*J. Mol. Biol.* (1971) **55**, 379–400

# The Interpretation of Protein Structures: Estimation of Static Accessibility

B. Lee and F. M. Richards

*Department of Molecular Biophysics and Biochemistry, Yale University New Haven, Conn., U.S.A.*

(*Received 14 September 1970*)

A program is described for drawing the van der Waal's surface of a protein molecule. An extension of the program permits the accessibility of atoms, or groups of atoms, to solvent or solute molecules of specified size to be quantitatively assessed. As defined in this study, the accessibility is proportional to surface area. The accessibility of all atoms in the twenty common amino acids in model tripeptides of the type Ala-X-Ala are given for defined conformation. The accessibilities are also given for all atoms in ribonuclease-S, lysozyme and myogoblin. Internal cavities are defined and discussed. Various summaries of these data are provided. Forty to fifty per cent of the surface area of each protein is occupied by non-polar atoms. The actual numerical results are sensitive to the values chosen for the van der Waal's radii of the various groups. Since there is uncertainty over the correct values for these radii, the derived numbers should only be used as a qualitative guide at this stage.

The average change in accessibility for the atoms of all three proteins in going from a hypothetical extended chain to the folded conformation of the native protein is about a factor of 3. This number applies to both polar (nitrogen and oxygen) and non-polar (carbon and sulfur) atoms considered separately. The larger non-polar amino acids tend to be more "buried" in the native form of all three proteins. However, for all classes and for residues within a given class the accessibility changes on folding tend to be highly variable.

## 1. Introduction

The successful elucidation of the structure of a protein by single-crystal diffraction procedures provides a list of atomic co-ordinates whose reliability will vary in different parts of the molecule. The presentation and analysis of such three-dimensional information in terms of chemical detail or biological function is only just beginning. The most successful general starting point is an accurate physical model based on the Kendrew skeleton parts. However, the details of side-chain packing are often not easy to comprehend in skeleton models without very detailed and lengthy examination. Physical space-filling models are another form of presentation, but they are very difficult to construct accurately when structures with hundreds of atoms are involved, and the interior of the macromolecule is usually invisible because of the opaque parts. In this report we suggest that a computer-produced space-filling model prepared on a stack of clear plastic sheets may be a useful presentation complementary to the various physical models. In the course of working with this graphics program, it became clear that certain useful summary information could be obtained easily.

The topology of the surface of a protein is intimately related to its function; parts of the surface are directly involved in interactions with other molecules; the solvent–protein interface is almost certainly related to the structure of the native molecule; and the chemical reactivity of the various functional groups will depend on their relation to this interface. An initial detailed description of such an irregular surface might be described by a list of the accessibility of each atom or group of atoms in the structure to a solvent (or solute) molecule of defined size. Such accessibility estimates can be derived directly from the graphics program. Since the calculations are based on the co-ordinate list for the protein which at present reflects no information on vibrations or flexibility, the numbers derived are static accessibilities. It is hoped that such numbers will be more directly useful for the interpretation of chemical data than the co-ordinates themselves.

## 2. Procedures

### (a) *The drawing program*

Van der Waal's radii are assigned to each atom or group of atoms. The set of radii chosen is that given by Bondi (1964) and quoted by Scott & Scheraga (1966) and by Ramachandran & Sasisekharan (1968), Table 1. Hydrogen atoms are not considered individually but are included where appropriate in the estimate for the group radius. This approximation is warranted by the present rather low accuracy of the X-ray co-ordinates, and the general nature of the information to be derived in this work.

The structure is thus represented by a set of interlocking spheres of appropriate van der Waal's radii. The continuous structure is sectioned by a set of parallel planes with a predetermined spacing. The intersection of each sphere with a given plane appears as a circle. The overlapping arcs of circles representing atoms of the same molecule are eliminated. The drawing in any one plane thus becomes the trace of the envelope of the van der Waal's surface for the molecule. Overlapping circles representing atoms from different molecules are not eliminated in order to distinguish one molecule from another and to facilitate easy recognition of serious overlap between symmetry-related neighboring molecules. The polar atoms, oxygen and nitrogen, are dotted and labeled, the non-polar atoms, carbon and sulfur, appear as solid lines. The sequence number of a given residue is written at the center of both the $\alpha$- and $\beta$- carbon atoms. The skeleton covalent bonds and hydrogen bonds are shown between atom centers to assist viewing. A picture of some stacked sections of ribonuclease-S (Wyckoff *et al.*, 1970) is shown in Plate I, where various types of interactions and packing can be seen.

TABLE 1

*Assumed van der Waal's radii from Bondi (1964)*

| | |
|---|---|
| Main-chain $\alpha$-carbon atom | 1·70 Å |
| Main-chain carbonyl oxygen atom | 1·52 Å |
| Main-chain amide NH group | 1·55 Å |
| Main-chain carbonyl carbon atom | 1·80 Å |
| All side-chain atoms and groups | 1·80 Å |
| Iron atom in the heme of myoglobin | 0·64 Å † |

† This is the value for the crystal radii of $Fe^{3+}$ given by Pauling (1960).

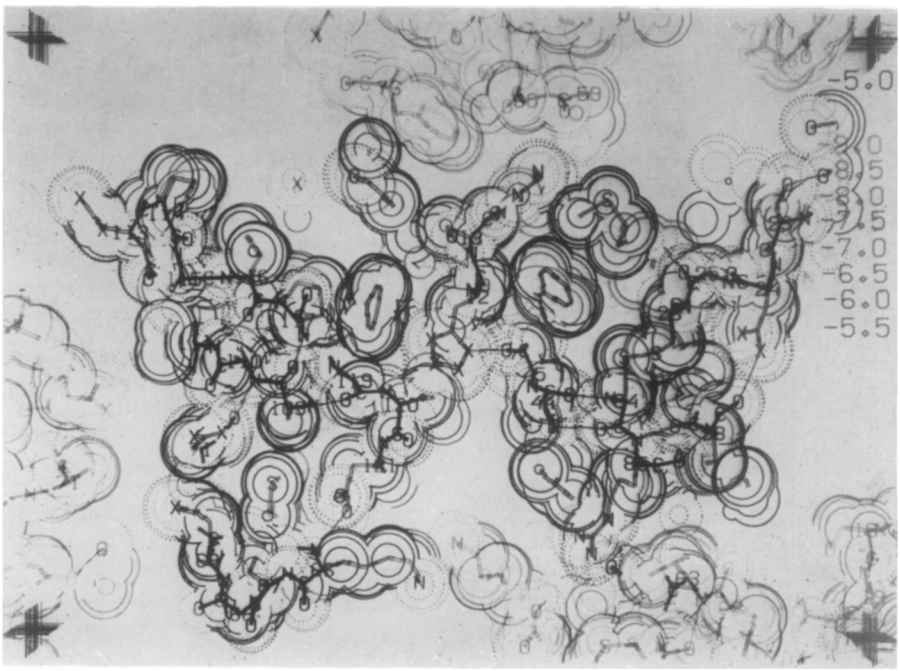


PLATE I. Some stacked sections of the van der Waal's contact drawing for pancreatic ribonuclease-S. The central molecule is shown in black while parts of the symmetry-related molecules appear in a lighter shade of gray. The construction and labelling of these drawings are discussed in the text.

[*facing p. 380*

### (b) *Estimation of static accessibility*

An atom, or group of atoms, is defined as accessible if a solvent molecule of specified size can be brought into van der Waal's contact. In this study the solvent molecule has been assumed to be a sphere of radius 1·4 Å. For calculation a sphere is centered at each atomic position in the co-ordinate list and is assigned a radius equal to the sum of that of the atom and that of the solvent molecule. The drawing program is now entered (with or without the graphics output). The surface computed will be the locus of the center of a solvent molecule as it rolls along the protein making the maximum permitted contact. If any part of an arc around a given protein atom is "drawn" then that atom is accessible. The length of the arc will be a measure of accessibility in that plane. The total accessibility will be proportional to the summed length of all arcs drawn for that atom.

The terms "accessible surface area" and "accessibility" are used in subsequent sections of this paper and defined as follows. Accessible surface area, A, of an atom is the area on the surface of a sphere of radius $R$, on each point of which the center of a solvent molecule can be placed in contact with this atom without penetrating any other atoms of the molecule. The radius $R$ is given by the sum of the van der Waal's radius of the atom and the chosen radius of the solvent molecule. An approximation to this area is computed by this program using the formula:

$$\text{accessible surface area} = A = \sum (R/\sqrt{R^2 - Z_i^2}) \cdot D \cdot L_i, \tag{1}$$

$$D = \Delta Z/2 + \Delta' Z \tag{2}$$

where $L_i$ is the length of the arc drawn on a given section $i$, $Z_i$ is the perpendicular distance from the center of the sphere to the section $i$, $\Delta Z$ is the spacing between the sections, and $\Delta' Z$ is $\Delta Z/2$ or $R$-$Z_i$, whichever is smaller. Summation is over all of the arcs drawn for the given atom. The accessiblity is defined simply by the accessible surface area divided by $4\pi R^2$ and multiplied by 100.

$$\text{Accessibility} = 100\ A/4\pi R^2. \tag{3}$$

It must be emphasized that the number really refers to *static* accessibility or *static* accessible surface area since no account has been taken of potential flexibility or movement of groups in the structure.

### (c) *Identification of cavities*

Cavities inside a protein molecule that are large enough to accommodate at least one solvent molecule can be identified with the computer programs developed. This is done by isolating all the accessibility contours with concave curvature on each section and eliminating those that have a channel leading from the inside of this concave contour to the outside of the molecule. The elimination is done with the aid of a stack of the graphics output of accessibility contours recorded on films of the cathode ray tube display. An example of a section that shows such a cavity is given in Figure 1.

The "volume" of a cavity was calculated by computing the area of the region inside of the cavity contour, multiplying it by the spacing between the sections, and summing the result over all the planes that show the cavity. This number represents that volume of space which can be occupied by the center of a solvent molecule. A cavity, large enough to accommodate one solvent molecule but not large to leave any room for the molecule to move about, has a volume of just zero by this definition.

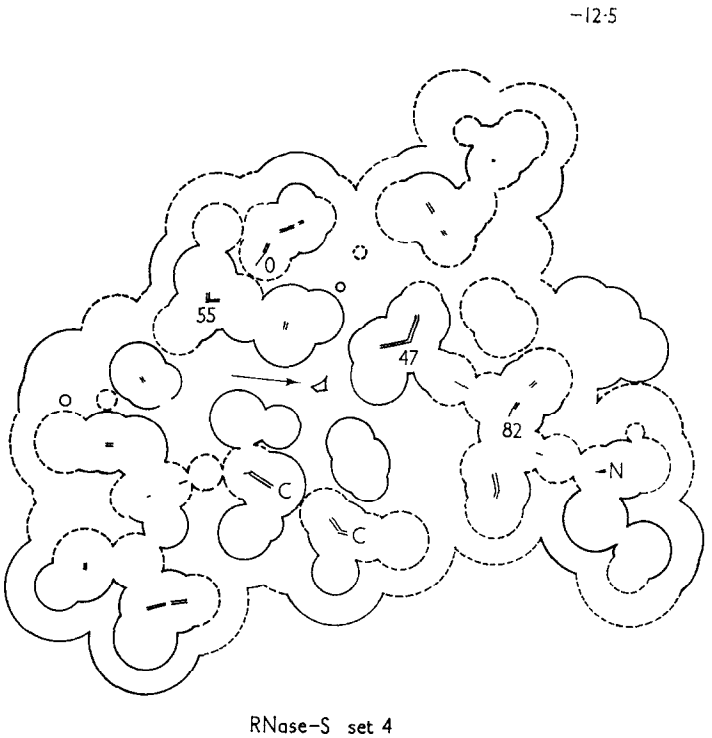


FIG. 1. Superposition of van der Waal's contours and accessibility contours of a section of the ribonuclease-S molecule. The arrow indicates the cavity inside the molecule large enough to accommodate a solvent molecule of radius 1·4 Å.

## 3. Results of Static Accessibility Calculations

### (a) *Model systems*

The computer programs described were used to compute the accessibilities of a number of model systems prior to their application to macromolecules. The model systems were generated using standard bond lengths and angles, except for the side chains of proline, histidine and tryptophan for which values reported in the literature from the crystallographic studies were used (Leung & Marsh, 1958; Donohue & Caron, 1964; Pasternak, 1956).

The first model systems were tripeptides of the form Gly-X-Gly and Ala-X-Ala, where X is the residue whose accessibility was to be computed. The main-chain dihedral angles were chosen to be those of the extended $\beta$ conformation of silk ($\phi = -140°$, $\psi = 135°$)† except when X was a proline residue, in which case the conformation of poly-L-proline II ($\phi = -77{\cdot}2°$, $\psi = 145{\cdot}9°$) was adopted. Two sets of dihedral angles for side chains were chosen. For the first set (the $\beta$, *trans* set) all the variable dihedral angles were chosen to be 180° or *trans*. Whenever there was a branch, the first branch

† For the conventions of dihedral angles, see the editors appendix to the article by Ramachandran & Sasisekharan (1968), and the statement by IUPAC-IUB Commission on Biochemical Nomenclature on Description of Conformation of Polypeptide Chains, *J Mol. Biol.* **55**, 299. According to this convention, a dihedral angle of zero corresponds to the *cis*-conformation for both the main chain and side chains. This convention is used throughout this paper.

## TABLE 2

*Static accessibility of amino-acid residues, X, in model tripeptides Ala-X-Ala*

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GLY | CA | N | C | O | | | | | | | | | | | |
| 1 | 21.7 | 7.7 | 7.5 | 35.1 | | | | | | | | | | | |
| ALA | CA | N | C | O | CB | | | | | | | | | | |
| 1 | 3.0 | 5.7 | 3.2 | 32.6 | 48.2 | 48.2 | | | | | | | | | |
| VAL | CA | N | C | O | CB | CG1 | CG2 | | | | | | | | |
| 1 | 2.2 | 5.5 | 2.2 | 32.3 | 6.5 | 38.2 | 37.8 | 27.5 | | | | | | | |
| LEU | CA | N | C | O | CB | CG | CD1 | CD2 | | | | | | | |
| 1 | 2.3 | 5.7 | 2.0 | 29.2 | 16.4 | 10.0 | 50.5 | 19.3 | 24.0 | | | | | | |
| 2 | 0.8 | 5.6 | 5.1 | 26.5 | 16.6 | 4.2 | 34.1 | 50.6 | 26.4 | | | | | | |
| ILE | CA | N | C | O | CB | CG2 | CG1 | CD | | | | | | | |
| 1 | 2.3 | 2.4 | 2.2 | 27.1 | 5.9 | 33.7 | 15.1 | 50.2 | 26.2 | | | | | | |
| 2 | 2.9 | 1.6 | 3.4 | 25.1 | 7.7 | 32.8 | 17.4 | 50.1 | 27.0 | | | | | | |
| PRO | CA | N | C | O | CB | CG | CD | | | | | | | | |
| 1 | 5.7 | 0.0 | 0.5 | 22.8 | 25.2 | 32.8 | 19.8 | 25.9 | | | | | | | |
| CYS | CA | N | C | O | CB | SG | SD | | | | | | | | |
| 1 | 1.6 | 5.7 | 2.1 | 32.4 | 18.4 | 18.9 | 15.9 | 17.7 | | | | | | | |
| 2 | 2.3 | 5.5 | 6.4 | 26.5 | 20.8 | 20.0 | 22.1 | 20.9 | | | | | | | |
| MET | CA | N | C | O | CB | CG | SD | CE | | | | | | | |
| 1 | 2.3 | 5.7 | 2.2 | 32.3 | 17.9 | 11.4 | 31.8 | 58.5 | 29.9 | | | | | | |
| PHE | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | | | | |
| 1 | 1.4 | 5.7 | 2.0 | 32.3 | 18.3 | 3.4 | 24.7 | 28.9 | 29.1 | 12.4 | 6.6 | 17.6 | | | |
| 2 | 1.4 | 5.6 | 3.0 | 26.5 | 18.8 | 2.5 | 20.4 | 26.4 | 25.6 | 29.0 | 14.7 | 18.3 | | | |
| TRY | CA | N | C | O | CB | CG | CD1 | NE | CE1 | CZ1 | CH | CZ2 | CE2 | CD2 | |
| 1 | 1.6 | 5.9 | 2.1 | 32.8 | 15.2 | 3.5 | 7.2 | 19.4 | 3.8 | 26.4 | 28.5 | 28.2 | 23.4 | 4.3 | 15.6 19.4 16.0 |
| 2 | 2.4 | 5.3 | 6.6 | 26.5 | 16.7 | 2.4 | 10.8 | 25.2 | 4.9 | 26.9 | 26.7 | 27.9 | 23.1 | 4.9 | 16.0 25.2 17.0 |
| SER | CA | N | C | O | CB | OH | | | | | | | | | |
| 1 | 2.3 | 5.7 | 2.2 | 32.3 | 21.4 | 43.9 | 21.4 | 43.9 | 32.6 | | | | | | |
| THR | CA | N | C | O | CB | CG | OH | | | | | | | | |
| 1 | 2.2 | 5.5 | 2.2 | 32.3 | 6.4 | 38.3 | 36.5 | 22.3 | 36.5 | 27.1 | | | | | |
| TYR | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | OH | | | |
| 1 | 1.4 | 5.7 | 2.0 | 32.3 | 18.3 | 3.4 | 24.7 | 24.7 | 3.9 | 8.5 | 6.6 | 51.6 | 12.9 | 51.6 | 17.7 |
| 2 | 1.4 | 5.6 | 3.0 | 26.5 | 18.8 | 2.5 | 20.4 | 22.1 | 2.1 | 16.0 | 14.7 | 48.6 | 13.8 | 48.6 | 18.1 |
| ASP | CA | N | C | O | CB | CG | OD1 | OD2 | | | | | | | |
| 1 | 1.4 | 5.7 | 2.0 | 32.3 | 17.9 | 2.9 | 50.7 | 19.9 | 10.4 | 35.3 | 22.8 | | | | |
| 2 | 3.0 | 1.0 | 0.9 | 22.9 | 22.0 | 1.9 | 39.4 | 39.7 | 12.0 | 39.5 | 25.7 | | | | |
| GLU | CA | N | C | O | CB | CG | CD | OE1 | OE2 | | | | | | |
| 1 | 2.3 | 5.7 | 2.2 | 32.3 | 9.9 | 11.4 | 3.7 | 50.9 | 42.6 | 8.3 | 46.8 | 23.7 | | | |
| 2 | 2.9 | 5.4 | 6.4 | 26.5 | 12.5 | 10.9 | 3.7 | 50.5 | 42.6 | 9.1 | 46.5 | 24.1 | | | |
| HIS | CA | N | C | O | CB | CG | ND | CE | NE | CD | | | | | |
| 1 | 1.7 | 5.8 | 2.1 | 33.4 | 18.2 | 3.5 | 27.6 | 30.9 | 24.8 | 7.6 | 15.0 | 26.2 | 18.8 | | |
| 2 | 7.6 | 2.1 | 0.8 | 22.6 | 21.6 | 2.1 | 24.2 | 30.9 | 31.2 | 16.0 | 17.6 | 27.7 | 21.0 | | |
| LYS | CA | N | C | O | CB | CG | CD | CE | NZ | | | | | | |
| 1 | 2.3 | 5.7 | 2.2 | 32.3 | 16.8 | 10.5 | 21.7 | 26.0 | 55.8 | 18.7 | 55.8 | 26.1 | | | |
| ARG | CA | N | C | O | CB | CG | CD | NE | CZ | NH1 | NH2 | | | | |
| 1 | 2.3 | 5.7 | 2.2 | 32.3 | 16.8 | 10.5 | 17.8 | 16.9 | 3.6 | 51.3 | 45.5 | 12.2 | 37.9 | 23.2 | |

The first row for each amino acid is for the $\beta$, *trans* set of conformations. The second row, when given, is for the $\beta$, alternate set (see text). The last column of each row gives the average over the side-chain atoms. For polar residues, separate averages over the side-chain non-polar and polar atoms are given in the third and the second from the last columns, respectively. The seventh amino acid in the Table is for cystine/2 residue; the atom S$\delta$ is included in computing the average given in the last column.

was used for the criterion for *trans* except for tryptophan, for which the second branch (the branch of the atom Cδ2) was used. For the second set (the β alternate set), the side-chain dihedral angles were altered for some of the amino acids as follows—cystine: $\chi^1 = -60°, \chi^2 = -90°, \chi^3 = -90°$; leucine: $\chi^1 = 180°, \chi^{2,1} = 60°, \chi^{2,2} = 180°$: isoleucine $\chi^{1,1} = 60°$, $\chi^{1,2} = -60°$, $\chi^2 = 180°$; aspartate: $\chi^1 = 60°$, $\chi^{2,1} = 90°$; glutamate; $\chi^1 = -60°, \chi^2 = 180°, \chi^{3,1} = 0°$; histidine: $\chi^1 = 72{\cdot}0°$, $\chi^{2,1} = -120{\cdot}5°$ (with respect to Nδ); phenylalanine and tyrosine: $\chi^1 = 180°$, $\chi^{2,1} = 90°$; tryptophan: $\chi^1 = -66{\cdot}84°$, $\chi^{2,1} = 60{\cdot}59°$ (with respect to Cδ1). The side-chain conformations cited above for histidine and tryptophan are those observed in the crystal structures (*loc. cit.*).

The results of calculations for the α-carbon and the side-chain atoms of the residue X in Gly-X-Gly were essentially identical to those for Ala-X-Ala. For other main-chain atoms of the central residue, the accessibilities tended to increase by varying amounts that were less than 3 when glycines were used in place of alanines to surround the residue, The accessibilities of the atoms of residue X in Ala-X-Ala are listed in Table 2.

In order to investigate the effect of variations in the main-chain conformation, the model system Gly-Ala-Gly was constructed in nine different conformations. The main-chain dihedral angles and the accessibilities in each of these conformations are listed in Table 3.

TABLE 3

*Static accessibility of the alanyl residue in Gly-Ala-Gly in different conformations*

| φ† | ψ† | Cα | N | C | O | Cβ | Comment |
|---|---|---|---|---|---|---|---|
| −140 | 135 | 3·2 | 6·3 | 4·9 | 34·2 | 47·9 | β-sheet of silk |
| −120 | 160 | 3·1 | 5·5 | 4·4 | 34·1 | 47·0 | |
| −100 | 135 | 3·6 | 6·9 | 1·8 | 31·7 | 49·8 | |
| − 70 | 150 | 6·5 | 8·8 | 0·2 | 26·1 | 47·9 | near collagen fold |
| −120 | 80 | 4·0 | 6·7 | 4·2 | 31·7 | 48·4 | |
| −100 | 0 | 4·9 | 1·1 | 3·4 | 35·0 | 51·0 | |
| − 50 | − 26 | 8·6 | 4·8 | 1·3 | 30·3 | 51·4 | $3_{10}$-helix |
| − 57 | − 48 | 8·0 | 5·9 | 1·6 | 30·0 | 51·8 | right-handed α-helix |
| 57 | 48 | 11·2 | 9·5 | 5·1 | 23·9 | 43·8 | left-handed α-helix |

† Main-chain dihedral angles in degrees. For convention, see Footnote on p. 382.

The third set of model systems investigated were α-helical structures with main-chain dihedral angles of $\phi = -57{\cdot}8°$ and $\psi = -47{\cdot}0°$. The results for polyglycine, polyalanine, and polyserine in two different side-chain conformations are given in Table 4.

The spacing $\Delta Z$ between the slicing planes used to obtain figures reported herein was 0·25 Å. Accuracy of the computation was estimated by repeating the computation for many Gly-X-Gly systems after slicing the structure in several different directions. The root-mean-square and maximum deviations in accessibility were 0·3 and 0·5, respectively. For α-helical polyalanine, a nine residue polymer was constructed and sectioned in two different directions. The results quoted are for the central 5th residue, averaged over the two independent computations. The root-mean-square and maximum deviations for these two computations were 0·1 and 0·2, respectively. For α-helical polyglycine and polyserine, an eleven-residue polymer was constructed and computation was done once. The results for the central three residues (the 5th, 6th and 7th) were

TABLE 4

*Static accessibilities in the α-helical conformation*

| | Cα | N | C | O | Cβ | OH |
|---|---|---|---|---|---|---|
| Polyglycine | 23·8 | 0·9 | 5·3 | 12·1 | | |
| Polyalanine | 5·2 | 0·5 | 1·6 | 5·5 | 36·0 | |
| Polyserine (II) † | 4·1 | 0·3 | 0·1 | 1·6 | 8·9 | 48·0 |
| Polyserine (III)‡ | 2·3 | 0·1 | 0·3 | 1·2 | 15·7 | 35·3 |

† The side-chain dihedral angle $\chi^1 = 180°$
‡ The side-chain dihedral angle $\chi^1 = -60°$

averaged. The root-mean-square and maximum deviations in this case were 0·2 and 0·6, respectively, for the polyglycine and 0·2 and 0·4, respectively, for the two polyserines. For Gly-Gly-Gly, computations were also carried out with $\Delta Z$=0·5 Å. The root-mean-square and maximum deviations between the results with $\Delta Z$=0·5 Å and 0·25 Å were 0·4 and 0·5.

In summary the method of estimating accessibilities gives results which are reliable to about ±0·5 regardless of sectioning direction if the interplanar spacing is 0·25 Å or less. The error increases to ±1 to 2 if the sectioning interval is increased to 0·5 Å. For many purposes the increased computing costs required by the finer sectioning will not be warranted.

(b) *Proteins*

Accessibility calculations were performed for ribonuclease-S (Wyckoff *et al.*, 1970), lysozyme (D. C. Phillips, personal communication), and myoglobin (Watson, 1969). Co-ordinates for individual atoms of these proteins are available from the single crystal X-ray diffraction studies. The solvent radius and the spacing between the slicing planes were 1·4 Å and 0·25 Å, respectively. The calculated accessibilities are given in Table 5. For ribonuclease-S, computations were also performed with 0·88 Å and 0·5 Å intervals of the slicing planes. The root-mean-square and maximum deviations between the calculated accessibilities with the spacings of 0·88 Å and 0·5 Å are 1·1 and 5·0; between 0·5 Å and 0·25 Å, 0·5 and 2·2, respectively. It should be emphasized again that the calculated accessibilities are static accessibilities; no account was taken of the potentially important local or massive flexibility of the molecule. In addition, the average accuracy of the X-ray crystallographic determination of atomic co-ordinates is regarded to be not much better than ~0·25 Å and probably much worse for many atoms on the surface of the protein. Individual numbers given in the Table must, therefore, be used with caution.

The cavities found in the three proteins are listed in Table 6. The electron density map of ribonuclease-S shows that one large and two small cavities found in this molecule are genuine empty spaces. On the other hand, three "buried" water molecules are found in lysozyme (D. C. Phillips, personal communication). Calculations were performed for lysozyme with and without these molecules. One major and six minor cavities are found without these molecules, of which two minor ones disappear when the three water molecules are included in the computation. The remaining five cavities are given in Table 6. In myoglobin, 13 large and small cavities are found. The cavities G1 and G2 are very close to one another and could almost be regarded as

TABLE 5

*Static accessibilities for ribonuclease-S, lysozyme, and myoglobin*

(a) *RNAse-S*

| GLY | CA | N | C | O |
|---|---|---|---|---|
| 68 | 20.7 | 0.0 | 2.5 | 41.4 |
| 88 | 27.1 | 11.7 | 4.7 | 30.0 |
| 112 | 22.0 | 6.0 | 4.1 | 5.0 |
| | 23.3 | 5.9 | 3.8 | 25.5 |

| ALA | CA | N | C | O | CB | |
|---|---|---|---|---|---|---|
| 4 | 1.6 | 6.7 | 0.0 | 0.0 | 26.5 | 26.5 |
| 5 | 0.0 | 0.7 | 0.8 | 1.7 | 25.0 | 25.0 |
| 6 | 0.5 | 0.7 | 0.0 | 0.6 | 25.5 | 25.5 |
| 19 | 1.4 | 0.0 | 2.1 | 17.8 | 40.8 | 40.8 |
| 20 | 0.0 | 0.0 | 5.8 | 33.2 | 25.8 | 25.8 |
| | | | | 21.8 | | |
| 52 | 0.5 | 1.6 | 0.4 | 0.9 | 39.3 | 39.3 |
| 56 | 6.1 | 0.4 | 0.0 | 1.0 | 24.0 | 24.0 |
| 64 | 0.0 | 3.6 | 0.2 | 17.5 | 38.4 | 38.4 |
| 96 | 0.4 | 0.0 | 0.0 | 0.0 | 23.6 | 23.6 |
| 102 | 0.6 | 0.0 | 0.0 | 0.0 | 27.0 | 27.0 |
| 109 | 0.0 | 0.0 | 0.0 | 0.0 | 8.1 | 8.1 |
| 122 | 4.3 | 0.9 | 0.0 | 0.0 | 16.0 | 16.0 |
| | 1.3 | 1.2 | 0.8 | 7.3 | 26.7 | 26.7 |

| VAL | CA | N | C | O | CB | CG1 | CG2 | |
|---|---|---|---|---|---|---|---|---|
| 43 | 0.0 | 0.0 | 0.1 | 5.0 | 9.1 | 3.4 | 21.9 | 11.5 |
| 47 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 54 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.3 | 0.1 |
| 57 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 |
| 63 | 3.8 | 0.0 | 0.3 | 0.0 | 9.3 | 12.5 | 2.7 | 8.2 |
| 108 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 116 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 11.1 | 4.3 | 5.1 |
| 118 | 0.0 | 0.0 | 0.0 | 5.9 | 0.1 | 1.8 | 3.9 | 1.9 |
| 124 | 5.1 | 0.0 | 0.8 | 22.4 | 9.7 | 19.3 | 14.2 | 14.4 |
| | | | | 0.0 | | | | |
| | 1.0 | 0.0 | 0.1 | 3.3 | 3.1 | 5.4 | 5.3 | 4.6 |

| LEU | CA | N | C | O | CB | CG | CD1 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|
| 35 | 0.0 | 0.0 | 0.0 | 0.0 | 3.4 | 0.0 | 0.0 | 1.8 | 1.3 |
| 51 | 0.0 | 4.2 | 0.0 | 0.0 | 20.3 | 1.7 | 3.6 | 30.3 | 14.0 |
| | 0.0 | 2.1 | 0.0 | 0.0 | 11.9 | 0.9 | 1.8 | 16.1 | 7.6 |

| ILE | CA | N | C | O | CB | CG2 | CG1 | CD | |
|---|---|---|---|---|---|---|---|---|---|
| 81 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 1.1 | 3.2 | 1.1 |
| 106 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | -0.0 | -0.0 | 0.0 |
| 107 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 19.9 | 5.1 |
| | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.3 | 0.4 | 7.7 | 2.1 |

| PRO | CA | N | C | O | CB | CG | CD | |
|---|---|---|---|---|---|---|---|---|
| 42 | 0.1 | 0.0 | 0.0 | 2.0 | 22.6 | 28.8 | 9.4 | 20.3 |
| 93 | 0.7 | 0.0 | 1.3 | 0.0 | 19.1 | 23.9 | 16.3 | 19.8 |
| 114 | 2.8 | 0.0 | 0.0 | 4.8 | 21.4 | 23.7 | 26.5 | 23.9 |
| 117 | 0.0 | 0.0 | 0.0 | 1.4 | 0.0 | 4.0 | 0.0 | 1.3 |
| | 0.9 | 0.0 | 0.3 | 2.1 | 15.8 | 20.1 | 13.1 | 16.3 |

| CYS | CA | N | C | O | CB | SG | |
|---|---|---|---|---|---|---|---|
| 26 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 40 | 0.0 | 2.8 | 0.0 | 14.6 | 1.7 | 0.5 | 1.1 |
| 58 | 0.0 | 0.0 | 0.0 | 0.2 | 0.1 | -0.0 | 0.0 |
| 65 | 0.0 | 0.0 | 0.7 | 2.8 | 0.0 | 2.1 | 1.1 |
| 72 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 84 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 95 | 0.0 | 2.4 | 0.1 | 2.1 | 0.0 | 5.5 | 2.8 |
| 110 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.0 | 0.6 | 0.1 | 2.5 | 0.2 | 1.0 | 0.6 |

| MET | CA | N | C | O | CB | CG | SD | CE | |
|---|---|---|---|---|---|---|---|---|---|
| 13 | 0.0 | 0.0 | 2.0 | 2.3 | 4.1 | 2.0 | 0.0 | 0.0 | 1.5 |
| 29 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 3.1 | 16.3 | 0.0 | 4.9 |
| 30 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 79 | 0.0 | 0.0 | 0.6 | 0.0 | 0.0 | 5.0 | 0.6 | 1.6 | 1.8 |
| | 0.0 | 0.0 | 0.7 | 0.6 | 1.0 | 2.5 | 4.2 | 0.4 | 2.0 |

| PHE | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 1.6 | 1.5 | 0.4 |
| 46 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 120 | 0.0 | 0.0 | 0.0 | 13.8 | 0.9 | 0.0 | 4.8 | 0.6 | 0.6 | -0.0 | 0.0 | 1.0 |
| | 0.0 | 0.0 | 0.0 | 4.6 | 0.3 | 0.0 | 1.6 | 0.2 | 0.2 | 0.5 | 0.5 | 0.5 |

| SER | CA | N | C | O | CB | OH | | | |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 0.0 | 0.0 | 0.0 | 0.0 | 3.5 | 17.1 | 3.5 | 17.1 | 10.3 |
| 16 | 0.2 | 0.0 | 0.7 | 28.8 | 16.0 | 24.2 | 16.0 | 24.2 | 20.1 |
| 18 | 5.4 | 0.0 | 0.9 | 27.9 | 23.0 | 23.7 | 23.0 | 23.7 | 23.3 |
| 21 | 12.7 | 42.2 | 3.0 | 7.0 | 24.9 | 31.5 | 24.9 | 31.5 | 28.2 |
| 22 | 8.3 | 7.9 | 0.0 | 10.5 | 28.7 | 32.7 | 28.7 | 32.7 | 30.7 |
| 23 | 0.0 | 5.4 | 3.5 | 18.0 | 16.3 | 18.7 | 16.3 | 18.7 | 17.5 |
| 32 | 5.3 | 0.1 | 0.4 | 17.8 | 28.4 | 7.8 | 28.4 | 7.8 | 18.1 |
| 50 | 2.4 | 0.0 | 0.0 | 0.0 | 19.1 | 4.2 | 19.1 | 4.2 | 11.6 |
| 59 | 5.8 | 0.2 | 0.0 | 25.2 | 26.9 | 29.4 | 26.9 | 29.4 | 28.2 |
| 75 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.7 | 0.0 | 1.7 | 0.8 |
| 77 | 2.6 | 0.0 | 0.0 | 0.6 | 22.2 | 29.2 | 22.2 | 29.2 | 25.7 |
| 80 | 0.0 | 0.0 | 0.0 | 0.0 | 7.4 | 16.4 | 7.4 | 16.4 | 11.9 |
| 89 | 7.2 | 5.3 | 0.0 | 2.7 | 10.7 | 36.9 | 10.7 | 36.9 | 23.8 |
| 90 | 0.0 | 0.0 | 0.0 | 12.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 123 | 0.0 | 0.0 | 0.1 | 11.4 | 6.9 | 23.9 | 6.9 | 23.9 | 15.4 |
| | 3.3 | 4.1 | 0.6 | 10.8 | 15.6 | 19.8 | 15.6 | 19.8 | 17.7 |

| THR | CA | N | C | O | CB | CG | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 3.6 | 0.0 | 2.9 | 0.0 | 12.4 | 29.4 | 13.7 | 20.9 | 13.7 | 18.5 |
| 17 | 0.8 | 0.0 | 0.8 | 31.3 | 5.0 | 32.0 | 12.2 | 18.5 | 12.2 | 16.4 |
| 36 | 0.0 | 0.0 | 0.0 | 0.9 | 0.7 | 3.2 | 0.0 | 2.0 | 0.0 | 1.3 |
| 45 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 9.9 | 1.3 | 4.9 | 1.3 | 3.7 |
| 70 | 0.0 | 1.3 | 0.2 | 14.5 | 0.6 | 39.2 | 3.8 | 19.9 | 3.8 | 14.5 |
| 78 | 0.0 | 0.4 | 0.0 | 12.1 | 0.0 | 10.4 | 26.3 | 5.2 | 26.3 | 12.2 |
| 82 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3 | 9.4 | 0.1 | 9.4 | 3.2 |
| 87 | 2.7 | 0.0 | 3.5 | 12.8 | 6.5 | 12.5 | 6.7 | 9.5 | 6.7 | 8.6 |
| 99 | 0.0 | 1.0 | 0.0 | 23.4 | 0.0 | 2.0 | 17.4 | 1.0 | 17.4 | 6.5 |
| 100 | 2.8 | 0.0 | 0.8 | 0.0 | 9.3 | 10.3 | 25.1 | 9.8 | 25.1 | 14.9 |
| | 1.1 | 0.3 | 0.8 | 9.5 | 3.5 | 14.9 | 11.6 | 9.2 | 11.6 | 10.0 |

| TYR | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 25 | 1.7 | 2.1 | 0.0 | 0.0 | 15.9 | 0.5 | 1.8 | 0.1 | 1.1 | 16.0 | 17.3 | 0.1 | 7.5 | 0.1 | 6.6 |
| 73 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 5.8 | 0.0 | 17.8 | 0.8 | 17.8 | 2.9 |
| 76 | 1.4 | 0.0 | 0.3 | 21.8 | 11.7 | 1.5 | 8.7 | 8.0 | 2.0 | 15.4 | 10.8 | 37.0 | 8.3 | 37.0 | 11.9 |
| 92 | 0.0 | 1.9 | 0.0 | 20.2 | 12.5 | 1.5 | 13.0 | 13.5 | 0.6 | 1.4 | 0.2 | 32.0 | 6.1 | 32.0 | 9.3 |
| 97 | 0.0 | 0.0 | 0.0 | 7.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 115 | 0.0 | 1.6 | 0.1 | 9.9 | 12.4 | 1.0 | 5.1 | 11.4 | 2.7 | 7.3 | 11.2 | 13.8 | 7.3 | 13.8 | 8.1 |
| | 0.5 | 0.9 | 0.1 | 9.8 | 8.8 | 0.7 | 4.8 | 5.5 | 1.1 | 7.7 | 6.6 | 16.8 | 5.0 | 16.8 | 6.5 |

| ASP | CA | N | C | O | CB | CG | OD1 | OD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 14 | 0.0 | 0.0 | 0.0 | 11.6 | 0.0 | 0.0 | 3.1 | 7.6 | 0.0 | 5.3 | 2.7 |
| 38 | 0.1 | 0.0 | 0.0 | 25.7 | 22.9 | 0.8 | 0.9 | 37.8 | 11.9 | 19.3 | 15.6 |
| 53 | 0.2 | 0.0 | 0.0 | 2.1 | 0.8 | 1.8 | 14.9 | 36.4 | 1.3 | 25.6 | 13.8 |
| 83 | 0.0 | 0.0 | 0.0 | 0.0 | 1.6 | 0.0 | 28.4 | 0.1 | 0.8 | 14.3 | 7.5 |
| 121 | 1.2 | 0.0 | 1.4 | 2.3 | 0.0 | 0.0 | 1.9 | 18.2 | 0.0 | 10.1 | 5.0 |
| | 0.3 | 0.0 | 0.3 | 8.3 | 5.1 | 0.5 | 9.8 | 20.0 | 2.8 | 14.9 | 8.9 |

| ASN | CA | N | C | O | CB | CG | ND1 | ND2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 24 | 8.3 | 1.0 | 0.0 | 0.0 | 7.8 | 1.5 | 16.6 | 7.9 | 4.6 | 12.3 | 8.5 |
| 27 | 0.0 | 0.0 | 0.0 | 0.4 | 4.4 | 0.5 | 11.0 | 0.2 | 2.4 | 5.6 | 4.0 |
| 34 | 0.0 | 0.0 | 0.0 | 5.5 | 2.0 | 4.0 | 20.5 | 24.8 | 3.0 | 22.7 | 12.9 |
| 44 | 1.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 0.0 | 0.3 | 0.2 |
| 62 | 0.0 | 1.2 | 0.1 | 20.3 | 16.7 | 2.3 | 25.6 | 3.8 | 9.5 | 14.7 | 12.1 |
| 67 | 3.1 | 0.0 | 0.2 | 35.2 | 6.5 | 1.8 | 34.5 | 15.6 | 4.1 | 25.0 | 14.6 |
| 71 | 0.0 | 0.0 | 0.0 | 0.0 | 6.3 | 0.0 | 14.0 | 1.1 | 3.2 | 7.5 | 5.3 |
| 94 | 11.4 | 4.6 | 2.7 | 10.2 | 19.1 | 8.7 | 43.7 | 0.0 | 9.9 | 21.9 | 15.9 |
| 103 | 0.0 | 0.0 | 0.0 | 8.2 | 5.1 | 2.4 | 38.3 | 18.9 | 3.7 | 28.6 | 16.2 |
| 113 | 0.0 | 5.5 | 0.3 | 31.1 | 19.8 | 1.3 | 31.8 | 49.5 | 10.5 | 40.7 | 25.6 |
| | 2.4 | 1.2 | 0.3 | 11.1 | 8.8 | 1.5 | 23.7 | 12.2 | 5.1 | 17.9 | 11.5 |

| GLU | CA | N | C | O | CB | CG | CD | OE1 | OE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 0.0 | 0.1 | 1.3 | 29.4 | 12.2 | 16.0 | 0.6 | 5.2 | 0.6 | 9.6 | 2.9 | 6.9 |
| 9 | 0.0 | 0.0 | 2.0 | 0.4 | 1.5 | 6.4 | 0.4 | 31.7 | 20.0 | 2.8 | 25.8 | 12.0 |
| 49 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.7 | 19.8 | 0.0 | 17.8 | 7.1 |
| 86 | 0.0 | 0.0 | 0.2 | 23.7 | 3.4 | 0.0 | 0.6 | 7.2 | 21.2 | 1.3 | 14.2 | 6.5 |
| 111 | 0.8 | 0.0 | 2.7 | 8.5 | 14.1 | 0.3 | 0.8 | 29.1 | 19.3 | 5.0 | 24.2 | 12.7 |
| | 0.2 | 0.0 | 1.2 | 12.4 | 6.2 | 4.5 | 0.5 | 17.8 | 16.2 | 3.7 | 17.0 | 9.0 |

| GLN | CA | N | C | O | CB | CG | CD | NOE1 | NOE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 3.9 | 1.2 | 0.2 | 2.6 | 1.2 |
| 28 | 0.5 | 0.0 | 0.0 | 0.5 | 3.3 | 11.3 | 2.6 | 6.3 | 44.2 | 5.7 | 25.2 | 13.5 |
| 55 | 0.0 | 0.0 | 1.0 | 2.9 | 6.7 | 1.0 | 1.0 | 17.3 | 24.8 | 2.9 | 21.0 | 10.2 |
| 60 | 2.9 | 0.0 | 0.3 | 3.8 | 0.0 | 2.6 | 1.1 | 14.4 | 11.4 | 1.2 | 12.9 | 5.9 |
| 69 | 4.5 | 0.0 | 0.1 | 0.0 | 6.0 | 0.0 | 0.1 | 26.2 | 27.1 | 2.0 | 26.6 | 11.9 |
| 74 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 22.4 | 0.0 | 0.0 | 11.2 | 4.5 |
| 101 | 0.0 | 0.8 | 0.2 | 11.4 | 3.5 | 20.7 | 8.2 | 17.5 | 33.9 | 10.8 | 25.7 | 16.8 |
| | 1.1 | 0.1 | 0.2 | 2.7 | 2.8 | 5.1 | 1.9 | 15.4 | 20.4 | 3.3 | 17.9 | 9.1 |

| HIS | CA | N | C | O | CB | CG | ND1 | CE1 | NE2 | CD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 3.4 | 0.0 | 0.0 | 0.8 | 0.0 | 0.6 |
| 48 | 0.0 | 0.0 | 0.0 | 3.4 | 0.3 | 0.7 | 0.2 | 5.5 | 3.9 | 0.0 | 1.6 | 2.1 | 1.8 |
| 105 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.2 | 27.4 | 6.4 | 0.0 | 3.6 | 13.7 | 7.0 |
| 119 | 0.0 | 0.0 | 0.0 | 0.1 | 12.9 | 1.8 | 3.8 | 3.9 | 19.2 | 12.4 | 9.4 | 8.2 | 9.0 |
| | 0.0 | 0.0 | 0.0 | 0.9 | 3.3 | 0.6 | 3.1 | 10.1 | 7.4 | 3.1 | 3.9 | 6.0 | 4.6 |

| LYS | CA | N | C | O | CB | CG | CD | CE | NZ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 14.4 | 29.7 | 2.5 | 30.4 | 25.3 | 12.5 | 13.8 | 18.6 | 57.0 | 17.6 | 57.0 | 25.5 |
| 7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 17.7 | 9.3 | 39.2 | 6.8 | 39.2 | 13.3 |
| 31 | 0.0 | 0.0 | 0.1 | 0.0 | 0.2 | 12.4 | 22.3 | 11.7 | 26.9 | 11.7 | 26.9 | 14.7 |
| 37 | 6.2 | 2.1 | 0.8 | 6.4 | 15.2 | 9.4 | 0.0 | 15.5 | 28.1 | 10.0 | 28.1 | 13.6 |
| 41 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.4 | 0.0 | 23.4 | 10.6 | 6.5 | 10.6 | 7.3 |
| 61 | 5.8 | 0.2 | 0.0 | 0.0 | 0.9 | 11.8 | 0.0 | 14.6 | 30.5 | 6.8 | 30.5 | 11.6 |
| 66 | 5.1 | 0.9 | 1.1 | 25.4 | 8.4 | 13.8 | 14.1 | 13.0 | 52.0 | 12.3 | 52.0 | 20.3 |
| 91 | 0.0 | 0.0 | 0.0 | 4.0 | 10.9 | 22.2 | 7.6 | 30.4 | 42.3 | 17.8 | 42.3 | 22.7 |
| 98 | 2.9 | 0.0 | 0.0 | 0.0 | 6.1 | 1.3 | 22.5 | 14.7 | 38.2 | 11.1 | 38.2 | 16.5 |
| 104 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 9.7 | 17.0 | 2.1 | 23.9 | 7.3 | 23.9 | 10.6 |
| | 3.4 | 3.3 | 0.4 | 6.6 | 6.7 | 9.6 | 11.5 | 15.3 | 34.9 | 10.8 | 34.9 | 15.6 |

| ARG | CA | N | C | O | CB | CG | CD | NE | CZ | NT1 | NT2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 4.3 | 0.2 | 0.0 | 0.0 | 12.9 | 5.8 | 7.3 | 0.0 | 0.0 | 9.0 | 18.8 | 6.5 | 9.3 | 7.7 |
| 33 | 0.5 | 0.0 | 0.0 | 6.6 | 0.0 | 0.0 | 8.7 | 14.1 | 1.6 | 7.8 | 17.4 | 2.6 | 13.1 | 7.1 |
| 39 | 3.4 | 0.0 | 0.0 | 0.0 | 9.3 | 0.0 | 5.3 | 16.6 | 1.0 | 44.1 | 16.8 | 3.9 | 25.8 | 13.3 |
| 85 | 0.0 | 0.0 | 0.0 | 0.0 | 2.9 | 0.3 | 15.3 | 5.7 | 1.1 | 42.7 | 40.9 | 4.9 | 29.7 | 15.5 |
| | 2.1 | 0.0 | 0.0 | 1.6 | 6.3 | 1.5 | 9.1 | 9.1 | 0.9 | 25.9 | 23.4 | 4.5 | 19.5 | 10.9 |

## (b) *Lyzozyme*

| GLY | CA | N | C | O |
|---|---|---|---|---|
| 4 | 20.4 | 0.4 | 0.1 | 0.0 |
| 16 | 18.6 | 1.2 | 1.4 | 10.5 |
| 22 | 17.7 | 3.0 | 0.1 | 31.6 |
| 26 | 0.0 | 0.0 | 0.0 | 0.0 |
| 49 | 10.8 | 3.6 | 0.0 | 0.0 |
| 54 | 0.0 | 0.0 | 0.0 | 0.0 |
| 67 | 16.3 | 0.4 | 5.8 | 40.0 |
| 71 | 10.9 | 6.0 | 0.4 | 6.4 |
| 102 | 14.2 | 0.0 | 3.1 | 35.7 |
| 104 | 2.4 | 3.0 | 0.0 | 0.0 |
| 117 | 22.3 | 7.5 | 2.8 | 38.0 |
| 126 | 25.2 | 11.0 | 2.8 | 23.0 |
| | 13.2 | 3.0 | 1.4 | 15.4 |

| ALA | CA | N | C | O | CB | |
|---|---|---|---|---|---|---|
| 9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 10 | 1.1 | 0.0 | 1.0 | 2.9 | 28.3 | 28.3 |
| 11 | 6.2 | 0.6 | 0.0 | 0.0 | 9.8 | 9.8 |
| 31 | -0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.0 |
| 32 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 |
| 42 | 7.7 | 0.2 | 0.0 | 0.0 | 9.8 | 9.8 |
| 82 | 6.7 | 0.0 | 0.0 | 1.2 | 12.0 | 12.0 |
| 90 | 4.5 | 0.7 | 0.0 | 3.8 | 10.0 | 10.0 |
| 95 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 107 | 6.6 | 0.0 | 0.8 | 32.6 | 10.7 | 10.7 |
| 110 | 3.0 | 0.0 | -0.0 | 0.0 | 4.5 | 4.5 |
| 122 | 0.0 | 0.0 | 0.2 | 10.2 | 7.7 | 7.7 |
| | 3.0 | 0.1 | 0.2 | 4.2 | 7.7 | 7.7 |

| VAL | CA | N | C | O | CB | CG1 | CG2 | |
|---|---|---|---|---|---|---|---|---|
| 2 | 0.0 | 3.2 | 0.0 | 18.7 | 6.2 | 21.4 | 24.9 | 17.5 |
| 29 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 92 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 |
| 99 | 0.0 | 0.0 | 0.0 | 2.8 | 0.0 | 0.0 | 0.0 | 0.0 |
| 109 | 0.0 | 3.0 | 0.0 | 0.0 | 8.8 | 29.8 | 27.7 | 22.1 |
| 120 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 8.2 | 0.4 | 2.9 |
| | 0.0 | 1.1 | 0.0 | 3.6 | 2.5 | 9.9 | 8.8 | 7.1 |

| LEU | CA | N | C | O | CB | CG | CD1 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|
| 8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 |
| 17 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 25 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 0.0 | 0.0 | 0.7 | 0.3 |
| 56 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 |
| 75 | 1.0 | 0.0 | 0.0 | 11.3 | 1.1 | 6.1 | 6.5 | 45.0 | 14.7 |
| 83 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 84 | 0.0 | 0.0 | 1.2 | 3.2 | 5.1 | 0.0 | 0.0 | 21.0 | 6.5 |
| 129 | 3.4 | 1.7 | 9.5 | 39.6 | 5.7 | 0.0 | 0.6 | 7.6 | 3.5 |
| | | | | 9.4 | | | | | |
| | 0.5 | 0.2 | 1.3 | 7.1 | 1.5 | 0.8 | 0.9 | 9.3 | 3.1 |

| ILE | CA | N | C | O | CB | CG2 | CG1 | CD1 | |
|---|---|---|---|---|---|---|---|---|---|
| 55 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | 0.0 |
| 58 | 1.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.1 |
| 78 | 0.0 | 0.0 | 0.0 | 0.0 | 3.3 | 0.7 | 0.6 | 34.8 | 9.9 |
| 88 | 0.0 | 1.1 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | -0.0 | 0.6 |
| 98 | 0.0 | 0.0 | 0.0 | 2.7 | 0.0 | 3.7 | 0.0 | 1.4 | 1.3 |
| 124 | 0.0 | 0.0 | 0.2 | 5.8 | 0.0 | 0.0 | 2.9 | 21.1 | 6.0 |
| | 0.3 | 0.2 | 0.0 | 1.4 | 0.5 | 0.8 | 1.0 | 9.6 | 3.0 |

| PRO | CA | N | C | O | CB | CG | CD | |
|---|---|---|---|---|---|---|---|---|
| 70 | 6.4 | 0.7 | 1.7 | 26.2 | 25.8 | 32.1 | 7.9 | 21.9 |
| 79 | 0.0 | 0.0 | 0.1 | 0.0 | 15.0 | 27.7 | 8.5 | 17.1 |
| | 3.2 | 0.3 | 0.9 | 13.1 | 20.4 | 29.9 | 8.2 | 19.5 |

| CYS | CA | N | C | O | CB | SG | |
|---|---|---|---|---|---|---|---|
| 6 | 0.0 | 2.4 | 2.1 | 0.7 | 22.3 | 10.2 | 16.2 |
| 30 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 0.2 |
| 64 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 76 | 0.1 | 0.0 | 0.9 | 15.7 | 4.6 | 0.2 | 2.4 |
| 80 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.0 | 0.2 |
| 94 | 1.3 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 |
| 115 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | -0.0 | 0.0 |
| 127 | 0.5 | 0.0 | 0.1 | 22.8 | 1.3 | 0.0 | 0.7 |
| | 0.2 | 0.3 | 0.4 | 4.9 | 3.6 | 1.3 | 2.5 |

| MET | CA | N | C | O | CB | CG | SD | CE | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 |
| 105 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 |
| | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

| PHE | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.0 | 0.0 | 0.8 | 5.8 | 0.0 | 0.0 | 0.0 | 3.9 | 5.6 | 0.0 | 0.0 | 1.3 |
| 34 | 0.4 | 0.0 | 0.1 | 24.0 | -0.0 | 0.0 | 0.0 | 5.1 | 14.8 | 9.3 | 4.4 | 4.8 |
| 38 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 3.3 | 7.9 | 1.6 |
| | 0.1 | 0.0 | 0.3 | 9.9 | 0.1 | 0.0 | 0.0 | 3.0 | 6.8 | 4.2 | 4.1 | 2.6 |

| TRY | CA | N | C | O | CB | CG | CD1 | NE | CE1 | CZ1 | CH | CZ2 | CE2 | CD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 28 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 62 | 0.0 | 0.0 | 0.0 | 0.0 | 0.9 | 1.0 | 6.0 | 18.0 | 2.0 | 18.6 | 18.5 | 10.3 | 8.7 | 2.1 | 7.6 | 18.0 | 8.6 |
| 63 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 11.2 | 0.0 | 11.4 | 5.3 | 2.3 | 0.0 | 0.0 | 2.2 | 11.2 | 3.1 |
| 108 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 3.9 | 2.9 | 0.7 | -0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.5 | 2.9 | 0.8 |
| 111 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 5.0 | 3.2 | 0.0 | 0.0 | 0.0 | 0.9 | 0.0 | 0.8 |
| 123 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 11.4 | 2.0 | 11.5 | 12.5 | 8.4 | 0.7 | 0.4 | 4.0 | 11.4 | 4.7 |
| | 0.2 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 1.8 | 7.2 | 0.8 | 7.8 | 6.6 | 3.5 | 1.6 | 0.4 | 2.5 | 7.2 | 3.0 |

| SER | CA | N | C | O | CB | OH | | | |
|---|---|---|---|---|---|---|---|---|---|
| 24 | 0.0 | 0.7 | 0.0 | 0.0 | 15.9 | 6.4 | 15.9 | 6.4 | 11.2 |
| 36 | 0.9 | 0.0 | 2.8 | 3.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 50 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.4 | 0.0 | 1.4 | 0.7 |
| 60 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 72 | 0.0 | 0.0 | 0.1 | 26.4 | 6.6 | 0.1 | 6.6 | 0.1 | 3.4 |
| 81 | 0.4 | 2.4 | 0.0 | 8.4 | 26.8 | 29.8 | 26.8 | 29.8 | 28.3 |
| 85 | 7.1 | 0.7 | 0.0 | 0.0 | 24.7 | 10.6 | 24.7 | 10.6 | 17.6 |
| 86 | 0.0 | 2.9 | 1.1 | 12.4 | 15.5 | 31.1 | 15.5 | 31.1 | 23.3 |
| 91 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 100 | 0.0 | 0.0 | 0.1 | 13.6 | 9.1 | 2.7 | 9.1 | 2.7 | 5.9 |
| | 0.8 | 0.7 | 0.4 | 6.5 | 9.9 | 8.2 | 9.9 | 8.2 | 9.0 |

| THR | CA | N | C | O | CB | CG | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 40 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 43 | 0.0 | 1.3 | 0.4 | 13.6 | 2.3 | 9.8 | 36.2 | 6.1 | 36.2 | 16.1 |
| 47 | 3.2 | 0.3 | 1.3 | 29.8 | 6.9 | 45.8 | 38.3 | 26.3 | 38.3 | 30.3 |
| 51 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.7 | 0.4 | 1.4 | 0.4 | 1.0 |
| 69 | 0.0 | 0.7 | 2.0 | 1.0 | 0.0 | 1.5 | 0.0 | 0.8 | 0.0 | 0.5 |
| 89 | 0.0 | 0.0 | 0.0 | 0.0 | 3.4 | 27.0 | 8.5 | 15.2 | 8.5 | 12.9 |
| 118 | 5.0 | 0.0 | 3.4 | 13.6 | 10.1 | 23.1 | 4.2 | 16.6 | 4.2 | 12.5 |
| | 1.2 | 0.3 | 1.0 | 8.3 | 3.2 | 15.7 | 12.5 | 9.5 | 12.5 | 10.5 |

| TYR | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 1.4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.1 | 17.8 | 0.2 | 0.0 | 0.0 | 26.2 | 3.7 | 26.2 | 6.5 |
| 23 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.4 | 10.0 | 1.7 | 3.5 | 0.0 | 21.6 | 2.5 | 21.6 | 4.9 |
| 53 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 0.0 | 0.0 | 0.0 | 6.3 | 1.9 | 6.3 | 2.4 |
| | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 3.5 | 13.6 | 0.7 | 1.2 | 0.0 | 18.0 | 2.7 | 18.0 | 4.6 |

| ASP | CA | N | C | O | CB | CG | OD1 | OD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | 0.0 | 0.0 | 0.0 | 13.0 | 7.1 | 0.4 | 5.9 | 21.8 | 3.7 | 13.9 | 8.8 |
| 52 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.1 | 18.4 | 0.0 | 9.8 | 4.9 |
| 66 | 3.8 | 0.0 | 5.1 | 23.1 | 1.6 | 0.0 | 0.0 | 0.0 | 0.8 | 0.0 | 0.4 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 87 | 1.0 | 0.1 | 0.0 | 1.1 | 13.2 | 2.2 | 8.0 | 46.3 | 7.7 | 27.1 | 17.4 | |
| 101 | 3.5 | 0.0 | 1.9 | 37.0 | 14.5 | 1.5 | 18.8 | 10.0 | 8.0 | 14.4 | 11.2 | |
| 103 | 2.0 | 0.3 | 3.4 | 0.2 | 0.6 | 1.8 | 42.1 | 38.2 | 1.2 | 40.1 | 20.7 | |
| 119 | 5.2 | 2.2 | 0.0 | 1.6 | 22.1 | 1.5 | 5.6 | 30.8 | 11.8 | 18.2 | 15.0 | |
| | 2.2 | 0.4 | 1.5 | 10.9 | 8.4 | 1.1 | 11.6 | 23.7 | 4.8 | 17.6 | 11.2 | |

| ASN | CA | N | C | O | CB | CG | ND1 | ND2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | 0.0 | 0.0 | 0.1 | 16.6 | 12.2 | 1.7 | 24.8 | 28.1 | 6.9 | 26.4 | 16.7 |
| 27 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 7.1 | 11.9 | 0.0 | 9.5 | 4.8 |
| 37 | 0.0 | 0.2 | 0.0 | 5.7 | 9.6 | 2.2 | 8.3 | 49.2 | 5.9 | 28.7 | 17.3 |
| 39 | 0.0 | 0.0 | 0.0 | 0.0 | 2.7 | 0.0 | 0.5 | 32.5 | 1.4 | 16.5 | 8.9 |
| 44 | 1.5 | 0.3 | 0.3 | 0.0 | 3.1 | 1.6 | 40.5 | 17.7 | 2.4 | 29.1 | 15.7 |
| 46 | 4.0 | 0.0 | 3.1 | 3.3 | 12.4 | 2.4 | 1.4 | 26.1 | 7.4 | 13.7 | 10.6 |
| 48 | 4.4 | 0.2 | 5.0 | 13.3 | 23.3 | 1.6 | 8.6 | 12.5 | 12.5 | 10.6 | 11.5 |
| 59 | 0.0 | 1.6 | 0.0 | 0.0 | 10.3 | 1.7 | 3.0 | 8.0 | 6.0 | 5.5 | 5.7 |
| 65 | -0.0 | 0.0 | 0.1 | 1.6 | 0.5 | 1.7 | 14.1 | 37.6 | 1.1 | 25.9 | 13.5 |
| 74 | 0.0 | 7.6 | 1.3 | 7.6 | 13.9 | 0.0 | 0.0 | 10.0 | 7.0 | 5.0 | 6.0 |
| 77 | 0.0 | 0.0 | 0.4 | 17.6 | 11.8 | 1.7 | 33.8 | 42.5 | 6.7 | 38.1 | 22.4 |
| 93 | 0.0 | 0.0 | 1.4 | 10.9 | 6.8 | 1.2 | 17.1 | 32.4 | 4.0 | 24.7 | 14.4 |
| 106 | 0.0 | 0.0 | 0.0 | 1.4 | 2.3 | 1.6 | 2.1 | 20.8 | 1.9 | 11.4 | 6.7 |
| 113 | 1.4 | 0.0 | 0.2 | 37.4 | 8.6 | 2.3 | 24.0 | 30.7 | 5.5 | 27.3 | 16.4 |
| | 0.8 | 0.7 | 0.8 | 8.2 | 8.4 | 1.4 | 13.2 | 25.7 | 4.9 | 19.5 | 12.2 | |

| GLU | CA | N | C | O | CB | CG | CD | OE1 | OE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 5.4 | 3.7 | 0.0 | 3.2 | 7.5 | 10.3 | 1.8 | 21.4 | 18.4 | 6.5 | 19.9 | 11.8 |
| 35 | 0.0 | 0.0 | 0.6 | 12.9 | 0.0 | -0.0 | 0.2 | 14.9 | 1.9 | 0.1 | 8.4 | 3.4 |
| | 2.7 | 1.8 | 0.3 | 8.1 | 3.7 | 5.1 | 1.0 | 18.1 | 10.1 | 3.3 | 14.1 | 7.6 |

| GLN | CA | N | C | O | CB | CG | CD | NOE1 | NOE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 41 | 0.0 | 0.0 | 1.4 | 3.3 | 4.5 | 2.6 | 2.5 | 37.4 | 46.7 | 3.2 | 42.1 | 18.8 |
| 57 | 0.0 | 0.0 | 0.0 | 8.9 | 0.2 | 0.0 | 0.0 | 1.8 | 0.0 | 0.1 | 0.9 | 0.4 |
| 121 | 0.0 | 0.0 | 0.0 | 0.8 | 4.6 | 3.9 | 3.3 | 49.4 | 35.1 | 3.9 | 42.3 | 19.3 |
| | 0.0 | 0.0 | 0.5 | 4.3 | 3.1 | 2.2 | 2.0 | 29.5 | 27.3 | 2.4 | 28.4 | 12.8 |

| HIS | CA | N | C | O | CB | CG | ND1 | CE1 | NE2 | CD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 2.4 | 0.0 | 4.6 | 2.4 | 0.0 | 0.0 | 0.0 | 7.5 | 2.7 | 10.2 | 4.4 | 1.4 | 3.4 |

| LYS | CA | N | C | O | CB | CG | CD | CE | NZ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 9.7 | 4.1 | 0.1 | 0.0 | 3.6 | 8.7 | 2.6 | 17.2 | 27.9 | 8.0 | 27.9 | 12.0 |
| 13 | 0.0 | 0.0 | 1.3 | 11.8 | 1.6 | 2.6 | 5.0 | 16.3 | 26.4 | 6.4 | 26.4 | 10.4 |
| 33 | 0.0 | 0.0 | 0.0 | 1.4 | 0.0 | 0.0 | 0.0 | 3.9 | 40.1 | 1.0 | 40.1 | 8.8 |
| 96 | 0.0 | 0.0 | 0.3 | 1.0 | 3.1 | 0.0 | 6.0 | 0.7 | 21.9 | 2.4 | 21.9 | 6.3 |
| 97 | 1.5 | 0.0 | 0.0 | 0.0 | 1.0 | 8.3 | 7.8 | 26.6 | 47.4 | 10.9 | 47.4 | 18.2 |
| 116 | 0.4 | 0.0 | 0.1 | 21.9 | 0.0 | 11.6 | 0.6 | 26.4 | 25.8 | 9.7 | 25.8 | 12.9 |
| | 1.9 | 0.7 | 0.3 | 6.0 | 1.6 | 5.2 | 3.7 | 15.2 | 31.6 | 6.4 | 31.6 | 11.4 |

| ARG | CA | N | C | O | CB | CG | CD | NE | CZ | NH1 | NH2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 0.0 | 4.7 | 0.0 | 0.0 | 10.3 | 0.0 | 5.3 | 16.1 | 0.0 | 0.0 | 26.4 | 3.9 | 14.2 | 8.3 |
| 14 | 5.7 | 0.7 | 0.0 | 26.6 | 0.0 | 23.9 | 22.5 | 11.8 | 1.7 | 36.5 | 33.4 | 12.0 | 27.2 | 18.5 |
| 21 | 4.2 | 5.8 | 5.1 | 19.3 | 0.8 | 0.2 | 12.9 | 10.5 | 1.8 | 34.1 | 31.3 | 3.9 | 25.3 | 13.1 |
| 45 | 0.0 | 5.9 | 0.1 | 27.1 | 15.8 | 1.5 | 12.4 | 3.5 | 1.9 | 25.9 | 29.7 | 7.9 | 19.7 | 13.0 |
| 61 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 3.5 | 0.0 | 1.7 | 40.8 | 26.8 | 1.3 | 22.6 | 10.4 |
| 68 | 3.5 | 0.0 | 1.6 | 10.4 | 18.4 | 1.5 | 7.9 | 7.0 | 1.8 | 15.7 | 24.6 | 7.4 | 15.8 | 11.0 |
| 73 | 4.8 | 0.0 | 0.0 | -0.0 | 6.5 | 14.7 | 7.9 | 15.1 | 2.9 | 18.2 | 51.0 | 8.0 | 28.1 | 16.6 |
| 112 | 0.0 | 0.0 | 0.0 | 9.2 | 0.0 | 6.3 | 13.2 | 1.2 | 1.8 | 36.7 | 16.0 | 5.3 | 18.0 | 10.7 |
| 114 | 2.6 | 0.0 | 1.2 | 13.0 | 3.6 | 6.1 | 7.7 | 10.9 | 1.9 | 32.7 | 31.4 | 4.8 | 25.0 | 13.5 |
| 125 | 5.2 | 0.0 | 1.4 | 2.2 | 16.9 | 10.8 | 11.3 | 7.7 | 1.9 | 24.0 | 9.9 | 10.3 | 13.9 | 11.8 |
| 128 | 4.7 | 2.6 | 2.1 | 18.0 | 19.5 | 13.1 | 16.9 | 15.0 | 3.7 | 46.1 | 50.9 | 13.3 | 37.3 | 23.6 |
| | 2.8 | 1.8 | 1.1 | 11.4 | 8.4 | 7.1 | 11.1 | 9.0 | 1.9 | 28.2 | 30.2 | 7.1 | 22.5 | 13.7 |

| H2O | OW1 | OW2 | OW3 |
|---|---|---|---|
| 130 | 0.0 | 0.0 | 0.0 |

## (c) *Myoglobin*

| GLY | CA | N | C | O |
|---|---|---|---|---|
| 5 | 26.0 | 5.2 | 3.7 | 4.9 |
| 23 | 7.3 | 0.3 | 0.0 | 0.6 |
| 25 | -0.0 | 0.0 | 0.0 | 0.0 |
| 65 | -0.0 | 0.0 | 0.0 | 0.0 |
| 73 | 0.3 | 0.0 | 0.1 | 0.0 |
| 80 | 2.8 | 0.0 | 0.7 | 12.8 |
| 121 | 25.4 | 4.0 | 4.1 | 27.6 |
| 124 | 23.9 | 2.8 | 0.8 | 0.0 |
| 129 | 20.1 | 0.6 | 0.5 | 0.1 |
| 150 | 19.8 | 0.0 | 2.6 | 33.5 |
| 153 | 26.3 | 15.0 | 10.4 | 33.7 |
| | | | | 46.2 |
| | 13.8 | 2.5 | 2.1 | 13.3 |

| ALA | CA | N | C | O | CB | |
|---|---|---|---|---|---|---|
| 15 | 2.9 | 0.8 | 0.4 | 15.0 | 32.2 | 32.2 |
| 19 | 5.3 | 0.0 | 0.5 | 33.8 | 35.8 | 35.8 |
| 22 | 0.0 | 0.3 | 0.0 | 0.0 | 17.2 | 17.2 |
| 53 | 2.8 | 1.5 | 1.1 | 4.0 | 43.1 | 43.1 |
| 57 | 7.0 | 0.1 | 1.0 | 21.2 | 33.7 | 33.7 |
| 71 | 4.6 | 0.9 | 0.0 | 0.0 | 12.3 | 12.3 |
| 74 | 4.9 | 0.6 | 0.0 | 1.7 | 19.0 | 19.0 |
| 84 | 6.5 | 0.6 | 0.0 | 14.1 | 31.7 | 31.7 |
| 90 | -0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.7 |
| 94 | 0.9 | 0.0 | 0.0 | 2.2 | 0.0 | 0.0 |
| 110 | 0.7 | 2.5 | 0.0 | 0.0 | 1.4 | 1.4 |
| 125 | 0.3 | 2.7 | 4.7 | 8.9 | 48.0 | 48.0 |
| 127 | 0.0 | 0.0 | 0.0 | -0.0 | 4.2 | 4.2 |
| 130 | 0.1 | 2.1 | 0.0 | 0.0 | 0.3 | 0.3 |
| 134 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 143 | 0.0 | 0.0 | 0.2 | 0.0 | 9.6 | 9.6 |
| 144 | 3.7 | 0.7 | 1.7 | 4.7 | 36.2 | 36.2 |
| | 2.3 | 0.8 | 0.6 | 6.2 | 19.1 | 19.1 |

| VAL | CA | N | C | O | CB | CG1 | CG2 | |
|---|---|---|---|---|---|---|---|---|
| 1 | 8.9 | 25.4 | 1.7 | 29.2 | 7.9 | 48.8 | 16.5 | 24.4 |
| 10 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 13 | 0.1 | 0.1 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 0.2 |
| 17 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.7 | 0.2 |
| 21 | 0.0 | 2.6 | 0.0 | 0.0 | 6.3 | 13.7 | 4.2 | 8.1 |
| 66 | 0.0 | 0.0 | 0.0 | 0.0 | 1.4 | 13.4 | 26.1 | 13.6 |
| 68 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 |
| 114 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 7.8 | 0.7 | 2.9 |
| | 1.1 | 3.5 | 0.2 | 3.7 | 2.0 | 10.6 | 6.0 | 6.2 |

| LEU | CA | N | C | O | CB | CG | CD1 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 2.7 | 0.0 | 0.3 | 8.5 | 0.0 | 0.6 | 0.0 | 0.5 | 0.3 |
| 9 | 0.3 | 0.0 | 0.0 | 1.1 | 0.0 | 0.0 | 9.8 | 29.2 | 9.8 |
| 11 | 0.0 | 0.0 | 1.7 | 6.2 | 8.1 | 0.0 | 4.2 | 18.2 | 7.6 |
| 29 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 |
| 32 | 0.3 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.3 | 0.1 |
| 46 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 13.5 | 0.0 | 3.4 |
| 49 | 0.0 | 0.0 | 0.0 | 2.3 | 0.0 | 0.0 | 0.0 | 2.4 | 0.6 |
| 61 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 |
| 69 | 0.0 | 0.0 | 1.5 | 0.0 | 0.5 | 0.0 | -0.0 | 0.1 | 0.2 |
| 72 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 76 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 86 | 0.0 | 0.1 | 0.2 | 0.0 | 5.8 | 0.1 | 0.2 | 0.0 | 1.5 |
| 99 | 0.4 | 0.8 | -0.0 | 0.0 | 0.7 | 0.0 | -0.0 | 4.0 | 1.2 |
| 104 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 |
| 115 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 135 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| 137 | 0.8 | 0.0 | 0.0 | 0.0 | 3.1 | 0.0 | 0.0 | 12.6 | 4.0 |
| 149 | 1.9 | 0.0 | 3.2 | 22.2 | 0.0 | 0.0 | 22.5 | 0.0 | 5.6 |
| | 0.4 | 0.0 | 0.4 | 2.2 | 1.0 | 0.0 | 2.8 | 3.7 | 1.9 |

| ILE | CA | N | C | O | CB | CG2 | CG1 | CD1 | |
|---|---|---|---|---|---|---|---|---|---|
| 28 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | -0.0 | 0.0 |
| 30 | 0.0 | 0.0 | 0.4 | 0.0 | 0.1 | 5.0 | 0.0 | 1.2 | 1.6 |
| 75 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | -0.0 | 0.0 | 0.8 | 0.2 |
| 99 | 0.0 | 0.1 | 1.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 101 | 0.0 | 0.8 | 1.3 | 0.2 | 7.3 | 12.7 | 10.9 | -0.0 | 7.7 |
| 107 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.0 |
| 111 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.0 |
| 112 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.9 | 0.0 | 1.5 | 2.6 |
| 142 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | -0.0 | 0.0 |
| | 0.0 | 0.1 | 0.4 | 0.0 | 0.8 | 3.0 | 1.2 | 0.4 | 1.3 |

| PRO | CA | N | C | O | CB | CG | CD | |
|---|---|---|---|---|---|---|---|---|
| 37 | 0.0 | 0.0 | 3.4 | 4.7 | 24.0 | 27.2 | 4.4 | 18.5 |
| 88 | 0.0 | 0.0 | 1.5 | 5.3 | 24.5 | 20.8 | 1.8 | 15.7 |
| 100 | 0.3 | 0.1 | 0.0 | 0.0 | 19.2 | 18.0 | 0.3 | 12.5 |
| 120 | 0.5 | 0.0 | 3.0 | 14.2 | 25.6 | 28.5 | 7.3 | 20.4 |
| | 0.9 | 0.0 | 2.0 | 6.1 | 23.3 | 23.6 | 3.4 | 16.8 |

| MET | CA | N | C | O | CB | CG | SD | CE | |
|---|---|---|---|---|---|---|---|---|---|
| 55 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 5.0 | 1.2 |
| 131 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 |
| | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 0.6 |

| PHE | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 33 | 0.0 | 0.0 | 0.1 | 5.5 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 | 0.0 |
| 43 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | 0.0 |
| 46 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 |
| 106 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 0.0 | 0.0 | 0.4 | 11.1 | 7.1 | 2.3 | 3.0 |
| 123 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 138 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.0 | 0.0 | 0.0 | 1.1 | 0.0 | 0.0 | 0.1 | 0.1 | 1.9 | 1.2 | 0.4 | 0.5 |

| TRY | CA | N | C | O | CB | CG | CD1 | NE | CE1 | CZ1 | OH | CZ2 | CE2 | CD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 0.0 | 0.0 | 0.6 | 0.0 | 0.1 | 0.0 | 1.9 | 3.3 | 0.0 | 0.1 | -0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 3.3 | 0.5 |
| 14 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.9 | 0.0 | 0.0 | 1.7 | 0.0 | -0.0 | -0.0 | 0.0 | 0.3 | 0.0 | 0.3 |
| | 0.0 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 1.4 | 1.7 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3 | 1.7 | 0.4 |

| SER | CA | N | C | O | CB | OH | | | |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 5.5 | 4.1 | 0.0 | 0.0 | 21.7 | 9.0 | 21.7 | 9.0 | 15.3 |
| 35 | 3.8 | 0.0 | 0.0 | 19.3 | 5.2 | 12.5 | 5.2 | 12.5 | 8.8 |
| 58 | 8.0 | 0.3 | 0.0 | 0.0 | 3.2 | 7.7 | 3.2 | 7.7 | 5.4 |
| 92 | 3.3 | 0.5 | 0.0 | 0.0 | 10.4 | 0.0 | 10.4 | 0.0 | 5.2 |
| 108 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 |
| 117 | 2.4 | 0.0 | 0.4 | 27.4 | 21.6 | 14.9 | 21.6 | 14.9 | 18.3 |
| | 3.8 | 0.8 | 0.1 | 7.8 | 10.4 | 7.3 | 10.4 | 7.3 | 8.8 |

| THR | CA | N | C | O | CB | CG | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 39 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 51 | 1.3 | 0.0 | 0.0 | 0.0 | 5.1 | 2.7 | 36.4 | 3.9 | 36.4 | 14.7 |
| 67 | 1.7 | 0.0 | 0.0 | 4.1 | 0.0 | 31.7 | 14.5 | 15.9 | 14.5 | 15.4 |
| 70 | 2.4 | 0.0 | 0.7 | 5.0 | 3.4 | 8.9 | 46.4 | 6.1 | 46.4 | 19.5 |
| 95 | 0.0 | 0.0 | 2.4 | 27.3 | 5.3 | 16.0 | 16.1 | 10.7 | 16.1 | 12.5 |
| | 1.1 | 0.0 | 0.6 | 7.3 | 2.7 | 11.9 | 22.7 | 7.3 | 22.7 | 12.4 |

| TYR | CA | N | C | O | CB | CG | CD1 | CE1 | CZ | CE2 | CD2 | OH | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 103 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 4.9 | 0.0 | 22.0 | 0.8 | 22.0 | 3.4 |
| 146 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 5.4 | 0.0 | 0.0 | 0.0 | 0.2 | 0.8 | 0.2 | 0.8 |
| 151 | 2.7 | 0.0 | 2.3 | 6.5 | 4.0 | 0.0 | 16.2 | 14.3 | 1.8 | 1.0 | 0.0 | 31.1 | 5.3 | 31.1 | 8.5 |
| | 0.9 | 0.0 | 0.8 | 2.2 | 1.3 | 0.0 | 5.6 | 6.8 | 0.6 | 2.0 | 0.0 | 17.7 | 2.3 | 17.7 | 4.3 |

| ASP | CA | N | C | O | CB | CG | OD1 | OD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 1.5 | 0.3 | 0.0 | 0.0 | 14.9 | 2.1 | 15.5 | 18.4 | 8.5 | 16.9 | 12.7 |
| 27 | 0.0 | 0.0 | 0.1 | 2.7 | 0.0 | 1.8 | 32.5 | 11.6 | 0.9 | 22.1 | 11.5 |
| 44 | 0.0 | 1.7 | 1.3 | 13.0 | 16.4 | 3.1 | 14.7 | 43.4 | 9.8 | 29.0 | 19.4 |
| 60 | 1.0 | 1.7 | 0.0 | 0.0 | 14.2 | 1.3 | 13.6 | 14.8 | 7.7 | 14.2 | 11.0 |
| 126 | 3.4 | 0.9 | 0.0 | 5.9 | 22.2 | 1.3 | 40.0 | 21.9 | 11.7 | 30.9 | 21.3 |
| 141 | 0.2 | 0.0 | 0.1 | 7.8 | 5.7 | 2.1 | 18.9 | 12.1 | 3.9 | 15.5 | 9.7 |
| | 1.0 | 0.8 | 0.3 | 4.9 | 12.2 | 1.9 | 22.5 | 20.4 | 7.1 | 21.5 | 14.3 |

| ASN | CA | N | C | O | CB | CG | NOD1 | NOD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 122 | 3.1 | 0.0 | 0.0 | 14.8 | -0.0 | 0.0 | 32.4 | 12.3 | 0.0 | 22.4 | 11.2 |
| 132 | 0.0 | 0.0 | 0.0 | 0.0 | 3.4 | 1.8 | 14.8 | 22.8 | 2.6 | 18.8 | 10.7 |
| | 1.5 | 0.0 | 0.0 | 7.4 | 1.7 | 0.9 | 23.6 | 17.5 | 1.3 | 20.6 | 10.9 |

| GLU | CA | N | C | O | CB | CG | CD | OE1 | OE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 0.0 | 7.4 | 0.6 | 1.9 | 21.3 | 26.1 | 2.5 | 26.2 | 14.3 | 16.7 | 20.3 | 18.1 |
| 6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.1 | 0.0 | 3.4 | 23.7 | 2.7 | 13.5 | 7.0 |
| 18 | 0.0 | 0.0 | 1.2 | 20.2 | 3.0 | 9.6 | 2.1 | 16.0 | 21.8 | 4.9 | 18.9 | 10.5 |
| 38 | 4.3 | 1.5 | 0.0 | 6.0 | 18.5 | 0.1 | 1.0 | 19.7 | 10.4 | 6.5 | 15.1 | 9.9 |
| 41 | 0.1 | 0.0 | 0.4 | 8.1 | 4.6 | 20.0 | 2.0 | 22.1 | 48.9 | 8.9 | 35.5 | 19.5 |
| 52 | 0.0 | 1.7 | 0.0 | 0.0 | 10.7 | 2.1 | 2.3 | 31.8 | 17.5 | 5.0 | 24.6 | 12.9 |
| 54 | 1.6 | 0.2 | 0.0 | 3.8 | 5.0 | 10.8 | 1.3 | 43.2 | 4.2 | 5.7 | 23.7 | 12.9 |
| 59 | 0.0 | 4.6 | 0.2 | 0.0 | 16.0 | 21.8 | 1.9 | 32.2 | 9.5 | 13.3 | 20.8 | 16.3 |
| 83 | 2.5 | 0.0 | 3.2 | 2.5 | 19.1 | 20.1 | 2.0 | 8.6 | 37.8 | 13.8 | 23.2 | 17.5 |
| 85 | 1.1 | 0.0 | 0.3 | 8.2 | 6.9 | 0.0 | 1.0 | 13.1 | 29.5 | 2.6 | 21.3 | 10.1 |
| 105 | 0.0 | 0.0 | 0.0 | 0.0 | 2.9 | 0.4 | 1.8 | 26.8 | 14.5 | 1.7 | 20.6 | 9.3 |
| 109 | 0.0 | 0.0 | 1.6 | 2.0 | 9.5 | 15.7 | 1.1 | 41.1 | 0.9 | 9.1 | 21.0 | 13.8 |
| 136 | 0.5 | 0.0 | 0.6 | 3.5 | 7.1 | 0.0 | 0.8 | 33.0 | 29.8 | 2.6 | 31.4 | 14.1 |
| 148 | 9.8 | 1.3 | 0.1 | 31.0 | 7.2 | 13.3 | 1.2 | 27.7 | 19.6 | 7.2 | 23.6 | 13.8 |
| | 1.4 | 1.2 | 0.6 | 6.2 | 9.4 | 10.6 | 1.5 | 24.6 | 20.2 | 7.2 | 22.4 | 13.3 |

| GLN | CA | N | C | O | CB | CG | CD | NOE1 | NOE2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 0.0 | 0.5 | 1.2 | 0.2 | 11.6 | 7.8 | 3.8 | 45.9 | 44.3 | 7.7 | 45.1 | 22.7 |
| 26 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.4 | 0.7 | 15.9 | 0.8 | 8.3 | 3.8 |
| 91 | 0.0 | 0.0 | 0.2 | 0.0 | 6.9 | 2.4 | 0.3 | 17.6 | 25.2 | 3.2 | 21.4 | 10.5 |
| 128 | 0.0 | 0.0 | 0.2 | 0.0 | 0.9 | 9.7 | 1.8 | 31.2 | 12.7 | 4.2 | 21.9 | 11.3 |
| 152 | 1.8 | 7.2 | 1.9 | 17.0 | 2.2 | 1.0 | 1.8 | 10.1 | 19.0 | 1.7 | 14.6 | 6.8 |
| | 0.4 | 1.5 | 0.7 | 3.5 | 4.3 | 4.4 | 1.8 | 21.1 | 23.4 | 3.5 | 22.3 | 11.0 |

| HIS | CA | N | C | O | CB | CG | ND1 | CE1 | NE2 | CD2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | 4.0 | 0.8 | 0.0 | 0.0 | 3.3 | 2.0 | 19.8 | 23.2 | 9.6 | 10.6 | 9.8 | 14.7 | 11.4 |
| 24 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.9 | 0.0 | -0.0 | 0.5 | 0.0 | 0.3 |
| 36 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 4.0 | 18.1 | 11.8 | 1.2 | 4.8 | 7.9 | 5.9 |
| 48 | 6.0 | 1.8 | 0.0 | 10.3 | 24.5 | 2.9 | 8.9 | 11.8 | 19.1 | 18.0 | 14.3 | 14.0 | 14.2 |
| 64 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.8 | 0.1 | 0.0 | -0.0 | 0.0 | 0.4 | 0.2 |
| 81 | 0.0 | 0.0 | 0.0 | 7.0 | 6.6 | 1.9 | 11.3 | 28.6 | 30.7 | 28.0 | 16.3 | 21.0 | 17.9 |
| 82 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | 1.7 | 0.4 | 0.0 | 0.3 |
| 93 | 0.0 | 0.0 | 0.0 | 0.2 | -0.0 | -0.0 | -0.0 | -0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 |
| 97 | 1.7 | 0.0 | 0.0 | 4.7 | 0.0 | 0.0 | 7.5 | 8.0 | 0.1 | 0.0 | 2.0 | 3.8 | 2.6 |
| 113 | 3.3 | 0.5 | 0.0 | 0.0 | 1.5 | 0.1 | 21.0 | 23.4 | 14.1 | 3.9 | 7.2 | 17.5 | 10.6 |
| 116 | 0.0 | 0.0 | 0.0 | 0.9 | 3.3 | 1.8 | 1.2 | 21.6 | 28.3 | 17.0 | 10.9 | 14.8 | 12.2 |
| 119 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 11.7 | 5.0 | 1.4 | 5.9 | 2.9 |
| | 1.3 | 0.3 | 0.0 | 1.9 | 3.3 | 0.7 | 6.2 | 11.4 | 10.5 | 7.1 | 5.6 | 8.3 | 6.5 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LYS | CA | N | C | O | CB | CG | CD | CE | NZ | | | | |
| 16 | 1.7 | 0.0 | 0.0 | 2.0 | 0.0 | 15.0 | 0.6 | 12.2 | 9.8 | 6.9 | 9.8 | 7.5 | |
| 34 | 0.4 | 0.0 | 2.5 | 22.9 | 8.6 | 0.0 | 3.8 | 10.8 | 33.8 | 5.8 | 33.8 | 11.4 | |
| 42 | 3.3 | 0.4 | 0.0 | 14.8 | 0.0 | 0.3 | 2.2 | 24.6 | 9.7 | 6.8 | 9.7 | 7.4 | |
| 47 | 0.1 | 0.0 | 0.5 | 0.0 | 6.5 | 1.5 | 10.1 | 4.3 | 40.2 | 5.6 | 40.2 | 12.5 | |
| 50 | 0.1 | 0.0 | 0.0 | 27.7 | 12.3 | 18.6 | 21.7 | 1.1 | 28.1 | 13.4 | 28.1 | 16.4 | |
| 56 | 0.0 | 0.0 | 0.6 | 11.5 | 9.5 | 3.4 | 18.5 | 5.6 | 51.4 | 9.2 | 51.4 | 17.7 | |
| 62 | 0.0 | 0.0 | 0.0 | 0.0 | 2.2 | 0.0 | 7.6 | 0.1 | 42.6 | 2.5 | 42.6 | 10.5 | |
| 63 | 2.5 | 0.5 | 0.1 | 0.9 | 3.3 | 10.1 | 13.0 | 20.3 | 55.2 | 11.7 | 55.2 | 20.4 | |
| 77 | 0.0 | 0.0 | 0.0 | 19.7 | 12.1 | 1.6 | 17.4 | 4.9 | 29.5 | 9.0 | 29.5 | 13.1 | |
| 78 | 1.6 | 0.0 | 1.3 | 7.6 | 0.0 | 13.3 | 2.4 | 20.8 | 14.3 | 9.1 | 14.3 | 10.2 | |
| 79 | 0.0 | 0.0 | 0.0 | 11.1 | 5.1 | 0.0 | 13.2 | 2.4 | 29.8 | 5.2 | 29.8 | 10.1 | |
| 87 | 1.2 | 0.5 | 0.0 | 0.0 | 9.9 | 4.2 | 17.2 | 6.9 | 34.0 | 9.5 | 34.0 | 14.4 | |
| 96 | 3.5 | 0.0 | 0.8 | 23.7 | 3.5 | 10.6 | 15.1 | 25.7 | 55.3 | 13.7 | 55.3 | 22.0 | |
| 98 | 1.1 | 0.0 | 3.3 | 4.3 | 4.1 | 6.8 | 18.3 | 25.3 | 55.5 | 13.6 | 55.5 | 22.0 | |
| 102 | 0.5 | 1.9 | 0.0 | 0.0 | 10.4 | 14.5 | 8.4 | 11.6 | 35.8 | 11.2 | 35.8 | 16.1 | |
| 133 | 0.1 | 0.0 | 0.0 | 2.0 | 1.4 | 9.5 | 4.7 | 22.4 | 30.4 | 9.5 | 30.4 | 13.7 | |
| 140 | 2.9 | 0.6 | 0.1 | 3.7 | 2.4 | 8.7 | 26.2 | 26.5 | 43.3 | 15.9 | 43.3 | 21.4 | |
| 145 | 0.7 | 0.7 | 0.0 | 0.0 | 3.0 | 2.7 | 3.9 | 0.0 | 2.0 | 2.4 | 2.0 | 2.3 | |
| 147 | 0.3 | 0.0 | 0.9 | 9.4 | 0.3 | 16.8 | 11.7 | 25.7 | 56.4 | 13.6 | 56.4 | 22.2 | |
| | 1.0 | 0.2 | 0.5 | 8.5 | 5.0 | 7.2 | 11.4 | 13.2 | 34.6 | 9.2 | 34.6 | 14.3 | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ARG | CA | N | C | O | CB | CG | CD | NE | CZ | NH1 | NH2 | | | |
| 31 | 1.5 | 1.4 | 0.1 | 0.0 | 5.1 | 1.2 | 2.7 | 4.3 | 1.6 | 33.6 | 31.4 | 2.6 | 23.1 | 11.4 |
| 45 | 4.0 | 1.3 | 0.0 | 3.5 | 15.4 | 0.0 | 11.6 | 3.0 | 1.8 | 12.4 | 10.3 | 7.2 | 8.5 | 7.8 |
| 118 | 1.5 | 0.0 | 0.5 | 18.7 | 0.0 | 16.5 | 0.1 | 2.7 | 1.2 | 19.5 | 12.6 | 4.4 | 11.6 | 7.5 |
| 139 | 0.0 | 0.0 | 1.2 | 0.0 | 5.0 | 0.0 | 3.2 | 0.0 | 0.0 | 16.5 | 1.7 | 2.1 | 6.1 | 3.8 |
| | 1.7 | 0.7 | 0.4 | 5.6 | 6.4 | 4.4 | 4.4 | 2.5 | 1.2 | 20.5 | 14.0 | 4.1 | 12.3 | 7.6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| HEM | FE | OW | CH1 | CH2 | CH3 | CH4 | |
| 154 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PPR | N | C1 | C2 | C3 | C4 | CM | CA | CB | CG | O1 | O2 | |
| 155 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 13.5 | 5.5 | 29.2 | 1.9 | 10.6 | 10.9 | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VPR | N | C1 | C2 | C3 | C4 | CM | CA | CB |
| 156 | 0.0 | -0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VPL | N | C1 | C2 | C3 | C4 | CM | CA | CB |
| 157 | 0.0 | 0.0 | -0.0 | -0.0 | -0.0 | 0.0 | -0.0 | 0.0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PPL | N | C1 | C2 | C3 | C4 | CM | CA | CB | CG | O1 | O2 | |
| 158 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 4.5 | 2.9 | 0.0 | 0.0 | 23.3 | 18.0 | |

Residues of the same kind are grouped together. The last column in each row gives the average over the side-chain atoms. For polar residues, separate averages over the side-chain non-polar and polar atoms are given in the third and the second from the last columns, respectively. Column averages are given in the last row for each class of residues. For lysozyme, the three "buried" water molecules are given the residue name $H_2O$ and included in the computation. The heme group of myoglobin is included in the computation. The atoms of the heme are divided into 5 groups and given the names HEM, PPR, VPR, VPL, and PPL. The groups PPR, VPR, VPL and PPL are those labeled 1, 2, 3 and 4, respectively, in Watson (1969).

one elongated cavity. The cavity J, on the other hand, is a split double cavity joined by a narrow "pass". It should be noted that the atoms listed for a given cavity may not be a complete list of all the atoms that form the boundary of the cavity. If an atom forms a portion of the boundary of a cavity that lies entirely in between two slicing sections, that atom will not be found by the method described in this paper. It should also be noted that the cavities that are confined entirely to a region between two slicing planes are not detected by the present procedure.

The total accessible surface areas are calculated to be 7010 Å$^2$, 6710 Å$^2$, and 8020 Å$^2$ respectively for ribonuclease-S, lysozyme, and myoglobin. If each atom in a macromolecule occupies the same volume, the total area, $A$, is expected to be related to the number of atoms, $N$, by $A = KN^{2/3}$, where $K$ depends on the shape of the molecule. When the numbers of non-hydrogen atoms are substituted for $N$, $K$ becomes 72 Å$^2$, 67 Å$^2$, and 69 Å$^2$ respectively for the above three proteins. The agreement between these numbers is a reflection of the compact over-all shapes that they have.

The contributions made to the total accessible surface area by four different types

TABLE 6

*Cavities in ribonuclease-S, lysozyme and myoglobin*

*Ribonuclease-S*

*Cavity A*

| | | |
|---|---|---|
| *VAL 54 | CG2 | 0·22 |
| VAL 57 | CG1 | 0·13 |
| ILE 106 | CG1 | 0·04 |
| ILE 106 | CD | 0·18 |
| VAL 108 | CG1 | 0·14 |

vol. = 0·029 Å$^3$

*Cavity B*

| | | |
|---|---|---|
| PHE 8 | CZ | 0·13 |
| VAL 47 | CG1 | 0·84 |
| VAL 47 | CG2 | 0·19 |
| VAL 54 | CG1 | 0·03 |
| *VAL 54 | CG2 | 1·05 |
| ILE 106 | CG1 | 0·81 |
| VAL 108 | CG1 | 0·60 |
| PHE 120 | CE2 | 0·11 |

vol. = 0·375 Å$^3$

*Cavity C*

| | | |
|---|---|---|
| VAL 54 | O | 0·03 |
| CYS 58 | SG | 0·01 |
| VAL 108 | CG2 | 0·01 |

vol. = 0·0003 Å$^3$

*Lysozyme*

*Cavity A*

| | | |
|---|---|---|
| ALA 31 | CA | 0·14 |
| PHE 34 | CB | 0·19 |
| ALA 110 | C | 0·13 |
| CYS 115 | SG | 0·08 |

vol. = 0·017 Å$^3$

*Cavity B*

| | | |
|---|---|---|
| LEU 8 | CD2 | 0·70 |
| MET 12 | CG | 0·16 |
| MET 12 | SD | 1·49 |
| MET 12 | CE | 0·55 |
| ALA 32 | CB | 0·10 |
| ILE 55 | CG1 | 0·39 |
| ILE 55 | CG2 | 2·44 |
| LEU 56 | CD1 | 0·76 |
| LEU 56 | CD2 | 0·27 |
| ILE 88 | CD1 | 0·96 |
| VAL 92 | CG1 | 0·62 |

vol. = 1·327 Å$^3$

*Cavity C*

| | | |
|---|---|---|
| ALA 31 | C | 0·01 |
| ALA 31 | CB | 0·02 |
| GLU 35 | CG | 0·004 |
| LEU 56 | CD2 | 0·02 |
| TRP 108 | CD1 | 0·07 |

vol. = 00·01 Å$^3$

*Cavity D*

| | | |
|---|---|---|
| TRP 28 | CE2 | 0·01 |
| TRP 28 | CZ2 | 0·02 |
| ALA 31 | CB | 0·03 |
| LEU 56 | CG | 0·01 |
| LEU 56 | CD2 | 0·18 |
| MET 105 | CE | 0·21 |
| TRP 108 | CE1 | 0·14 |
| TRP 108 | CZ1 | 0·05 |

vol. = 0·017 Å$^3$

*Cavity E*

| | | |
|---|---|---|
| ASN 65 | CA | 0·03 |
| *ASN 74 | CB | 0·02 |
| ARG 73 | O | 0·03 |
| *SER 72 | OH | 0·02 |

vol. = 0·002 Å$^3$

*Myoglobin*

*Cavity A*

| | | |
|---|---|---|
| LEU 89 | CD1 | 0·93 |
| HIS 93 | CB | 1·31 |
| HIS 93 | CG | 0·95 |
| HIS 93 | ND1 | 0·27 |
| HIS 93 | CE1 | 0·05 |
| HIS 93 | CD2 | 0·06 |
| LEU 104 | CDI | 1·35 |
| LEU 104 | CD2 | 1·07 |
| ILE 142 | CG2 | 0·11 |
| ILE 142 | CD1 | 0·96 |
| *TYR 146 | OH | 0·03 |
| VPR 156 | Cl | 0·46 |
| VPR 156 | C2 | 0·14 |

vol. = 0·853 Å$^3$

*Cavity B*

| | | |
|---|---|---|
| ILE 101 | CD1 | 0·01 |
| LEU 104 | CG | 0·02 |
| *TYR 146 | OH | 0·01 |

vol. = 0·0002 Å$^3$

*Cavity C*

| | | |
|---|---|---|
| LEU 89 | C | |
| LEU 89 | CD1 | 0·02 |
| ALA 90 | CA | 0·04 |
| HIS 93 | CB | 0·14 |
| ILE 142 | CG2 | 0·03 |
| ILE 142 | CD1 | 0·12 |

vol. = 0·007 Å$^3$

*Cavity D*

| | | |
|---|---|---|
| LEU 29 | CD1 | 2·50 |
| LEU 32 | CD1 | 2·01 |
| PHE 33 | CE1 | 0·95 |
| PHE 33 | CZ | 0·01 |
| THR 39 | CB | 0·01 |
| PHE 43 | CE2 | 2·27 |
| PHE 43 | CZ | 0·07 |
| ILE 107 | CD1 | 0·51 |
| HEM 154 | OW | 0·49 |
| VPL 157 | C2 | 0·27 |
| VPL 157 | C3 | 0·96 |
| VPL 157 | C4 | 0·08 |
| VPL 157 | CA | 0·18 |

vol. = 1·781 Å$^3$

TABLE 6—*continued*

Myoglobin—*continued*

*Cavity E*

| | | |
|---|---|---|
| LEU 29 | CD1 | 0·19 |
| LEU 29 | CD2 | 1·49 |
| PHE 33 | CE2 | 0·54 |
| PHE 33 | CZ | 0·73 |
| PHE 43 | CZ | 0·002 |
| PHE 46 | CE2 | 1·40 |
| LEU 61 | CG | 0·38 |
| HIS 64 | CD2 | 0·53 |
| vol. = 0·470 Å$^3$ | | |

*Cavity F*

| | | |
|---|---|---|
| LEU 29 | O | 0·01 |
| LEU 29 | CD1 | 0·01 |
| LEU 32 | CB | 0·01 |
| vol. = 0·0002 Å$^3$ | | |

*Cavity G1*

| | | |
|---|---|---|
| GLY 25 | CA | 0·42 |
| ILE 28 | CB | 0·32 |
| ILE 28 | CG2 | 0·45 |
| LEU 29 | CG | 0·39 |
| LEU 29 | CD1 | 0·23 |
| GLY 65 | CA | 0·04 |
| VAL 68 | CG2 | 0·79 |
| LEU 69 | CD1 | 0·51 |
| vol. = 0·371 Å$^3$ | | |

*Cavity G2*

| | | |
|---|---|---|
| ILE 28 | CG2 | 0·44 |
| VAL 68 | CG2 | 0·61 |
| LEU 69 | CD1 | 0·30 |
| ILE 107 | CD1 | 0·18 |
| ILE 111 | CD1 | 0·60 |
| VPR 156 | CB | 0·01 |
| vol. = 0·178 Å$^3$ | | |

*Cavity H*

| | | |
|---|---|---|
| LEU 72 | CDI | 0·19 |
| LEU 72 | CD2 | 0·40 |
| ILE 107 | CG2 | 0·36 |
| SER 108 | CA | 0·08 |
| SER 108 | OH | 2·19 |
| ILE 111 | CD1 | 1·50 |
| LEU 135 | CD1 | 1·69 |
| PHE 138 | CE1 | 1·09 |
| VPR 156 | CB | 1·62 |
| vol. = 1·754 Å$^3$ | | |

*Cavity I*

| | | |
|---|---|---|
| TRP 7 | CH | 1·75 |
| ILE 75 | C | 0·22 |
| ILE 75 | O | 0·07 |
| ILE 75 | CG2 | 1·20 |
| LEU 76 | CA | 0·05 |
| LEU 76 | CD1 | 2·08 |
| HIS 82 | ND1 | 1·38 |
| HIS 82 | CE1 | 0·08 |
| ALA 134 | O | 1·99 |
| ALA 134 | CB | 0·03 |
| *LEU 137 | CD1 | 0·58 |
| PHE 138 | CB | 0·83 |
| vol. = 2·229 Å$^3$ | | |

*Cavity J*

| | | |
|---|---|---|
| TRP 14 | CZ2 | 5·48 |
| TRP 14 | CE2 | 0·42 |
| VAL 17 | CG1 | 0·44 |
| *VAL 17 | CG2 | 1·07 |
| HIS 24 | CD2 | 0·12 |
| ILE 28 | CD1 | 2·87 |
| LEU 69 | CD1 | 0·08 |
| *LEU 69 | CD2 | 0·24 |
| LEU 72 | CD1 | 1·02 |
| LEU 76 | CD2 | 0·28 |
| ILE 111 | CG2 | 2·85 |
| ILE 111 | CD1 | 0·33 |
| LEU 115 | CD1 | 2·15 |
| LEU 115 | CD2 | 0·73 |
| MET 131 | SD | 1·83 |
| MET 131 | CE | 0·12 |
| LEU 135 | CE | 0·28 |
| LEU 135 | CD2 | 0·01 |
| vol. = 2·854 Å$^3$ | | |

*Cavity K*

| | | |
|---|---|---|
| *LEU 9 | O | 0·04 |
| VAL 10 | CA | 0·11 |
| *VAL 13 | CG1 | 0·12 |
| ALA 127 | O | 0·18 |
| vol. = 0·016 Å$^3$ | | |

*Cavity L*

| | | |
|---|---|---|
| *HIS 12 | CD2 | 0·06 |
| *VAL 13 | CG1 | 0·23 |
| *LYS 16 | CD | 0·12 |
| *LYS 16 | NZ | 0·05 |
| *ASN 122 | O | 0·11 |
| ASN 122 | CB | 0·12 |
| vol. = 0·031 Å$^3$ | | |

For each cavity, the names of atoms that form the boundary of the cavity, their accessibility to a hypothetical solvent molecule inside the cavity, and the volume of the cavity are given. Atoms that are accessible to the outside of the protein as well as to the interior of the cavity are marked with a *. The residue names HEM, VPR and VPL are explained in the legend to Table 5.

of atoms are shown in Figure 2. The three "buried" water molecules of lysozyme and the heme group of myoglobin are included as side-chain atoms. It should be noted that non-polar atoms make up 40 to 50% of the total accessible surface area.

The extent to which different types of atoms are shielded from the bulk solvent by virtue of the fact that they are embedded in the protein matrix was estimated in the following manner. The accessible surface area of individual atoms for each of the 20 amino acid residues was computed for the model $\beta$, *trans* set of Ala-X-Ala systems as described in the previous section. These numbers were multiplied by the number

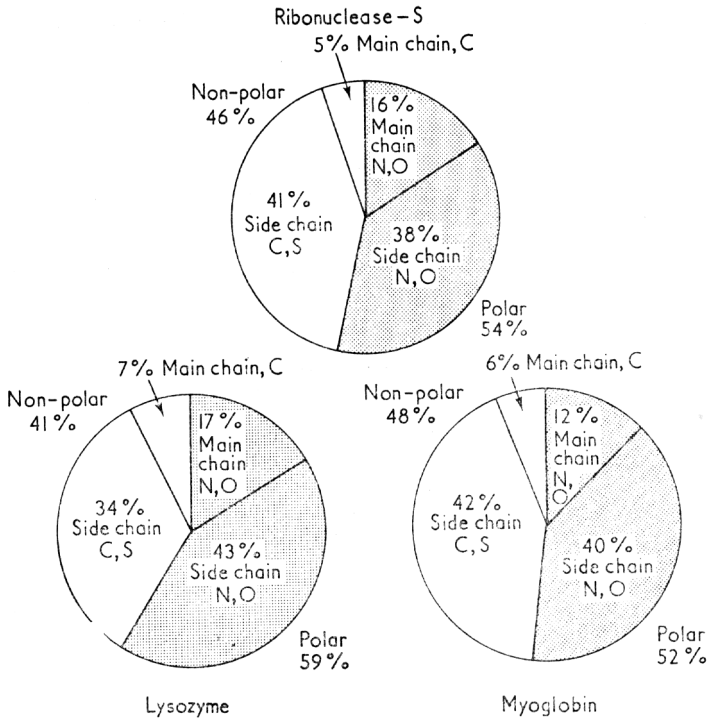


FIG. 2. Contributions made to the total accessible surface area by different types of atoms. Contributions from the polar atoms are shaded. The numbers should be considered only approximate as they are sensitive to the choice of van der Waal's radii for the various types of groups. The appropriate radii are not yet well established.

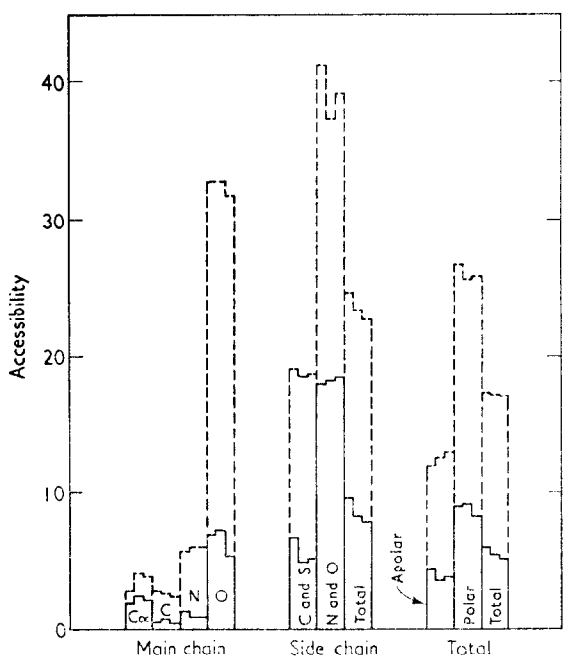


FIG. 3. Average accessibilities for different types of atoms. Solid lines for the actual proteins and dotted lines for the hypothetical extended chains of the same composition as the native proteins. Each major bar in the histogram is divided into three parts. The left-hand third refers to ribonuclease-S, the middle to lysozyme, and the right-hand third to myoglobin.

TABLE 7

*Relative accessibility of various classes of atoms in native proteins based on the fully extended chains as unity*

| | *Main chain* | | | | *Side chain* | | | *Total* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Cα | N | C | O | C,S | N,O | Total | Apolar | Polar | Total |
| RNase-S | 0·71 | 0·23 | 0·21 | 0·21 | 0·35 | 0·44 | 0·39 | 0·36 | 0·34 | 0·35 |
| Lysozyme | 0·60 | 0·14 | 0·26 | 0·22 | 0·26 | 0·49 | 0·35 | 0·28 | 0·36 | 0·32 |
| Myoglobin | 0·55 | 0·14 | 0·21 | 0·17 | 0·27 | 0·47 | 0·34 | 0·29 | 0·32 | 0·30 |
| Average | 0·62 | 0·17 | 0·23 | 0·20 | 0·29 | 0·47 | 0·36 | 0·31 | 0·34 | 0·32 |

of residues occurring in a given protein and summed over all the atoms of a given type. The resulting numbers represented the maximum accessible area in a fully extended polypeptide chain. The ratios of the equivalent sums for the native protein to these numbers give a set of factors by which the accessibility of the various classes of atoms are reduced in going from a fully extended chain to the native folded configuration. Table 7 give the results of this calculation. The contributions made by the three "buried" water molecules of lysozyme and the heme group of myoglobin were excluded in computing the sum of the accessible surface areas for the protein.

The computed accessibilities averaged over different groups of atoms are shown in Figure 3. The three "buried" water molecules of lysozyme and the heme group of myoglobin are included as side chain atoms. The average and the range of accessibilities of side chain atoms of different amino acid species are shown in Figure 4. In both of the Figures, corresponding numbers are plotted for model $\beta$, *trans* set of Ala-X-Ala as estimates of maximum accessibility.

## 4. Discussion

The numerical values of accessibilities of individual atoms must be used with caution, as stated earlier. We suggest these values be used in conjunction with the stack of transparent space-filling drawings so that the possible effects of local flexibility of the molecule can be considered.

In the cavity information given in Table 6, the calculation for ribonuclease-S is based upon the amino-acid sequence and the three-dimensional structure of the bovine enzyme. The sequence of the ribonuclease from rat is known (Beintema & Gruber, 1967) and cogent reasons have been put forward (Wyckoff, 1968) for the plausibility of identical structures for the two molecules. The residues Val 57 and Val 108 in the bovine enzyme are changed to isoleucines in the rat enzyme. Examination of a physical model indicates that, if valine 57 is replaced by an isoleucine, observing correct geometry around the asymmetric carbon atom C$\beta$, the newly added methyl group C$\delta$ will be placed nicely in cavity A. Hardly any alteration of the rest of the structure is required for the insertion. Cavity C will disappear in exactly the same manner upon replacing valine 108 by a leucine. It has been shown (Marchiori, Rocchi, Moroder & Scoffone, 1966) that phenylalanine 8 could be replaced by a tyrosine in ribonuclease-S while retaining 80% of the enzymic activity. The added OH group will partly fill the cavity B. The phenyl group of Phe 120 can be rotated around the C$\alpha$—C$\beta$ bond in such a way that the volume of cavity B is adjustable without required movement of any other side chains.

Many large cavities are found in myoglobin. A number of them are around the heme group. Cavity A is the site where a xenon atom can be inserted (Schoenborn,

---

FIG. 4. Average side-chain accessibilities for different amino-acid residues. The amino acids are arranged in 4 groups—non-polar residues, tryptophan and basic residues, amides and acidic residues, and the hydroxyl residues. In each group residues are placed in ascending order of the accessibility of the corresponding model tripeptides, which are shown as thick horizontal bars. The actual accessibilities averaged over all side-chain atoms and over all the amino-acid residues of that class are given as vertical columns. For polar residues, separate averages over polar and non-polar side chain atoms are plotted as broken lines, the generally upper curve being for polar atoms. Vertical lines are drawn to indicate the minimum and maximum accessiblities for a given class—left line for non-polar, right line for polar, and the middle line for all side-chain atoms. The number of amino-acids residues in each class are also shown across the bottom of the Figure.

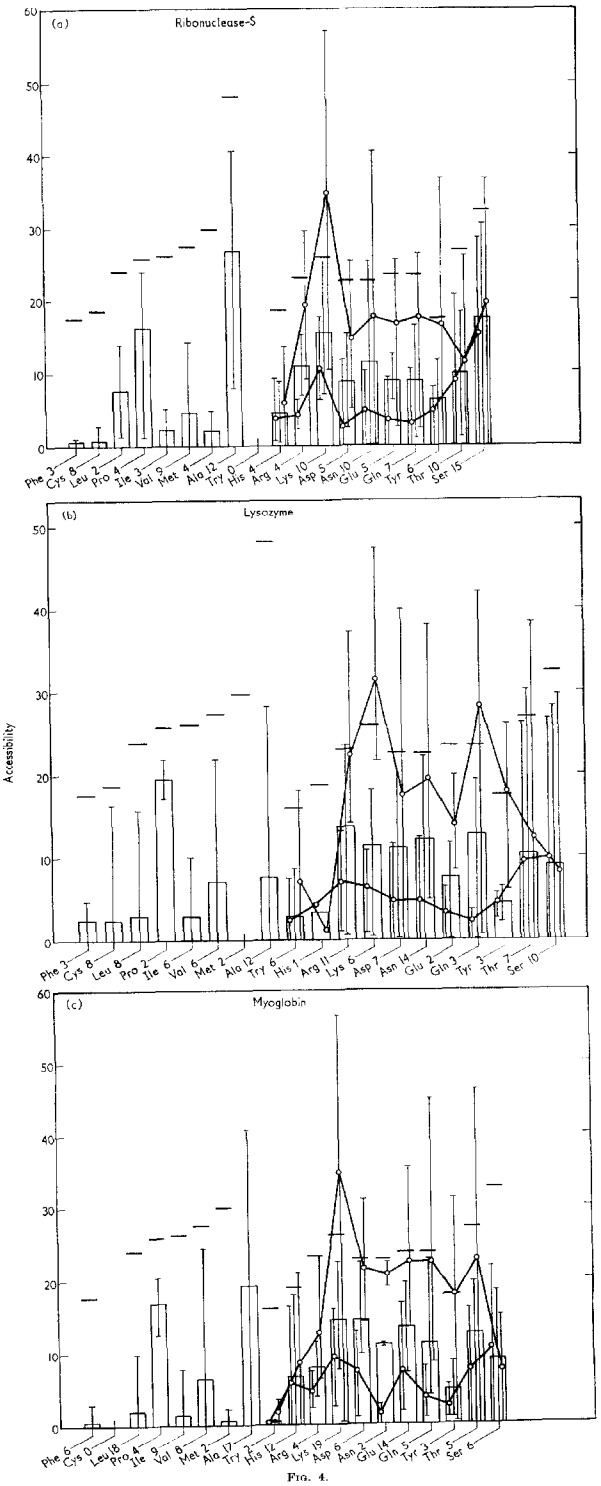
(a)
Ribonuclease-S
(b)
Lysozyme
(c)
Myoglobin
Accessibility
0
10
20
30
40
50
60
Phe 3
Cys 8
Leu 2
Pro 4
Ile 3
Val 9
Met 4
Ala 12
Try 0
His 4
Arg 4
Lys 10
Asp 5
Asn 10
Glu 5
Gln 7
Tyr 6
Thr 10
Ser 15
Phe 3
Cys 8
Leu 8
Pro 2
Ile 6
Val 6
Met 2
Ala 12
Try 6
His 1
Arg 11
Lys 6
Asp 7
Asn 14
Glu 2
Gln 3
Tyr 3
Thr 7
Ser 10
Phe 6
Cys 0
Leu 18
Pro 4
Ile 9
Val 8
Met 2
Ala 17
Try 2
His 12
Arg 4
Lys 19
Asp 6
Asn 2
Glu 14
Gln 5
Tyr 3
Thr 5
Ser 6

FIG. 4.

Watson & Kendrew, 1965). The heme bound water molecule forms part of the boundary of cavity D. The actual numbers are also uncertain to the extent that the appropriate van der Waal's radii of the various groups are not well known.

It is seen in Figure 2 that a large fraction of the accessible surface area is non-polar. Table 7 shows that when an amino-acid residue in an extended polypeptide chain is incorporated into a small protein molecule its accessible surface area is reduced by a factor of nearly 3. The amount of reduction in the accessible surface area is almost the same for the polar and the non-polar atoms. If one looks at the side-chain atoms alone, there is a definite tendency for non-polar atoms to be more "buried" than the polar atoms. This generalization does not apply to the main chain largely because of the special character of the carbonyl oxygen atom. This particular atom is highly exposed in an extended chain and shows a very marked reduction in average accessibility in the native protein. The main-chain amide nitrogen and carbonyl carbon atoms also show marked reduction but their accessibilities are so small even in the extended conformation that they do not affect the total average significantly.

As seen in Figure 3, an average side-chain polar atom is nearly 3·5 times as accessible as an average non-polar side chain atom in the native protein. This factor is about 2 in the extended chain conformation. Thus the *changes in accessibility* in going from the extended chain to the folded conformation for polar and non-polar atoms differ by less than a factor of 2 as seen in Table 7. The solvent contribution to the driving force leading from a random chain to the native protein must be related to these changes in accessibility. There may be a much closer balance of the opposing polar and non-polar contributions to this term than might hitherto have been supposed. Such an effect might in turn bear on the peculiar resistance of recent data to fit the concept of "hydrophobic bonding" (Brandts, Oliveira & Westort, 1970).

In Figure 4 it can be seen that proline is an outstanding exception to the tendency of non-polar residues to be "buried". The accessibility of alanine is highly variable.

For both lysozyme and myoglobin, C$\beta$ of the serine residues is, on the average, more accessible than the OH. One must note also that the range of accessibility of individual amino-acid residues of a given type is rather large for all classes as indicated by the long vertical lines in Figure 4. Because of these wide variations simple concepts, such as that of "buried" and "exposed" residues, cannot be used safely as guides in attempts at structure prediction. Simple generalizations of any sort are remarkably difficult to see in the summaries so far produced.

We wish to express our sincere appreciation to all members of the laboratory for many discussions on the subject of this paper and especially to H. W. Wyckoff.

This work was supported by research grants from the National Institute of General Medical Sciences. One of us (B. L.) held a postdoctoral fellowship from the Arthritis and Metabolic Institute of the National Institutes of Health.

## REFERENCES

Beintema, J. J. & Gruber, M. (1967). *Biochim. biophys. Acta*, **147**, 612.
Bondi, A. (1964). *J. Phys. Chem.* **68**, 441.
Brandts, J. F., Oliveira, R. J. & Westort, C. (1970). *Biochemistry*, **9**, 1038.
Donohue, J. & Caron, A. (1964). *Acta Cryst.* **17**, 1178.
Leung, Y. C. & Marsh, R. E. (1958). *Acta Cryst.* **11**, 17.
Marchiori, F., Rocchi, R. Moroder, L. & Scoffone, E., (1966). *Gazz Chim. Ital.* **96**, 1549.
Pasternak, R. A. (1956). *Acta Cryst.* **9**, 341.
Pauling, L., (1960). In *The Nature of the Chemical Bond*, 3rd ed. p. 518. Ithaca, New York: Cornell University Press.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Advanc. Protein Chem.* **23**, 283.
Schoenborn, D. P., Watson, H. C. & Kendrew, J. C. (1965). *Nature*, **207**, 28.
Scott, R. A. & Scheraga, H. A. (1966). *J. Chem. Phys.* **45**, 2091.
Watson, H. C. (1969). *Progress in Stereochemistry*, **4**, 299.
Wyckoff, H. W. (1968). *Brookhaven Symp. Biol.* **21**, 252.
Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B. & Richards, F. M. (1970). *J. Mol. Biol.* **245**, 305.