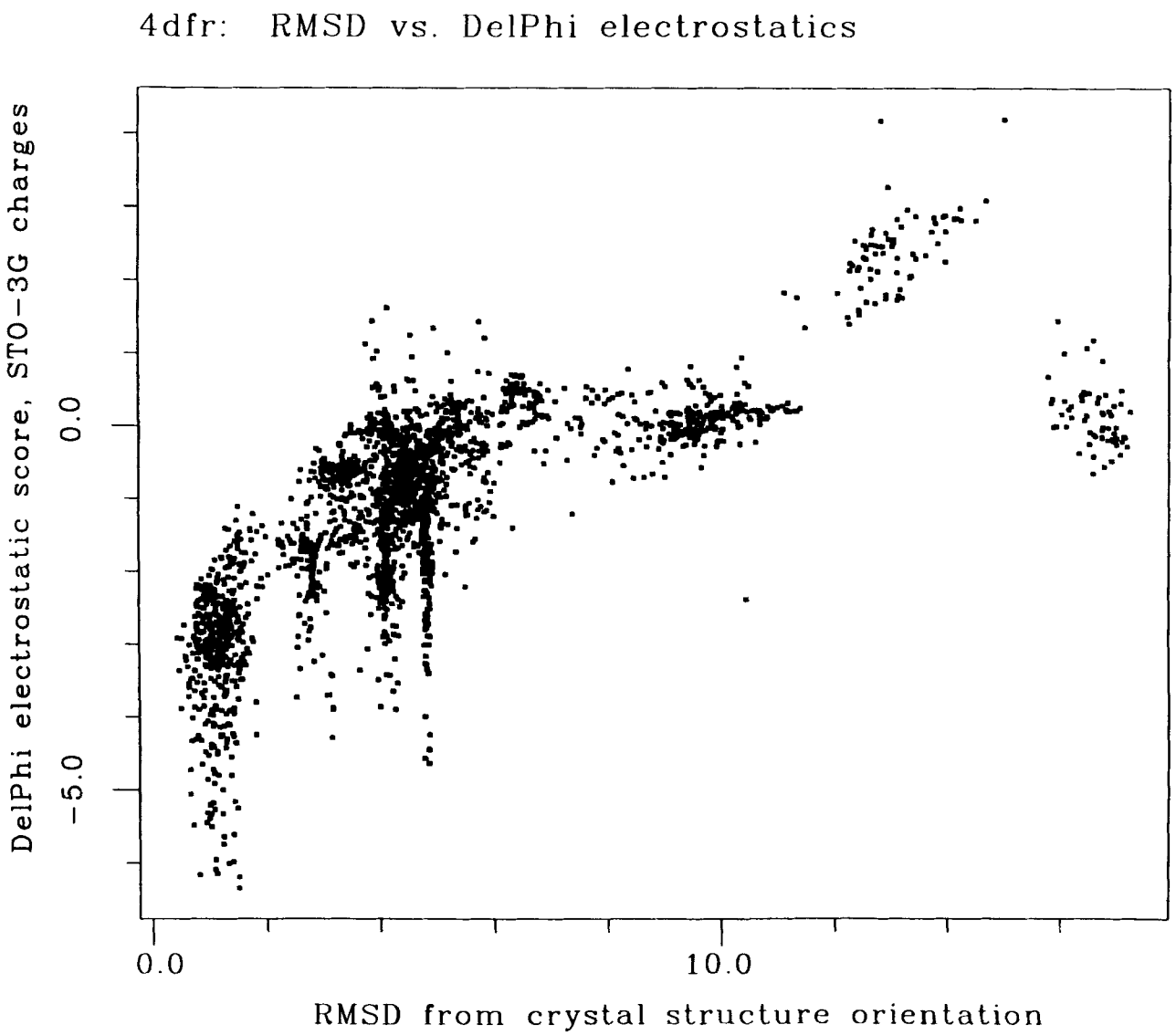
Sampling Configuration Space

DOCK was designed to sample the six degrees of freedom involved in the relative placement of two rigid three-dimensional objects. We find that of the few thousand configurations examined for each test case, tens to hundreds of orientations are close to the experimental result (rmsds under 1.0 Å). This suggests that, in these systems where the ligand and receptor conformations are known, configurational sampling has been performed adequately.

We do not mean to imply that conformational issues are unimportant. In fact, prediction of the conformations of ligand and receptor may be a limiting aspect of the procedure. We have addressed internal degrees of freedom by docking and then linking rigid fragments,[40] adding multiple conformations of molecules to databases for searching, and combining docking with conformational search strategies.[41]

Scoring

A primary focus of this article has been the comparison of methods for assessing orientations of a ligand within a receptor site. We have found that a



(d)

Figure 4. (continued).

molecular mechanics function consistently identifies configurations resembling the crystal structure of the complex for four diverse systems, while the other functions tested do not.

Contact Scores

A simple contact score, as implemented in DOCK version 2.0,[16,17] is used as a measure of shape complementarity. The highest contact scores are associated with low-rmsd dockings in three out of the four cases. However, the number of low-rmsd/high-score orientations and their separation in score from alternative modes are smaller for contact scores than for the other options tested. These problems are most severe when the ligand has a roughly symmetric shape, such as that of a cylinder (glycyl-L-tyrosine) or oblate ellipsoid (β-D-glucose). In any case, shape fitting is only a part of molecular recognition; even a perfect measure of shape complementarity cannot, in general, be expected to identify the preferred geometry of a ligand–receptor complex. We conclude that contact scoring is most useful for discarding orientations that overlap receptor atoms and finding templates for lead compounds.

Electrostatic Scores

We have used two different functions to calculate electrostatic potentials: a simple Coulombic form with a distance-dependent dielectric function and the linearized Poisson–Boltzmann equation, solved numerically with a finite-difference algorithm as implemented in DelPhi.[20,21] In addition, both quantum-mechanical and connectivity-based partial charge sets have been evaluated.

To a first approximation, the electrostatic scoring options are equally successful when the ligand bears

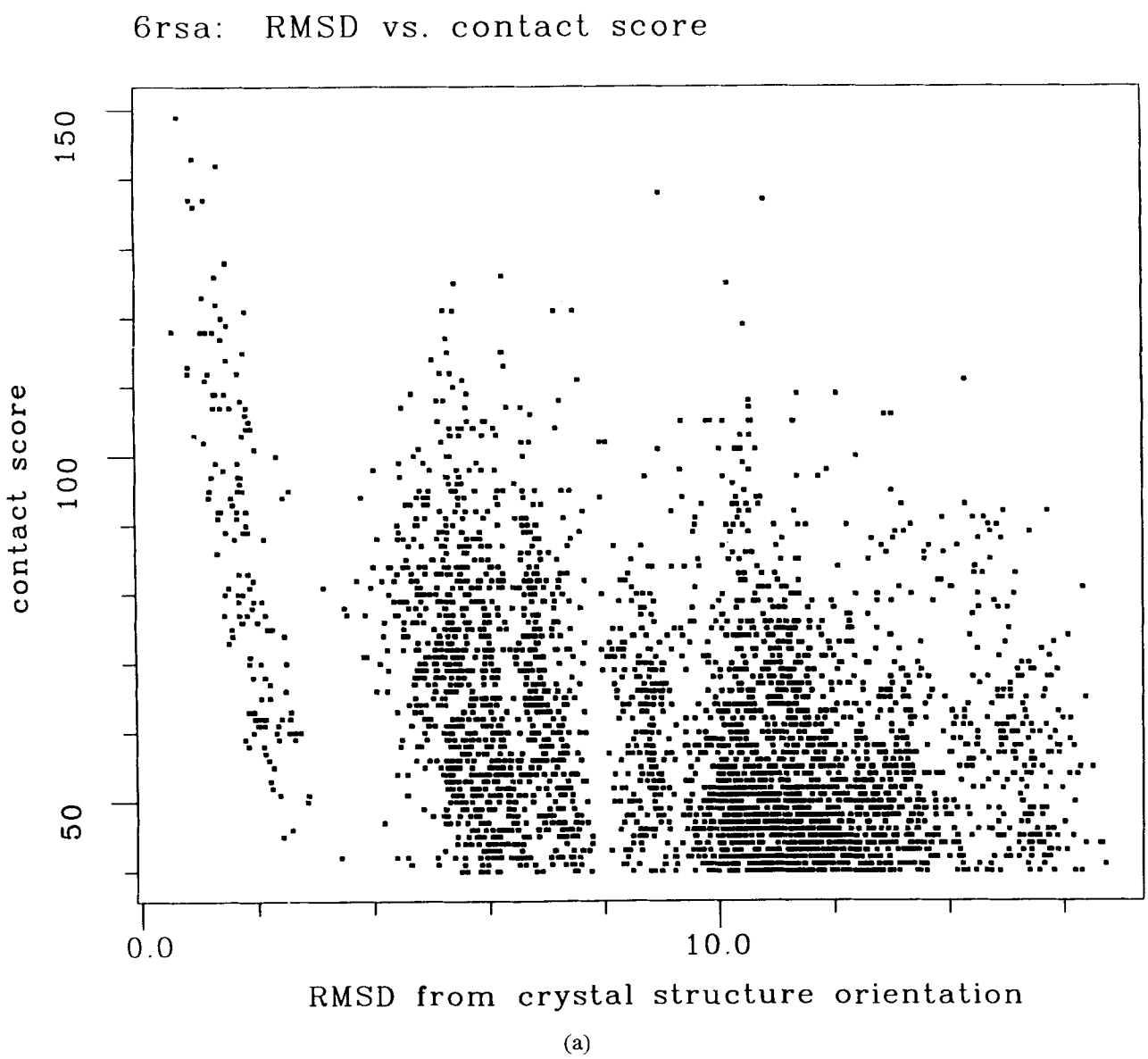


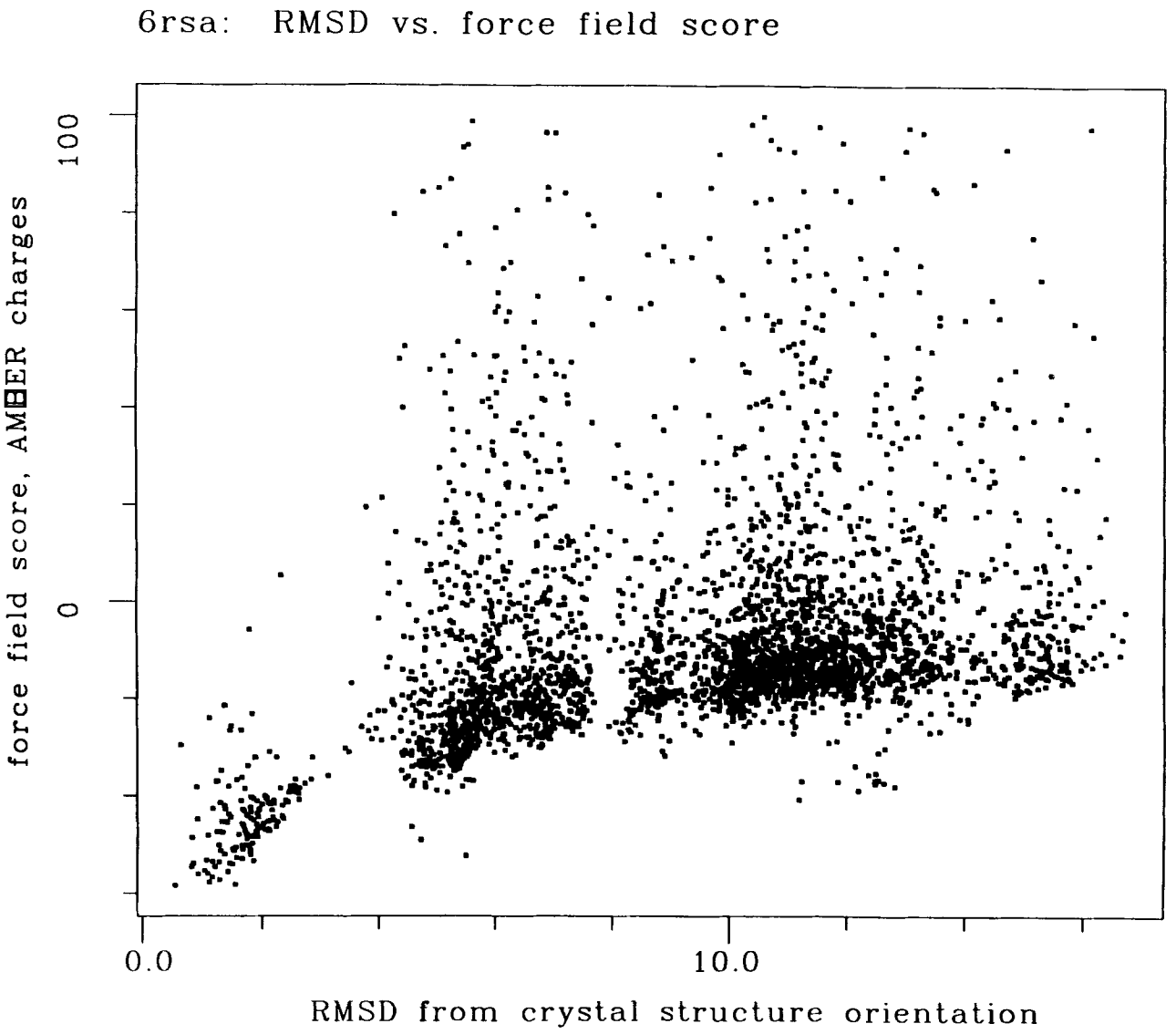
Figure 5. 6rsa test case using AMBER charges: rmsd vs. score. (a), contact score, all 3738 orientations with scores of 40 or greater; (b), force field score, the 3489 orientations with energies below 100.0 kcal/mol.

a formal charge. When the ligand is not formally charged, differences are more evident, in part because the connectivity-based charges tend to be smaller than those derived quantum-mechanically. Each of the electrostatic scoring options is less helpful than the total force field score, however.

It must be stressed that the way in which we have computed DelPhi electrostatic interaction energies is not rigorously correct and that using different parameters for calculating the potential map may affect the results. A more rigorous application of DelPhi to calculate electrostatic interaction energies in solution has been described,[42] that involves evaluation of a full thermodynamic cycle including the bound and unbound states of the molecules. This requires the DelPhi program to be run for each ligand–receptor geometry, an option not feasible in our application. The approach we used leads to underestimation of favorable interactions because dipoles induced upon binding and solvent exclusion by the ligand are not modeled.[43] It may be helpful to approximate solvent exclusion by considering the spheres to be regions of low dielectric when calculating the receptor potential (B.K. Shoichet, unpublished results).

Force Field Scores

Of the options investigated, the force field score is the most successful in identifying ligand-receptor configurations that resemble the experimental geometry (Table V). Several points deserve mention. First, while each DOCK run produced numerous low-rmsd orientations, these receive a wide range of force field scores due to the sensitivity of the 6–12



(b)

Figure 5. (continued).

potential to small displacements. This is not a limitation in practice, as only the most favorable orientations are kept for further study.

Second, we find that grid-based scoring preserves, to a large extent, the rank-ordering of orientations yielded by continuum (nongrid) calculations. To examine the effects of the grid approximation, interaction energies calculated within the AMBER analysis module were compared to grid-based interaction energies (4dfr test system; results not shown). Using grid spacings of 0.5 Å or less and trilinear interpolation, electrostatic interaction energies were reproduced to within a kcal/mol. Net favorable VDW energies were also matched reasonably well (within a few kcal/mol); however, net unfavorable (positive) VDW energies differed by as much as several thousand kcal/mol. This is not surprising, as the VDW potential surface rises steeply as contacts become too close. Interpolation preserves the rank-ordering of orientations found with the continuum calculations better than does simply using the values for the grid point nearest to each ligand atom. Overall, it is apparent that the grid approximation does not degrade the results for the most favorable orientations.

Third, for all cases and for all scoring methods but one, the DOCK procedure found structures that scored better than the crystallographically determined position of the ligand. Simply stated, free energy determines the binding, but free energy is not entirely represented by the scoring functions we have used. As described in the computational methods section, the contact score is a rough measure of shape complementarity; charge–charge interactions are not considered. Conversely, the DelPhi score and the electrostatic component of the force field score do not include steric contributions. The total force field score, while combining these two important

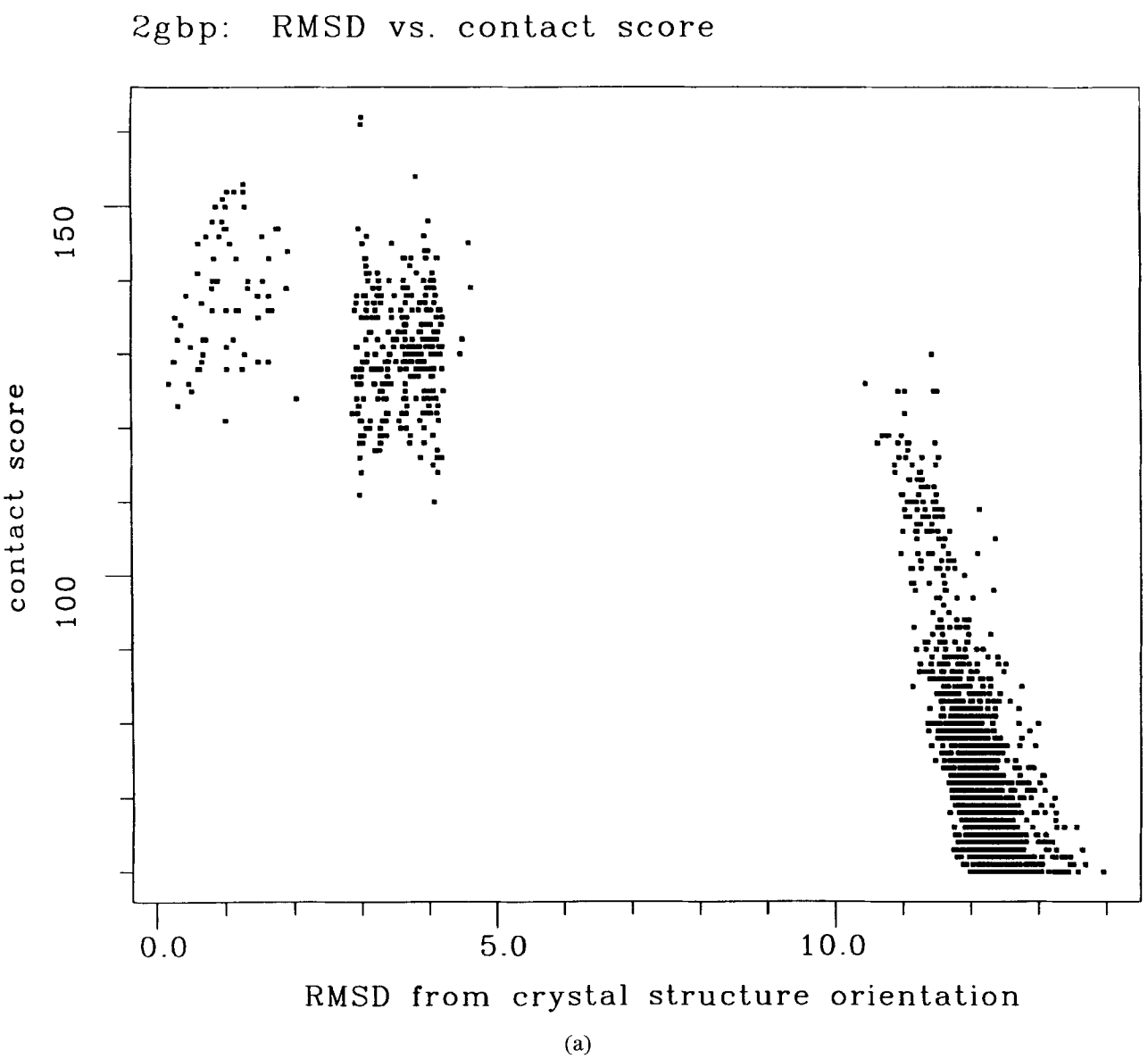
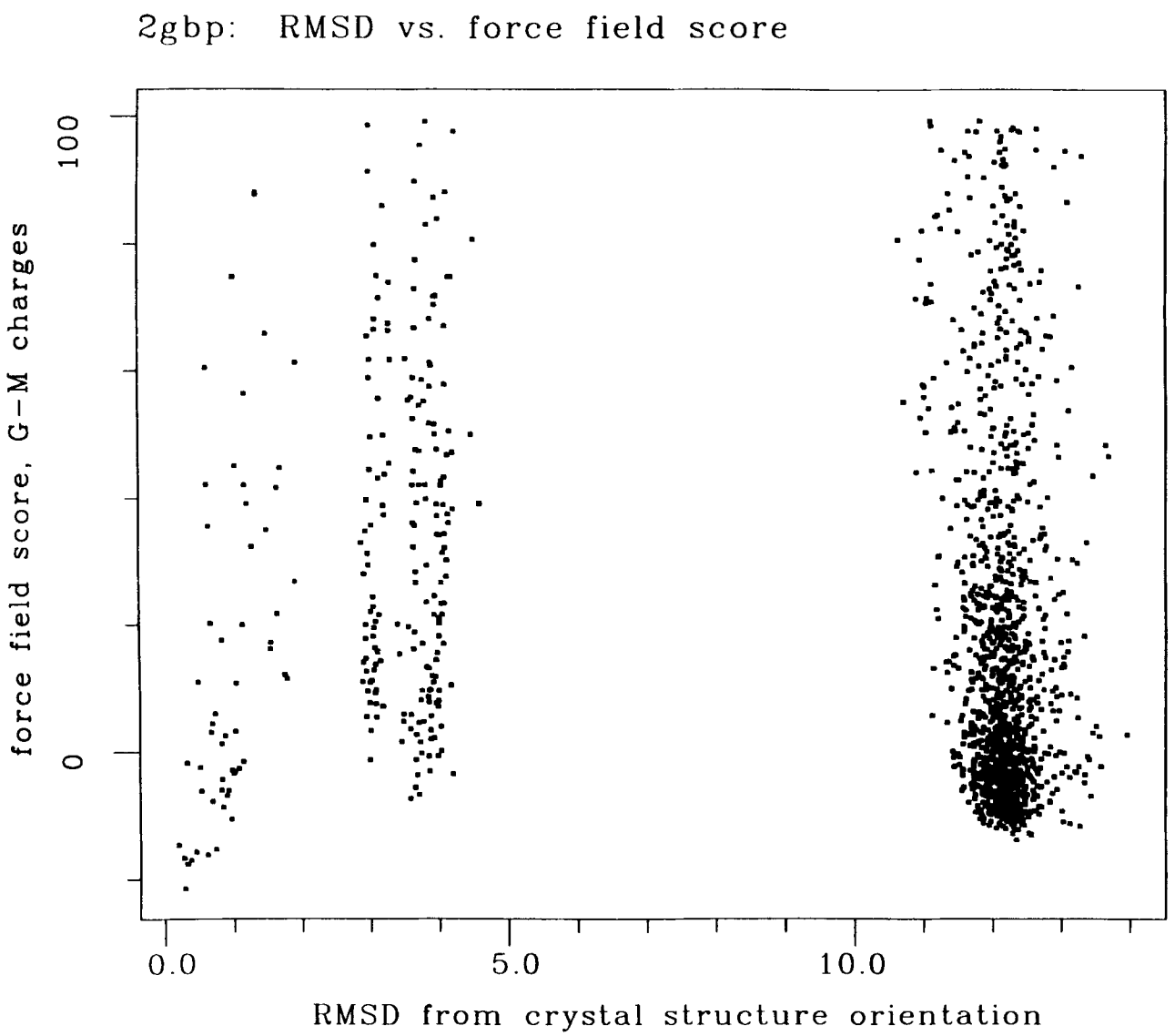


Figure 6. 2gbp test case using Gasteiger–Marsili charges: rmsd vs. score. (a), contact score, all 2265 orientations with scores of 60 or greater; (b), force field score, the 1680 orientations with energies below 100.0 kcal/mol.

aspects of molecular recognition, represents (at best) an estimate of the enthalpy of interaction. The calculation of entropy, and thus of free energy, requires sampling of a statistical ensemble of system configurations, as may be obtained through molecular dynamics or Monte Carlo simulations. Furthermore, a rigorous representation of events in solution must include explicit solvent molecules, and calculation of the free energy of binding requires evaluation of the unbound as well as of the bound structures. These considerations apply to any molecular mechanics study of complexation energetics. Approximations in addition to those of standard force field calculations include discretization of space (the grid approximation) and neglect of intramolecular terms.

There are limitations inherent in the structure of the receptor, as well as in the scoring functions. Any crystallographic study yields a structure that contains the effects of static and thermal disorder. Only average atomic positions can be derived from the diffraction data, and these do not necessarily match the coordinates for any single molecule within the crystal lattice. Slight bias may result from the use of potential energy terms during refinement. Finally, the placement of hydrogens in "standard geometries" introduces some uncertainty.

For the reasons stated above, it is important not to overinterpret the various types of scores. Even when the units returned are kcal/mol, the results cannot be converted into absolute or relative binding affinities. For this reason, we prefer to call the calculated quantities "scores" rather than "energies," even when a molecular mechanics function is used.



(b)

Figure 6. (continued).

Time Requirements

The calculations necessary for docking studies, as described above, can be divided into two phases. The first phase includes calculation of a molecular surface, generation of site-filling sphere clusters, and creation of grids for scoring. These steps are done once per receptor, and in our systems took approximately 1 h on a Silicon Graphics IRIS 4D/25 (Table III). The time spent calculating the force field grids depends not only on the number of grid points but also on the cutoff distance, the total number of receptor atoms, the number of receptor atoms within the cutoff distance of each grid point (i.e., the shape of the receptor and the location of the grid box), and whether the dielectric function is distance-dependent or constant.

The second phase of the process is docking itself. The times spent in DOCK (Table IV) depend strongly on the size of the ligand, the number of spheres used, and the distance matching tolerances. Notably, the penalty for performing force field or DelPhi scoring in addition to contact scoring is relatively small. Thus, the time costs of more sophisticated calculations can, in part, be shifted to the predocking stage and traded for increased usage of physical memory.

Limitations

It is important to point out that the method described here does not address internal degrees of freedom,[1,10,40] pay special attention to hydrogen bonds,[3,10,11] keep track of surface area burial[1,7,11] or solvation energy,[4] or include energy minimization.[1–4,9,10] These capabilities are present to various

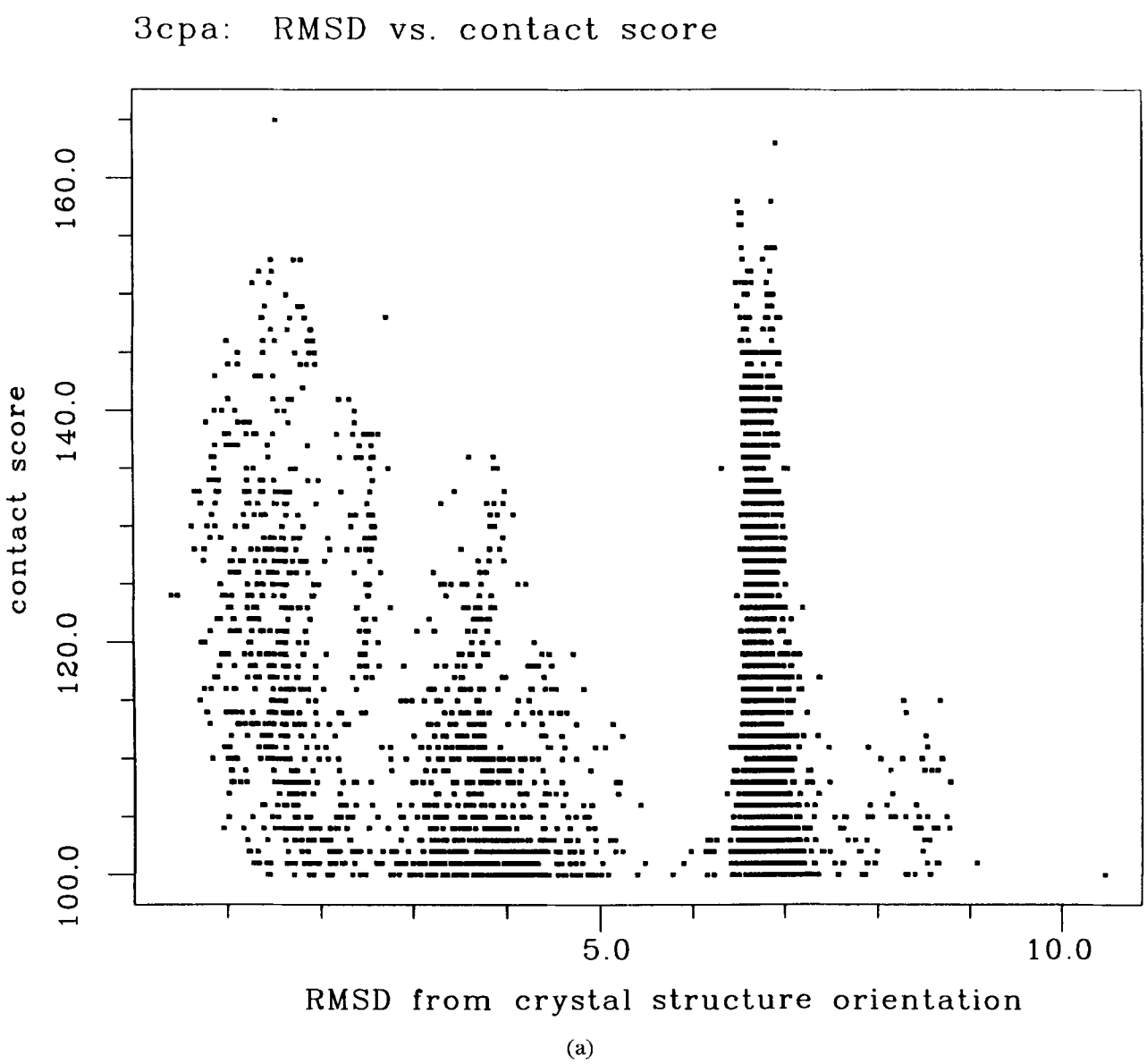


Figure 7. 3cpa test case using AMBER charges: rmsd vs. score. (a), contact score, all 4327 orientations with scores of 100 or greater; (b), force field score, the 2684 orientations with energies below 100.0 kcal/mol.

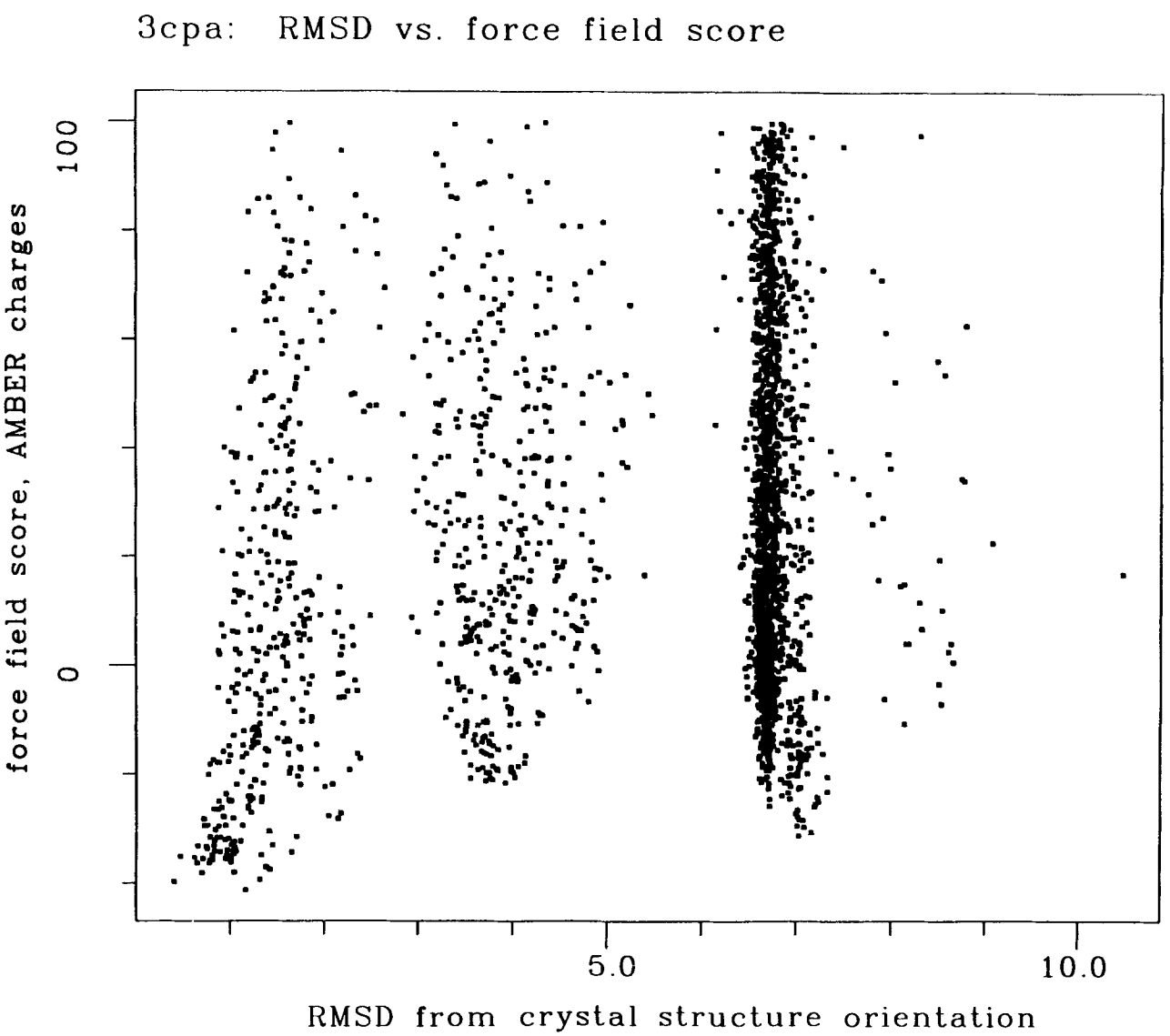
extents in some of the other docking algorithms, albeit at a computational cost.

Our goal in ligand-design applications has been to find lead compounds in an efficient way rather than trying to find every molecule in the database that might bind to the receptor. Some potential leads may be missed because they are in the wrong conformation for binding.

Full energy-minimization of docked complexes requires parameterization of each "ligand" molecule. This is difficult when large numbers of compounds are to be evaluated. Some docking methods[2-4] employ rigid-body minimization, in which there are no intramolecular degrees of freedom; only nonbonded parameters are required. Although minimization is useful for finding local optima, it adds to the costs of computation. Docking alone can be much faster but is prone to missing optimal geometries. The algorithm presented here is functionally equivalent to rigid-body minimization from multiple starting configurations, as long as the sampling of orientations within the site is sufficiently dense. When thousands of orientations per molecule are produced, as in the test cases above, each is related to many others by very slight rigid-body movements. The optimum can be singled out according to score.

Comparison to Other Methods

The force field grid is not conceptually novel. Goodford's GRID program calculates interaction energies for probes of several types at grid points within a site.[12] Most similar to our method, however, are programs that use force field grids to evaluate docked structures. We will first address interactive docking algorithms. Pattabiraman et al. used the geometric



(b)

Figure 7. (continued).

mean approximation and the same evaluation function we use, allowing for either a distance-dependent or constant dielectric function[2]; an advantage of their approach is that the grid resolution may differ for the electrostatic and steric parts of the calculation. This allows one to grid space more finely for the evaluation of VDW energies, which are very sensitive to small displacements. Tomioka and coworkers[3] used a similar molecular mechanics function but did not employ the geometric mean approximation for VDW parameters; the interaction energies for multiple types of probe atoms are stored, as in the GRID program. In addition, their method allows ligand bonds to be rotated and counts the number of intermolecular hydrogen bonds that are formed.

Not fully addressed in some of these articles are the issues of user control over the grid calculation and wide applicability to a number of receptor structures. It is most convenient when, as in our approach, the user can specify the grid location, dimensions, and resolution, the cutoff distance for interactions, and the dielectric function without changing the code itself. Because the grid values are stored in one-dimensional arrays, any combination of spacing and x, y, and z extents may be used as long as the total number of points does not exceed the array size (10^6 in the current work).

Our work to date suggests that the geometric mean approximation is useful and that it is not necessary to store values for probes of several types. Given the steepness of the VDW potential, memory is better spent on grids with finer spacing. Since hydrogen-bonding groups are generally not allowed to respond to the docked ligand, it seems that it is beyond the resolution of the method to place any particular emphasis on hydrogen bonds. In addition, such calculations would increase computational time since angles, as well as distances, must be taken

Table V. Comparison of crystallographic and best-scoring orientations.

System	Contact		Force field (FF)		FF electrostatic		DelPhi electrostatic	
	Score	rmsd[a]	Score[b]	rmsd[a]	Score[b]	rmsd[a]	Score[b]	rmsd[a]
4dfr[c]								
Crystal orientation	81	—	−29.085	—	−11.501	—	−3.001	—
Best-scoring orientation	136	1.42	−30.829	0.64	−18.456	1.04	−6.347	1.50
6rsa[d]								
Crystal orientation	134	—	−61.389	—	−42.446	—	−13.719	—
Best-scoring orientation	149	0.65	−58.429	0.56	−49.108	4.42	−16.214	0.92
2gbp[e]								
Crystal orientation	130	—	−16.995	—	−4.452	—	−0.234	—
Best-scoring orientation	162	3.01	−21.483	0.29	−8.397	11.81	−4.252	3.34
3cpa[d]								
Crystal orientation	94[f]	—	−25.167[g]	—	−19.640	—	−3.565	—
Best-scoring orientation	165	1.52	−41.302[h]	1.16	−32.418	0.98	−11.914	1.86

[a]Å, relative to crystal structure orientation.
[b]kcal/mol.
[c]STO-3G charges for ligand.
[d]AMBER all-atom charges for ligand.
[e]Gasteiger–Marsili charges for ligand.
[f]Does not include the scores of two atoms that violate the close contact limits set in DISTMAP; see text.
[g]AMBER minimization with a 1000.0-Å cutoff, allowing only the ligand atoms to move, results in an interaction energy of −57.254 kcal/mol and an rmsd of 0.45 Å relative to the unminimized crystal structure orientation.
[h]AMBER minimization with a 1000.0-Å cutoff, allowing only the ligand atoms to move, results in an interaction energy of −54.599 kcal/mol and an rmsd of 0.94 Å relative to the unminimized crystal structure orientation.

into account. As mentioned previously, we feel it is most efficient to perform rigid-body docking not only because of the time required to consider torsional degrees of freedom but also because specification of which bonds are rotatable and generation of the corresponding parameters are neither trivial nor easily automated.

Automated docking is preferable to interactive docking for database searching. Furthermore, the results of automated docking are not so strongly dependent on the preconceptions of the user and in general a greater region of orientation space is explored. Many of the automated methods, however, have only been applied to systems smaller than macromolecule–ligand complexes,[9] or to reduced representations of molecules, so detail at the level of individual atoms is not considered.[4]

Goodsell and Olson used Monte Carlo simulated annealing with grid-based energy evaluation to dock molecules automatically.[10] Interaction energies were calculated for different probe types using the AMBER function[13] but with the 10-12 term scaled by a factor of 10 to give a potential well of −4 kcal/mol and $D = 40$. Complexes of known structure were examined as test cases. Each simulation began with the ligand in the rough vicinity of the site and proceeded with incremental rigid-body movements and bond rotations. Several simulations were carried out for each complex, with the correct structure being found and given the best energy in nearly all cases. Advantages of this method include consideration of ligand flexibility, reasonable computational demands, and insights that may be afforded by the simulation trajectories.

There is a fundamental difference between our docking method and the Monte Carlo simulated an-

Table VI. Separation of best-scoring orientational families in force field score and rank.

System	Δ(force field score)[a](kcal/mol)	Δ(rank)[a]
4dfr[b]	5.8	113
6rsa[c]	6.2	23
2gbp[d]	7.7	9
3cpa[c]	10.3	46

[a]Difference between the top-ranked orientations of the two best-scoring families.
[b]STO-3G charges for ligand used to calculate force field score.
[c]AMBER all-atom charges for ligand used to calculate force field score.
[d]Gasteiger–Marsili charges for ligand used to calculate force field score.

Table VII. Average rmsds for the 10 best-scoring orientations.

System: average rmsd[a] (range[a])	Type of score	
	Contact	Force field
4dfr[b]	5.0 (1.4–9.2)	0.9 (0.5–1.3)
6rsa[c]	2.8 (0.7–10.8)	1.1 (0.6–1.6)
2gbp[d]	1.7 (0.9–3.8)	2.8 (0.2–12.3)[e]
3cpa[c]	6.1 (1.5–6.9)	1.0 (0.4–1.4)

[a]Å, relative to crystal structure orientation.
[b]STO-3G charges for ligand used to calculate force field score.
[c]AMBER all-atom charges for ligand used to calculate force field score.
[d]Gasteiger–Marsili charges for ligand used to calculate force field score.
[e]The average and range for the eight best force field scores are 0.4 and 0.2–0.7 Å, respectively.

nealing approach. Our procedure is not carried out within a representation of Cartesian space; instead, it depends only on internal distance matching. Thus, there is no dependence on the starting locations of the molecules, and there are no effects due to steric hindrance or unfavorable charge–charge interactions en route to the site. Molecules may be docked successfully even when there is no low-energy pathway from the outside of the protein to the binding site.

Some aspects of the energy function used for the simulated annealing merit discussion. Since hydrogen bond energies in AMBER are included primarily in the electrostatic term, scaling the 10-12 term by a factor of 10 is equivalent to weighting hydrogen bonds doubly in the overall interaction energy. In addition, although bond rotations are allowed, internal energies are not included in the calculation.

Docking by internal distance matching is quite rapid. While our algorithm may require more computational time than the simulated annealing procedure for tasks that are done once per site, prior to docking, the time per low-energy orientation generated is encouragingly small. The distance-matching algorithm is flexible as well as powerful in that the user may easily vary the thoroughness of the procedure and the number of sterically allowed orientations that will be found.

CONCLUSION

In summary, we added molecular mechanics scoring capabilities to a rapid, geometric docking algorithm. Computational costs are kept to a minimum by precalculating values on three-dimensional grids. Four crystallographic complexes are used as test cases, in which the small molecule component is docked back into the receptor; the results are encouraging as the force field score is able to identify the correct family of orientations in each case. Scoring methods that consider solely sterics or solely electrostatics are less successful. Many approximations are inherent in the method; however, we feel that a reasonable balance between rigor and computational tractability has been achieved. Since the results of database searching are highly dependent on the scoring function, improving the evaluation of orientations of a single molecule is an important step in improving the effectiveness of searching for lead compounds.

The authors gratefully acknowledge support from NIH grants GM-31497 (I.D.K.) and GM-39552 (G.L.K.), DARPA grant MDA-91-J-1013 (F.E.C.), and Glaxo Inc.; thanks are due also to K.A. Sharp, A.R. Leach, D.A. Pearlman, P.A. Kollman, R. Langridge, and the UCSF Computer Graphics Laboratory for their advice and assistance. The SYBYL molecular modeling package has been useful in many aspects of our work; we thank Tripos Associates, Inc. (St. Louis, MO), for making this package available to us.

SPHGEN, DISTMAP, and contact scoring are included in DOCK version 2.0[16,17]; CHEMGRID and force field scoring will be included in the release of DOCK version 3.0. DOCK and associated programs are implemented in Fortran77 and available from I.D.K.

References

1. B. Busetta, I.J. Tickle, and T.L. Blundell, *J. Appl. Cryst.*, **16,** 432 (1983).
2. N. Pattabiraman, M. Levitt, T.E. Ferrin, and R. Langridge, *J. Comp. Chem.*, **6,** 432 (1985).
3. N. Tomioka, A. Itai, and Y. Iitaka, *J. Comp.-Aided Mol. Design*, **1,** 197 (1987).
4. S.J. Wodak and J. Janin, *J. Mol. Biol.*, **124,** 323 (1978).
5. I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, and T.E. Ferrin, *J. Mol. Biol.*, **161,** 269 (1982).
6. D. Goodsell and R.E. Dickerson, *J. Med. Chem.*, **29,** 727 (1986).
7. M.L. Connolly, *Biopolymers*, **25,** 1229 (1986).
8. M. Billeter, T.F. Havel, and I.D. Kuntz, *Biopolymers*, **26,** 777 (1987).
9. K.B. Lipkowitz and R. Zegarra, *J. Comp. Chem.*, **10,** 595 (1989).
10. D.S. Goodsell and A.J. Olson, *Proteins*, **8,** 195 (1990).
11. F. Jian and S.-H. Kim, *J. Mol. Biol.*, **219,** 79 (1991).
12. P.J. Goodford, *J. Med. Chem.*, **28,** 849 (1985).
13. S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., and P. Weiner, *J. Am. Chem. Soc.*, **106,** 765 (1984).
14. S.J. Weiner, P.A. Kollman, D. T. Nguyen, and D.A. Case, *J. Comp. Chem.*, **7,** 230 (1986).
15. R.L. DesJarlais, R.P. Sheridan, G.L. Seibel, J.S. Dixon, I.D. Kuntz, and R. Venkataraghavan, *J. Med. Chem.*, **31,** 722 (1988).
16. B.K. Shoichet and I.D. Kuntz, *J. Mol. Biol.*, **221,** 327 (1991).
17. B.K. Shoichet, D.L. Bodian, and I.D. Kuntz, *J. Comp. Chem.* (in press).
18. M.L. Connolly, *J. Appl. Crystallogr.*, **16,** 548 (1983).
19. M.L. Connolly, *Science*, **221,** 709 (1983).
20. I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, *Proteins*, **1,** 47 (1986).
21. M.K. Gilson, K.A. Sharp, and B.H. Honig, *J. Comp. Chem.*, **9,** 327 (1987).
22. F.M. Richards, *Annu. Rev. Biophys. Bioeng.*, **6,** 151 (1977).
23. A.R. Fersht and M.J.E. Sternberg, *Prot. Eng.*, **2,** 527 (1989).
24. A.T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96,** 5319 (1977).
25. D.R. Ferro and J. Hermans, *Acta Crystallogr.*, **A33,** 345 (1977).
26. F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, *J. Mol. Biol.*, **112,** 535 (1977).
27. E.E. Abola, F.C. Bernstein, S.H. Bryant, T.F. Koetzle, and J. Weng, in *Crystallographic Databases: Information Content, Software Systems, Scientific Applications*, F.H. Allen, G. Bergerhoff, and R. Seivers, eds., Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, pp. 107–132.
28. J.T. Bolin, D.J. Filman, D.A. Matthews, R.C. Hamlin, and J. Kraut, *J. Biol. Chem.*, **257,** 13650 (1982).
29. B. Borah, C.-W. Chen, W. Egan, M. Miller, A. Wlodawer, and J.S. Cohen, *Biochemistry*, **24,** 2058 (1985).

30. N.K. Vyas, M.N. Vyas, and F.A. Quiocho, *Science*, **242,** 1290 (1988).
31. W.N. Lipscomb, private communication.
32. M. Whitlow and M.M. Teeter, *J. Am. Chem. Soc.*, **108,** 7164 (1986).
33. U.C. Singh and P.A. Kollman, *J. Comp. Chem.*, **5,** 129 (1984).
34. J. Gasteiger and M. Marsili, *Tetrahedron*, **36,** 3219 (1980).
35. M. Marsili and J. Gasteiger, *Croat. Chem. Acta*, **53,** 601 (1980).
36. J. Gasteiger and M. Marsili, *Organ. Magn. Reson.*, **15,** 353 (1981).
37. A. Streitweiser, *Molecular Orbital Theory for Organic Chemists*, Wiley, New York, 1961.
38. W.P. Purcel and J.A. Singer, *J. Chem. Eng. Data*, **12,** 235 (1967).
39. Molecular Modeling System SYBYL, Version 5.4, TRIPOS Associates, Inc., St. Louis, MO, January 1991.
40. R.L. DesJarlais, R.P. Sheridan, J.S. Dixon, I.D. Kuntz, and R. Venkataraghavan, *J. Med. Chem.*, **29,** 2149 (1986).
41. A.R. Leach and I.D. Kuntz, *J. Comp. Chem.* (in press).
42. M.K. Gilson and B. Honig, *Proteins*, **4,** 7 (1988).
43. M.E. Davis and J.A. McCammon, *J. Comp. Chem.*, **11,** 401 (1990).