Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Features of homotetrameric molecular association in protein crystals

Uma V. Katre and C. G. Suresh*

Division of Biochemical Sciences, National Chemical Laboratory, Pune 411 008, India

Correspondence e-mail: cg.suresh@ncl.res.in

Received 4 March 2008
Accepted 17 September 2008

The crystal structures of proteins showing homotetrameric association, a common feature observed in many lectins, have been analyzed in order to understand the characteristics of tetrameric association in terms of the arrangement of subunits and their biological significance. The analysis could group the tetramer units into the following four categories. (i) Tetrahedral molecules, in which the four monomers form a nearly perfect tetrahedral arrangement. The angle between the axes of any two monomers is $\sim 109^{\circ}$. (ii) Molecules that form a sandwiched dimer of dimers in which the two dimers are arranged perpendicular to each other, one upon the other. (iii) Planar molecules, in which the four monomers lie in one plane and the corresponding sides of adjacent monomers face in opposite directions. This can be considered as a flattened tetrahedral shape. (iv) Planar closed molecules, in which all four monomers lie in one plane arranged in a head-to-tail fashion in a square. The first group and its variant, the third group, are the most commonly found arrangements in crystal structures. Each arrangement has its own importance for biological function. Some tetrameric assemblies that deviate from the majority described above also have relevance to their biological function.

1. Introduction

Researchers have approached the problem of protein assembly in various contexts. A variety of reasons have been proposed for the formation of oligomers and complexes. For example, in the case of virus assembly, genetic economy was identified as the advantage for subunit interactions and arrangement (Dokland, 2000; Phelps *et al.*, 2000). The association of several copies of the same gene product can minimize errors or alternatively can uniformly distribute mutations throughout the assembly. For proteins such as haemoglobin, the requirement for tetramers is for the allosteric control of oxygen binding (Monod *et al.*, 1965). In certain membrane proteins, subunit assembly helps to create an external hydrophobic and internal polar surface to help ion transport (Manting *et al.*, 2000; Sakaguchi *et al.*, 1997). Oligomerization has been proposed in situations where regulation of the concentration levels of the constituent subunits is required (Bray & Lay, 1997), which also results in increased stability and reduced surface area of the constituent molecules (Larsen *et al.*, 1997; Zaremba *et al.*, 2005, 2006). In certain enzymes subunits assemble to form a symmetric substrate-binding cleft, such as in HIV protease (Wlodawer & Erickson, 1993). Oligomerization has been identified as one of the ways to achieve thermostability of proteins in thermophilic organisms (Walden *et al.*, 2001). As may be expected, in most cases oligomerization leads to increased complexity.

Many plant lectins, such as the *Artocarpus hirsuta* lectin reported by our group (Rao *et al.*, 2004), and enzymes such as penicillin V acylase (Suresh *et al.*, 1999) and conjugated bile acid hydrolase (Kumar *et al.*, 2006), also reported by our laboratory, showed tetrameric association of their subunits. From the analysis of the structure of the *A. hirsuta* lectin, one possibility that emerged is that a pattern of subunit association can help to keep the functional sites of subunits at a maximum distance away from each other. This helps to form networks of cells during the agglutination of red blood cells. Based on these observations, an analysis of the structures of proteins that are known to form homotetramers has been undertaken.

Many reports on the various aspects of protein oligomerization and their relevance in biology have appeared (Ali & Imperiali, 2005). One such study estimated that more than one third of cellular proteins form oligomers (Goodsell & Olson, 2000). Although oligomers can be composed of multiple subunits of the same polypeptide (homo-oligomers) or different polypeptides (hetero-oligomers), it is hypothesized that proteins preferably form homo-oligomers in cells (Goodsell, 1991). Similarly, it has been argued that even though oligomeric proteins can be formed from any number of subunits, the average oligomeric state of cellular proteins expected is tetrameric (Goodsell, 1991). In homo-oligomeric proteins, since the constituents of the assembly are identical, formation of an oligomer can introduce simple point-group symmetry. Goodsell & Olson (2000) have estimated that cyclic, dihedral and cubic point-group symmetries are most frequently observed.

The stability of an oligomer will directly depend on the strength of association of the subunits, their affinity and duration. Thus, a subunit with strong subunit interactions will invariably be found as an oligomer, while the formation of oligomers by subunits with weaker interactions may depend on the concentration and other conditions such as pH and temperature or may occur in response to some other stimuli (Nooren & Thornton, 2003*a*,*b*; Nagano *et al.*, 2008).

Many researchers have attempted to rationalize and quantify protein–protein recognition, the type of interactions and the nature of the interfaces involved in protein oligomerization in a wider sense (Chothia & Janin, 1975; Miller *et al.*, 1987; Argos, 1988; Janin *et al.*, 1988; Miller, 1989; Jones & Thornton, 1996; Elgavish & Shaanan, 2001). Rationalization and prediction of quaternary-structure formation in the case of legume lectins has been carried out in terms of buried hydrophobic surface, interaction energy and shape complementarity at the interface of subunits, structure and sequence analysis (Prabu *et al.*, 1999; Chandra *et al.*, 2001; Brinda *et al.*, 2004, 2005; Del Sol *et al.*, 2007). Similarly, attempts have been made to study protein oligomerization and stability in the case of legume lectins by analyzing the unfolding of their quaternary organization (Srinivas *et al.*, 2001). Here, an attempt is made to analyze the quaternary structures of tetrameric lectins and other proteins showing tetrameric association in terms of their symmetry of organization and biological relevance.

2. Materials and methods

All computational work was carried out on a Silicon Graphics workstation (Octane) with Irix 6.5 as the operating system as well as on an IBM PC with Fedora Core 6 as the operating system. The atomic coordinates of various homotetrameric proteins were downloaded from the Protein Data Bank (PDB; http://www.rcsb.org/pdb/) according to their space groups, with a sequence-identity cutoff of 70%. The information in the PDB file (`REMARK 350`) was used to decide whether or not the biological unit of the protein was a homotetramer. Although an attempt has been made to include all unique tetrameric

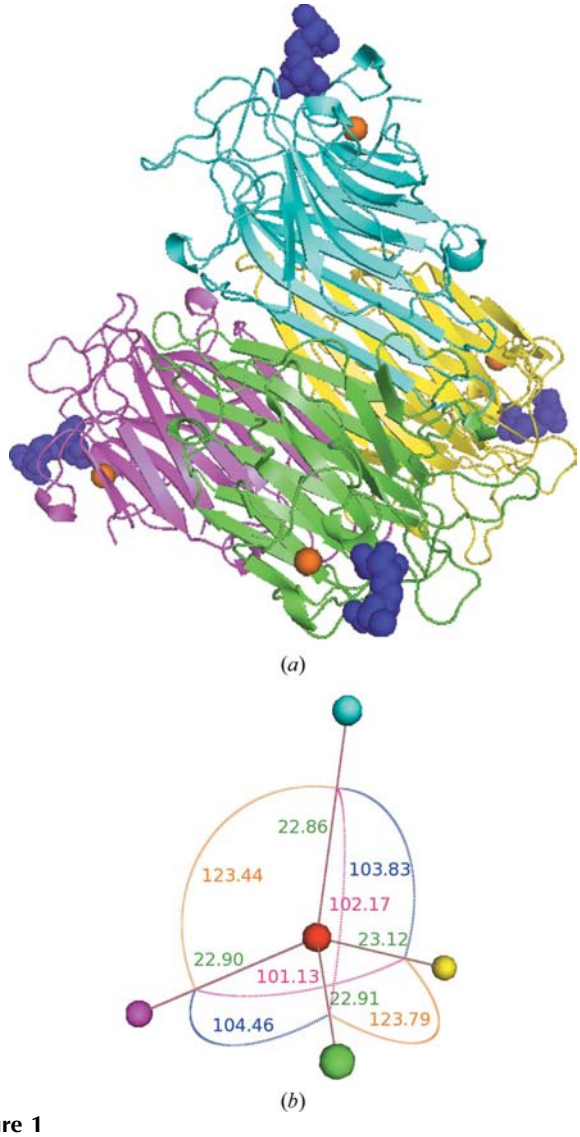

Figure 1
(*a*) A perfect tetrahedral arrangement of subunits as observed in the tetramer of ConA (PDB code 1qdc). The carbohydrate-ligand molecules, methyl-6-*O*-(α-D-mannopyranosyl)-α-D-mannopyranoside (shown as space-filling models) bind at the four corners of the tetrahedron. Ca^{2+} ions are shown as orange spheres, whereas Mn^{2+} ions are shown in purple. (*b*) Distances and angles between the centres of mass of four subunits in the case of 1qdc. The distances are shown in green. Pink, angles between subunits of the same dimer; blue, angles between the adjacent subunits of different dimers; orange, angles between diagonally opposite subunits. The same colour-coding has been used for all subsequent diagrams of centres of mass.

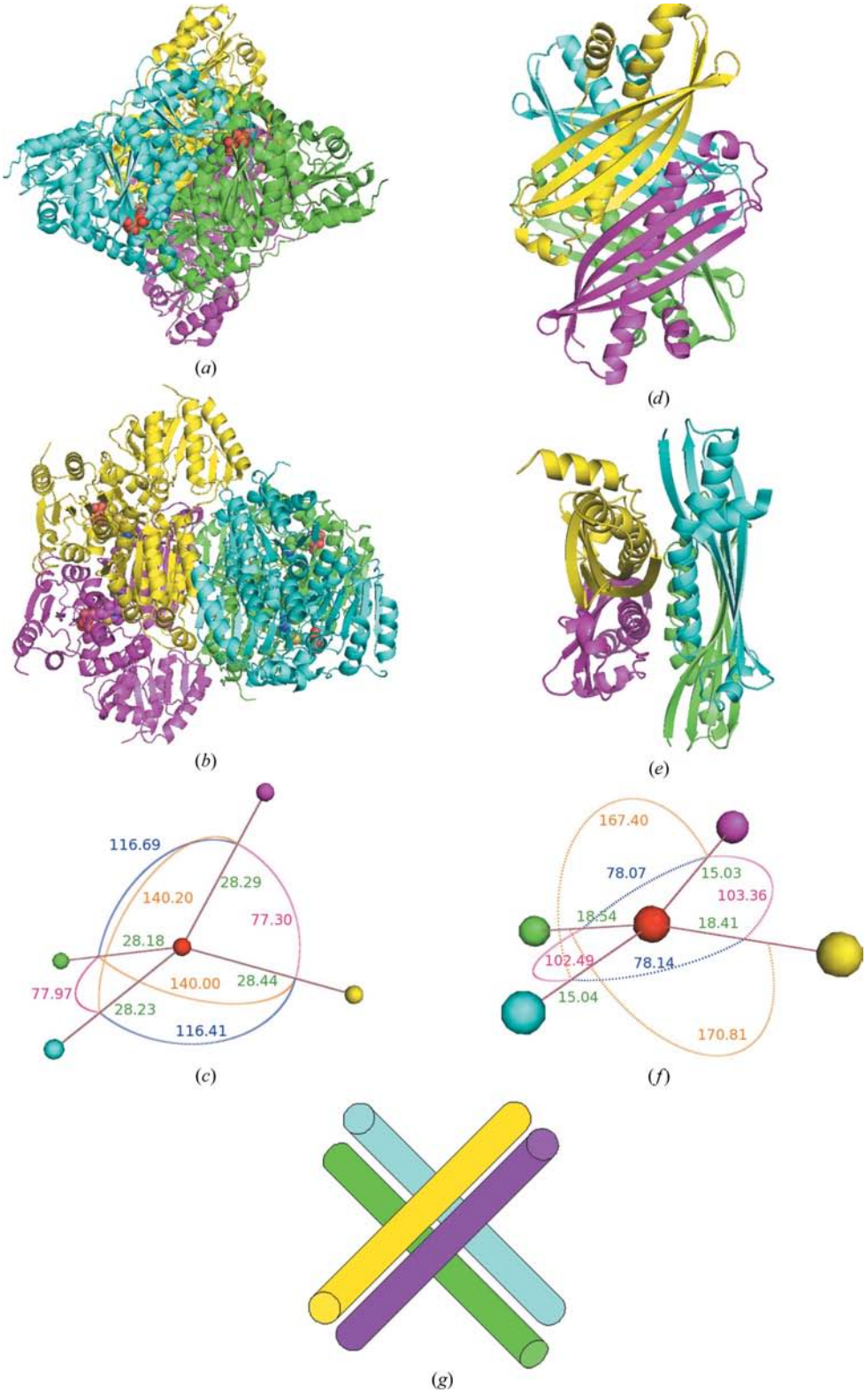

Figure 2
(*a*, *b*) Indole-pyruvate decarboxylase from *Enterobacter cloacae* (PDB code 1ovm) shown in two different orientations. The ligand molecules, thiamine diphosphate, are shown as space-filling models. The subunits coloured green and cyan form one dimer and those coloured magenta and yellow form another. (*c*) Distances and angles between the centres of mass of the four subunits of 1ovm. (*d*, *e*) SecB from *Escherichia coli* (PDB code 1qyn) shown in two different orientations. These two proteins are dimers of dimers and the two dimers are roughly perpendicular to each other. (*f*) Distances and angles between the centres of mass of the four subunits of 1qyn. (*g*) A schematic representation of the dimers placed perpendicular to each other.

protein structures in the analysis, it is possible that some structures might have been excluded if they were not categorized as tetramers by the PDB. If the subunit consisted of two or more different polypeptide chains and if four such subunits formed the biological molecule then they were also considered for analysis.

Wherever the asymmetric unit differed from the biological unit, the coordinates of the biological unit were downloaded from the PDB website (`*.pdb1` files), or sometimes the biological unit was generated by calculating the coordinates of symmetry-related subunits. The protein structures were visualized using the graphical software *QUANTA* (Accelrys Inc.) and grouped into various classes. The secondary-structure content of the protein was also roughly estimated in order to check whether there was any dependence of the overall assembly of the molecule on the secondary structure. To calculate the centres of mass (COMs) for each molecule and its subunits, the (centre-of-mass) command in *Coot* (Emsley & Cowtan, 2004) was used. The coordinates of the COMs of all the four subunits as well as that of the tetramer were converted into a PDB file by placing a hypothetical water molecule at each of the five positions. This PDB file was displayed in *Coot* or *PyMOL* (DeLano, 2000) and the distances of the COMs of each subunit from the COM of the tetramer as well as the angles between them were calculated. The diagrams of quaternary structures of proteins and those of centres of mass were prepared using *PyMOL* (DeLano, 2000).

3. Results and discussion

650 unique homotetrameric protein structures selected at a sequence-identity cutoff of 70% were analyzed. The structures

were grouped according to space group in order to find any trend in preferred quaternary structure associated with a particular space group. These protein structures could be grouped into four major categories.

3.1. Dihedral/tetrahedral-type assemblies

In this type of assembly, the four subunits of a protein are arranged pointing towards the four corners of an approximate tetrahedron. Concanavalin A (ConA; PDB code 1qdc; Bouckaert *et al.*, 1999), a legume lectin, exhibits a near-perfect tetrahedral shape (Fig. 1*a*) as shown by the measurement of tetrahedral angles between the centres of mass of the four subunits (Fig. 1*b*). Other tetrameric legume lectins, for example *Dioclea grandiflora* (PDB code 1dgl; Rozwarski *et al.*, 1998) and *D. guianensis* lectins (PDB codes 1h9p and 1h9w; Wah *et al.*, 2001) as well as the tetramer of the heterodimeric subunit *Dolichos lablab* lectin (FRIL; PDB code 1qmo; Hamelryck *et al.*, 2000) also show a near-perfect tetrahedral arrangement of subunits (point group 23). A variation of this type of arrangement is observed in many other proteins, including enzymes, and the arrangement gives an internal dihedral symmetry to the tetramer (point group 222).

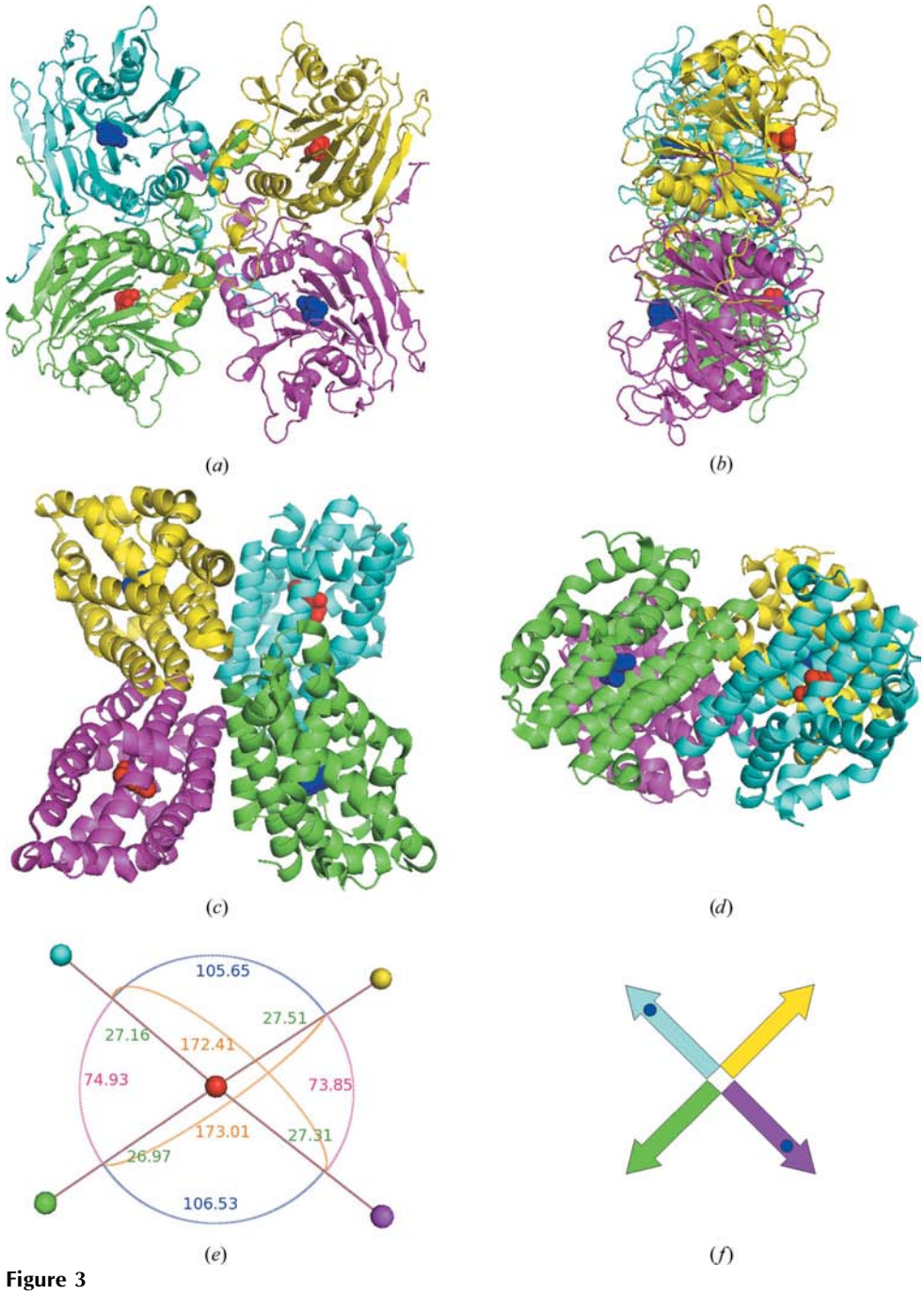

Figure 3
(*a*) Top view and (*b*) side view of the enzyme penicillin V acylase from *B. sphaericus* (PDB code 2pva). The bound dithiane diol molecules are shown as space-filling models. The ligand-binding sites of any two adjacent subunits, which are same as substrate-binding sites, lie on opposite faces. (*c*) Top view and (*d*) side view of TenA homologue from *Pyrobaculum aerophilum* (PDB code 2gm8), an all-α protein. The ligand (4-amino-5-hydroxymethyl-2-methylpyrimidine) molecules are shown as space-filling models. (*e*) Distances and angles between the centres of mass of the four subunits of 2pva. (*f*) A schematic representation of 2pva, showing the substrate-binding sites face up.

Most of the tetrameric proteins are reported to show two steps of oligomerization. Firstly, two monomers associate to form a dimer and two dimers in turn associate to form a tetramer (Powers & Powers, 2003). Owing to this, even in the near-perfect tetrahedral-shaped molecules the angles between any two monomers deviate at least slightly from the normal tetrahedral angle of 109° 28′. The angles between two monomers of the same dimer have the lowest value, ∼100°, followed by the angles between adjacent monomers of two different dimers, which are at least 3–4° more than the angle between monomers of the same dimer. The remaining two angles between the diagonally opposite subunits have a larger value of ∼125°, or sometimes even more, to compensate for the other four reduced angles.

Not all homotetrameric proteins in this category show such a near-perfect tetrahedral shape. Most of them have angles between their subunits that deviate much more, distorting the tetrahedron, and this gives the molecule a twisted shape rather than a perfect tetrahedron.

3.2. Sandwiched dimer of dimers: two perpendicularly placed dimers

In this type of assembly, two protein monomers again associate to form a dimer and two dimers associate to form a tetramer. However, these dimers are placed roughly perpendicular to each other; that is, if each dimer is considered to be enclosed in a box the two boxes appear perpendicular to each other. This was thought to be a distortion of the tetrahedral shape. The angles between monomers of the same dimer are reduced to ∼80° and those between the adjacent monomers of two dimers range between 100 and 120°. The angles between two oppositely placed monomers have values in the range 140–160°. An example of such an arrangement is the indole-pyruvate decarboxylase from *Enterobacter cloacae* (PDB code 1ovm; Schutz *et al.*, 2003) shown in Figs. 2(*a*), 2(*b*) and 2(*c*). However, in another such protein, SecB from *Escherichia coli* (PDB code 1qyn; Dekker *et al.*, 2003), the angles between monomers of the same dimer are ∼100°, while those between the adjacent monomers of two dimers are ∼80° (Figs. 2*d*, 2*e* and 2*f*). A schematic representation of this type of arrangement is shown in Fig. 2(*g*). This kind of arrangement also produces a 222 point symmetry for the tetramer and may have the same biological significance as that of the tetrahedral arrangement.

3.3. Planar assembly (planar 1)

This type of arrangement is characterized by the presence of all four subunits of the protein in a single plane, with any two adjacent subunits facing in opposite directions. This can be considered as a flattened tetrahedral shape. From solvent-accessibility calculations, it is assured that even in this type of arrangement the protein first dimerizes and the dimers associate to form tetramers. The angles between two monomers of the same dimer range from 70° to 80°, while those between the adjacent monomers of different dimers have values of ∼100°. However, the angles between diagonally opposite monomers have values very close to 180° and are responsible for the flattened or planar shape of the molecule. This type of arrangement also shows point-group 222 symmetry between the subunits. One such example is the enzyme penicillin V acylase from *Bacillus sphaericus* (PDB code 2pva; Suresh *et al.*, 1999) shown in Figs. 3(*a*) and 3(*b*). Figs. 3(*c*) and 3(*d*) show a similar kind of arrangement in the TenA homologue from *Pyrobaculum aerophilum* (PDB code 2gm8), an all-α protein. Fig. 3(*e*) shows the distances and angles between the centres of mass of the four subunits of penicillin V acylase and Fig. 3(*f*) displays a schematic representation of the molecules, showing two substrate-binding sites facing up and the other two down. Certain legume lectins, for example *Phaseolus vulgaris* lectin (PHA-L; PDB code 1fat; Hamelryck *et al.*, 1996) and soybean agglutinin (SBA; PDB code 2sba; Dessen *et al.*, 1995), also belong to this class.

3.4. Planar closed molecules (planar 2)

The three types of subunit arrangements in tetrameric proteins described so far show point-group 222 symmetry. The fourth kind of arrangement, in which all the four subunits lie in a plane and facing in the same direction, shows a fourfold symmetry between the subunits. Owing to this, a 'closed' tetramer is formed. The angles between any two adjacent monomers are nearly 90° and those between oppositely placed subunits are almost equal to 180° and in two opposite directions. Many membrane-bound proteins, such as in ion channels and cell-surface enzymes, show this kind of arrangement; for example, the potassium channel from *Streptomyces lividans* (PDB code 1bl8; Doyle *et al.*, 1998), which is shown in Figs. 4(*a*) and 4(*b*). In Fig. 4(*c*) the distances and angles between the

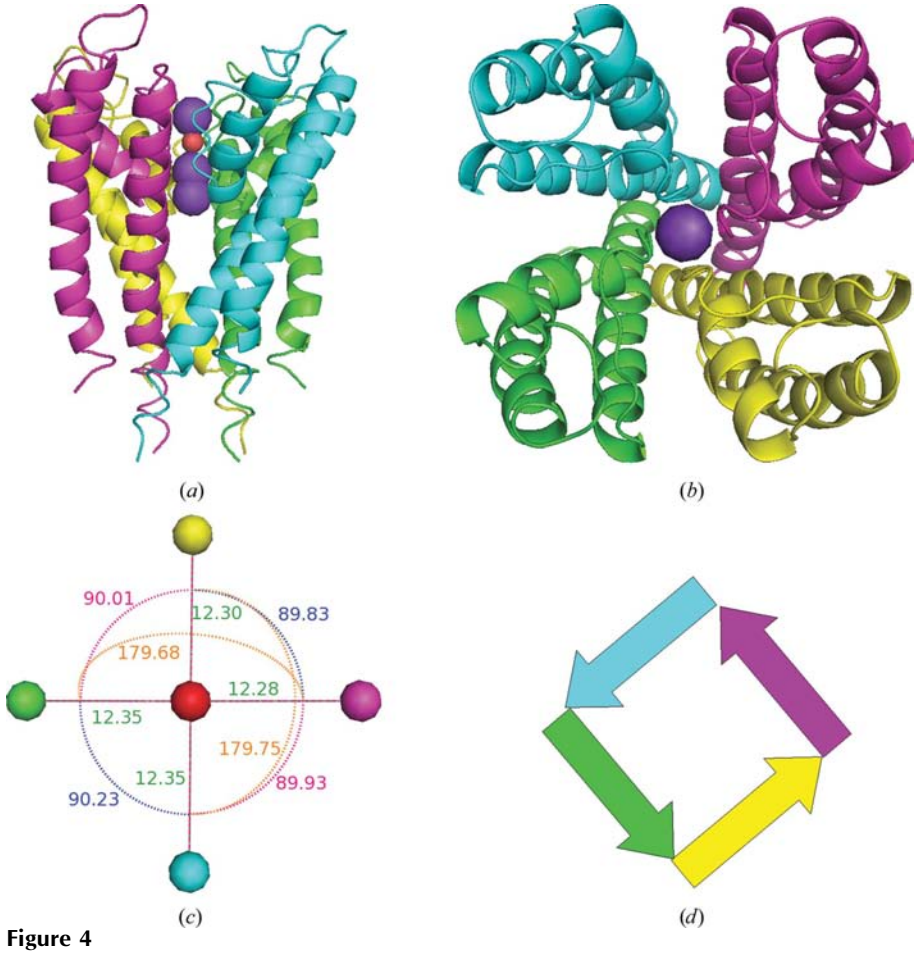

Figure 4
(*a*) Side view and (*b*) top view of potassium channel from *Streptomyces lividans* (PDB code 1bl8). The potassium ions and a water molecule are shown as spheres. (*c*) Distances and angles between the centres of mass of the four subunits. (*d*) Schematic representation of the molecule, exhibiting fourfold symmetry.

centres of mass are shown; a schematic representation of this kind of arrangement is shown in Fig. 4(*d*). A drawing of a potassium channel protein of this category, embedded in the membrane lipid bilayer, can be found at http://biop.ox.ac.uk/www/lj2001/sansom/sansom_1.jpg.

3.5. Tetrameric arrangements not belonging to the patterns described

Although most homotetrameric molecules could be grouped into one of the four above-mentioned categories, some molecules could not be included. Most such molecules, when analyzed in detail, were found to be wrongly labelled as homotetramers in their respective PDB files and hence were discarded from further analysis. Examples were human acidic fibroblast growth factor (PDB code 2afg; Blaber *et al.*, 1996), *Escherichia coli* SufC (PDB code 2d3w; Kitaoka *et al.*, 2006), rat liver dihydropteridine reductase (PDB code 1dir; Su *et al.*, 1994), the MPPN domain of mouse Nup35 (PDB code 1wwh; Handa *et al.*, 2006) and salicylic acid-binding protein 2 (SABP2) from *Nicotiana tabacum* (PDB code 1xkl; Forouhar *et al.*, 2005). Such likely errors in quaternary-structure assignments in the PDB have been reported previously (Levy *et al.*, 2006). However, some of the molecules were found to be genuinely homotetrameric but displayed a quaternary structure which could not be fitted into any of the above-mentioned categories. They did not show any other recognizable pattern either. The arrangements of subunits in some such molecules are described below.

3.5.1. Peanut lectin (PDB code 2pel). This legume lectin, despite sharing sequence as well as secondary-structure and tertiary-structure similarity with other legume lectins, shows a very peculiar 'open' quaternary structure. It contains two identical dimers, each having a twofold symmetry between its subunits; however, at the quaternary-structure level the molecule does not show any 222 or fourfold symmetry (Figs. 5*a* and 5*b*; Banerjee *et al.*, 1994, 1996). This unusual and unexpected structure was also reported to be responsible for the difficulty in solving its structure (Vijayan, 2007).

3.5.2. DNA-binding proteins. Several DNA-binding proteins such as lactose operon repressor protein (LacR) from *E. coli* (PDB code 1tlf; Friedman *et al.*, 1995) and tumour-suppressor protein p53 (PDB code 2ac0; Kitayner *et al.*, 2006) form a very peculiar quaternary structure which consists of two dyad-symmetric dimers which are nearly parallel to each

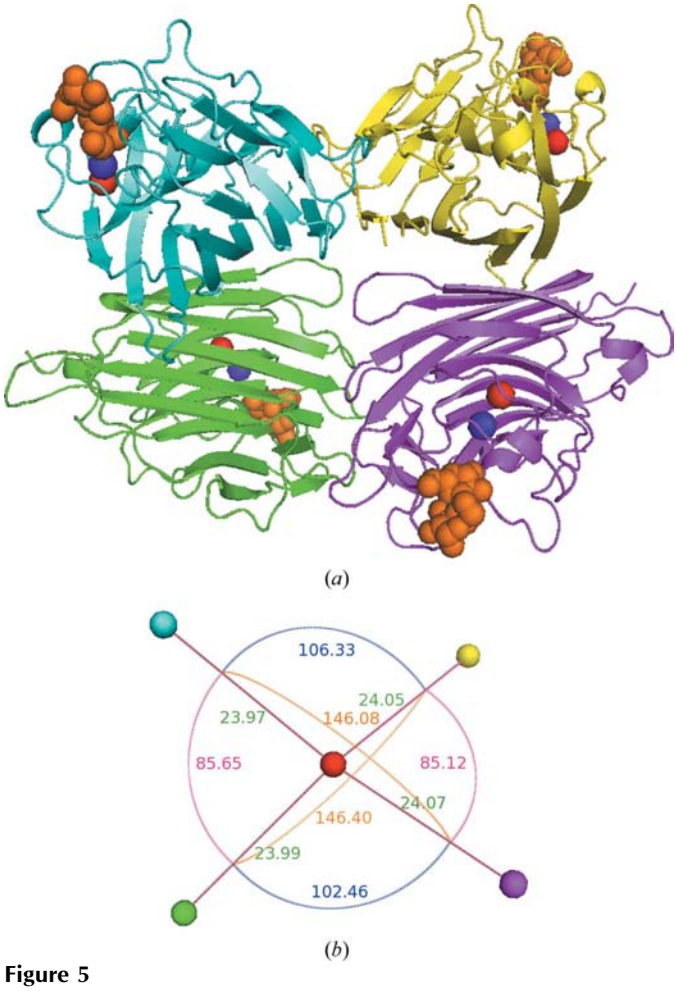

Figure 5
(*a*) Quaternary structure of peanut lectin (PDB code 2pel). The bound lactose molecules are shown as space-filling models. Purple spheres are Ca^{2+} ions, while red spheres are Mn^{2+} ions. (*b*) Distances and angles between the centres of mass of the four subunits.

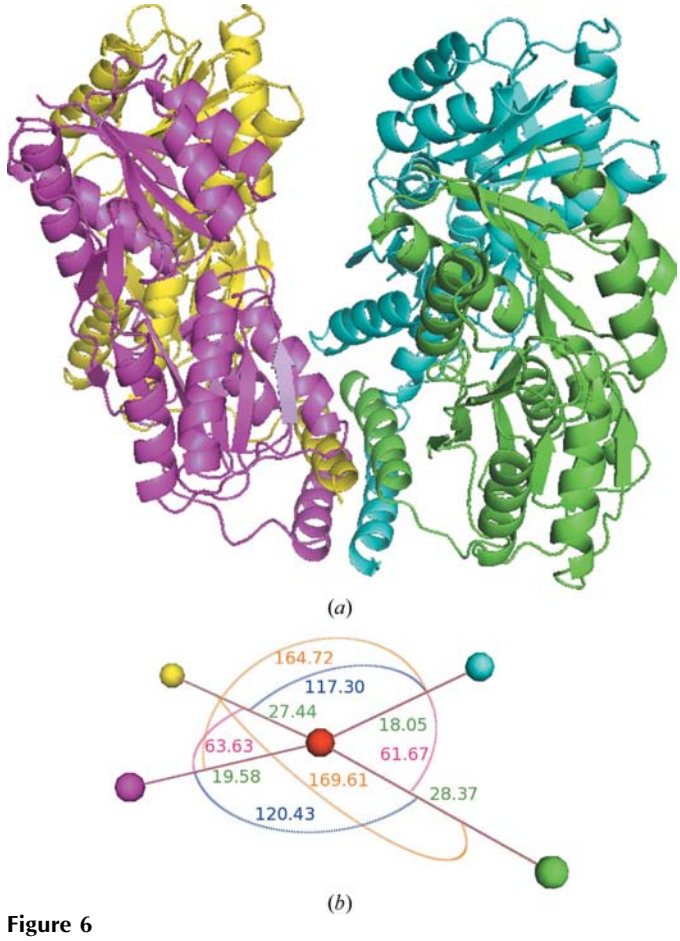

Figure 6
(*a*) Quaternary structure of the lactose operon repressor protein (PDB code 1tlf). The C-terminal helices of all four monomers are involved in tetramerization of the molecule. (*b*) Distances and angles between the centres of mass of the four subunits.

other. As a consequence of this, all four DNA-binding domains of intact LacR are placed on the same side of the tetramer. This results in a deep V-shaped cleft between the two dimers. An antiparallel four-helix bundle which is formed from four C-terminal helices, one contributed by each monomer, functions as a tetramerization domain (Figs. 6*a* and 6*b*). On binding to DNA, the tethered dimers of this protein broaden by ∼12° and the dimers twist by ∼8° (Lewis *et al.*, 1996). On removing the C-terminal helix, which assists in oligomerization of the protein, the affinity of the dimer towards the operator DNA sequence remains the same (Brenowitz *et al.*, 1991); however, the induction ratio (the ratio of the level of transcription in the presence and the absence of inducer) decreases (Oehler *et al.*, 1990).

3.5.3. λ phage-transcription activator protein CII (PDB code 1xwr). This is another DNA-binding protein which binds to a unique direct-repeat sequence $TTGCN_6TTGC$, which is observed in three phage promoters it activates. The tetramer is formed from dimers but does not exhibit any closed symmetry (Fig. 7*a*). The arrangement of the centres of mass of the four subunits is also peculiar in this protein (Fig. 7*b*). Here also, tetramerization is achieved by the formation of a four-helix bundle, each helix contributed by a monomer. The unusual quaternary structure of this protein allows it to position the helix–turn–helix motifs of two of the four CII subunits for interaction with successive major grooves of B-DNA from one face of DNA and helps to identify a direct-repeat DNA sequence rather than the inverted-repeat sequence (Datta *et al.*, 2005).

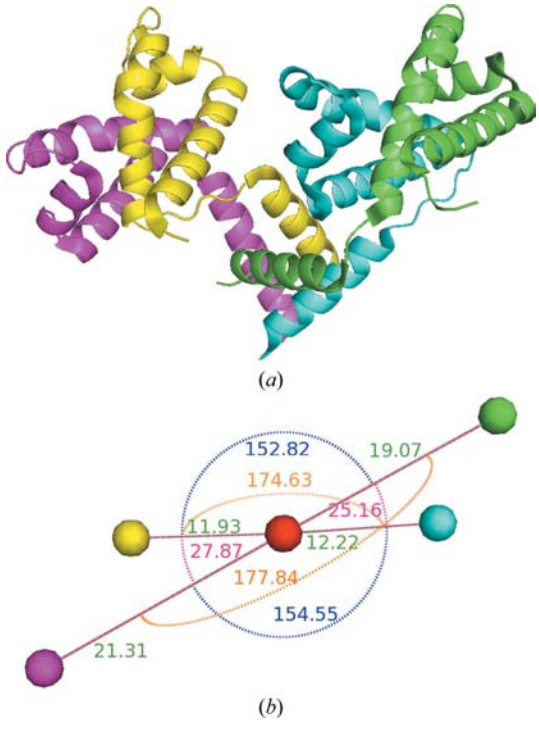

Figure 7
(*a*) λ phage CII protein (PDB code 1xwr). This protein also has four helices involved in tetramerization. (*b*) Distances and angles between the centres of mass of the four subunits, depicting their unusual arrangement.

3.5.4. *Mycobacterium tuberculosis* D-3-phosphoglycerate dehydrogenase (PGDH). The crystal structure of this enzyme (PDB code 1ygy; Dey *et al.*, 2005) consists of a dimeric asymmetric unit made of two identical subunits, each consisting of four domains. However, in one of the two subunits there is a rotation of ∼180° around a hinge region connecting two of the four domains. This introduces significant asymmetry in the dimer. Two such asymmetric units associate to form a biologically active tetramer (Fig. 8*a*). The distances and angles between the centres of mass of the four subunits are shown in Fig. 8(*b*). This asymmetric arrangement leads to the formation of two different and distinct domain interfaces between identical domains in the asymmetric unit as well as introducing asymmetry in the substrate-binding sites, which might have a role in the activity and regulation of the enzyme (Dey *et al.*, 2005).

3.6. Biological significance

In most of the proteins exhibiting a tetrahedral shape the binding sites are located at the four corners of the tetrahedron (Fig. 1*a*). This reduces the steric hindrance between the ligand molecules and hence increases the binding efficiency of the molecule. This could be the possible reason for the tetrahedral or distorted tetrahedral shape being the most commonly observed feature of homotetrameric molecules.

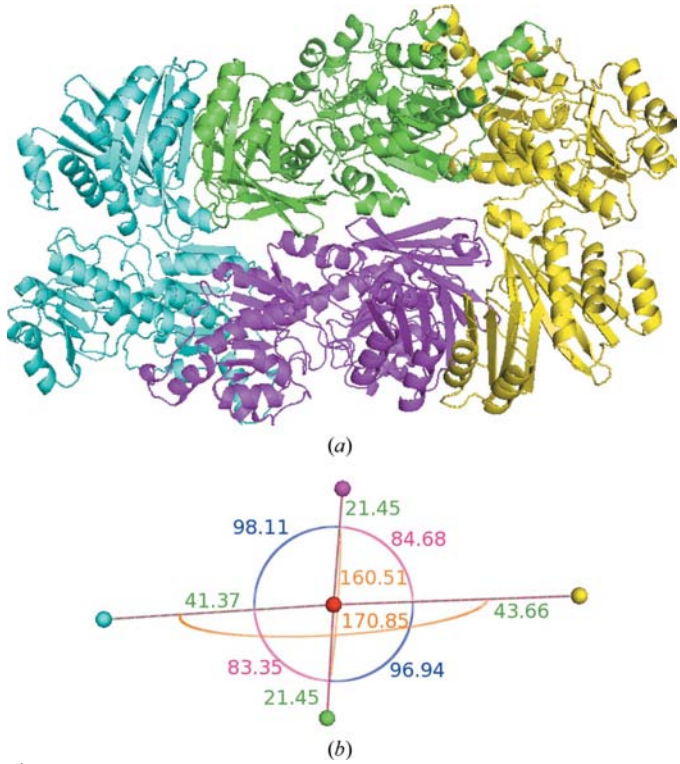

Figure 8
(*a*) Quaternary structure of *M. tuberculosis* PGDH (PDB code 1ygy). All four subunits have identical primary structure and consist of four domains each; however, in the subunits coloured cyan and yellow two of the four domains are flipped by almost 180° around a hinge region. This introduces the asymmetry in the dimer. (*b*) Distances and angles between the centres of mass of the four subunits, which form the corners of an approximate rhombus.

Table 1
The different tetrameric arrangements of subunits observed in the asymmetric unit of crystals in seven crystal systems.

Crystal system	Symmetry in tetramer	No. of subunits in ASU	Prevalent type of subunit arrangement	No. of structures in the tetramer type	Tetramers in each crystal system
Triclinic	Mostly 222	4	Tetrahedral/distorted tetrahedral	12	21
			Twisted perpendicular	3	
			Planar 1	5	
			Planar 2	1	
Monoclinic	Mostly 222	4, 8	Tetrahedral/distorted tetrahedral	83	138
			Planar 1	52	
			Planar 2	2	
			Unclassified	1	
Orthorhombic	Mostly 222	1, 2, 4, 8	Tetrahedral/distorted tetrahedral	136	258
			Twisted perpendicular	38	
			Planar 1	74	
			Planar 2	8	
			Unclassified	2	
Trigonal	222	4 in space groups without twofold axes, 2, 4 in space groups with twofold axes	Tetrahedral/distorted tetrahedral	35	59
			Twisted perpendicular	1	
			Planar 1	23	
Hexagonal	222	4 in space groups without twofold axes, 1, 2, 4 in space groups with twofold axes	Tetrahedral/distorted tetrahedral	33	52
			Twisted perpendicular	6	
			Planar 1	11	
			Planar 2	1	
			Unclassified	1	
Tetragonal	222 or fourfold	1, 2, 4	Tetrahedral/distorted tetrahedral	44	113
			Twisted perpendicular	4	
			Planar 1	23	
			Planar 2	42	
Cubic	222 or fourfold	1, 2, 4	Tetrahedral/distorted tetrahedral	4	9
			Planar 1	2	
			Planar 2	3	
Total					650

As observed in the case of molecules with tetrahedral shape, planar molecules which have their adjacent subunits facing in opposite directions also have their binding sites placed at maximum distance from each other, which reduces the steric hindrance between ligand molecules. Enzymes belonging to this category have an additional feature suitable for ligand binding. Since their binding sites are placed on two opposite sides, substrate approaching from any side encounters the active site and hence the binding efficiency is increased.

The arrangement of monomers involving a fourfold symmetry and hence a closed planar pattern seems to be favoured by membrane-bound proteins such as aquaporins, plastocyanins and potassium channel proteins or DNA-binding proteins such as RUVA. The reason could be that this type of arrangement gives polarity to the molecule as a consequence of which all the hydrophobic part of the molecule is buried in the membrane and the hydrophilic part remains exposed inside the channel.

3.7. Correlation of subunit arrangement with crystal system

As the homotetrameric protein structures were also grouped according to their space groups, the prevalence of each type of arrangement in particular space groups was evident. As expected, the largest number of molecules displayed a tetrahedral or distorted tetrahedral shape in almost all of the crystal systems. Other types of quaternary arrangements which display 222 symmetry, namely planar molecules with adjacent subunits facing in opposite directions, were also observed in almost all crystal systems. Molecules that are sandwiched dimers of dimers, with both dimers roughly perpendicular to each other, were mainly observed in orthorhombic space groups.

In triclinic and monoclinic space groups the asymmetric unit is always a tetramer or sometimes even two tetramers. In orthorhombic space groups monomeric or dimeric asymmetric units are also observed which are located at special positions and the crystallographic symmetry operations give rise to the functional tetramer. Occurrence of monomeric/dimeric asymmetric units is also common in space groups with two twofold axes belonging to tetragonal and hexagonal crystal systems, such as $P4_x22$ and $P6_x22$, where x denotes the screw axis. Trigonal space groups with a twofold axis, such as $P3_x21$ or $P3_x12$, may show the presence of a dimeric asymmetric unit placed at the special position. All these conditions are the result of the presence of 222 symmetry in the molecule.

A planar closed arrangement of subunits involving fourfold symmetry is relatively rare. This type of arrangement is mainly observed in tetragonal or cubic space groups, where in most cases the asymmetric unit is a monomer and the biological tetrameric molecule can be generated using the symmetry operations. Rarely, this kind of arrangement is seen in orthorhombic or monoclinic space groups, but when present the asymmetric unit will always be a tetramer.

Table 1 lists all crystal systems with the prevalent types of quaternary arrangements observed and the number of PDB structures belonging to each.

4. Conclusions

A study of unique homotetrameric protein structures reported in the PDB revealed four prevalent types of arrangement of subunits in the tetramer. Three of these four types showed point-group 222 symmetry in their subunits, while the fourth type exhibited fourfold symmetry. While the 222 symmetry is commonly observed in many proteins, including lectins and most enzymes, the fourfold symmetry is restricted to mostly membrane-bound proteins such as ion channels and certain membrane-bound enzymes. Some other unusual quaternary structures were also observed in this study which did not conform to any of the four groups. Such unusual arrangements may be correlated with the specific biological activity of the proteins concerned.

UVK thanks CSIR, New Delhi, India for a research fellowship. The authors thank Professor M. Vijayan for discussions and encouragement. Vidyut Kalwit and Avanti Kulkarni helped in arranging the data obtained from the PDB entries.

References

Ali, M. H. & Imperiali, B. (2005). *Bioorg. Med. Chem.* **13**, 5013–5020.
Argos, P. (1988). *Protein Eng.* **2**, 101–113.
Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). *J. Mol. Biol.* **259**, 281–296.
Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanaraj, V., Mahanta, S. K., Suguna, K., Surolia, A. & Vijayan, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.
Blaber, M., DiSalvo, J. & Thomas, K. A. (1996). *Biochemistry*, **35**, 2086–2094.
Bouckaert, J., Hamelryck, T. W., Wyns, L., Loris, R. (1999). *J. Biol. Chem.* **274**, 29188–29195.
Bray, D. & Lay, S. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 13493–13498.
Brenowitz, M., Mandal, N., Pickar, A., Jamison, E. & Adhya, S. (1991). *J. Biol. Chem.* **266**, 1281–1288.
Brinda, K. V., Mitra, N., Surolia, A. & Vishveshwara, S. (2004). *Protein Sci.* **13**, 1735–1749.
Brinda, K. V., Surolia, A. & Vishveshwara, S. (2005). *Biochem. J.* **391**, 1–15.
Chandra, N. R., Prabu, M. M., Suguna, K. & Vijayan, M. (2001). *Protein Eng.* **14**, 857–866.
Chothia, C. & Janin, J. (1975). *Nature (London)*, **256**, 705–708.
Datta, A. B., Panjikar, S., Weiss, M. S., Chakrabarti, P., Parrack, P. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 11242–11247.
Dekker, C., de Kruijff, B. & Gros, P. (2003). *J. Struct. Biol.* **144**, 313–319.
Del Sol, F. G., Cavada, B. S. & Calvete, J. J. (2007). *J. Struct. Biol.* **158**, 1–9.
DeLano, W. L. (2000). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Dessen, A., Gupta, D., Sabesan, S., Brewer, C. F. & Sacchettini, J. C. (1995). *Biochemistry*, **34**, 4933–4942.
Dey, S., Grant, G. A. & Sacchettini, J. C. (2005). *J. Biol. Chem.* **280**, 14892–14899.
Dokland, T. (2000). *Structure*, **8**, R157–R162.
Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). *Science*, **280**, 69–77.
Elgavish, S. & Shaanan, B. (2001). *Protein Sci.* **10**, 753–761.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Forouhar, F., Yang, Y., Kumar, D., Chen, Y. & Fridman, E. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 1773–1778.
Friedman, A. M., Fischmann, T. O. & Steitz, T. A. (1995). *Science*, **268**, 1721–1727.
Goodsell, D. S. (1991). *Trends Biochem. Sci.* **16**, 203–206.
Goodsell, D. S. & Olson, A. J. (2000). *Annu. Rev. Biophys. Biomol. Struct.* **29**, 105–153.
Hamelryck, T. W., Dao-Thi, M.-H., Poortmans, F., Chrispeels, M. J., Wyns, L. & Loris, R. (1996). *J. Biol. Chem.* **271**, 20479–20485.
Hamelryck, T. W., Moore, J. G., Chrispeels, M. J., Loris, R. & Wyns, L. (2000). *J. Mol. Biol.* **299**, 875–883.
Handa, N., Kukimoto-Niino, M., Akasaka, R., Kishishita, S., Murayama, K., Terada, T., Inoue, M., Kigawa, T., Kose, S., Imamoto, N., Tanaka, A., Hayashizaki, Y., Shirouzu, M. & Yokoyama, S. (2006). *J. Mol. Biol.* **363**, 114–124.
Janin, J., Milles, S. & Chothia, C. (1988). *J. Mol. Biol.* **204**, 155–164.
Jones, S. & Thornton, J. M. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
Kitaoka, S., Wada, K., Hasegawa, Y., Minami, Y., Fukuyama, K. & Takahashi, Y. (2006). *FEBS Lett.* **580**, 137–143.
Kitayner, M., Rozenberg, H., Kessier, N., Rabinovich, D., Shaulov, L., Haran, T. E. & Shakked, Z. (2006). *Mol. Cell*, **22**, 741–753.
Kumar, R. S., Brannigan, J. A., Prabhune, A. A., Pundle, A. V., Dodson, G. G., Dodson, E. J. & Suresh, C. G. (2006). *J. Biol. Chem.* **281**, 32516–32525.
Larsen, T. A., Olson, A. J. & Goodsell, D. S. (1997). *Structure*, **6**, 421–427.
Levy, E. D., Pereira-Leal, J. B., Chothia, C. & Teichmann, S. A. (2006). *PLoS Comput. Biol.* **2**, 1395–1406.
Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. & Lu, P. (1996). *Science*, **271**, 1247–1254.
Manting, E. H., van der Does, C., Remigy, H., Engel, A. & Driessen, A. J. (2000). *EMBO J.* **19**, 852–861.
Miller, S. (1989). *Protein Eng.* **3**, 77–83.
Miller, S., Lesk, A. M., Janin, J. & Chothia, C. (1987). *Nature (London)*, **328**, 834–836.
Monod, J., Wyman, J. & Changeux, J.-P. (1965). *J. Mol. Biol.* **12**, 88–118.
Nagano, C. S., Calvete, J. J., Barettino, D., Perez, A., Cavada, B. S. & Sanz, L. (2008). *Biochem. J.* **409**, 417–428.
Nooren, I. M. & Thornton, J. M. (2003*a*). *EMBO J.* **22**, 3486–3492.
Nooren, I. M. & Thornton, J. M. (2003*b*). *J. Mol. Biol.* **325**, 991–1018.
Oehler, S., Eismann, E. R., Krämer, H. & Müller-Hill, B. (1990). *EMBO J.* **9**, 973–979.
Phelps, D. K., Speelman, B. & Post, C. B. (2000). *Curr. Opin. Struct. Biol.* **10**, 170–173.
Powers, E. T. & Powers, D. L. (2003). *Biophys. J.* **85**, 3587–3599.
Prabu, M. M., Suguna, K. & Vijayan, M. (1999). *Proteins*, **35**, 58–69.
Rao, K. N., Suresh, C. G., Katre, U. V., Gaikwad, S. M. & Khan, M. I. (2004). *Acta Cryst.* D**60**, 1404–1412.
Rozwarski, D. A., Swami, B. M., Brewer, C. F. & Sacchettini, J. C. (1998). *J. Biol. Chem.* **273**, 32818–32825.
Sakaguchi, T., Tu, Q., Pinto, L. H. & Lamb, R. A. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5000–5005.
Schutz, A., Sandalova, T., Ricagno, S., Hubner, G., Konig, S. & Schneider, G. (2003). *Eur. J. Biochem.* **270**, 2312–2321.
Srinivas, V. R., Bhanuprakash Reddy, G., Ahmad, N., Swaminathan, C. P., Mitra, N. & Surolia, A. (2001). *Biochim. Biophys. Acta*, **1527**, 102–111.
Su, Y., Skinner, M. M., Xuong, N. H., Matthews, D. A., Whiteley, J. M. & Varughese, K. I. (1994). *Acta Cryst.* D**50**, 884–888.
Suresh, C. G., Pundle, A. V., SivaRaman, H., Rao, K. N., Brannigan, J. A., McVey, C. E., Verma, C. S., Dauter, Z., Dodson, E. J. & Dodson, G. G. (1999). *Nature Struct. Biol.* **6**, 414–416.

Vijayan, M. (2007). *J. Biosci.* **32**, 1059–1066.
Wah, D. A., Romero, A., del Sol, F. G., Cavada, B. S., Ramos, M. V., Grangeiro, T. B., Sampaio, A. H. & Calvete, J. J. (2001). *J. Mol. Biol.* **310**, 885–894.
Walden, H., Bell, G. S., Russell, R. J. M., Siebers, B., Hensel, R. & Taylor, G. L. (2001). *J. Mol. Biol.* **306**, 745–757.
Wlodawer, A. & Erickson, J. W. (1993). *Annu. Rev. Biochem.* **62**, 543–585.
Zaremba, M., Sasnauskas, G., Urbanke, C. & Siksnys, V. (2005). *J. Mol. Biol.* **348**, 459–478.
Zaremba, M., Sasnauskas, G., Urbanke, C. & Siksnys, V. (2006). *J. Mol. Biol.* **363**, 800–812.