A computergraphical method of describing the shapes of subunit interfaces of oligomers

Analysis of the quaternary structure of concanavalin A and of prealbumin

E. James MILNER-WHITE

Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, U.K.

(Received 8 February 1982/Accepted 14 May 1982)

A new method is described for mapping the shape of the intersubunit contacts of oligomers with identical subunits or pairs of subunits. Plots are generated for the dimer–dimer interactions of concanavalin A and prealbumin, which are both tetramers. Both contacts lie at the junction between the faces of two apposed β -sheets, but in most other respects the interactions are seen to be different.

It is inherently difficult to represent three-dimensional structures on paper, and especially so in the case of intersubunit interactions of oligomers. When the three-dimensional structure of an oligomeric protein is presented by X-ray crystallographers, the intersubunit contacts are usually described by means of lists of the residues in contact, accompanied by a diagram of the shape of the oligomer and perhaps some stereo pairs showing some of the residues in contact. In many cases the reader is left in considerable doubt as to the precise shape and geometry of the interaction. It is the purpose of the present work to suggest new methods of picturing these contact areas that will enable more structural information to be conveyed.

Several useful methods of picturing protein surfaces have already been described. There are those (Richards, 1977; Greer & Bush, 1978; Morgan *et al.*, 1979; Langridge *et al.*, 1981) based on the surface defined as the closest approach of solvent spheres to the macromolecule, first worked out by Lee & Richards (1971). There are also methods of plotting individual atoms of a macromolecule as spheres with their approximate van der Waal's radii together with hidden line removal (Knowlton & Cherry, 1977; Feldman *et al.*, 1978), which are now widely used both for proteins and nucleic acids. In the following approach atoms are considered initially simply as points. The computing involved in the method has been described by Nicol *et al.* (1982).

Concanavalin A (Hardman & Ainsworth, 1972; Reeke *et al.*, 1975, 1978) and prealbumin (Blake & Oatley, 1977; Blake *et al.*, 1978; Blake, 1981) were chosen for the initial study because they are both tetramers consisting mainly of four layers of

antiparallel sheet; the dimer–dimer contact lies between the middle two sheets, whereas a feature of the monomer–monomer contact is the set of hydrogen bonds linking anti-parallel strands at the centre of each sheet. In a study of the quaternary structure of a range of tetrameric proteins (Milner-White, 1980) it was noted that prealbumin had certain structural features different from those of the others, including concanavalin A. The present paper is an investigation of the dimer–dimer contacts of these two proteins.

Methods

The co-ordinates of the atoms are taken from the Protein Data Bank (Bernstein *et al.*, 1977). The co-ordinates were then transformed, when necessary, so that their dyad axes corresponded with the dyad axes of the tetramer. For prealbumin, in which two out of the three axes are not true axes, the positions of the pseudo-axes were estimated by assuming that they pass through a point midway between the centres of mass of adjacent subunits. This involves a rotation of the *X*- and *Y*-axes about the *Z*-axis (the true dyad). Since the rotations are somewhat different for the two axes, the average rotation was calculated and used to calculate the optimum position of the new axes.

The molecular space, for computational purposes, was divided into columns 0.1 nm $\times$ 0.1 nm square in cross-section. The computer program then considers one dimer and finds the atom in each column nearest to one end of it. This generates the 'surface' of the dimer. By a simple rotation the corresponding 'surface' of the symmetry-related dimer is generated. In the next step each 'surface atom' is considered in

and give their distance from it in nm (± 0.05 nm). Letters (A, B, C etc.) indicate surface atoms that lie at distances (0.1, 0.2, 0.3 nm etc., ± 0.05 nm) on the other side of the dyad axis. (b) The same data are plotted but with squares of differential shading instead of symbols. The darkness of the shading indicates the degree to which the atoms protrude towards the other dimer. Where they overlap the darker atoms, which lie nearer to the reader, obscure the lighter ones. (c) The sequence number and type of amino acid of each interacting atom or group of such atoms is plotted. (d) The outline of both dimers is drawn. That with the thicker outline lies to the front. The shading marks the areas of inter-dimer contact. The different areas of intersubunit contact are indicated by different kinds of hatching.

subunit nomenclature of Reeke *et al.* (1975) is used. Note that in Fig. 5 of that paper the numerals III and IV should be reversed (G. N. Reeke, personal communication). x and y are dyad axes. (a), (b), (c), and (d) are plotted to the same scale and in the same way as those of Fig. 1. The symbols used and other details are given in the legend to Fig. 1.

turn. If it lies at a distance of 0.4 nm or less from any one of the 'surface atoms' of the symmetry-related dimer (not necessarily in the same column), it is placed in a new array of 'contact surface atoms'. If not, it is discarded.

The dimer-dimer interfaces are related by two dyad axes (for prealbumin one is a pseudo-dyad), which define a plane. This corresponds to the plane of the paper in Figs. 1 and 2. The third dimension is illustrated by symbols in Figs. 1(a) and 2(a). Atoms at the surface that protrude beyond the plane including the two dyad axes are indicated by numbers (+0.05 nm to +0.15 nm = '1', +0.15 nm to +0.25 nm = '2' etc.). Surface atoms that lie on the other side of the plane (nearer to the centre of the dimer) are specified by letters (−0.05 nm to −0.15 nm = 'A', −0.15 nm to −0.25 nm = 'B' etc.). Atoms within 0.05 nm of the plane are indicated by '0'.

In Figs. 1(b) and 2(b) the co-ordinates of the contact surface atoms are plotted in a different way, with the use of shading and hidden surface removal to illustrate the shape of each interface. Each atom is represented by a shaded square 0.42 nm × 0.42 nm. Atoms that protrude towards the other dimer are darker. Where two atoms of different shades overlap, the whole of the darkest atom is plotted, leaving only that part of the other atom that is not obscured. In the current graphics program there are a maximum of 16 different shades, each one representing a distance of 0.073 nm in these Figures. Figs. 1(c), 1(d), 2(c) and 2(d) are not computer-generated. They are drawings traced from the plots of Figs. 1(a) and 2(a). Figs. 1(c) and 2(c) indicate the amino acid and residue number of each contact surface atom or group of such atoms. Figs. 1(d) and 2(d) give views of the interfaces between individual subunits in the dimer-dimer contact.

Most of the computing was performed on an ICL 2976 computer. The symbols in Figs. 1(a) and 2(a) were printed by an LA120 Decwriter III. Figs. 1(b) and 2(b) were plotted on a matrix printer attached to a Hewlett-Packard 2647A graphics terminal in the raster-scan mode. Further details are described by Nicol *et al.* (1982).

Results and discussion

Figs. 1(a) and 2(a) show plots of all the atoms at the surface of one dimer that lie within 0.4 nm of one of the surface atoms of the other dimer. All the atoms or residues that were previously found [Table 5 and Fig. 12 in Blake *et al.* (1978) for prealbumin; Table IX in Reeke *et al.* (1975) for concanavalin A] to be involved in the interactions are included in the Figures. It should be noted that, if a part of the subunit-subunit interface is very irregular or is interfolded, some of the surface might be expected to

be missed in this procedure. However, as the size of cross-section of the columns into which each dimer is divided is decreased, any atoms that have been missed are expected to re-appear. The usual size of each column was 0.1 nm × 0.1 nm; the plotting procedures were therefore performed with columns of dimensions 0.075 nm × 0.075 nm in cross-section. Only a few further atoms were detected, as presented in Table 1.

The precise position of each interacting atom is indicated in Figs. 1(a) and 2(a) with the help of symbols noted in the legends. Appreciation of the three-dimensional structure is easier, however, in Figs. 1(b) and 2(b), where interacting atoms in one dimer are more heavily shaded, depending on the degree to which they protrude towards the other dimer. The symmetry of each dimer is also apparent, because protrusions and concavities are related by each of the dyad axes.

In Figs. 1(c) and 2(c) the type and residue number of all the amino acids in the dimer that contribute interacting atoms are plotted. For prealbumin residues from six out of the eight strands of the innermost β -sheet are included. However, the major dimer-dimer contacts derive from two loops at the edge of the β -sheets. Residues 112–115 form a β -turn, and residues 17–23 contain a hydrogen-bonding network folded into a loop. The interactions between the dimers consist mainly of hydrogen bonds and hydrophobic interactions. However, in concanavalin A all 12 strands of the inner β -sheet in each dimer contain residues that interact, and there are several electrostatic interactions as well as hydrogen bonds and hydrophobic interactions. These features are described by Blake *et al.* (1978) and by Reeke *et al.* (1975).

Figs. 1(d) and 2(d) show the outlines of the associated dimers with the areas of contact between different subunits shaded differently. These areas were estimated by superimposing tracings of the two sets of contact atoms in the correct orientations. There are water-filled cavities at the centre of both tetramers. In the case of prealbumin the cavity is part of a channel that extends through the whole molecule. As a result the whole dimer-dimer interacting surface is divided into two separate and complementary areas. It is in this channel that the ligand thyroid hormone binds (Blake, 1981), with its aromatic end to the centre. Only one molecule has been shown to bind per tetramer, although there appears to be room for two. For concanavalin A, on the other hand, there is a central cavity plus, probably, two smaller cavities, shown as diamond-shaped areas in Fig. 2(d), but there is no obvious channel leading from these to the outer solvent.

Considering the data of Table 1 for prealbumin in conjunction with Fig. 1(d), it is seen that the contacts across one axis (*X*) are somewhat more

Table 1. *Extent of inter-subunit contacts across different axes measured by the numbers of such atoms in one dimer within 0.4 nm of an atom in the other dimer*

The numbers in parentheses refer to the numbers of interacting atoms found when the molecular space was divided into columns $0.075 \text{ nm} \times 0.075 \text{ nm}$ in cross-section instead of the usual $0.10 \text{ nm} \times 0.10 \text{ nm}$.

Prealbumin		
I-IV and II-III, across <i>X</i> -axis	(61)	59
Across both axes (<i>X</i> and <i>Z</i>)	(3)	3 (α -C Gly-22 and δ -C Leu-110)
I-III and II-IV, across <i>Z</i> -axis	(34)	34
Total	(98)	96
Concanavalin A		
I-III and II-IV, across <i>y</i> -axis	(144)	136
Across both axes (<i>y</i> and <i>x</i>)	(6)	6
I-IV and II-III, across <i>x</i> -axis	(4)	4
Total	(154)	146

numerous than those related by the other axis (*Z*). For concanavalin A, however, nearly all the many contacts are across one axis (*y*) and only a few lie across the other axis (*x*). In other words, there are few contacts between subunits I and IV or between subunits II and III. This is related to the fact that the intersubunit boundaries within each dimer lie approximately parallel to each other (and to the *x*-axis). For clarity, these boundaries have been illustrated in only parts of the shaded areas of Fig. 1(d).

The study of primary structures (Jörnvall *et al.*, 1981) has indicated that prealbumin is one of a family of proteins that includes retinol-binding protein, glucagon, vasoactive intestinal peptide, secretin and glicentin. Concanavalin A, on the other hand, has a supersecondary structure with similarities to immunoglobulin G and to superoxide dismutase (Richardson *et al.*, 1976). On this basis similarities between concanavalin A and prealbumin are not to be expected. My analysis of quaternary structure confirms this. Although the dimers of both tetramers form contacts at the faces of their inner β -sheets, the shapes of these interactions differ.

A major purpose of this work is to present a new way of visualizing subunit-subunit interfaces. It has certain advantages compared with other methods of portraying surfaces quantitatively in three dimensions. It is very simple to use. It avoids too much detail because only the relevant interacting parts of the surface are portrayed. This means that it is potentially suitable for picturing the largest (often the most interesting) proteins. It should be useful for proteins that change their conformation at the quaternary structure level and also for making evolutionary comparisons between proteins.

I thank the Computing Service Department at the University of Glasgow, Glasgow, Scotland, U.K., and the Protein Data Bank at the Department of Chemistry,

Brookhaven National Laboratory, Upton, NY, U.S.A., for their help.

References

- Bernstein, F. C., Koetzle, T. K., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rogers, J. R., Kennard, O., Simanouchi, T. & Mitsuo, T. (1977) *J. Mol. Biol.* **112**, 535–542
- Blake, C. C. F. (1981) *Proc. R. Soc. London Ser. B* **211**, 413–431
- Blake, C. C. F. & Oatley, S. J., (1977) *Nature (London)* **268**, 115–120
- Blake, C. C. F., Geisow, M. J., Oatley, S. J., Rerat, B. & Rerat, C. (1978) *J. Mol. Biol.* **121**, 339–356
- Feldman, R. J., Bing, D. H., Furie, B. L. & Furie, B. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5409–5412
- Greer, J. & Bush, B. L. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 303–307
- Hardman, K. D. & Ainsworth, C. F. (1972) *Biochemistry* **11**, 4910–4919
- Jörnvall, H., Carlström, A., Pettersson, T., Jacobsson, B., Persson, M. & Mutt, V. (1981) *Nature (London)* **291**, 261–263
- Knowlton, K. & Cherry, L. (1977) *Comput. Chem.* **1**, 161–166
- Langridge, R., Ferrin, T. E., Kuntz, I. D. & Connolly, M. L. (1981) *Science* **211**, 661–666
- Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379–400
- Milner-White, E. J. (1980) *Biochem. J.* **187**, 297–302
- Morgan, R. S., Miller, S. L. & McAdam, J. M. (1979) *J. Mol. Biol.* **127**, 31–39
- Nicol, C. J., Kilgour, A. C. & Milner-White, E. J. (1982) *Inter-Universities Computing Centre Bulletin* in the press
- Reeke, G. N., Becker, J. W. & Edelman, G. M. (1975) *J. Biol. Chem.* **250**, 1525–1547
- Reeke, G. N., Becker, J. W. & Edelman, G. M. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2286–2290
- Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng.* **6**, 151–156
- Richardson, J. S., Richardson, D. C., Thomas, K. A., Silverton, E. W. & Davies, D. R. (1976) *J. Mol. Biol.* **102**, 221–235