"Iso-depth contour map" of a molecular surface

Anne Badel-Chagnon, Julie Nessi, Laurent Buffat and Serge Hazout

Unité de Recherches Biomathématiques et Biostatistiques, INSERM U263, Université Paris, Paris, France

The representation and display of protein surfaces are useful in many areas of molecular modeling, and surface shape study is particularly important in the analysis of protein–ligand interactions. We introduce here the notion of the molecular surface convex hull, allowing the depth of any molecular surface point to be defined. A two-dimensional (2D) map, the iso-depth contour map, and a three-dimensional (3D) representation, the iso-depth lines, allow the topography of a molecular surface to be displayed in terms of knobs (high depth) and holes (low depth).

Keywords: molecular surface representation, contour map, iso-depth curves, crambin

INTRODUCTION

The representation and display of protein surfaces are useful in many areas of molecular modeling, and surface shape is particularly important in the analysis of protein–ligand interactions.[1] The objectives of existing computer procedures, depending on the type of biological study (3D-representation, molecule matching, or docking), are various: 3D-representation of the molecule surface, global description of the surface by analytical functions to reduce the amount of information, local representation of the surface or extraction of local geometric characteristics of the surface. In the first class (3D-representation), several algorithms have been developed. For example, the solvent accessible surface may be defined by intersection contours of the protein surface with a set of parallel planes, or by a polyhedral surface construction using a procedure of hidden edge elimination.[2] A complementary process consists in describing and smoothing those planar contours by B-splines.[3] In the second class, simplified representations based on expansion functions (Legendre polynomials[4] or Fourier series[5]) have also been used, and have the advantage that the accuracy of the approximation can be tested within the same mathematical framework. In the third class, several methods have attempted to approximate surfaces and compute geometric shape characteristics. Various studies have tackled the computation of molecular surface fractality.[6,7] Therefore, a local geometric measure can be performed on the whole molecule, however the problem of the surface unfolding is not easy. An approach consists in projecting the geometric characteristic of the molecular surface onto a "sea-level ellipsoid" and building the iso-level lines by contouring all points of same characteristics. The characteristic may quantify the molecule relief, or represent an energetic potential.[8,9]

Our main problem is the docking of two molecules using a geometric approach. Holes and knobs constitute the geometric patterns of a molecular surface, and an active site is generally an association of holes and knobs. So, the docking problem is equivalent to finding holes and knobs of the same characteristics. In this paper, we use the notion of the molecular surface convex hull to extract these patterns. The aim of the study is, in the context of the docking problem, to define a preprocessing tool to characterize the holes in a molecular surface. So, our approach consists in visualizing the molecular surface either by a 2D graph: an *iso-depth contour map*, or by a 3D display, *iso-depth lines*, each showing the point set associated with a given depth relative to the molecular surface convex hull. Such descriptions can help in finding the geometric patterns and the active site.

METHODS

The basic principle of building iso-depth lines lies on the notion of convex hull. We will first define the convex hull of a three-dimensional point set, and describe the algorithm of its construction. Then we will define the depth of a molecular surface point relative to the convex hull, and describe how to obtain iso-depth contour map or lines.

Convex hull: definition and building method

The convex hull of a protein simply represents a wrapping of the molecule. Relative to the docking problem, when two molecules are in contact one against the other, the surface molecules are complementary. At the active site, the "depth" of the hole is similar to the "height" of the corresponding knob. The set of tangent planes to this hole constitutes the basis of the knob. As the convex hull is defined as the wrapping of the tangent planes (or the intersection of all half-spaces that contain the object), the notion of hole "depth" relative to the possible contacts can be deduced

Color Plates for this article are on page 193.

Address reprint requests to Dr. Badel-Chagnon at Unité de Recherches Biomathématiques et Biostatistiques, INSERM U263, Université Paris 7, 2, Place Jussieu, 75251 Paris cedex 05, France.

Received 22 September 1993; revised 24 December 1993; accepted 4 January 1994

from the convex hull. Moreover, the notion of sea-level ellipsoid used by certain authors does not correspond to a physical reality, when there is only one convex hull for a given object. Other measures can be deduced from the knowledge of the convex hull, such as an *occupational volume* defined as the volume of the convex hull, which is more precise than the volume of an inertia ellipsoid, or a *molecular surface complexity,* the ratio between the convex hull and the molecular surface area (or the convex hull volume relative to that of the molecular surface).

Convex hull definition. The convex hull is defined by the wrapping of all planes tangent to the object (Figures 1a and 1b). It constitutes simply a close object wrapping. The result is a polyhedron characterized by a subset of molecular surface points. Every triangular face is put either on the top of the knobs or between three adjacent knobs. According to the triangle shapes, one can deduce the presence or absence of large holes. In fact, triangles of small size are located on knobs, and holes are hidden by large triangles.

We have carried out a method called the *growing tetrahedral crystallization method* that defines the convex hull of a spatial set of points. In our study, the initial point set is composed of the dots of Connolly's surface.[10]

Convex hull construction. The principle of our approach consists in defining an increasing series of convex polyhedrons by juxtaposition of tetrahedrons. The processing is recurrent, and each cycle is composed of two stages: the first stage, called the *extension phase,* consists of the addition of a "maximal extension" tetrahedron to the intermediate convex hull. The second stage, called the *completion phase,* allows the juxtaposition of tetrahedrons to keep the polyhedron convexity. The object points inside the tetrahedron are discarded; hence the number of object points rapidly decreases. These two transformations constitute the *crystallization process,* which is repeated until the final pattern is found, i.e., when no more points are outside the polyhedron. The algorithm is described as follows:

Step 0: Choose an initial tetrahedron of maximal size included in the object. All the points inside this tetrahedron are eliminated since they cannot belong to the convex hull.

Step k: At the previous step ($k-1$), the intermediate polyhedron is convex. It is a series of linked triangulated oriented faces. A face of the convex polyhedron is examined (Figure 1c).

The extension phase begins: the point of maximal extension relative to the current face (i.e., the point outside the object and the farthest from this face) is determined (point P in Figure 1c). In the case where no such point exists, this face will belong to the final polyhedron (i.e., the convex hull). Otherwise, a new tetrahedron is constructed from this point and this current face, and the object's points inside the new tetrahedron are eliminated (result in Figure 1d).

The completion phase then begins: the convexity of the new polyhedron is tested. A procedure allows the faces close to the current face, that have the previous point of maximal extension above them, to be detected (face F' in Figure 1d). So new tetrahedrons defined by the triangular faces and the point of maximal extension are added to the polyhedron to maintain the convexity. Figure 1e shows a completion ex-

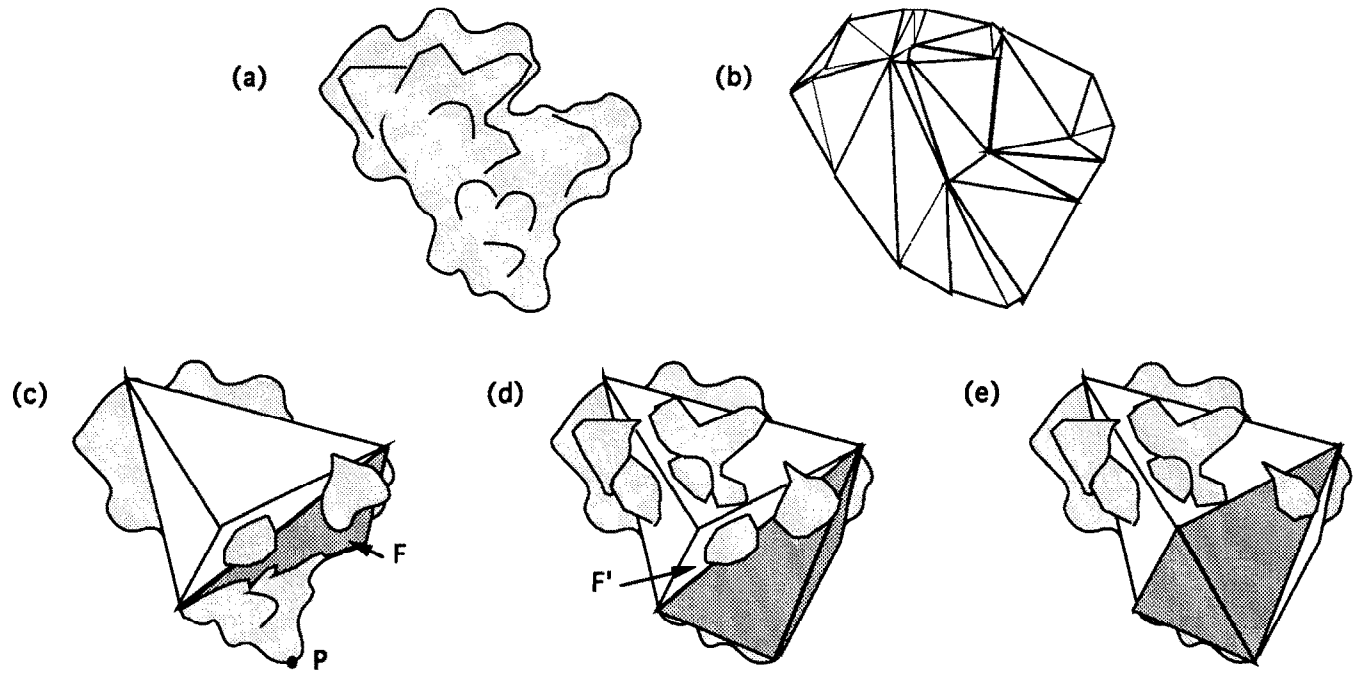

Figure 1. Construction of the convex hull of a molecular surface. (a) Molecular surface. Only a point set is used for the convex hull building. (b) Convex hull. The result is a polyhedron composed of triangular faces. (c) Intermediate shape of the convex hull: a convex polyhedron composed of several juxtaposed tetrahedra. For face F, P *is the maximal extension point. (d) Adding of a tetrahedron (defined by point* P *and face* F*) in the extension phase. In this case* P *is also above* F'. *(e) Adding a tetrahedron (defined by point* P *and face* F') *in the completion phase to maintain the polyhedron convexity.*

ample by addition of one tetrahedron. The faces hidden by those new faces are discarded.

The convex hull construction is stopped when no more points are outside the polyhedron. The final polyhedron, composed of triangular faces, is the convex hull of the point set.

Our approach presents two important advantages: the method is efficient in terms of consuming time and can treat any spatial set of points, and an approximate form of the convex hull may be obtained by stopping the iterative process when points outside the polyhedron are very close to the faces (i.e., when the distance of the external point relative to the polyhedron face is less than a given threshold).

Iso-depth lines: definition and building method

An iso-depth line is a series of connected molecular surface points that are located at the same distance (called *depth*) from the convex hull surface. Initially, the molecular surface can be divided into several point subsets, each corresponding to a given depth value. The objective is to transform a point subset into a series of connected points, i.e., an iso-depth line. Before describing this line construction in a 3D space, it is necessary to define precisely the notion of depth of a molecular surface point.

Definitions: relative and absolute depth of a molecular surface point. Each point M of the molecular surface is associated with a triangular face F of the convex hull (see Figure 2). In fact, it necessarily belongs to one tetrahedron defined by a triangular face of the convex hull and the centroïd C of the convex hull. The absolute depth D_M of the point M is defined as the distance between this point and the associated face, and is expressed by the scalar product:

$$D_M = \overrightarrow{n_F} \cdot \overrightarrow{MC_F} \qquad (1)$$

where $\overrightarrow{n_F}$ denotes the normal (unitary) vector of face F, and C_F the center of the face F. In Figure 2, D_M is represented by the length of the segment $[MH]$, H denoting the projection of M onto F.

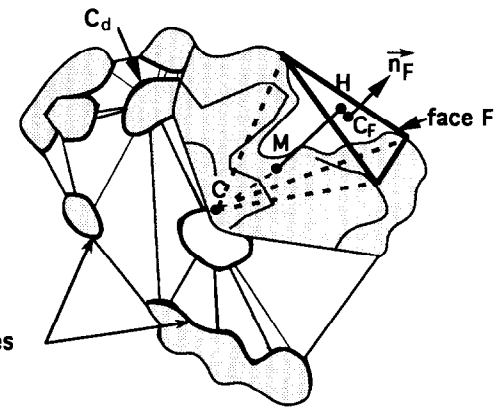

Figure 2. Calculation of the depth of a molecular surface point and determination of the iso-depth lines. The iso-depth lines are deduced from the intersection between the molecular surface and the iso-depth surface C_d *(obtained by homothesis from the convex hull). Only a part of* C_d *is shown.*

Using this formula, the relative depth d_M is defined as the ratio between the distance of M to the face F and the distance of the centroïd C (of the convex hull) to the same face F:

$$d_M = \overrightarrow{n_F} \cdot \overrightarrow{MC_F} / \overrightarrow{n_F} \cdot \overrightarrow{CC_F} \qquad (2)$$

The quantity d_M varies between the values 0 (M is common to the molecular surface and the convex hull; i.e., it belongs to the knobs) and $(d_M)_{max}$ (M is the point of maximal depth).

Definitions: iso-depth surface and iso-depth lines. Consider the convex hull transformation by a homothesis of center C and coefficient k ($0 \leq k \leq 1$; 0 corresponds to the reduction of the convex hull to its centroïd C, 1 to the initial convex hull), the result is the iso-depth surface C_d corresponding to a relative depth $d = 1-k$. This surface contains the iso-depth lines. In fact, the closed lines of depth d are the intersection between the molecular surface and the iso-depth surface C_d (Figure 2). The iso-depth surfaces (or lines) can be built from the absolute or relative depth. Result may be very different using absolute or relative depth when the form of the convex hull is not similar to a spheroid.

Construction of iso-depth contour map. Our purpose is to build a 2D representation of the iso-depth lines, called an *iso-depth contour map.* The main stages of the preprocessing are:

(1) Computation of the absolute depth D_M, relative depth d_M, and the radius ρ (distance between C and M) of every molecular surface point.

(2) Definition of a depth matrix D in which each element associated to a direction gives the depth of a surface point. The directions in space to select a molecular surface point subset are defined by unitary vectors in spherical coordinates: $\overrightarrow{u_{ij}}$ ($\sin\theta_i \cos\phi_j$, $\sin\theta_i \sin\phi_j$, $\cos\theta_i$) where $\theta_i = i\Delta$ and $\phi_j = -\pi + j\Delta$, $i = 0, 1, \ldots, n$, $j = 0, 1, \ldots, 2n$, and the term $\Delta = \pi/n$ corresponds to a user-defined angular variation. To define the element d_{ij} of the depth matrix D of size $(n+1) \times (2n+1)$, the surface point M is found so that the vector $\overrightarrow{CM}$ has a direction close to the given direction $\overrightarrow{u_{ij}}$. In fact, a point M of the molecular surface is selected if and only if its distance to the straight line oriented by the vector $\overrightarrow{u_{ij}}$ (equal to the modulus of the vector product $\overrightarrow{u_{ij}} \times \overrightarrow{CM}$) is minimal, and the directions of the vectors $\overrightarrow{CM}$) close to $\overrightarrow{u_{ij}}$ belong to the same half-space; i.e., a positive scalar product $\overrightarrow{CM} \cdot \overrightarrow{u_{ij}} > 0$ is necessary. After selecting the surface point in a given direction, the value of its absolute or relative depth is given to the element d_{ij} of the matrix D. Another matrix R, called radius matrix, indicates the associated modulus ρ_{ij} of the vector $\overrightarrow{CM}$.

(3) Visualization of the iso-depth contour map, the iso-depth contour map is obtained by using the S Plus software.[11] To reduce the important distortions at the poles (i.e., for θ_i close to 0 and π) due to the plane representation, the matrix D normally described in a rectangle is shown into an ellipse.

Construction of iso-depth lines in 3D space. The algorithm consists in drawing the iso-depth lines by a series of

straight lines between points spatially close having the same depth. For a given depth d_0 (absolute or relative), we scan every "cell" of the matrix D (a cell of the matrix D is defined by the four elements $(d_{i,j}, d_{i+1,j}, d_{i+1,j+1}, d_{i,j+1})$), and two processes are performed (Figure 3): in the first step, each cell is divided into two triangles by selecting one of the two diagonals according to the minimal depth variation criterion ($|d_{i+1,j+1}-d_{i,j}|$ or $|d_{i,j+1} - d_{i,j+1}|$). The second step consists in determining for each triangle, the points of depth d_0 belonging to one of the three possible intervals (i.e., triangle sides) delimited by two of the following depth pairs $(d_{i,j}, d_{i+1,j})$, $(d_{i+1,j}, d_{i+1,j+1})$, $(d_{i+1,j+1}, d_{i,j+1})$ and $(d_{i,j+1}, d_{i,j})$ and the depth pair of the diagonal, $(d_{i+1,j+1}, d_{i,j})$ or $(d_{i,j+1}, d_{i,j+1})$. If d_0 belongs to one of these intervals, (d_1, d_2) denoting the extremities, then the ratio α equal $(d_0-d_1)/(d_2-d_1)$ is computed (with $0 \leq \alpha \leq 1$). As the angular variation Δ is small, this ratio is assumed to be a common factor for a linear interpolation of the spherical coordinates $(\rho_0, \theta_0, \phi_0)$ of a point belonging to an iso-depth line, i.e., $\rho_0 = \rho_1 + \alpha \cdot (\rho_2-\rho_1)$, $\theta_0 = \theta_1 + \alpha \cdot (\theta_2-\theta_1)$, and $\phi_0 = \phi_1 + \alpha \cdot (\phi_2 - \phi_1)$, where $(\rho_1, \theta_1, \phi_1)$ and $(\rho_2, \theta_2, \phi_2)$ denote the spherical coordinates of the extremities of the interval (or triangle side).

So *a fortiori* for each triangle, either no point or two distinct points of depth d_0 are found (a particular case can occur: points are identical to a triangle vertex when d_0 is equal to one of the four values of the cell). In the last case, the straight line between these points (defined by their spherical coordinates) constitute a fragment of the iso-depth line in the 3D space. Moreover, an approximate tessellation of the molecular surface composed of the whole set of the triangle edges in the 3D space is made up. A point of this polyhedron has the spherical coordinates $[\rho_{ij}, \theta_i, \phi_j]$. The iso-depth lines correspond to a set of straight lines drawn on the polyhedron faces.

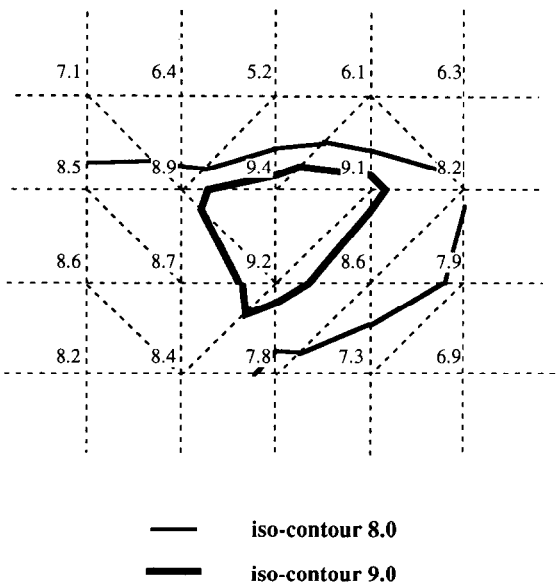

Figure 3. Definition of an iso-depth line fragment in 3D space from the depth matrix D. *The iso-depth contours are defined by linear interpolation and then are projected in the 3D space to become iso-depth lines.*

RESULTS

We will illustrate our approach by the study of the crambin. After a comparison of the contour maps defined by the radius ρ, the relative depth d_M and the absolute depth D_M, the 3D representation of the iso-depth lines on the triangulated molecular surface is shown.

Color Plate 1 shows the convex hull of the molecular surface of the crambin. The convex hull represented as a set of triangular faces is a close wrapping of the molecule. Each face belongs to one of the three classes of triangle patterns: small, large, or stretched. This classification is useful for a future docking research: the large faces hide holes, and the small faces define knobs of the molecular surface, the stretched faces are either knobs or holes, according to the shape of contiguous faces.

Comparison between different contour maps

Figure 4 shows how to understand the iso-depth contour maps. The area represented with dashed lines is the visible part of the molecule according to the X-axis. Figure 5 shows three types of contour maps corresponding to the matrix R of the radius ρ_{ij} of molecular surface points (Figure 5a), the matrix of the relative depth (Figure 5b) and of the absolute depth (Figure 5c) for the protein crambin (in the Protein Data Bank[12]). Usually, a protein is represented by the contour map based on the radius after a projection of the molecular surface onto a sphere or an ellipsoid. No reason justifies the choice of these surfaces. In our approach, the convex hull is considered as a surface of nil depth, and the depth is computed from this reference surface. We observe the lack of sensitivity of the surface description by the radius to analyze the local aspects of the surface (Figure 5a). It is to be noted that the details of the surface are important elements for verifying the docking between two molecules. However, the two iso-depth contour maps allow the real structure of the surface to be displayed through a 2D representation (Figures 5b and 5c). The iso-depth contour map gives precisely at each point the depths of the holes (the knobs in contact with the molecular surface have a depth of 0). We can observe approximately the presence of four holes; a large zone is composed by a large hole (a dark area in Figures 5b and 5c) and a channel leading to a small hole. Small differences appear between the contour maps of Figures 5b and 5c since the radius of the crambin varies between 10 and 15 Å. The small hole at the bottom of the contour map appears less deep in Figure 5b since it is located farther from the center of the convex hull than the large hole. Figure 6 shows different views of the contour maps of the crambin in which the different holes are approximately located at the center of the maps. To obtain these maps, the depth matrix D is computed again by determining the new spherical coordinates of the vectors $\overrightarrow{u_{ij}}$ after a given rotation.

Iso-depth lines

Color Plate 2 shows the iso-lines of the crambin. The colors of the lines vary between the red and the green according to the proximity of surface points to the convex hull. A large hole (colored in deep green) is observed in the center of the molecule surrounded by knobs (in red). The molecular sur-

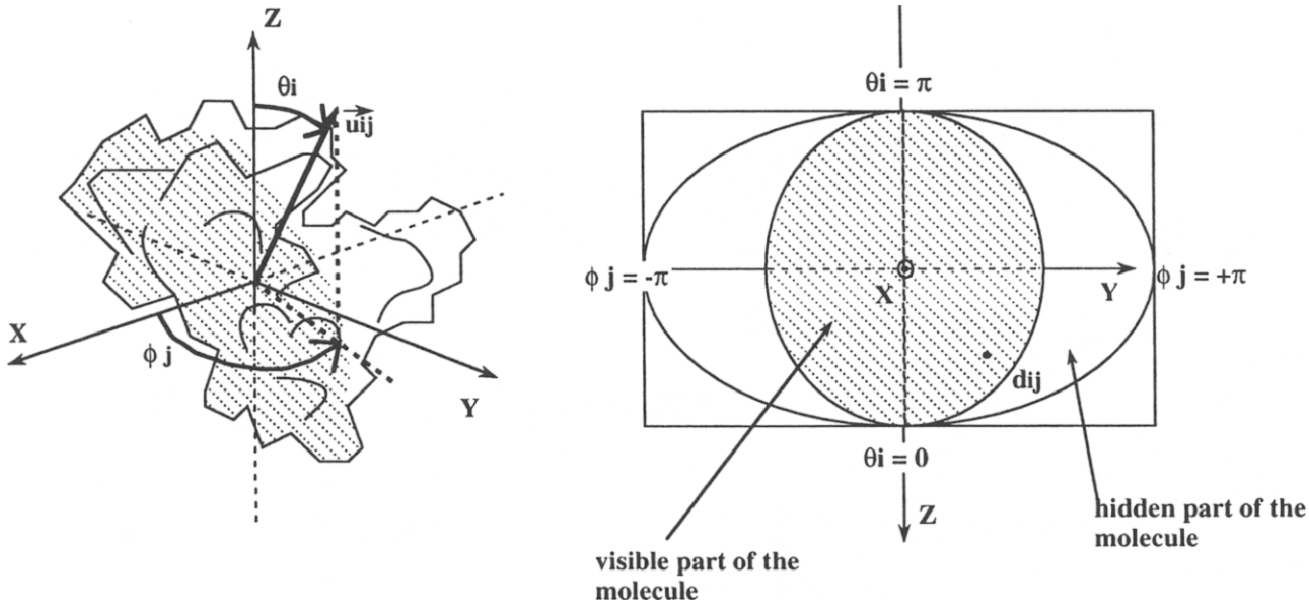

Figure 4. Correspondence between the reference of the molecule in the 3D space and that of the iso-depth contour map.

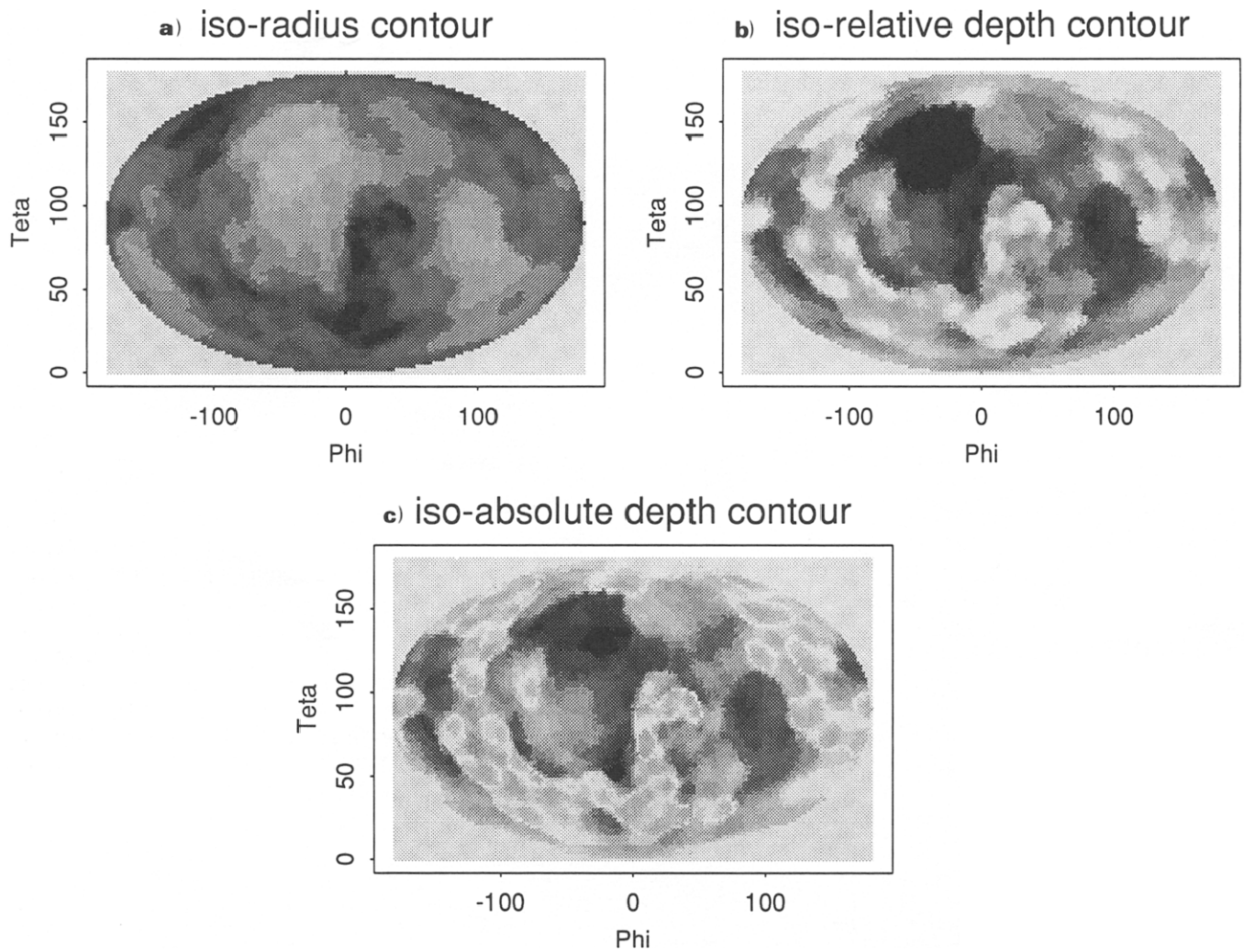

Figure 5. Contour maps of the molecular surface of the crambin from (a) the radius, (b) the relative depth, and (c) the absolute depth.

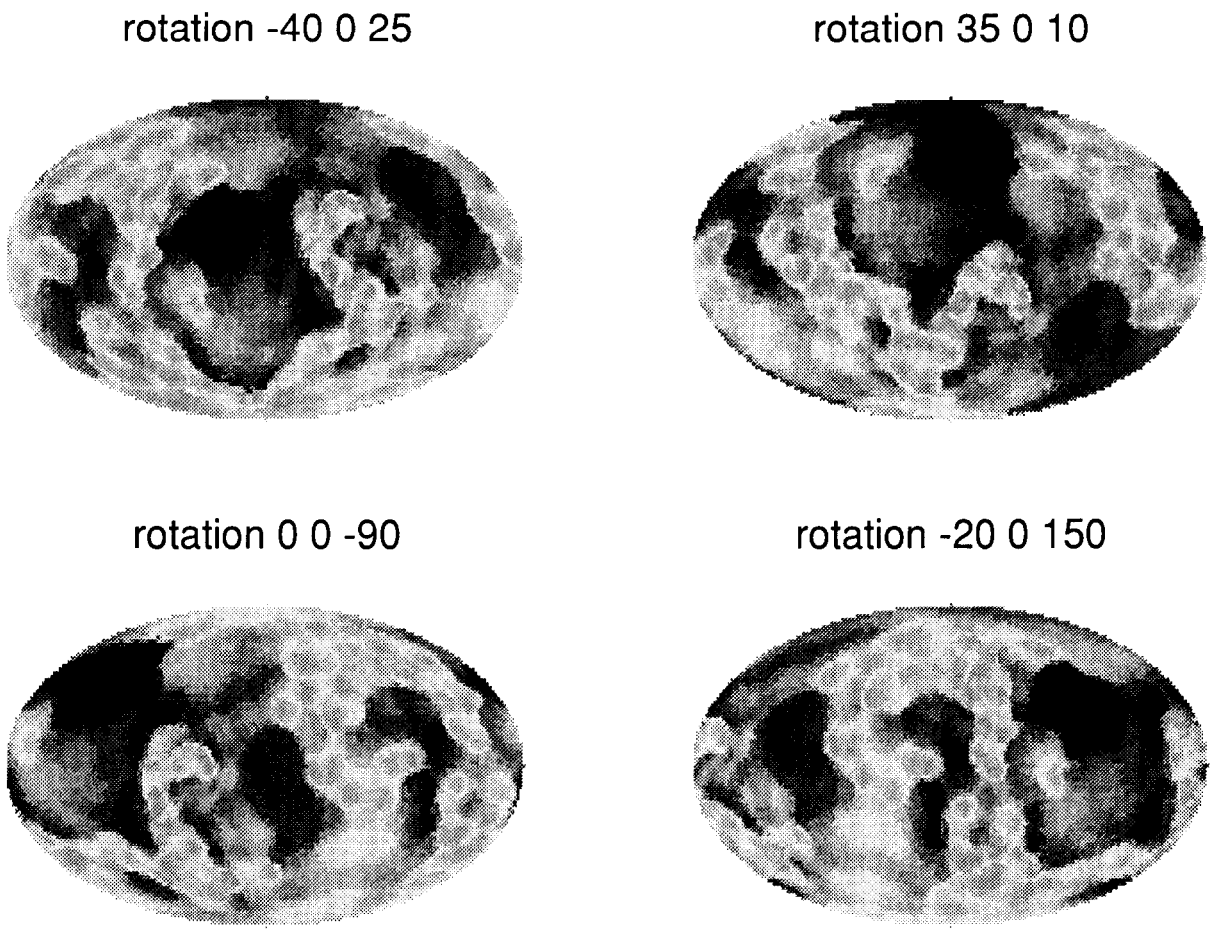

Figure 6. Views of the iso-depth contour map after different rotations of the molecule. The four large holes are shown according to the X*-axis.*

face is not even, a lot of small roughnesses appear along the iso-depth lines.

CONCLUSION

The approach described in this paper allows the characterization of every point of the molecular surface by its depth, and the representation of a molecular surface by iso-depth lines. We have shown its interest for visualizing the molecule relief in a precise and efficient way.

This approach needs the definition of the convex hull of a 3D set of points, that is more suitable than an arbitrary ellipsoid as projection space for building the depth contour, since it constitutes a close wrapping of the molecule. It allows the definition of two quantities, the relative and absolute depths, that present some interesting properties. Indeed, the relative depth has the advantage of being scale invariant and is adequate for describing the global shape of the molecule. The absolute depth quantifies the real difference in level of the molecular surface and allows the identification of individual features on the molecular surface and the quantification of the surface complexity in docking studies. Relative depth is equivalent to absolute depth if the object is uniformly scaled in size.

In our study, we have used them to display either contour maps (2D representation) or iso-lines (3D representation) of the molecular surface. The contour map representation is a convenient tool for analyzing the different structures and for describing globally the surface topology in terms of boundaries and zones. The problem of these representations is the use of transformation of the molecule in spherical coordinates, that leads to some bias in the display of the iso-depth maps. A solution to this problem is to give a series of depth values for certain map point, to represent separately the whole set of iso-depth contours.

The improvements of the method by complementary quantification (as the notions of occupational volume and molecular surface complexity) and its application to the docking problem will be developed in a further paper.

REFERENCES

1 Connolly, M.L. *Biopolymers.* 1986, **25,** 1229–1247
2 Connolly, M.L. *J. Mol. Graph.* 1986, **4,** 93–96
3 Colloc'h, N. and Mornon, J.P., *J. Mol. Graph.* 1990, **8,** 133–140
4 Saira Mian, I., Bradwell, A.R., and Olson, A.J. *J. Mol. Biol.* 1991, **217,** 133–151

5 Leicester, S.E., Finney, J.L., and Bywater, R.P. *J. Mol. Graph.* 1988, **6,** 104–108
6 Lewis, M. and Rees, D.C. *Science.* 1985, **230,** 1163–1165
7 Avnir, D., Farin, D., and Pfeifer, P. *Nature.* 1984, **308,** 261–263
8 Fanning, D.W., Smith, J.A., and Rose, D.D., *Biopolymers.* 1986, **25,** 863–883
9 Chirgadze, Y., Kurochkina, N., and Nikonov, S. *Protein Eng.* 1989, **3,** 105–110
10 Connolly, M.L. *Science.* 1983, **221,** 709–713.
11 Becker, R.A., Chambers, J.M., and Wilks, A.R. *The new S language,* Wadsworth & Brooks/Cole Computer Science Series, Pacific Grove, California, 1988
12 Bernstein, F.C., et al. *J. Mol. Biol.* 1977, **112,** 535–542