Protein Science (1993), *2*, 459–469. Cambridge University Press. Printed in the USA.

Mapping the surface properties of macromolecules

MICHAEL S. CHAPMAN
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907
(RECEIVED August 13, 1992; REVISED MANUSCRIPT RECEIVED November 16, 1992)

Abstract

Methods are presented for the rapid computation of schematic projections of the surfaces of macromolecules, similar to the "roadmaps" used to illustrate the surfaces of viruses (Rossmann, M.G. & Palmenberg, A.C., 1988, *Virology 164*, 373–382). Several types of projections are described, extending the application of "roadmaps" to the external surfaces of all macromolecules and their interior binding pockets and pores. The surface projections, showing the positions of residues, can be colored, shaded, contoured, and annotated to show physical, sequence, or functional properties such as surface topology, hydrophobicity, or sequence conservation, for example. The automated procedures are useful for surveys of the surface features of proteins sharing similar functional properties.

Keywords: graphics; macromolecules; surface properties

Many functions of macromolecules are endowed by their surface properties. Examples of such functions include substrate and cofactor binding, multimeric assembly, antigen recognition, and receptor binding. Methods are presented here that facilitate the correlation of function with the physical properties of the amino acids that line the surface of a known structure. They complement molecular graphics programs such as FRODO (Jones, 1978), Hydra (Hubbard, University of York), "O" (Jones et al., 1991), or Pluto (Motherwell, Dodson, & Evans) that are well suited for the examination of detailed atomic interactions and those that render Van der Waals surfaces (e.g., Connolly, 1983). A more schematic representation is often useful in looking for trends in the physical properties or sequence conservation over wide areas, in comparing the surfaces of different proteins, and for examining surfaces without the aid of sophisticated interactive graphics computers.

The antecedent to this work was the maps of Rossmann and Palmenberg (1988), in which the surface was projected onto a plane, with boundaries between labeled residues. Although the projected positions of residues were calculated computationally, manual conversion into a figure took days. These types of figures were helpful in understanding the correlation of sequence and viral capsid structure important in antibody interactions and receptor binding (Rossmann & Palmenberg, 1988; Kim et al., 1989; Tsao et al., 1991). They have also been of use in planning site-directed mutagenesis, through which such functionality can be experimentally dissected. The applications are now extended from viral capsids to all macromolecules. Drawing of the surface and map illustration of physical properties through coloring, shading, and/or contouring has been automated, each map taking only a few minutes to prepare. It is, therefore, practical to examine rapidly such surface topology, sequence similarity, hydrophobicity, or charge, and to compare such properties between members of a structural database that have common functional properties.

Reprint requests to: Michael S. Chapman, Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306.

Methods

Overview

Atomic coordinates, in the Brookhaven Protein Data Bank format, are expanded according to the molecular symmetry. The surface may be projected in several ways. In all cases, each of a set of projection vectors is extended until it intersects the Van der Waals sphere of one of the atoms (or, optionally, the solvent-accessible sphere). These points are mapped onto the projection plane. The boundary of each projected residue is traced. The map can be colored, shaded, or annotated according to the physical, chemical, or functional attributes of the residues of their constituent atoms. Maps are output using the PostScript® page description language (Adobe Systems, Inc., 1990) and can be viewed using a wide variety of printers and workstations.

Projections

The simplest is a parallel projection in which the atomic surface is projected along a user-defined direction onto a plane orthogonal to that direction. This projection suffers no spatial distortions and is most suitable for examining parts of a molecular surface. The entire surfaces of globular proteins are viewed with the cylindrical projections commonly used in atlases. The molecular surface is projected outward and radially from the center of mass using an equi-angular grid in spherical-polar coordinates. The surface grid points are then mapped onto a cylinder whose axis is the same as the polar axis. The cylinder is cut at a specified point and unrolled flat for display. The projected surface of a spherical molecule is spatially distorted for directions close to the cylindrical (polar) axis. This is similar to the longitudinal distortion near the poles of the globe when viewed in cylindrical projection. The inside surfaces of pores, channels, and the binding sites of substrates and cofactors are displayed by projecting outward upon a cylinder that surrounds the pore or site. The surface is projected at regular intervals in cylindrical-polar coordinates. The cylindrical axis is chosen to follow the channel or binding site and can be bent by selecting a series of guidepoints, although this results in slight distortions.

For all of these projections, for each grid point, the intersection is found of the projection vector and the Van der Waals surface. This calculation is split into several steps, so that it can be done within a couple of minutes, even for a large molecule. Let $\vec{x}_a$ be the position vector of atom a with reference to the origin of the unit projection vector $\vec{p}$. The perpendicular distance c_a from $\vec{p}$ to atom a is given by:

$$c_a^2 = |\vec{x}_a|^2 - (\vec{x}_a \cdot \vec{p})^2. \quad (1)$$

If v_a is the Van der Waals radius of atom a, atoms can be sorted quickly according to a criterion that keeps only those atoms whose Van der Waals sphere is intersected somewhere by $\vec{p}$:

$$c_a \leq v_a. \quad (2)$$

The distance, d_a, along $\vec{p}$ to the point of intersection with the Van der Waals sphere is

$$d_a = \vec{x}_a \cdot \vec{p} - \sqrt{v_a^2 - c_a^2}. \quad (3)$$

The atom of lowest d_a is the surface atom. d_a need not be calculated for any atom for which

$$\vec{x}_a \cdot \vec{p} - v_a > d_{\min}, \quad (4)$$

where $d_{\min}$ is the current estimate of the minimal d_a.

Residue boundaries

For each grid point, the distance (d_a), residue name, type, and atom type are stored in memory. Each grid point is at the center of a rectangular pixel in the projected map. A list is made of the vectors needed to delineate areas sharing the same residue name in the following way. For each pixel not previously assigned to a residue, neighboring pixels in the horizontal and vertical directions are examined to see if they are of the same residue. If so, the examination is recursively repeated for the neighbors of the pixel. When a different residue is encountered, coordinates are saved for the ends of a line along the common edge of the adjacent pixels. When no more neighboring pixels of the same residue are found, the list of coordinates is sorted to provide a clockwise tracing of the residue and reduced to include only those points at corners. One residue may cover several areas, some of which may be connected diagonally through common corner points. The boundary points are merged and modified to create a diagonal connection between the two areas. The positions for residue labels are the pixel centers closest to the center of mass for the area of each residue.

Mapping physical and sequence attributes

At this stage, the projection is stored in memory both as a pixel mapping and as a list of vectors defining the boundaries of residues. Either may be used as a basis for coloring, shading, or contouring. Height, used for plotting surface relief, is calculated directly for each pixel from d_a. Residues may be colored according to input properties that are specific for each residue (such as sequence similarity) or that are specific only for the type of residue (such as hydrophobicity). Pixels may be colored according to input properties that are specific for each atom, or specific only for the atom/residue type (such as atomic solvation; Table 1). In addition to shading and coloring, mapped properties may be outlined or contoured crudely (without interpolation) using the algorithm for residue tracing. A line separating two pixels is drawn if their properties straddle the chosen contour level. Contouring works best for slowly varying properties and can be used with coloring to show two properties simultaneously, such as to show surface topology in addition to sequence conservation.

Implementation

The method has been programmed in standard FORTRAN 77. It can be run interactively in a few minutes on a graphics workstation and has been implemented on those of several manufacturers. The ASCII output file can be printed on a PostScript printer, displayed on a workstation with a PostScript previewer, included

Table 1. *Atomic solvation parameters*[a]

```
*    C*   16   !  Atomic solvation parameters (cal A-2 mol-1)
*    N*   -6   !  Ref: Eisenberg & McLachlan (1986)
*    O*   -6
*    S*   21
ARG  NH*  -28  !  average of -6 and -50 for 2 NH's
ASP  OD*  -15  !  average of -6 and -24 for 2 OD's
GLU  OE*  -15  !  average of -6 and -24 for 2 OE's
HIS  NE*  -50
LYS  NZ*  -50
```

[a] This example, which creates a surface property based on atomic solvation parameters, illustrates the intuitive program input required to color the surface according to properties that depend on the type of residue and/or their constituent atoms. The four columns specify residue type, atom type, parameter value, and comments. Thus, the first four lines give the default values for carbon, nitrogen, oxygen, and sulfur atoms. The last five lines give values to side-chain atoms of specific residues. The values have been modified from Eisenberg and McLachlan (1986) to average the values of atoms whose identities are often crystallographically ambiguous. The attribute for all atoms of histidine residues could be set to 5 with the line "HIS * 5."

within other formatted documents, or modified with an editor.

Program control

The many options controlling the graphical output are called as required using an option driver that, in use, resembles that of "O" (Jones et al., 1991). It supports nested redirection of input, facilitating the modularization of frequently used command sequences. Each option sets a large number of default parameters that are fully documented, ranging from those controlling the projection, to page layout, line types, and the palette of colors. Program input is simple if defaults suffice, but also flexible, because the defaults may be overruled by specifying keywords and their values after calling an option. Thus, even for customized output, it is rarely necessary to edit the PostScript file directly. The example given in Table 2 was used to produce Figure 2.

Examples

Surface topology of human rhinovirus 14 (HRV14)

There are four known HRV14 epitopes that have distinct antigenic NIm sites, which elicit a neutralizing immunogenic response (Sherry et al., 1986). The atomic structure of HRV14 (Rossmann et al., 1985) showed that escape mutations to neutralization were located in exterior loops whose sequences, when aligned to other rhinoviruses (Palmenberg, 1989, and references therein) were hypervariable. A depressed canyon encircles the fivefold axis and was proposed to be the binding site for attachment to cellular receptors (Rossmann et al., 1985). A recent electron microscope study (Olson et al., 1993) confirmed that the binding site for rhinoviruses of the major receptor group did indeed occupy part of the canyon between hypervariable exposed NIm regions. These topological features are illustrated in Figure 1. The NIm sites are exposed with NIm1A, NIm1B, and NIm2 forming the highest features of the surface. The canyon is a valley encircling the fivefold axis and is 15–25 Å across and 10–20 Å deep. There is a second depression surrounding the diad, similar to the dimples of canine parvovirus (Tsao et al., 1991) and ϕX174 (McKenna et al., 1992), to which no function has yet been assigned.

Rhinoviral surface sequence variability

Before there was direct experimental evidence, support was given to the canyon hypothesis of rhinoviral receptor binding by the observation that the residues in the canyon are particularly conserved (Rossmann & Palmenberg, 1988). The conservation was apparent from a roadmap

Table 2. *A typical control file*[a]

```
Surface  ;  Coordinate_control
      Files  Input = hrv14.pdb  /* HRV14 native from Brookhaven PDB */
      Sphere  Center = 35.00000     00.00000    140.00000   Radius = 75.00000
      @icosahedral.symop /* Matrices for symmetry expansion in separate file */
      Plane  Icosahedron = 0.0,0.0  ;  End
Parallel_projection  Projection_vector = 0.0,0.0,-1.0  Origin = 0.0,0.0,200.0 \
   Horizontal_axis = 0.0,1.0,0.0  Vertical_axis = 1.0,0.0,0.0 \
   Use_plane = .false.  Projection_direction = .false. \
   Distance_origin = 0.0,0.0,0.0 \
   Start_horizontal = -55.0  End_horizontal = 55.0  Space_horizontal = 2.0 \
   Start_vertical = 0.0  End_vertical = 90.0  Space_vertical = 2.0 \
   Add_Hydrogen = .true.  Add_Water = .true. ;  End
New_page  PostScript file = rhino14_outer_sim.ps \
         Header = 'Human Rhinovirus 14 outer surface.'
Residue_properties
   Property  Number = 1  File = rhino_gap.similarity \
   Set_Color  Number = 1  Parameter = hue  Property_range = -0.1,1.5 \
          Key_title = 'Sequence conservation'  Parameter_range = 0.00,0.55
   Paint  ;  End
Contour
   Property  File = Height  ;  Header  Text = 'Distance from center (A)'
   Outline  Level = 145  Width = 2.2  ;  Outline Level = 155  Width = 1.5  ;  End
Icosahedron  Dash = 0.1,0.05
Trace_residues  Dash = 0.005,0.02
Label  Text = Canyon  Position = 1.40,4.70  Point = 4.82 3.34
Label  Text = 'NIm 1A'  Position = 6.69,5.59  Point = 4.92 4.57
Label  Text = 'NIm 1B'  Position = 2.45,6.25  Point = 4.25 5.79
Label  Text = 'NIm 2'  Position = 0.50,2.75  Point = 3.63,2.51
Label  Text = 'NIm 3'  Position = 8.46,2.44  Point = 6.71,1.41  ;  End
```

[a] The input file that was used to generate Figure 2 illustrates the flexibility and intuitiveness of the program input. Under "Coordinate_control," a coordinate file is read, and a sphere is filled with atoms, including those atoms brought into the sphere by the symmetry operators listed in a separate file, "icosahedral.symop." Residues are discarded if all of their atoms fall outside an optional icosahedral asymmetric unit defined by planes. The second section defines the type of projection and gives its direction and extent. "Add_water" requests that a solvent-accessible surface be calculated. "Residue_properties" reads in sequence similarity scores, sets up a color palette, and paints the roadmap according to similarity. "Trace_residues" draws boundaries around each residue in a finely dashed line. Much of the input is not required if defaults suffice. For example, "Trace_residues" would outline the residues in solid lines if "Dash . . ." were absent. The input ends with a series of labeling commands.

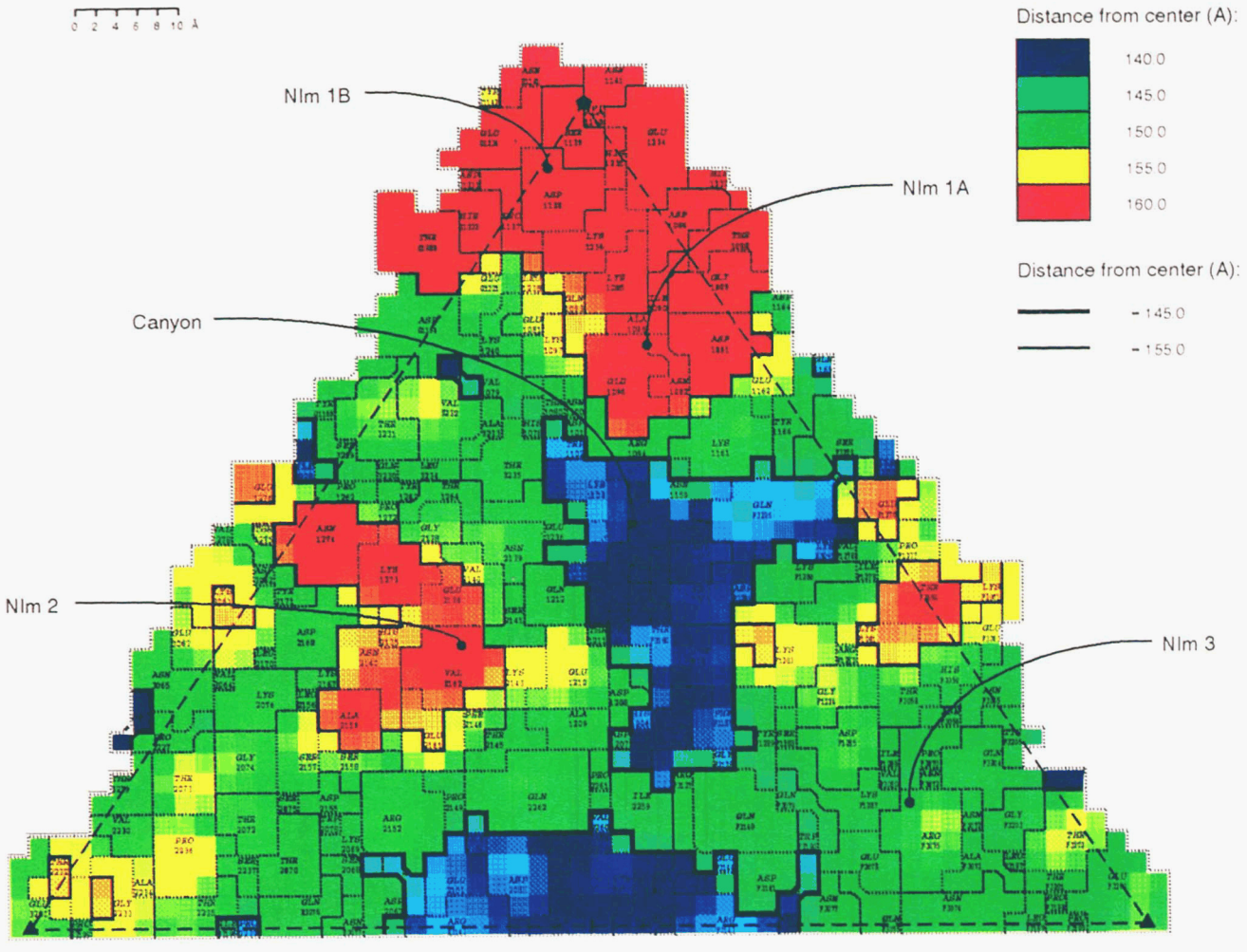


Fig. 1. Surface topology of HRV14: The spherical capsid of HRV14 can be split into 60 identical triangular units related by icosahedral symmetry. Symmetry axes on the boundary of one of these units are shown as ⬟ (fivefold), ▲ (threefold), and ⬮ (diad). The outer surface has been projected onto a plane orthogonal to the diad (z-axis). Contouring and coloring is according to the radial distance from the center of the virus to the solvent-accessible surface. The surface is comprised of viral proteins 1, 2, and 3, labeled according to the residue number + 1,000 × the protein number. The labels of residues from symmetry equivalent subunits start with a letter specific for each protomer. The significance of the canyon and the NIm sites is explained in the text. The shape of the canyon differs slightly from that previously reported (Rossmann & Palmenberg, 1988) because the height is calculated from the radial distance to the center of the virus, rather than its projection upon the twofold (z) axis. (This and Figures 2, 4, 5B,C, and 6 illustrate the output of the program RoadMap.)

shaded according to the minimal base change per codon of aligned picornaviral sequences (Rossmann & Palmenberg, 1988). Prior to the automation described here, roadmaps typically required days to prepare, discouraging their use in comparative surveys and in testing speculative ideas. Figure 2 was prepared in just a few minutes and shows that the canyon is more conserved than other parts of the surface. Coloring of each HRV14 residue is according to the average mutational distance (Dayhoff et al., 1979; Gribskov & Burgess, 1986) between all pairs of residues from aligned rhinoviral sequences (Palmenberg, 1989), calculated with the program Similarity (Chapman, in prep.). The regions of highest sequence variability (red) coincide with the exposed NIm sites, suggesting a strong selective pressure for mutations in those regions that might enable the virus to escape immune surveillance.

Amphilicity of melittin

Melittin is the 26-residue lipolytic agent of bee venom. As a monomer, it forms a highly amphiphilic helix that integrates into and disrupts lipid membranes, but at the high concentrations found in the bee venom sac, it is a tetramer with the hydrophobic residues packed at the center (Terwilliger & Eisenberg, 1982). The surface of a dimer was cylindrically projected as shown schematically in Figure 3 and colored to illustrate (Fig. 4) the preponderance of hydrophobic residues at the inter-dimer con-

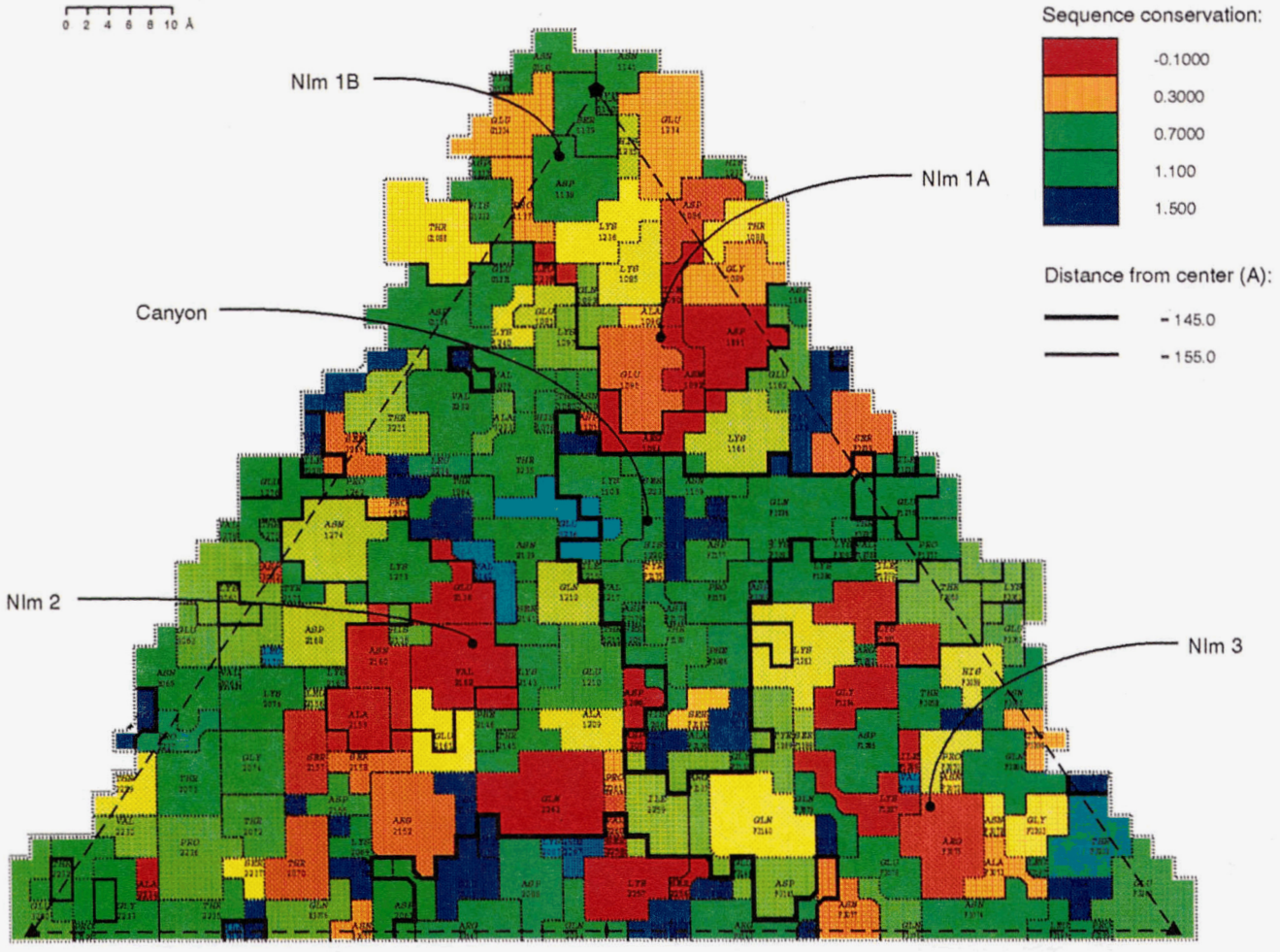


Fig. 2. Sequence similarity of rhinoviruses: The surface of HRV14 is projected as in Figure 1 with contours at the same radial distances. Dark blue areas are invariant loci, and red indicates highly variable amino acids. The sequence conservation scores were calculated using the program Similarity (Chapman, in prep.) from eight aligned rhinoviral capsid sequences (Palmenberg, 1989, and references therein). For every residue in each of all possible pairs of sequences, a score between -1.2 and $+1.5$ was allocated according to the aligned residues and a normalized mutational distance matrix (Dayhoff & Schwartz, 1979; Gribskov & Burgess, 1986). Within gaps of n residues, each unmatched residue was assigned a score of $(-3.0 - n \times 0.15)/n$. The scores that are plotted were averaged between all pairs of sequences and mapped to the structure of HRV14. The scores of regions that are insertions with respect to HRV14 were averaged with the nearest neighboring residue in the HRV14 sequence.

tact at the center of the tetramer. This hydrophobic region is surrounded by predominantly hydrophilic residues.

The apparently large surface areas of residues such as arginine B24 is an artifact of the cylindrical projection. The cause is apparent in Figure 3. With the surface projected radially on an equi-angular grid, the surface of the molecule near the poles of the sphere is sampled more frequently. Distortions in position and/or area are inherent in all projections of three-dimensional surfaces onto a plane. The advantage of the cylindrical projection is that it maintains the correct juxtapositions of neighboring residues. Occasionally there may be no parts of a protein that can be relegated to the area distortions of the poles, then it is necessary to illustrate the molecule using several maps that can be either cylindrical projections using different polar axes, or parallel projections viewed from different directions.

Lysozyme

Lysozyme is an example of a globular protein with a convoluted surface. Using its atomic coordinates (Blake et al., 1965; Diamond, 1974) the solvent-accessible surface is illustrated in Figure 5B, projected as shown in Figure 5A. The orientation of the projection was chosen to place the active-site cleft in a region free from area distortions, but as with melittin, the apparent surface areas of residues near the poles have enlarged areas. In this,

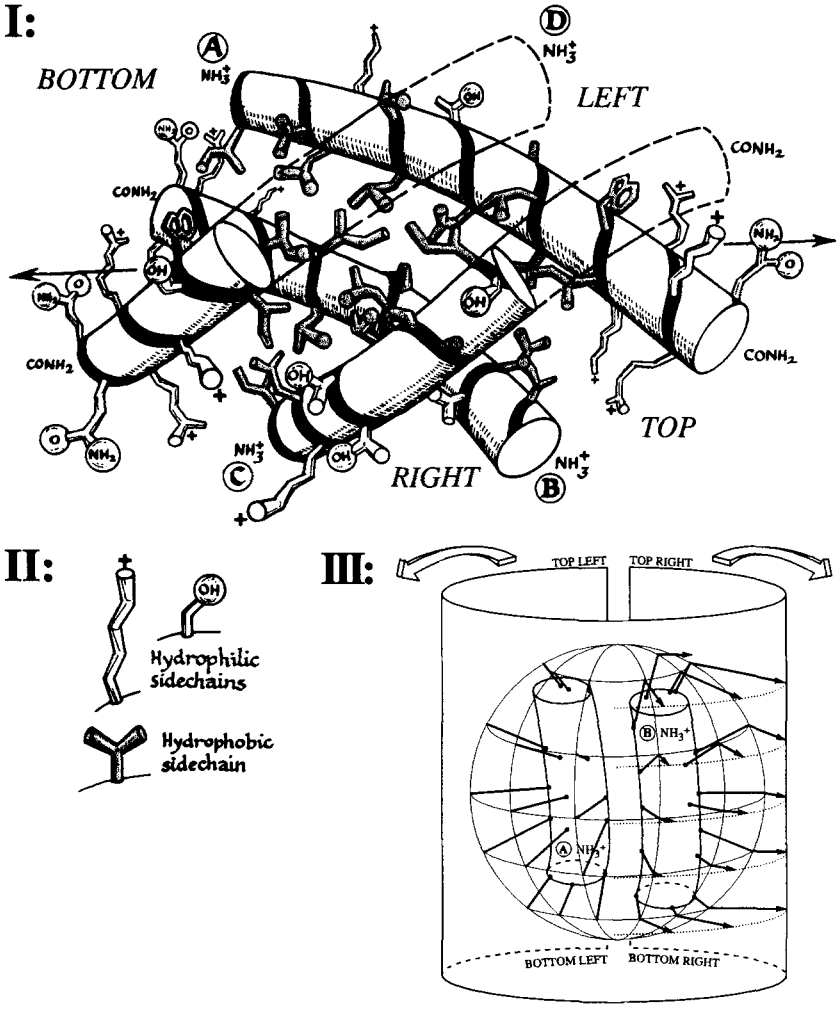


Fig. 3. **I:** Structure of the melittin tetramer (reproduced, with permission, from Terwilliger & Eisenberg, 1982). Hydrophilic residues are shaded lightly and hydrophobic residues darkly, as indicated in panel **II**. Prior to projection, chains C and D in the foreground, that are related to chains A and B by a crystallographic twofold axis (shown by arrows), were removed to leave a dimer. The labels "top," "bottom," "left," and "right" show the orientation of melittin in the projection of Figure 4. **III:** Chains A and B were surrounded by a sphere oriented with the polar axis approximately parallel to the helical axes of melittin. The meridian runs through the center of the tetramer and is the center of the projection in Figure 4. Illustrated on the left half of panel III is the radial projection of the surface from the center of mass onto a sphere surrounding the half-tetramer. For Figure 4 the sampling was every 10° in latitude and longitude, but here, for clarity, a sampling of 30° is shown. On the right of panel III is shown the projection of the sphere onto a coaxial cylinder such that the spherical angles are mapped linearly. The cylinder is then cut on the far side and unrolled as indicated with arrows. The corners of the cut cylinder are labeled according to their positions in Figure 4.

and most other cases, in spite of area distortions, cylindrical projections can be used to show overall surface features such as topology and the relative locations of surface residues.

Several surface residues are not seen near the convoluted surface of the cleft. For example, from a viewpoint outside the cleft, Asn 59 is obscured by Arg 61. To interpret the potential interactions with external molecules it is usually appropriate to plot (as in Fig. 5B) the solvent-accessible surface, defined by the closest approach of a water molecule to any given atom. Use of "solvent-accessible radii" increases the apparent volume of all residues, and only the closest to the projection surface are seen. However, occlusion can be reduced by using the smaller Van der Waals radii. Occlusion is often severest for the convoluted surfaces of binding pockets and clefts. These are better represented by the bent cylindrical-polar projection shown in Figure 5C. The Van der Waals surface was projected outward onto a curved cylinder from a cylindrical axis running down the center of the binding cleft of lysozyme (Imoto et al., 1972; Kelly et al., 1979; Fig. 5A). Such projections avoid occlusion, but, due to curvature of the cylindrical axis, have their own distortions parallel to the axis. Glutamine 57 is an extreme example with a projected length increased from 8.3 to 14 Å. The cylindrical axis wraps around the enzyme, changing direction by about 60° and Glu 57 is projected onto several successive segments of the curved cylinder. The high curvature of the guide points has given the lysozyme-binding cleft the largest distortions yet encountered. Even so, projections such as Figure 5C can be effective in schematically illustrating the positions of residues, the distribution of physical properties, and the closeness of interactions with substrates or cofactors.

Drug-binding pocket of HRV14

Figure 6 confirms that the WIN pocket is predominantly hydrophobic (dark shading). There are three regions where the pocket is more hydrophilic (light). Near the entrance (pore), Asn 1219 forms a hydrogen bond to the isoxazole group. About $\frac{1}{3}$ of the way in, a water mol-

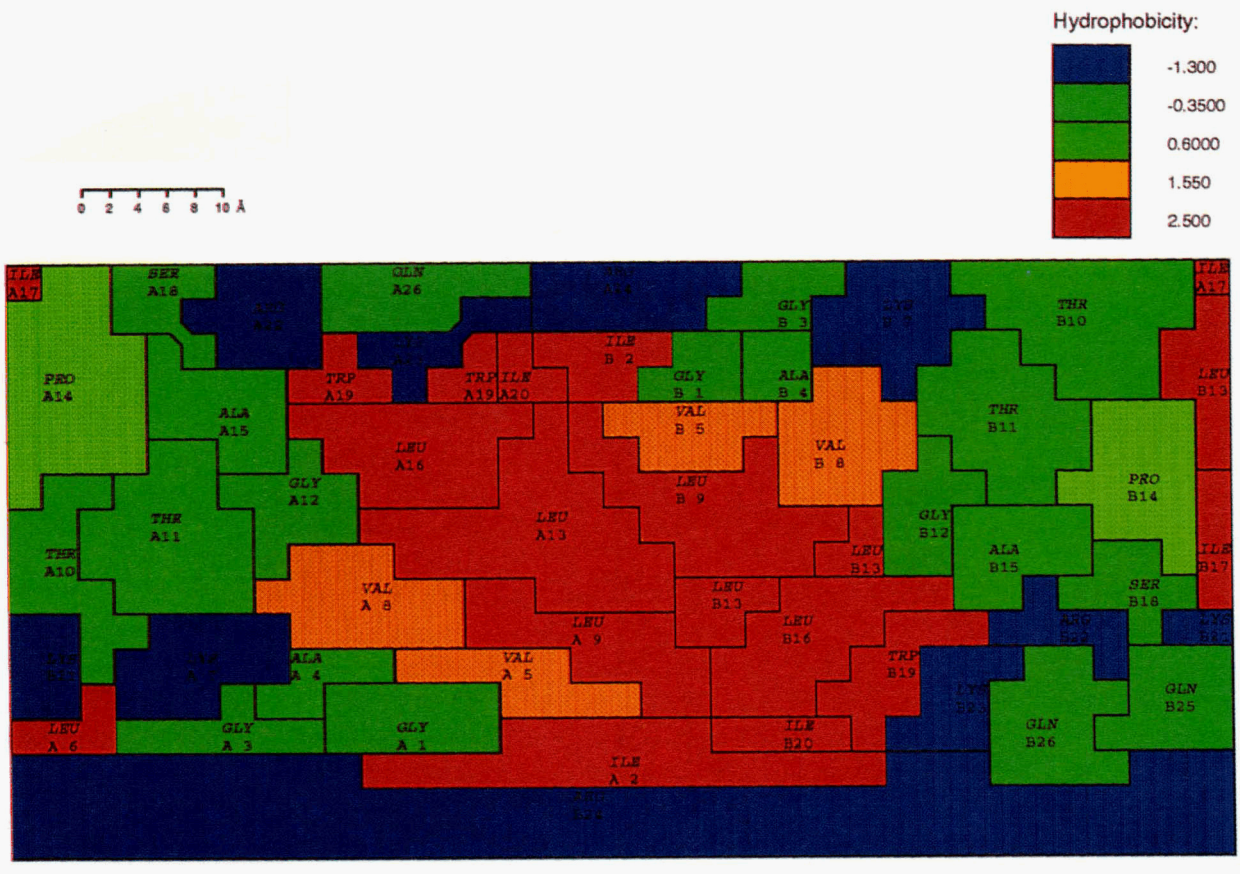


Fig. 4. Amphilicity of melittin: The whole of the solvent-accessible surface of a half-tetramer was cylindrically projected as shown in Figure 3. The hydrophobicity scores were set according to solvation energies (Eisenberg & McLachlan, 1986), calculated from the data of Fauchere and Pliska (1983). The diagram emphasizes that the hydrophobic residues are concentrated at the center of the tetramer where they are surrounded by neutral (green) and hydrophilic residues (blue).

ecule remains in place after drug binding, coordinated by the side chains of Ser 1126 and Tyr 1197 and the backbone carbonyl of Phe 1124. Surprisingly, although the end of the pocket is comprised of hydrophobic residues, this diagram drew attention to a slightly hydrophilic environment created by their carbonyl groups pointing into the pocket. The observed distribution of hydrophilicity matches roughly that of WIN compounds that are hydrophobic except for the terminal isoxazole and oxazoline groups and the phenoxy oxygen, approximately at the midpoint. Figures 5 and 6 illustrate the capabilities of the method with monochrome graphics devices. Considerably better gray-scale definition is generally obtained with monochrome video displays than on printers.

Conclusion

In addition to the examples discussed here, these methods have also been used to: (1) correlate the phenotypes of canine parvovirus (CPV) residues with their surface locations and conservation (Chapman & Rossmann, 1993) and plan mutagenesis experiments; (2) examine the interaction of DNA with the inside of the CPV capsid (Tsao et al., 1991); (3) examine the contact surfaces of subunits involved in several multimeric assemblies; (4) interpret the structure of bacteriophage ϕX174 (McKenna et al., 1992); (5) examine the properties of the rhinoviral receptor footprint (Olson et al., 1993) and look for similar properties in other viruses (Chapman & Rossmann, unpubl.); (6) examine the distribution of surface conformational changes induced by drug binding to HRV14 (Smith et al., 1986) relevant to the inhibition of host-cell receptor binding (Pevear et al., 1989); (7) map the footprint of a monoclonal antibody (Wikoff, Parrish, Baker, & Rossmann, pers. comm.); and (8) examine which residues of different rhinoviruses form the most intimate contacts with WIN drugs, and compare their chemical properties (Kim et al., 1993).

These schematic projections complement three-dimensional wire or space-filling representations and are most appropriate when a broader, less detailed picture is required. Their independence of sophisticated three-

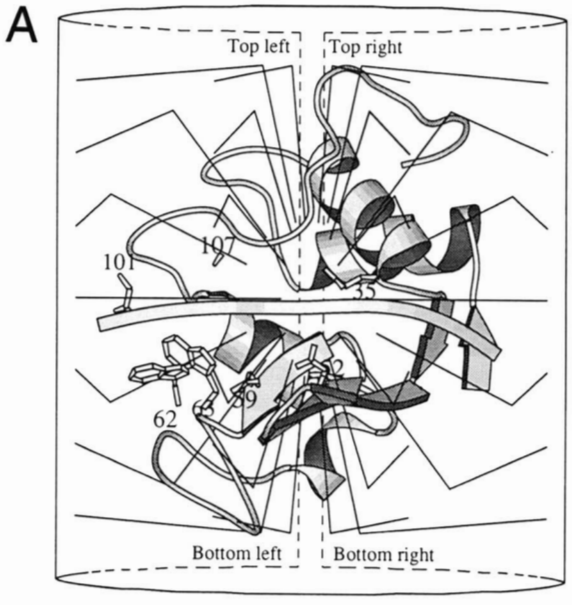
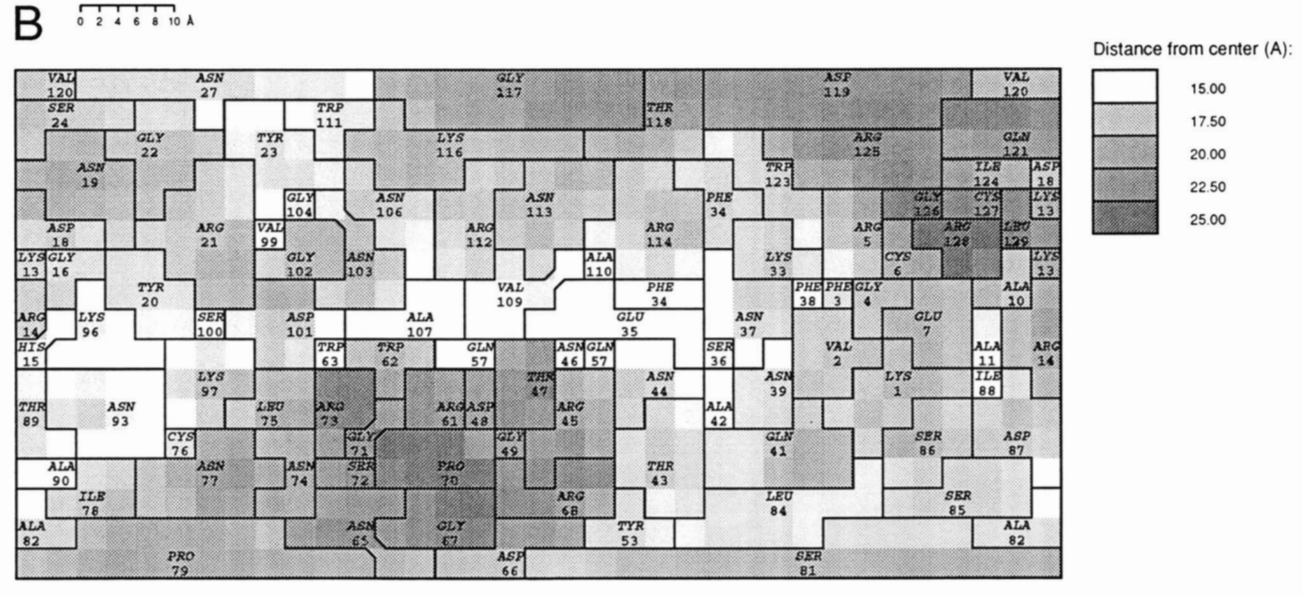
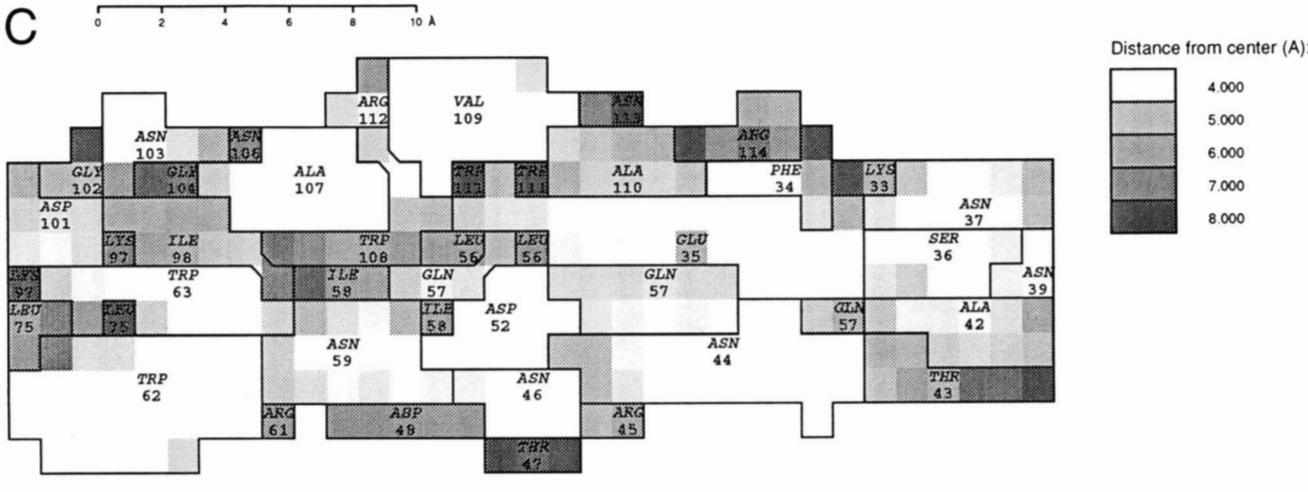


Fig. 5. A: The structure of lysozyme (Blake et al., 1965; Diamond, 1974) is drawn schematically using the program MolScript (Kraulis, 1991). The active-site cleft is indicated by the near-horizontal line at the center of the picture that uses 10 guide points to follow and extrapolate the path of the NAM-NAG-NAM trisaccharide (Kelly et al., 1979). The side chains of several catalytically important residues are shown (Imoto et al., 1972). Lines radiate from the center of mass to the surface of a sphere (not shown) and then to the surface of a cylinder to show schematically that the projection shown in B was calculated similarly to that of Figures 3 and 4. **B:** The cylindrical projection of the entire solvent-accessible surface is shaded according to its distance from the center of mass. The active-site cleft is obvious as the central white region. **C:** Details of the active-site cleft are shown by projecting the Van der Waals atomic surface outward onto a bent cylinder. The axis of the bent cylinder follows the guide points that were used to indicate the cleft in A. The cylinder was cut on the solvent side, closest to the viewer in A, and unrolled. Shading is according to the distance of the surface from the cylindrical axis.





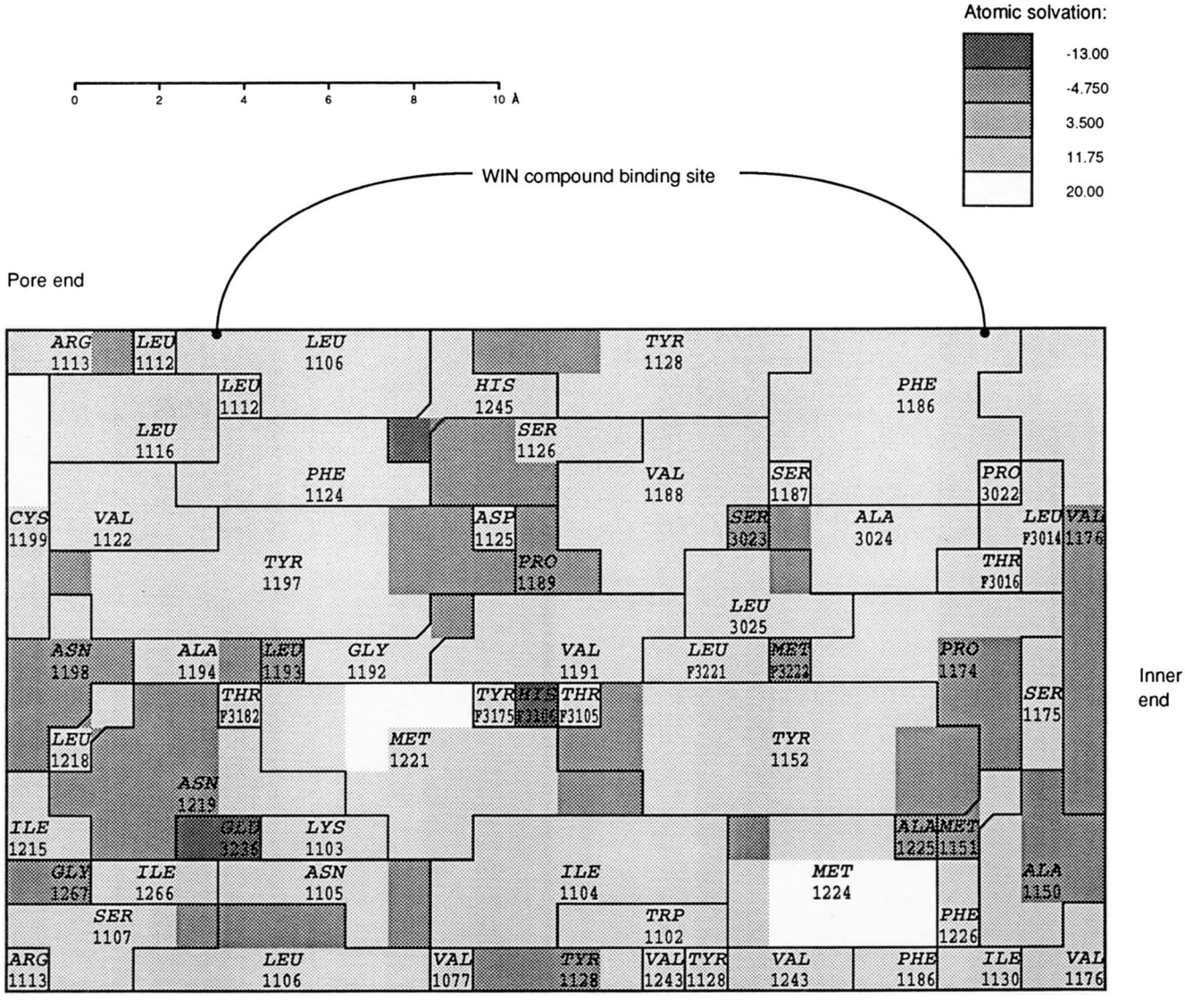


Fig. 6. Drug-binding pocket of HRV14: Six guide points were used to define a smooth curved line that follows the general path of the antiviral agent, WIN 51711 (Smith et al., 1986; McKinlay, 1985), and extends to the ends of the pocket. At 0.75-Å intervals, the Van der Waals surface of the viral pocket was projected outward onto a tube that followed the guide points. The tube was cut and unrolled. Shading of the atoms is according to atomic solvation parameters (Eisenberg & McLachlan, 1986), modified by averaging between atoms where the location of a charge is often crystallographically ambiguous.

dimensional graphics makes these maps useful for the dissemination of structural results to the wider scientific community. Trends in surface properties can be seen at a glance, while systematic labeling facilitates the planning of experiments to probe the functions of a molecule by site-directed mutagenesis. These advantages were apparent from the roadmaps of Rossmann and Palmenberg (1988). With the increased speed of map production, brought by the automation described here, through the easy representation of many physical, chemical, and sequence properties and with the support of projections suitable for the display not just of viruses but of other macromolecules, it is expected that these methods will find wide application to the illustration and comparison of many macromolecules.

The program "RoadMap" will be distributed under license from Purdue Research Foundation, free of charge for noncommercial use. Those interested should contact the author or the Purdue Research Foundation.

Acknowledgments

This work was inspired by the figures of Rossmann and Palmenberg (1988). Using microcomputer drawing programs, Jean-Yves

Sgro has illustrated, with much patience and artistry, the surface topology of several viruses (see, for example, Fig. 1 of Tsao et al. [1991]). This suggested that rapid and automatic calculation of roadmaps would prove to be a powerful tool in the interpretation of biological structures. I thank Ann Palmenberg for supplying the picornaviral sequence alignment used for the calculation of Figure 2. I gratefully acknowledge the help and support of Michael Rossmann in all aspects of this work, in whose laboratory this work was performed, with the assistance of grants from the National Science Foundation, National Institutes of Health, and Medical Research Council.

References

Adobe Systems, Inc. (1990). *PostScript Language Reference Manual*, 2nd Ed. Addison-Wesley, Reading, Massachusetts.

Blake, C.C.F., Koenig, D.F., Mair, G.A., North, A.C.T., Phillips, D.C., & Sarma, V.R. (1965). Structure of hen egg-white lysozyme, a three-dimensional Fourier synthesis at 2 Å resolution. *Nature (Lond.) 206*, 757-761.

Chapman, M.S. & Rossmann, M.G. (1993). Structure, sequence and function correlations among parvoviruses. *Virology*, in press.

Connolly, M.L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science 221*, 709-713.

Dayhoff, M.O., Schwartz, R.M., & Orcutt, B.C. (1979). A model of evolutionary change in proteins. Detecting distant relationships: Computer methods and results. In *Atlas of Protein Sequence and Structure*, (Dayhoff, M.O., Ed.) Vol. 5, Suppl. 3, pp. 353-358. National Biomedical Research Foundation, Washington, D.C.

Diamond, R. (1974). Real-space refinement of the structure of hen egg-white lysozyme. *J. Mol. Biol. 82*, 371-391.

Eisenberg, D. & McLachlan, A.D. (1986). Solvation energy in protein folding and binding. *Nature (Lond.) 319*, 199-203.

Fauchere, J.-L. & Pliska, V. (1983). Hydrophobic parameters π of amino-acid side chains from the partitioning of *N*-acetyl-amino-acid amides. *Eur. J. Med. Chem. Chim. Ther. 18*, 369-375.

Gribskov, R.R. & Burgess, M. (1986). Sigma factors from *E. coli, B. subtilis*, phage SP01, and phage T4 are homologous proteins. *Nucleic Acids Res. 14*, 6745-6763.

Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C., & Rupley, J.A. (1972). Vertebrate lysozymes. In *The Enzymes* (Boyer, P.D., Ed.), pp. 665-868. Academic Press, New York and London.

Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr. 11*, 268-272.

Jones, T.A., Zou, J.-Y., Cowan, S.W., & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A47*, 110-119.

Kelly, J.A., Sielecki, A.R., Sykes, B.D., James, M.N., & Phillips, D.C. (1979). X-ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme. *Nature (Lond.) 282*, 875-878.

Kim, K.H., Willingmann, P., Gong, Z.X., Kremer, M.J., Chapman, M.S., Minor, I., Oliveira, M.A., Rossmann, M.G., Andries, K., Diana, G.D., Dutko, F.J., McKinlay, M.A., & Pevear, D.C. (1993). A comparison of the anti-rhinoviral drug binding pocket in HRV14 and HRV1A. *J. Mol. Biol.*, in press.

Kim, S., Smith, T.J., Chapman, M.S., Rossmann, M.G., Pevear, D.C., Dutko, F.J., Felock, P.J., Diana, G.D., & McKinlay, M.A. (1989). The crystal structure of human rhinoviruses serotype 1A (HRV1A). *J. Mol. Biol. 210*, 91-111.

Kraulis, P. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 946-950.

McKenna, R., Xia, D., Willingmann, P., Ilag, L.L., Krishnaswamy, S., Rossmann, M.G., Olson, N.H., Baker, T.S., & Incardona, N.L. (1992). Atomic structure of single-stranded DNA bacteriophage ϕX174 and its functional implications. *Nature (Lond.) 355*, 137-143.

McKinlay, M.A. (1985). WIN51711: A new systematically active broad spectrum antipicornavirus agent. *J. Antimicrob. Chemother. 16*, 284-286.

Olson, N.H., Kolatkar, P.R., Oliveira, M.A., Cheng, R.H., Greve, J.M., McClelland, A., Baker, T.S., & Rossmann, M.G. (1993). Structure of a human rhinovirus complexed with its receptor molecule. *Proc. Natl. Acad. Sci. USA 90*, 507-511.

Palmenberg, A.C. (1989). Sequence alignments of picornaviral capsid proteins. In *Molecular Aspects of Picornavirus Infection and Detection. Semler* (Ehrenfeld, B.L., Ed.), pp. 211-241. American Society for Microbiology, Washington, D.C.

Pevear, D.C., Fancher, M.J., Felock, P.J., Rossmann, M.G., Miller, M.S., Diana, G., Treasurywala, A.M., McKinlay, M.A., & Dutko, F.J. (1989). Conformational change in the floor of the human rhinovirus canyon blocks adsorption to HeLa cell receptors. *J. Virol. 63*, 2002-2007.

Rossmann, M.G., Arnold, E., Erickson, J.W., Frankenberger, E.A., Griffith, J.P., Hecht, H.J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G., Rueckert, R.R., Sherry, B., & Vriend, G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature (Lond.) 317*, 145-153.

Rossmann, M.G. & Palmenberg, A.C. (1988). Conservation of the putative receptor attachment site in picornaviruses. *Virology 164*, 373-382.

Sherry, B., Mosser, A.G., Colonno, R.J., & Rueckert, R.R. (1986). Use of monoclonal antibodies to identify four neutralization immunogens on a common cold picornavirus, human rhinovirus 14. *J. Virol. 57*, 246-257.

Smith, T.J., Kremer, M.J., Luo, M., Vriend, G., Arnold, E., Kamer, G., Rossmann, M.G., McKinlay, M.A., Diana, G.D., & Otto, M.J. (1986). The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science 233*, 1286-1293.

Terwilliger, T.C. & Eisenberg, D. (1982). The structure of melittin. II. interpretation of the structure. *J. Biol. Chem. 257*, 6016-6022.

Tsao, J., Chapman, M.S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T.J., Rossmann, M.G., Compans, R.W., & Parrish, C.R. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science 251*, 1456-1464.