Molecular Surface Representations by Sparse Critical Points

Shuo Liang Lin,[1] Ruth Nussinov,[1,2,3] Daniel Fischer,[3,4] and Haim J. Wolfson[4]

[1]Laboratory of Mathematical Biology, National Cancer Institute-FCRF, and [2]PRI/Dynacorp, Frederick, Maryland 21702; [3]Sackler Institute of Molecular Medicine, and [4]Department of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel

ABSTRACT We have defined a molecular surface representation that describes precisely and concisely the complete molecular surface. The representation consists of a limited number of critical points disposed at key locations over the surface. These points adequately represent the shape and the important characteristics of the surface, despite the fact that they are modest in number. We expect the representation to be useful in areas such as molecular recognition and visualization. In particular, using this representation, we are able to achieve accurate and efficient protein–protein and protein–small molecule docking. © 1994 Wiley-Liss, Inc.*

INTRODUCTION

Shape complementarity plays an important role in molecular recognition. The shape of a molecule is often described operationally by a surface representation. A representation originating from the Lee and Richards' concept of solvent accessible surface and defined by Connolly generates a molecular surface by rolling a probe ball over the van der Waal's surfaces of the atoms of the molecule.[1–3] Protein surfaces generated this way have shown satisfactory complementarity at their interfaces,[4] and this representation has become popular for protein recognition problems. Connolly's surface can be computed analytically, and is continuous in essence. However, in applications such as protein docking, discrete dots are often used instead, in order to facilitate numerical schemes.[5–10] Using the *dot surface* as the surface representation, however, creates a dilemma for the applications: to accurately describe the surface, the dots should be at high density (typically tens per surface atom); on the other hand, the computational costs of the applications that use the dot surface rise with dot density nonlinearly, often from quadratically to exponentially, both in time and in memory. In addition, a change in dot density alters the locality of the dots, which in turn varies the surface properties associated with them. For example, the orientation of surface normals is highly sensitive to the location of the dots. While, unfortunately, the choice is often subjective, proper choice of dot density is vital to such applications. Due to limited computer power, reducing the number of dots is frequently a primary consideration. However, the way a poorer representation affects an application is hard to foresee and can be devastating. Undesirable performance in costs and/or results is discernible in the protein docking schemes which are based on the dot surface,[5] despite their success to various extents. Some of the schemes achieve accurate docking in affordable time (on the scale of hours) by confining the search to a likely binding site and removal of some surface atoms.[9] Such measures, however, can be susceptible to subjectivity. These difficulties are evidently associated with the surface representation the applications employ. In demand is hence a molecular surface representation which is both precise and concise.

Here we present a surface representation that is accurate (represents the true shape of the surface), complete (covers the whole molecule), concise (a few points per surface atom), rich in describing local surface properties (surface normal, area, curvature, connectivity, etc.), uniquely defined, and independent of the density of a dot surface. Using this representation, we have accurately docked proteins and small molecules to receptor proteins within a few minutes.

The representation describes a surface by a set of *critical points*, derived from Connolly's molecular surface. By definition, such a critical point is obtained by projecting the gravity center of a Connolly face (see methods) onto the surface. A critical point can be defined on either a convex, concave, or saddle-shaped face, correspondingly being dubbed as a *cap*, a *pit*, or a *belt*. These projected gravity centers represent the faces effectively both in their location

and in the orientation of their associated normals. The collection of these points covers the key locations of the extrusion and indentation of the molecular surface, effectively preserving the shape of the molecule. The locations of these points and their surface normals, as well as other associated properties, do not depend on a given dot density. Since each face produces exactly one critical point and each surface atom is associated with only a few faces, only a few critical points are generated per surface atom. This results in a substantial reduction from the usual number of dots describing a dot surface and hence in the costs of the applications, while at the same time still maintaining a healthy dose of redundancy to tolerate imperfect data and methods. This surface representation satisfies the demand of being accurate, concise, representative of the shape of the molecule, and loaded with useful features such as the normal, the curvature, and the area associated with a point. In addition, the way these points are generated readily facilitates a coarse-grained triangulation for fast visualization and for applications employing geodesic measures.* Since the critical points are defined upon the Connolly faces, their generation can adopt many of the relevant data structures and algorithms that have been already developed over the years.[3,11-13] The computation costs are expected to be comparable to such algorithms. By virtue of these advantages, we believe this representation should have general use in molecular recognition problems.

METHODS

A critical point is defined as the projection of the gravity center of a Connolly face on the molecular surface. To explain the above definition, let us begin with Connolly's mechanism of constructing the molecular surface.[3] Connolly's surface consists of connected domains of atomic size, called *faces*. There are convex, saddle-shaped, and concave faces, whose generation requires a rolling probe ball to touch tangentially only one, two, or three atoms, respectively. A convex face is the part of an atom's van der Waal's surface which the probe ball can touch. A concave face is the part of the probe ball surface which is bordered by the three atoms sustaining the ball. A saddle face is the part of the surface of a torus generated by the probe ball rolling around the groove between two atoms; the line drawn through the centers of the two atoms is called the *torus axis* and the trajectory of the probe ball center is called the *torus central circle*.

To find the gravity center of a face is straightfor-

ward. It is the integral of the coordinates of the surface area elements over the face, divided by the area of the face. Namely, the gravity center $\mathbf{g}$ is obtained by

$$\mathbf{g} = \int \mathbf{u} dS / \int dS \qquad (1)$$

where $\mathbf{u}$ is the vector to a surface element dS (for convenience we choose the origin, O, at the atomic center for a convex face, at the probe center for a concave face, and at the center of the torus central circle for a saddle-shaped face).

Our aim is to project the gravity center of a face onto the surface so that the direction of the projection would coincide with the direction of the normal to the surface at the projected point. First, it is necessary to find the projection center. A convex face is spherical, so that the projection center is simply the center of the generic atom. Similarly, a concave face is also spherical and the projection center is the center of the generic probe ball (Fig. 1a). For these faces the projection, $\mathbf{c}$, is obtained by extending the vector from the projection center to the gravity center onto the surface. Namely,

$$\mathbf{c} = R\mathbf{g} / |\mathbf{g}| \qquad (2)$$

where R is the radius of either the atom or the probe ball, for a convex or a concave face, respectively. This procedure produces the *cap* and the *pit*.

The projection center for a saddle-shaped face can be found on the torus central circle. Because the line from the projection center to the gravity center is a normal of the torus, it lies on a plane through the torus axis. To find the projection center, a plane is first defined with the gravity center and the torus axis (this plane is in fact the mirror-symmetric plane of the saddle-shaped face). The projection center is the intersection of this plane with the torus central circle (Fig. 1b). If $\mathbf{a}$ is a vector along the torus axis, and r is the radius of the torus central circle, the projection center $\mathbf{p}$ is obtained by

$$\mathbf{p} = r\mathbf{a} \times (\mathbf{g} \times \mathbf{a}) / |\mathbf{a} \times (\mathbf{g} \times \mathbf{a})|. \qquad (3)$$

Once the projection center is computed, the projection $\mathbf{c}$ is found along the vector from the projection center to the gravity center at the length of R, the probe radius. Namely,

$$\mathbf{c} = \mathbf{p} + R(\mathbf{g} - \mathbf{p}) / |\mathbf{g} - \mathbf{p}|. \qquad (4)$$

This gives us the *belt*.

There are two cases when the critical point cannot be properly determined by the above calculations. The first occurs when the denominator in either Eq. (2) or (3) is zero, which can happen if a face is either a complete sphere so that its gravity center coincides with the spherical center, or a complete torus so that the gravity center falls on the torus axis, respectively. In such situations, Eqs. (2) and (3) diverge and it becomes impossible to cast a projection unequivo-

---

*A straightforward and efficient triangulation takes advantage of the fact that every pit has three belts and three caps as neighbors, every belt has two caps and normally two pits as neighbors, and their neighboring relation has been established during the formation of the Connolly faces.
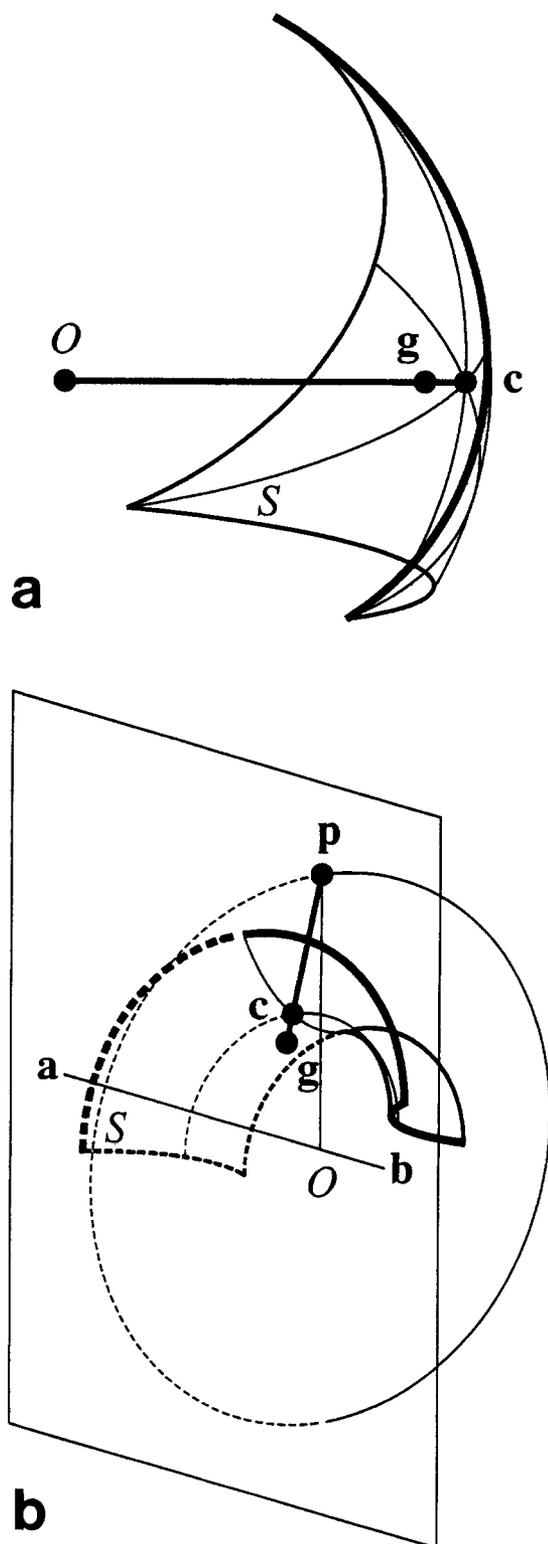
Fig. 1. An illustration of the generation of **(a)** *caps* and *pits*, and **(b)** *belts*, respectively. Symbols are as follows: *S* is a Connolly face, either convex (for caps) or concave (for pits) (a), and saddle-shaped (for belts) (b); *O* in (a) is either the atomic center (for caps) or the probe center (for pits). In b, *O* is the center of the torus central circle; in a and b, **g** and **c** are the gravity center of a face and its projection on the surface, respectively, and **a** and **b** are the two atoms along which the torus axis lies.

cally. The second case occurs when a convex or a concave face contains a hole into which the projection happens to fall instead of on the surface. The existence and the geometry of such holes are computable, for example, via the data structures prescribed by Connolly for analytical molecular surface computation.[3,12] These cases are associated with a few highly exposed or highly folded structural features.[3]

One possible solution of these singularities may be to subdivide such "irregular" faces. However, we leave the options open, since whether a solution is plausible depends on the application and on the structure of the molecule. In our experience, we have examined 18 molecules, including proteins and small ligands, and found that for any individual molecule the nonintegrable faces occupied less than 0.08% of all the faces and the faces with holes less than 0.69%. The latter number did not exclude those occurrences where the critical point did not fall into the hole. To the samples in our docking experiments, these infrequent occurrences add sporadic noise to the surfaces. However, the Computer Vision based technique (see below) we adopted is adequate to deal with these, since the technique is noise tolerant and capable of recognizing partially obscured objects. Consequently, no remedy was performed for such cases in preparing samples for the docking experiments, except for discarding the nonintegrable faces.

The coordinates of the sample structures examined in this study were obtained from the Brookhaven Protein Data Bank.[14] The atoms are given radii according to the extended atom model of CHARMM with polar hydrogens.[15] The radius of the probe ball was 1.8 Å, unless specified otherwise. We generated the critical points of the samples by modifying Connolly's MS program[11] so that it approximated the integration [Eq. (1)] by summing over the Connolly dots, without actually outputting the dots. Such an approach introduced artificial noise into the location and the normal orientation of the critical points. We have estimated how much noise was introduced by varying dot density at 1, 4, and 10 dots/Å^2 and evaluating the root-mean-squared deviations of critical point locations and normal orientation against a high density, 50 dots/Å^2, for each of the 18 molecules. The average location deviations for these molecules were 0.305, 0.154, and 0.107 Å at the three densities, respectively; the worst among them were 0.333, 0.165, and 0.113 Å. The corresponding average deviations in normal orientation were 10.4, 5.4, and 3.8°, the worst were 12.8, 6.9, and 4.7°. The data presented below were obtained with the density 10 dots/Å^2, which would give an uncertainty of critical point location of about 0.1 Å, and an uncertainty of normal orientation of about 4°.

RESULTS AND DISCUSSION

A small molecule, the prothetic heme group of a metmyoglobin,[16] is used as an example to demon-

strate the characteristics of this representation. Figure 2a and c illustrates the heme as van der Waal's balls and as a triangle mesh composed by the critical points, respectively. Comparing these two figures, it is obvious that the shape of the heme is nicely preserved by the critical points. It is also clear that the critical points span the entire molecular surface rather evenly. Figure 2b shows the Connolly dots and the faces of the heme, as well as the critical points on these faces. The critical points are located at the commanding spots of the Connolly faces. Comparing Figure 2a and 2c, it is clear that the critical points cover all the bumps and dents as well as the flat areas of the surface at key locations. Figure 2d compares the surface normals of a subset of the critical points, chosen by objective pruning criteria (see figure caption and below), that match at the interface of the heme–myoglobin complex. The pairing normals are all less than 43° from antiparallel, with the average angle and the root-mean-squared deviation at 157.5 ± 12.3°. We have also surveyed nine crystallographic complexes at 2.7 Å or better resolutions, including seven protein–protein complexes and two protein–small molecule complexes. At their interfaces there are 261 to 471 pairs of critical points matched to within 1 Å. The normals of these pairs are at the angles 156.1 ± 15.0°. Clearly, when the critical points are close spatially, their normals align fairly well. This dual complementarity demonstrates that the shapes of the surfaces are well represented. The dual location–orientation complementarity is a highly desirable characteristic for a surface representation, since it allows the surfaces to be compared simultaneously by spatial and directional relations for effective and efficient recognition. Currently, commonly used docking methods employ matching schemes which either do not use the normals at all, use them as a minor supplement, or employ a redundant number of normals to compensate for their inaccuracy.[5,7,9] One of the difficulties associated with using the normals of the Connolly dots is the uncertainty of the location of the dots and thus the orientation of the normals.

The remarkable reduction of the critical points from the Connolly dots is visible in Figure 2b and c. The heme is a 44—atom molecule. Using a 1.8 Å probe ball, 5,355 Connolly dots have been obtained for the heme at the density of 10 dots/Å². With our representation, the surface is represented by 59 caps, 120 pits, and 147 belts, totaling 326 critical points, achieving a 16-fold reduction against the number of Connolly dots. To reduce the Connolly dots to a comparable number, the dot density must be decreased below 0.64 dots/Å², a value impossible for most applications. In a survey of 19 proteins which contain between 200 and 3400 atoms, we find that on average a surface atom requires 125.4 Con-

nolly dots on the surface, at probe radius 1.8 Å and dot density 10/Å², while the surface atom needs only 1.1 caps, 2.3 pits, and 3.0 belts in our representation. In addition, our experience shows that with only a fraction of the critical points, obtained by objective pruning criteria, molecules in complexes can still recognize each other. The pruning mechanism includes (1) use only one set of the caps, pits, or belts, (2) weed points that are too close to each other, and (3) remove points that cover a face too small or too large. These measures are automatically performed on the entire surface for the sheer purpose of further reduction, taking no account of any particular feature of the molecule or the surface. They result in a coarser, less precise surface. However, we find that a subset containing one point or less per surface atom still suffices for an accurate docking (see below).

It should be emphasized that the Connolly dots cannot precisely dictate the surface properties at practical densities. Furthermore, these properties vary substantially with the change of density. In contrast, our representation does not depend on dot density, despite using a modest number of discrete points. Also, our representation differs essentially from the intermediate reduction of the Connolly dots performed by some docking applications.[3,9] The latter usually involves a time-demanding process to produce what is needed by the particular scheme the application employs; the resulting points usually do not constitute a representation of the complete molecular surface.

Our concise representation of the molecular surface facilitates employing combinatorial algorithms for molecular recognition problems. For example, pattern recognition techniques which have reached certain sophistication in association with computer vision and robotics[17,18] are useful with this representation. These techniques have shown high efficacy in the problems of protein substructure comparison,[19–21] where atoms are represented by isolated points. Similarly, the critical points in our representation are apart by nearly interatomic distances. The sparseness of the critical points allows adopting methods which deal with discrete objects. Conversely, other discrete-point representations usually proximate continuous surfaces;[3,10,22] their applications sometimes tend to derive schemes from the ideal, continuous limit, which find difficulty in the presence of imperfection.

Among the potential applications of the representation, protein docking methods are obvious candidates. First, the representation provides docking schemes with points which are limited in number while still cover the whole molecular surface at the same time. In terms of computational costs, the difference in the number of points compounds nonlinearly via the complexity of docking algorithms. In terms of the capacity of searching the surface, no blank points are left, not only on the extrusions and
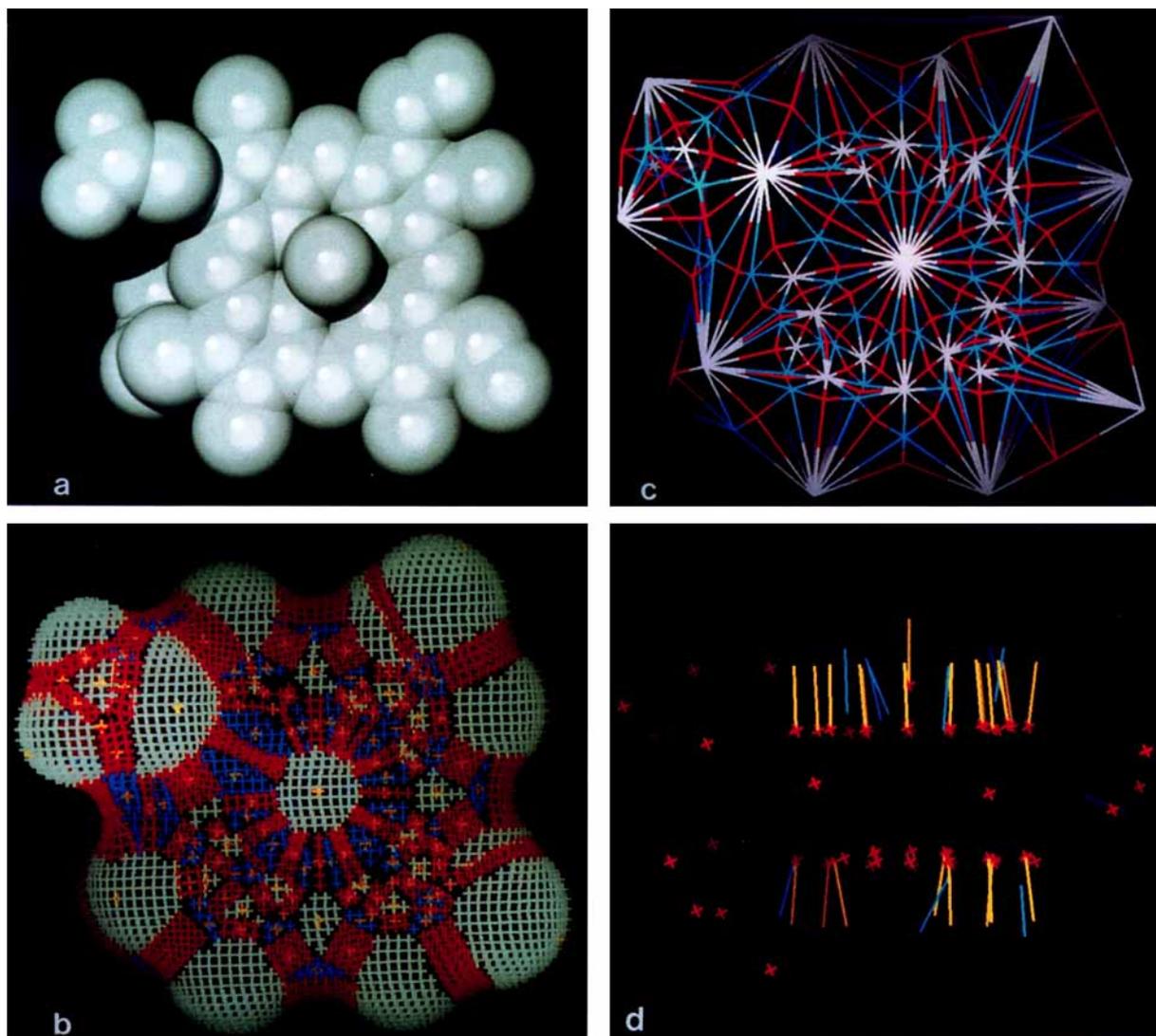
Fig. 2. An example of our molecular surface representation (the photographs are taken from screens displayed by QUANTA from the Molecular Simulation Inc., Waltham, MA.) (a) The heme displayed as the van der Waal's balls of the atoms (ray-tracing plot). (b) Connolly's dots (generated with the MS program[11] and the critical points on the heme surface. All points drawn as small crosses. Colors: light-green, convex faces; blue, concave faces; red, saddle-shaped faces; yellow, critical points. (c) The critical points connected in a triangle mesh. Colors: white, caps; blue, pits; red, belts. (d) Complementarity between the *caps* of the heme and the *pits* of the myoglobin in their location and normal orientation. This picture is a side view perpendicular to those in (a–c). The caps of the heme are shown in small crosses. The normals are shown in strikes, outward from the caps of the heme and inward from the pits of the myoglobin. The normals are drawn only for those caps and pits whose mutual distances are within 2 Å. The pits have been pruned as discussed in the text. Colors: yellow, caps; orange, normals of the caps; blue, normals of the pits.

invaginations but also on the flat areas. Second, with this representation docking methods need not pursue perfect point-to-point matches out of imperfect data and imperfect surfaces. Rather, the representation supports shape matching by a network of critical points, strategically disposed and adequately describing the shapes. An added benefit of adopting a multiple-point scheme is higher noise tolerance. Third, our representation supports more complementary properties at high precision. Among them is the surface normal, an important descriptor of the vectorial characteristic of local surface. Methods based on dot surface are, unfortunately, unable to fully employ surface normals due to an uncertainty in the locality of the dots, thus losing a valuable source of information. By including surface normals we have observed a dramatic performance enhancement over our previous docking experiments.[23–25] Docking methods can thus benefit from the conciseness, completeness, tolerance, and better complementarity featured by the representation.

For a brief demonstration of our representation's

actual performance in docking, we list in Table I some results from our recent docking experiments on known complexes taken from the Brookhaven Protein Data Bank.[14] They include protein–protein docking (2mhb, hemoglobin α subunit with β subunit; 2ptc, β-trypsin with pancreatic trypsin inhibitor; 1cho, α-chymotrypsin with ovomucoid third domain), and protein-small molecule docking (4mbn, myoglobin with heme; 3dfr, dihydrofolate reductase with methotrexate). We use these known complexes to examine whether the representation allows docking to regenerate the crystallographic results in a timely fashion, and thus whether it capably represents the molecular surface and benefits docking methods as we have expected.

The *Geometric Hashing* paradigm has been previously applied to protein structural matching,[19] to detection of surface motifs,[23] and to docking,[24,25] employing a different surface representation.[26] Following is a brief overview of our new docking method. The details will be reported separately.[27]

The *Geometric Hashing* technique has originated in Computer Vision as an object recognition paradigm.[17,18] Recently, we have adapted this approach to protein comparison as a highly efficient, amino acid sequence-order-independent technique,[19–21] for detecting a partial match of 3-D point sets. Proteins are represented as sets of unconnected 3-D *interest points*, e.g., their C_αs. We seek a 3-D rotation and translation which would superimpose a C_α atom-point set of one protein onto another so that a maximal number of C_α atom-points in one set would be "close enough" to points in the other set. When applied to comparisons of protein structures,[14] this method yields recurring "real" 3-D, substructural motifs, without a predefinition of the motif. The reader is referred to Bachar et al.[20] for a specific implementation of this approach. Obviously, the docking problem can be formulated in similar terms, given appropriate *interest points*. With ligand atoms and the Kuntz-styled negative image "spheres" on receptors[26] as the interest points, we have demonstrated the efficacy of this approach.[24,25] Our current surface representation provides sparse sets of *interest points* on the molecular surfaces for the use of the *Geometric Hashing* technique.

The idea behind the *Geometric Hashing* approach is to represent the 3-D points in a rotation and translation invariant manner. Each point is assigned a rotation and translation invariant signature (a "geometric color"). The signatures are efficiently compared via a lookup table (hash-table). The signatures are geometric characteristics generated in a reference frame defined by three points. All other points are represented by their coordinates in this frame.[19] In order to deal with partial matching the points are represented in all triplet-based frames.

There are several major advantages of bundling a

Geometric Hashing-based docking algorithm with our new, sparse surface representation of caps, pits, and belts as the *interest points*. First, it enables us to enhance the *Geometric Hashing* approach by incorporating the normal information. Geometrically, two points with their normals suffice to define a 3-D reference frame. This reduces the complexity of the algorithm by an order of magnitude. Also, under the current representation the signature of a point is not only its coordinates but also its normal direction. This increases significantly the method's discrimination power, leaving us with considerably fewer potential "false positives." Second, the new representation makes the use of *Geometric Hashing* a feasible and highly efficient approach for protein–protein recognition. Although the *Geometric Hashing* is among the most efficient methods for 3-D point set matching, it is still time consuming for very large point sets. The ideal number of *interest points* for the method to be efficient without loosing its discriminatory power is in the low hundreds. The sparse surface representation described here is ideally suited for this task, since the number of points that it produces is of the order of the number of atoms on the surface. This enables completely automated matching of entire proteins. There is no need to initially focus on the active site, as most other docking techniques do.

The experiments started with separating the molecules of known complexes and randomly rotating and translating the ligands (the smaller molecule) with respect to the receptor. Critical points were then generated for each of the molecules, and pruned if necessary to reduce their number to lower hundreds as described before. The points were subsequently submitted for docking to find all transformations that resulted in local-area shape match.

The computations were performed on an SGI Indigo XS24 workstation. The computations include generating the critical points, pruning the points to a subset and docking (finding all shape matches in local areas). The CPU times consumed for point generation and docking are listed in Table I. The pruning spent negligible CPU time (seconds) and thus is not reported in the table. Generating the critical points took at most $\sim$ 2.1 min for the examples. Docking was timed when the whole search was completed, regardless of when the best solution emerged. For all examples, including protein–protein matching, docking was completed within 2.0 min.

Accuracy of docking is measured by the root-mean-squared deviation (RMS) between the crystallographic positions of the ligand (the smaller molecule) atoms and those obtained by docking the detached ligand back to the receptor (the larger molecule). The results with the lowest RMS are entered in Table I. They are clearly at remarkably high accuracy: for all examples the RMS is around or below

TABLE I. Results of Docking Protein–Protein and Protein–Small Molecule Complexes, Obtained by Our Docking Method Using This Representation*

PDB	Receptor	Ligand	Receptor #Critical points (n)	Receptor #Surface atoms (n)	Ligand Critical points (n)	Ligand Surface atoms (n)	#Matching point pairs (n)	RMS (Å)	Receptor critical point generation time (min)	Ligand critical point generation time (min)	Docking time (min)
4mbn	Myoglobin	Heme	208	588	59	44	24	0.33	1.16	0.06	0.28
3dfr	Dihydrofolate reductase	Methotrexate	582	580	43	32	36	0.58	1.65	0.04	1.72
2mhb	Hemoglobin β-subunit	Hemoglobin α-subunit	263	477	245	428	20	0.99	1.47	1.40	1.42
2ptc	β-Trypsin	Pancreatic trypsin inhibitor	365	586	250	235	31	0.24	1.98	0.57	1.74
1cho	α-Chymotrypsin	Ovomucoid 3rd domain	343	636	237	227	24	0.38*	2.12	0.51	1.89

*Columns: (1) the Brookhaven Protein Data Bank code of the molecular complex; (2) the name of the receptor (the larger molecule) of the complex; (3) the name of the ligand (the smaller molecule) of the complex; (4) the number of critical points on the whole receptor surface, after pruning; (5) the number of surface atoms of the receptor; (6) the number of critical points on the whole ligand surface, after pruning; (7) the number of surface atoms of the ligand; (8) the number of the pairs of critical points from the ligand and the receptor which are away from each other within 2.0 Å in the crystallographic complex; (9) the lowest root-mean-squared deviation found among docking solutions, between the atomic positions of the ligand in the crystallographic data and the ones calculated from the transformation that brought the separated ligand back to the receptor; (10) the CPU time for generating the critical points of the receptor; (11) the CPU time for generating the critical points of the ligand; (12) the CPU time for the docking to complete all the searching, regardless of when the best solution emerged.

1.0 Å (to compare with some of the best protein–protein docking results).[9]

This table demonstrates why the representation allows efficient and accurate docking. Only a modest number of points serve as the input to the docking algorithm: only half to one points per surface atom. With these points the representation still finds 20 to more than 30 point-to-point matches at the ligand–receptor interface of the crystallographic complexes. The abundance of matches indicates that the concise, reduced representation preserves the shape of the surfaces to a satisfactory extent. For each individual molecule, the points involved in the original matches collectively describe the local surface shape in that area, regardless of how the ligand is separated and transformed. Successful docking occurs when the docking scheme is able to identify such clusters in the separated molecules and is able to take advantage of their spatial and directional complementarities. One of the examples, the myoglobin and the heme (4mbn), is illustrated in Figure 2. One can visually examine Figure 2d and perceive how the spatially matched points and their aligned normals allow one to bring the two molecules back to the docked complex structure after they are separated.

In summary, we have good reasons to expect that the representation presented here will find numerous applications besides biomolecular recognition and docking. To name a few, the representation can be useful for similarity comparisons between the surfaces of objects, and for the identification and classification of surface motifs.[23] As a byproduct of the surface representation, another complementarity has been found to exist between two sets of nonsurface points in a complex, namely, the centers of the rolling probe sphere corresponding to the critical points of one molecule, and the atomic centers of the second. These can serve as an additional tool for molecular recognition and rational molecular design. Finally, our representation gives molecular graphics and visualization a new and powerful tool that can be used in many different applications of computer-aided molecular design.

ACKNOWLEDGMENTS

REFERENCES

1. Lee, B., Richards, F.M. Solvent accessibility of groups in proteins. J. Mol. Biol. 55:379–400, 1971.
2. Richards, F.M. Areas, volumes, packing and protein structure. Annu. Rev. Biophys. Bioeng. 6:151–176, 1977.
3. Connolly, M.L. Analytical molecular surface calculation. J. Appl. Crystallogr. 16:548–558, 1983.
4. Langridge, R., Ferrin, T.E., Kuntz, I.D., Connolly, M.L. Real-time color graphics in studies of molecular interactions. Science 211:661–666, 1981.
5. Connolly, M.L. Shape complementarity at the hemoglobin $\alpha_1\beta_1$ subunit interface. Biopolymers 25:1229–1247, 1986.
6. Wang, H. Grid-search molecular accessible surface algorithm for solving the protein docking problem. J. Comp. Chem. 12:746–750, 1991.
7. Jiang, F., Kim S.-H. "Soft docking": Matching of molecular surface cubes. J. Mol. Biol. 219:79–102, 1991.
8. Bacon, D.J., Moult, J. Docking by least-squares fitting of molecular surface patterns. J. Mol. Biol. 225:849–858, 1992.
9. Shoichet, B.K., Kuntz, I.D. Protein docking and complementarity. J. Mol. Biol. 221:79–102, 1991.
10. Connolly, M.L. Shape distribution of protein topography. Biopolymers 32:1215–1236, 1992.
11. Connolly, M.L. Molecular surface-dot (MS, QCPE 429). QCPE Bull. 1:75, 1981.
12. Connolly, M.L. Molecular surface triangulation. J. Appl. Crystallogr. 18:499–505, 1985.
13. Connolly, M.L. The molecular surface package. J. Mol. Graphics 11:139–141, 1993.
14. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
15. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4(2):187–217, 1983.
16. Takano, T. Structure of myoglobin refined at 2.0 angstroms resolution. I. Crystallographic refinement of met-myoglobin from sperm whale. J. Mol. Biol. 110:537–568, 1977.
17. Lamdan, Y., Wolfson, H.J. Geometric hashing: A general and efficient model-based recognition scheme. Proc. IEEE Int. Conf. Computer Vision, pp. 238–249, Tampa, Florida, December 1988.
18. Lamdan, Y., Schwartz, J.T., Wolfson, H.J. Affine invariant model-based object recognition. IEEE Trans. Robot. Automation 6:578–589, 1990.
19. Nussinov, R., Wolfson H.J. Efficient detection of three-dimensional structural motifs in biological macromolecules by computer vision techniques. Proc. Natl. Acad. Sci. U.S.A. 88:10495–10499, 1991.
20. Bachar, O., Fischer, D., Nussinov, R., Wolfson, H.J. A Computer Vision based technique for 3-D sequence independent structural comparison of proteins. Prot. Eng. 6(3):279–288, 1993.
21. Fischer, D., Bachar, O., Nussinov, R., Wolfson, H.J. An Efficient Automated Computer Vision based technique for detection of three dimensional structural motifs in proteins. J. Biomolec. Struct. Dyn. 9(4):769–789, 1992.
22. Heiden, W., Goetze, T., Brickmann, J. Fast generation of molecular surfaces from 3D data fields with an enhanced "marching cube" algorithm. J. Comput. Chem. 14:246–250, 1993.
23. Fischer, D., Norel, R., Wolfson, H.J., Nussinov, R. Surface motifs by a computer vision techniques: Searches, detection and implications for protein-ligand recognition. Proteins 16:278–292, 1993.
24. Fischer, D., Norel, R., Nussinov, R., Wolfson, H.J. 3-D docking of protein molecules. Combinatorial Pattern Matching, Lecture notes in Computer Science 684, Springer Verlag, 20–34, 1993.
25. Norel, R., Fischer, D., Wolfson, H.J., Nussinov, R. Molecular surface recognition by a computer vision based technique. Protein Engineering, accepted.
26. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R., Ferrin, T. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 161:269–288, 1982.
27. Lin, S.L., Fischer, D., Norel, R., Wolfson, H.J., Nussinov, R., to be submitted.