Variability in Quaternary Association of Proteins With the Same Tertiary Fold: A Case Study and Rationalization Involving Legume Lectins

Moses M. Prabu, K. Suguna, and M. Vijayan*
Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

ABSTRACT **Legume lectins constitute a family of proteins in which small alterations arising from sequence variations in essentially the same tertiary structure lead to large changes in quaternary association. All of them are dimers or tetramers made up of dimers. Dimerization involves side-by-side or back-to-back association of the flat six-membered beta-sheets in the protomers. Variations within these modes of dimerization can be satisfactorily described in terms of angles defining the mutual disposition of the two subunits. In all tetrameric lectins, except peanut lectin, oligomerization involves the back-to-back association of side-by-side dimers. An attempt has been made to rationalize the observed modes of oligomerization in terms of hydrophobic surface area buried on association, interaction energy and shape complementarity, by constructing energy minimised models in each of which the subunit of one legume lectin is fitted in the quaternary structure of another. The results indicate that all the three indices favor and, thus, provide a rationale for the observed arrangements. However, the discrimination provided by buried hydrophobic surface area is marginal in a few instances. The same is true, to a lesser extent, about that provided by shape complementarity. The relative values of interaction energy turns out to be a still better discriminator than the other two indices. Variability in the quaternary association of homologous proteins is a widely observed phenomenon and the present study is relevant to the general problem of protein folding. Proteins 1999;35:58–69.**

Key words: legume lectins; quaternary association; buried surface area; shape complementarity; interaction energy

INTRODUCTION

Lectins are multi-valent carbohydrate binding proteins of non-immune origin.[1] In recent years, there has been a spurt in lectin research on account of their ability to specifically bind to cell surface carbohydrates and in view of the diverse applications that stem from this ability.[2–7] Although first isolated from plant sources, they are found in all classes and families of organisms. The richest sources of lectins are, however, the seeds of leguminous plants and legume lectins are among the most thoroughly studied families of proteins. Each legume lectin has a subunit made up of 237 to 253 amino acid residues. These lectins exhibit a high degree of homology, with a sequence identity ranging from 36 to 90% among those with known three-dimensional structure. The tertiary structures of all of them are the same except for some variations in the loops. Indeed, the legume lectin fold has now been recognized as an important carbohydrate-binding motif that occurs, with different degrees of variations, in other proteins as well.[8–11]

The mode of carbohydrate binding by legume lectins has been understood in terms of the sugar-combining site made up of four loops.[12,13] The diverse sugar specificities of these lectins have been explained, at least at the monosaccharide and the disaccharide levels, in terms of the variations in the length, composition, and interactions of one of these four loops.[14–24] However, sufficient attention has not been paid to the diversity of quaternary structures of legume lectins. The first few legume lectin structures to be solved[25–29] exhibited the "canonical" mode of dimerization involving the formation of a 12-stranded β-sheet, six strands belonging to each subunit. A different mode of dimerization was observed in *Griffonia simplicifolia* lectin IV (GSIV). The reason for this difference was attributed to interactions involving covalently bound carbohydrate[30] and those involving a buried glutamic acid residue.[19] Yet another mode of dimerization was observed in *Erythrina corallodendron* lectin (EcorL) and it was suggested that canonical dimerization is prevented on account of the glycosylation sites at the interface of such a dimer.[18] Peanut lectin (PNA), a tetramer the structure of which was solved in this laboratory,[22,31] is not glycosylated. Yet the main (dimeric) interface in it is non-canonical, suggesting thereby that the variability in quaternary association in legume lectins is not necessarily caused by interactions involving covalently linked sugar. The basic lectin from winged beans (WBAI) has 63% sequence identity with EcorL. However, unlike EcorL, the glycosylation sites in WBAI are far away from the region that would constitute the interface in the canonical dimer and, hence, do not

Grant sponsor: Department of Science and Technology, Government of India.
*Correspondence to: Prof. M. Vijayan, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. E-mail: mv@mbu.iisc.ernet.in
Received 2 October 1998; Accepted 8 December 1998

prevent the formation of such a dimer. Yet, WBAI forms EcorL-type dimers, demonstrating that the mode of dimerization is primarily dictated by factors intrinsic to the protein itself.[24] Thus the structures of PNA and WBAI conclusively show that legume lectins are an interesting family of proteins in which small alterations consequent to sequence variations, in essentially the same tertiary fold, lead to large changes in quaternary structure.

Having established that the observed variability in quaternary association in legume lectins results from factors intrinsic to the protein, it is important to precisely characterize this variability and to rationalize it. This paper is an attempt in that direction. The factors that are responsible for the oligomerization of proteins and the various elements that contribute to the stability of oligomers have been discussed extensively.[32–41] However, the predictive value of the suggestions contained therein is yet to be firmly established. In this context, a detailed theoretical study of oligomerization of legume lectins, an ideal family of proteins for such a study in view of the sensitivity of subunit association of its members to sequence variation, would hopefully help establish the relative importance of the various factors that have been suggested to be responsible for the stability of oligomers.

OLIGOMERIZATION IN LEGUME LECTINS

Legume lectins of known three-dimensional structure available in the Protein Data Bank (PDB)[42] are listed in Table I. As illustrated in Figure 1, the protomer of each is made up of a six-stranded nearly flat "back" β-sheet, a seven-stranded curved "front" β-sheet, a short five-membered β-sheet at the "top" of the molecule, and several loops that connect the sheets. All of them are dimers or tetramers believed to be made up of dimers. Each tetramer has mainly two types of interfaces. These interfaces have varying degrees of similarity, ranging from very close to broad, with those found in the dimeric proteins. All these interfaces involve the six-stranded back β-sheet of the monomer in one way or another and it is possible to describe each of them in terms of the mutual disposition of the β-sheets in the two participating subunits.

Dimerization in a majority of instances involves a side-by-side arrangement, resulting in a contiguous 12-stranded β-sheet with the dyad axis perpendicular to the β-sheet, as illustrated in Figure 2a. This kind of association was first observed in concanavalin A (Con A)[25,26,43] and may be described as an extended (II-type)[39] interface. Dimerization in other instances involves different kinds of back-to-back stacked (X-type) association of the six-stranded β-sheets. An example of one kind is shown in Figure 2b. Among the dimeric lectins Pea lectin (PSL),[27] Favin,[28] *Lathyrus ochrus* lectin I (LOLI),[29] and Lentil Lectin (LenL)[44] associate in a side-by-side fashion. GSIV exhibits one kind of back-to-back association while EcorL and WBAI exhibit another kind. In the four tetramers (Con A, PNA, Soybean agglutinin [SBA],[45] and Phytohemagglutinin-L [PHA-L],[46] illustrated in Fig. 3), subunits 1 and 2 and 3 and 4 (except in the case of PNA) associate in a side-by-side fashion. The 1–4 and the 2–3 interfaces are of one kind of back-to-back type or another. PNA presents a unique case of a tetramer without fourfold or 222 symmetry.[22,31] Consequently, the 1–2 and the 3–4 interfaces are not equivalent. It is believed that the 3–4 interface is an incidental consequence of the presence of 1–2 and 1–4 (and 2–3) interfaces.[22] Furthermore, although the 1–2 interface is of the side-by-side type, the two six-membered sheets do not form a contiguous 12-stranded β-sheet. The two sheets are connected through a number of interfacial water molecules.

Side-by-Side Association

Although the most dominant feature of the side-by-side association is the 12-stranded β-sheet encompassing both the subunits, several other hydrophobic and polar interactions are also involved in stabilizing the arrangement. In addition to the back β-sheet, the other secondary structural elements that take part in intersubunit interactions are (1) the loop that connects the first strand of the back β-sheet and the first strand of the curved front β-sheet, and (2) the loop between the second strand of the curved front β-sheet and the third strand of the back β-sheet.

The variations among the side-by-side dimers in different lectins are best described in terms of the twist angle (θ) and the propeller (ϕ) illustrated in Figure 4 and calculated as described in the Appendix. The values of these two angles in the concerned lectins are listed in Table II. Understandably, the values in PNA deviate substantially from those in other lectins. The deviations of these angles among the other lectins are small though not negligible. Interestingly, in all the dimers, the 12-stranded β-sheet presents a slightly concave surface. The concavity is clearly discernible in the views shown in Figure 5.

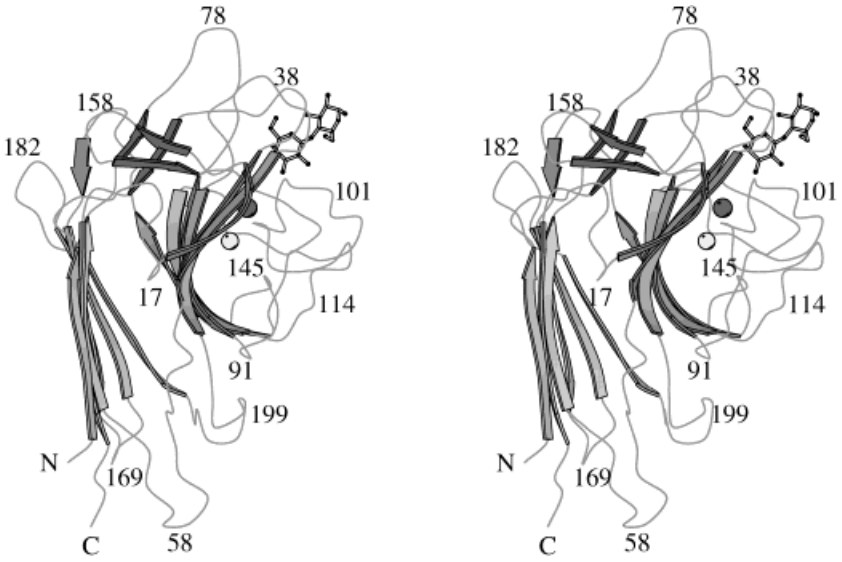
Furthermore, the propeller twist between the two individual β-sheets is quite substantial in all the cases.

The association in different dimers was compared in yet another way. An appropriate monomer (subunit 1) in each dimer was superposed on the corresponding monomer in the Con A dimer. The angles between the dyads of the different dimers were then calculated. The angle of inclination of the dyads of other lectin subunit pairs with respect to Con A are listed in Table II. Interestingly, in this orientation, the dyads of the different lectin dimers intersect at a common position.

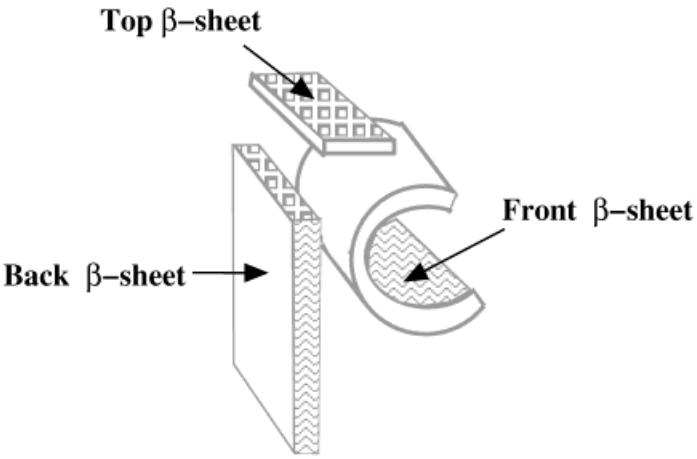
Back-to-Back Association

The back-to-back association found in EcorL, GSIV, and WBAI dimers and the 1–4 (2–3) interfaces in tetrameric Con A, PNA, SBA, and PHA-L, exhibit substantial variation. The dimeric association in GSIV and the 1–4 association in PNA are similar. So are the dimeric interfaces in EcorL and WBAI. SBA and PHA-L also have similar interfaces. On the other hand, the 1–4 interface in Con A is different from the other back-to-back interfaces. The four different kinds of back-to-back interfaces are illustrated in Figure 6.

In all the cases, the interface primarily involves the fourth, fifth, and sixth strands of the back β-sheet. However, in Con A, the loop that connects the first and the



(a)



(b)

Fig. 1. **a:** Stereo view of a typical legume lectin. The figure corresponds to a PNA subunit complexed with lactose. Stripled arrows represent the β-strands. This figure, Figures 2, 6, and 7 were generated using Molscript.[65] **b:** A schematic representation of a legume lectin subunit in terms of sheets. This and Figure 4 were generated using ShowCase™.

second strands of the back β-sheet, which is formed by a post-translational modification,[47–49] also participates in the association.

The angle (ψ) between the two sheets in the different interfaces are listed in Table III. As was done in the case of side-by-side association, the appropriate subunit of each of the other dimers or subunit pairs was superposed on the corresponding monomer in the WBAI dimer. The dyads of EcorL, WBAI, GSIV, and those of the 1–4 interfaces in PNA and Con A lie in the same plane and intersect nearly at the same location. The orientations of the other dyads with respect to that of WBAI are given in Table III. In fact, as far as the general main-chain features are concerned, one pair of subunits can be generated from another pair by a rotation of one subunit about an axis perpendicular to the plane of the dyads and passing through the point of their intersection. The rotation required is simply the difference in the angle between the pairs. The 1–4 interface in SBA and PHA-L is, however, of an entirely different kind. It cannot be obtained by a simple rotation of one subunit in an interface in another lectin.

Tetrameric Association

All the four tetrameric lectins whose structures are known (Fig. 3) involve side-by-side as well as back-to-back

TABLE I. Legume Lectins With Known Three-Dimensional Structure†

Lectin	Abbreviation	PDB code	Oligomeric state	Nature of interface(s)
Concanavalin A	Con A	2cna	4	Two II-type and two X-type
Pea Lectin	PSL	2ltn	2	II-type
Favin	Favin	NA	2	II-type
Lathyrus ochrus lectin I	LOLI	1loe	2	II-type
Griffonia simplicifolia lectin IV	GSIV	1lec	2	X-type
Erythrina corallodendron lectin	EcorL	1lte	2	X-type
Lentil lectin	LenL	1lem	2	II-type
Peanut agglutinin	PNA	2pel	4	One II-type, two X-type, and one unusual
Soybean agglutinin	SBA	1sba	4	Two II-type and two X-type
Phytohemagglutinin-L	PHA-L	1fat	4	Two II-type and two X-type
Winged bean agglutinin	WBAI	1wbl	2	X-type

†PDB*i* protein data bank. II-type, side-by-side interfaces; X-type, back-to-back interfaces.

interfaces. The disposition of the back β-sheets in the four structures is schematically illustrated in Figure 7.

The tetramers in Con A, SBA and PHA-L can be described as made up of the back-to-back association of two dimers, each resulting from the side-by-side association of two subunits. The (propeller) axis of the two dimers are inclined at 82.1° in Con A whereas the two axes are nearly parallel in SBA and PHA-L with interaxial angles of 10.0° and 10.4°, respectively. A rotation of 72.1° of one of the dimers about the common dyad of the two dimers, would bring the arrangement in Con A to that in SBA. The corresponding angle in the case of PHA-L is 71.7°.

The arrangement of subunits in PNA is substantially different. The molecule is best described as made up of two back-to-back dimers.[22,31] Each dimer is twofold symmetric. The two dimers in the tetramer are also related by a dyad. However, the different dyads do not intersect; nor are they mutually perpendicular. The molecule thus has an open quaternary arrangement. The interactions between the two dimers involve two different interfaces. One is a side-by-side interface mediated by a set of water bridges. The other interface is believed to be a fortuitous consequence of the presence of the two back-to-back and one side-by-side interfaces.[22,31]

MODELLING

Modelling essentially involved the fitting of a subunit of one legume lectin in the quaternary structure of another and refining the resulting model using molecular dynamics and energy minimization. Dimeric lectins exhibit three types of association. PSL was taken as the representative of side-by-side dimers. EcorL and WBAI form the same kind of back-to-back dimers. WBAI was chosen for the calculations. GSIV forms yet another type of back-to-back dimer. A WBAI subunit, for example, was superposed on each of the subunits in the PSL dimer. The resulting dimer may be called WBAI in PSL. Similarly, WBAI in GSIV, GSIV in WBAI and PSL, and PSL in WBAI and GSIV were constructed.

The short contacts in the models constructed by fitting a subunit of one legume lectin in another, were relieved to some extent using the ROTAMER-SEARCH module in the Biosym package INSIGHT-II©. A 5 Å water shell was generated around the dimers using the ASSEMBLY-SOAK in *Biosym*. The shell typically contained 2,000 water molecules. The hydrogen atoms were generated using the HBUILD option in X-PLOR.[50] The remaining bad contacts in the models were relieved using the REPEL function in X-PLOR. The models were subjected to conjugate gradient energy minimization for 100 cycles each. Non-crystallographic symmetry (NCS) and tethering restraints were imposed on the main-chain and the side-chain atoms. A dielectric constant of 1.0 was used throughout. The *cis* peptides in the structures were constrained at $\omega = 0$. Subsequently, simulated annealing protocol[51] present in the X-PLOR package was used for performing molecular dynamics in order to remove ambiguities about the preferences of side-chain and main-chain torsions among the available rotamers. The models were heated to 3,000 K and the molecular dynamics simulations were performed in steps of 25 K with each step containing 50 cycles spanning 5 femto seconds each. The main-chain and the side-chains were tethered to 30 and 5 cal/mol, respectively, while their corresponding NCS restraints were 50 and 20 cal/mol. At the end of the simulated annealing protocol, one more conjugate gradient minimization was undertaken. The minimizations were continued until the root mean square (r.m.s.) difference in energy difference in successive cycles was less than 0.1 kcal/mol. At convergence in each case, r.m.s. deviations from ideal values in bond lengths, bond angles, improper angles, and torsion angles were less than 0.004 Å, 1°, 1°, and 26°, respectively. For consistency, exactly the same calculations were carried out on the crystallographically obtained models as well.

Energy-minimized models of different tetramers were also obtained in a similar manner from dimers except that a 4 Å water shell was used when dealing with them. PSL dimers were superposed on the two dimers in Con A to produce a PSL in Con A tetramer. A PSL in SBA tetramer was similarly generated. Also constructed were a Con A in SBA tetramer and a SBA in Con A tetramer. A comparison involving PNA could not be carried out in terms of dimers.

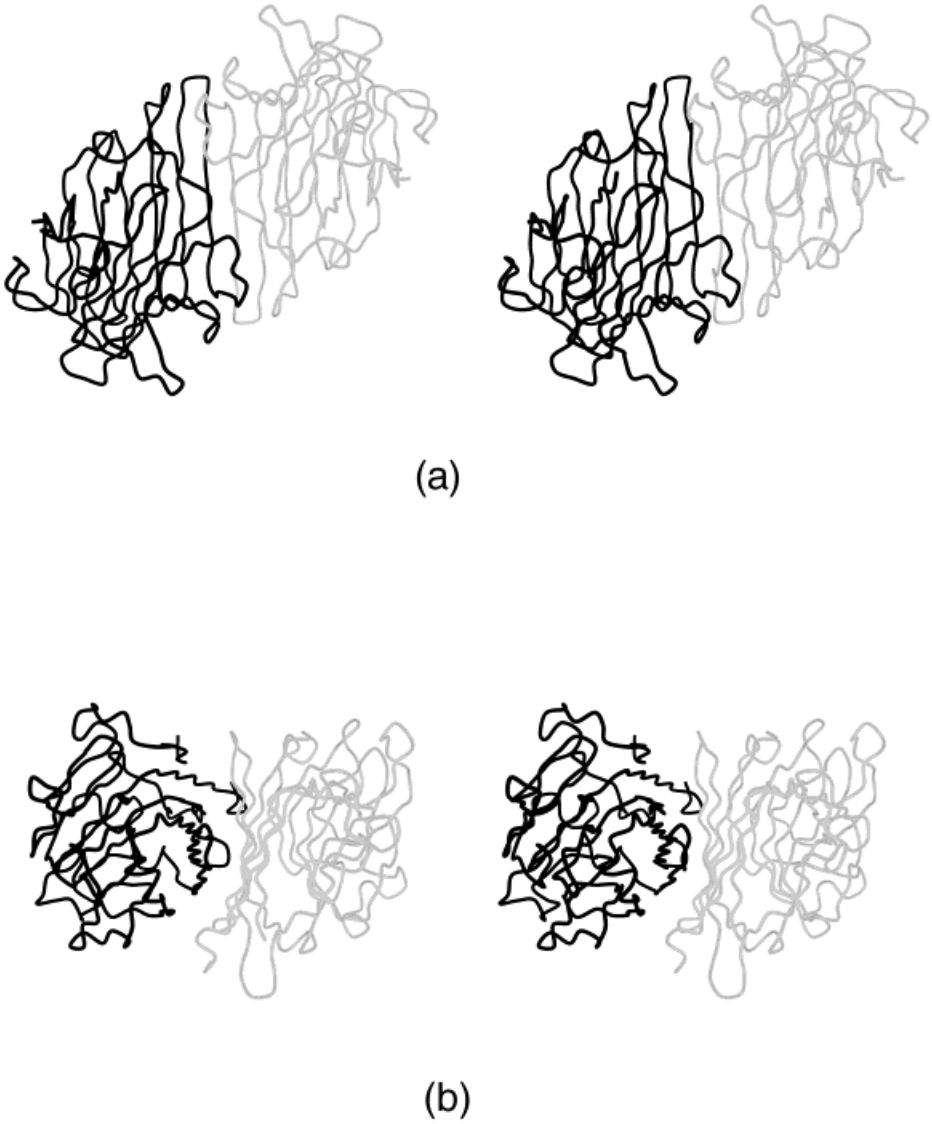


Fig. 2. Stereoviews of (**a**) the side-by-side dimer in Con A and (**b**) one kind of back-to-back dimer observed in GSIV. The two monomers are shaded differently.

Therefore, for this purpose, tetrameric models involving PNA, Con A, and SBA types of association were constructed using PNA, Con A, and SBA monomers. Again, exactly the same calculations were performed on crystallographically obtained tetrameric models as well.

Buried Hydrophobic Surface Area, Interaction Energy, and Shape Complementarity

The observed quaternary structures were sought to be understood in terms of not only hydrophobic surface area buried on dimerization or tetramerization and energy of interaction between subunits in different models, but also shape complementarity, which has been recently suggested to be important in stabilizing protein-protein interactions.[52–56] The focus here is not on the details of interactions at the interface(s) in each structure or model, as, for example, in some studies involving antigen-antibody interactions.[57] The effort here is to arrive at broad generalizations, if possible, and to assess the relative importance of the different factors suggested to be important for stability of oligomeric proteins.

The well-known Lee and Richards[58] algorithm was used to calculate accessible surface area buried on dimerization from monomers, tetramerization from dimers, and tetramerization from monomers. The X-PLOR package[50] was used for computing interaction energy employing distant dependent dielectric constant. The interaction energies between subunits in the constructed dimeric structures and between dimers in the tetrameric structures were computed appropriately. In comparisons involving PNA, interaction energies at each interface were computed individually and summed to yield a single value for each

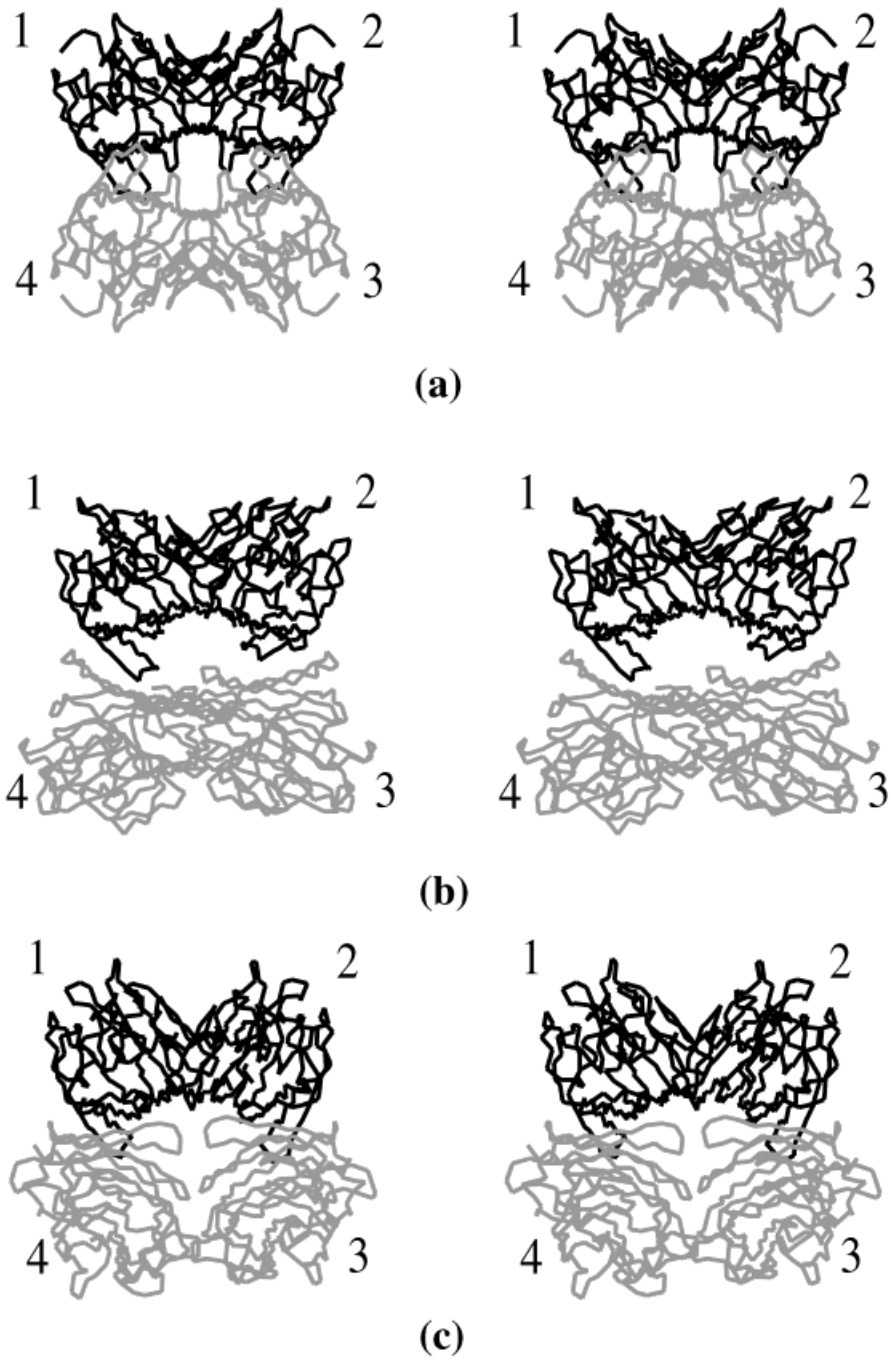


Fig. 3. Stereo views of the tetramers in (**a**) Con A, (**b**) SBA, and (**c**) PNA. The PHA-L tetramer is very similar to the SBA tetramer. This figure and Figure 5 were generated using the Biosym package INSIGHT-II©.

tetrameric structure. Shape complementarity was computed using the method of Lawrence and Colman[54] for the dimeric structures and tetrameric structures involving the dimers of Con A, SBA, and PSL. However, models involving the monomers of PNA, Con A, and SBA were dealt with separately. Shape complementarity can be determined between two surfaces only. More than one independent interface is present in these tetramers. Consequently, shape complementarity for each individual interface was computed separately. Histograms were generated using the range of shape complementarity values (refer to figs. 1 and 2 in ref. 54) against the number of product normals. This was done for all the interfaces and the number of product normals were summed within each range for the

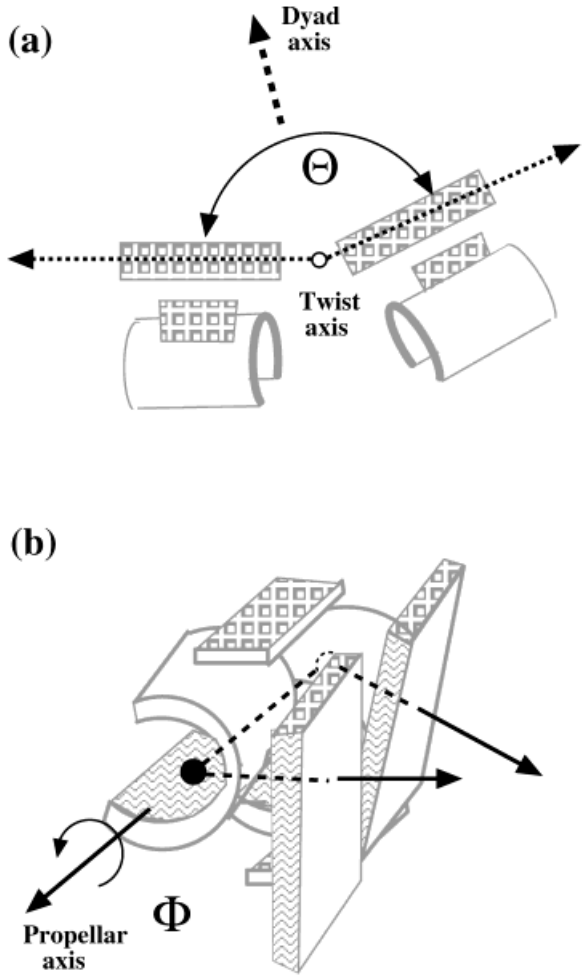


Fig. 4. Illustration of angles (**a**)θ and (**b**) ϕ that define the mutual orientations of the two subunits in a side-by-side dimer. Please refer to the Appendix for details.

TABLE II. Mutual Disposition of Subunits in the Side-by-Side Interfaces

	Twist (°)		Propeller (°)		Dyad
Protein	Θ	$\Delta\Theta_{ConA}$	Φ	$\Delta\Phi_{ConA}$	inclination
Con A	176.2	—	26.1	—	—
PNA	163.7	12.5	16.1	+10.0	−7.3
SBA	171.9	4.3	27.6	−1.5	+5.1
PHA-L	171.5	4.7	27.4	−1.3	+5.2
LenL	174.1	2.1	21.5	+4.6	+5.6
PSL	174.2	2.0	21.8	+4.3	+5.7
LOLI	174.4	1.8	21.6	+4.4	+5.9

whole tetramer. The median of the summed distribution was taken as the combined shape complementarity.

The buried surface area on dimerization, interaction energy, and shape complementarity in different dimer models are listed in Table IV. In the three sets of models, the hydrophobic surface area on dimerization is the highest in WBAI in WBAI, GSIV in GSIV, and PSL in PSL. However, in the first set, the area buried in WBAI in GSIV is very close to that in WBAI in WBAI. The observed models have substantially lower interaction energy than in the constructed models in each of the three cases.

PSL in PSL has higher shape complementarity than the other two PSL-based models. In the WBAI and GSIV based models also, the observed model has higher shape complementarity than the constructed models, but the discrimination is less emphatic. As indicated earlier, three types of tetramers are observed among legume lectins. Of these, that observed in PNA is of a special type and needs to be treated separately. The other two types involve the association of two side-by-side dimers. In Con A, the two dimers associate back-to-back in one way and they do in another way in SBA and PHA-L. Many lectins that form side-by-side dimers, as exemplified by PSL, do not associate into tetramers at all. The hydrophobic surface areas buried when Con A side-by-side dimers associate into a Con A type tetramer and a SBA type tetramer and those buried when SBA dimers associate into the two types of tetramers, are listed in Table V. Also given in Table V are the hydrophobic surface areas buried when PSL side-by-side dimers are used to construct Con A and SBA type tetramers. The interaction energy and the shape complementarity in the different models are also given in Table V. In consonance with the observation, buried hydrophobic surface area, interaction energy, and shape complementarity show that the arrangement of Con A dimers in a Con A tetramer is more stable than Con A dimers in a SBA tetramer and that of SBA dimers in a SBA tetramer is preferred over SBA dimers in a Con A tetramer. In both the sets, the interaction energies are decisively favourable to the observed arrangement. The shape complementarity in Con A dimers in Con A tetramer is only slightly better than that in Con A dimers in SBA tetramer, while the buried hydrophobic surface area is higher by about 1,500 Å^2 in the former than in the latter. On the other hand, the buried hydrophobic surface area in SBA dimers in SBA tetramer is only about 100 Å^2 larger than that in SBA dimers in Con A tetramer, while the shape complementarity is decisively in favour of the observed arrangement. The tetramers involving PSL dimers were constructed to seek a rationalization of the observed inability of these dimers to form tetramers. The buried surface areas and interaction energies in the models involving PSL are substantially lower than those in the favored arrangements involving Con A and SBA dimers. The shape complementarity in PSL dimers in Con A tetramer is close to, but lower than, that in Con A dimers in Con A tetramer.

The PNA tetramer consists of two back-to-back dimers in each of which the two back β-sheets are perpendicular to each other. One of the interfaces produced when the two dimers associate is similar to that in a side-by-side dimer except that the two sheets do not join together into a 12-stranded sheet; instead they are connected by a set of water bridges. Attempts to bring the two sheets closer to form a contiguous 12-stranded sheet resulted in unacceptable steric interactions. Likewise, rotation of the two side-by-side dimers in Con A and SBA in such a way as to create back-to-back interfaces of the type observed in PNA

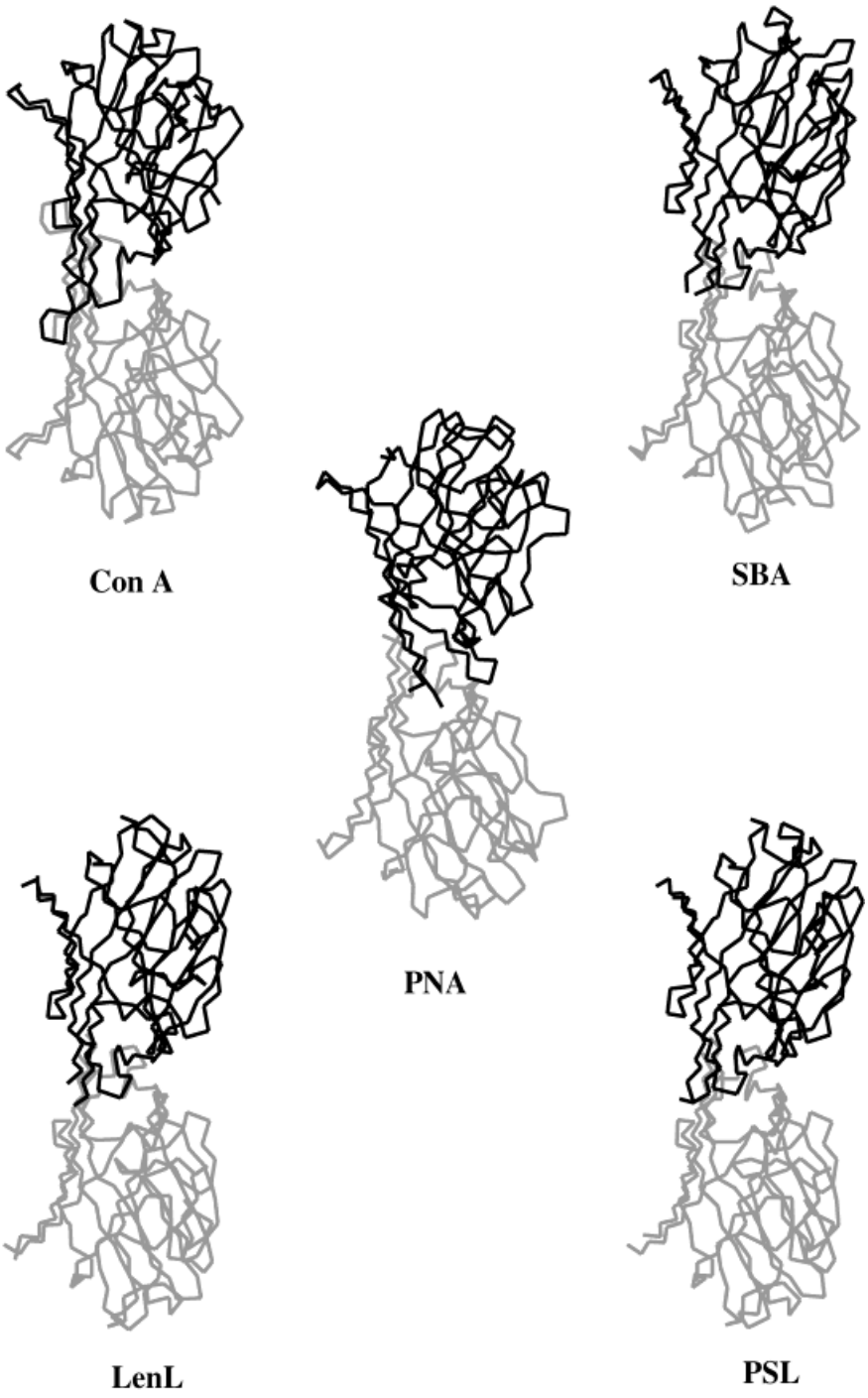


Fig. 5. C^{α} traces of different side-by-side legume lectin dimers.

also led to unacceptable steric interactions. Thus, side-by-side interface and interfaces of the type observed in PNA are mutually exclusive.

The question still remains as to why PNA subunits associate into an unusual tetramer instead of forming Con A type or SBA type tetramers. In an attempt to answer this question, the three type of tetramers constructed using the subunits of PNA, Con A and SBA were examined. The surface area buried on oligomerization in the nine energy minimized models along with the overall interaction energy and shape complementarity are listed in Table VI. Again, the three indices favour the observed arrangements, although in some instances only marginally. The SBA type tetramers involving Con A and PNA monomers

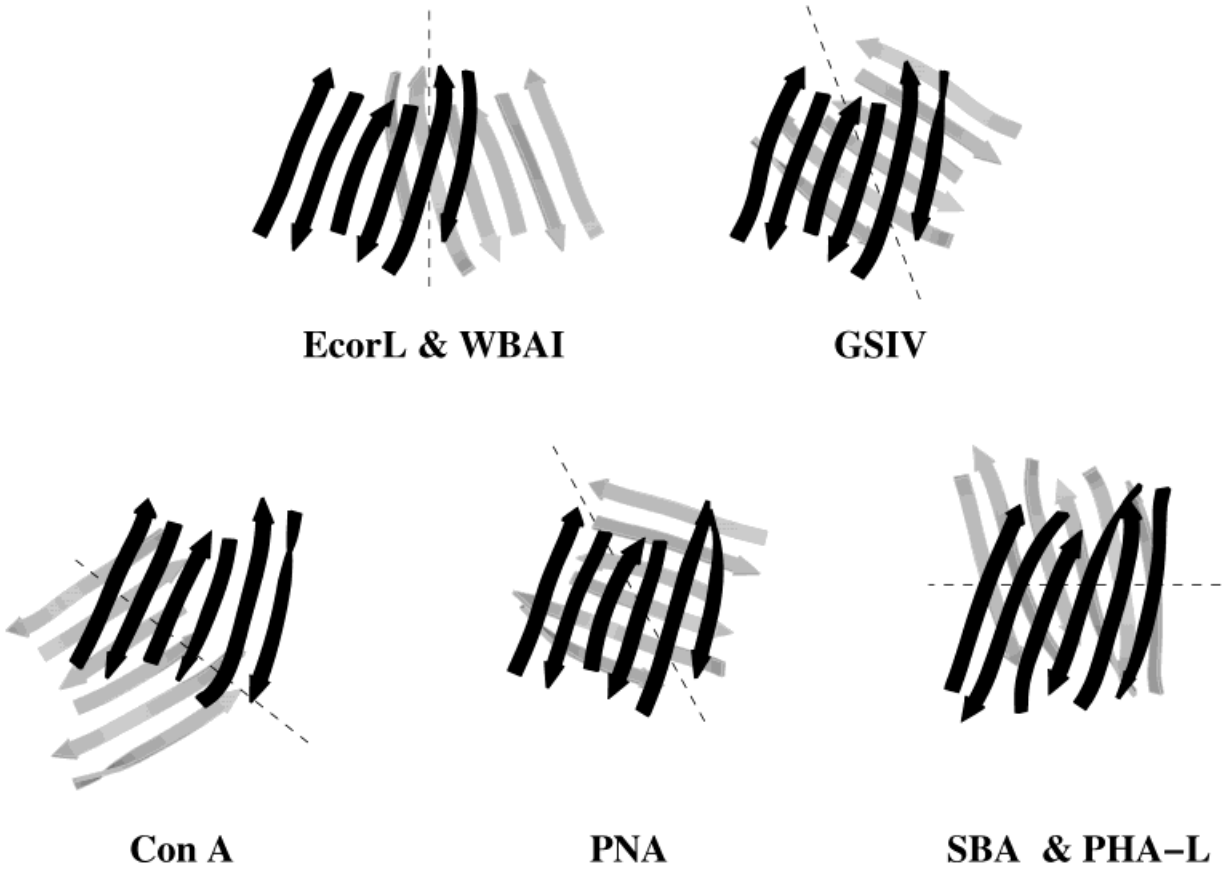


Fig. 6. Schematic representation of back β-sheets in the back-to-back dimers. One subunit has a common orientation in all the dimers. The dyads are indicated by broken lines.

TABLE III. Mutual Disposition of Subunits in the Back-to-Back Interfaces

Protein	Rotation (°) Ψ	Rotation (°) ΔΨWBAI	Dyad inclination
WBAI	49.4	—	—
EcorL	47.6	1.8	−0.9
GSIV	78.1	−28.7	19.5
PNA	96.1	−46.7	26.2
Con A	143.7	−94.3	53.7
SBA	−45.1	—	—
PHA-L	−45.7	—	—

are clearly not favored. The same appears to be the case with PNA or Con A type tetramers involving SBA monomers, although discrimination on the basis of shape complementarity is marginal. The situation is less clear in arrangements involving only PNA and Con A. The relevant differences in buried surface area are small. The same is true about the differences in interaction energy between PNA in PNA and PNA in Con A, and the differences in shape complementarity between Con A in PNA and Con A in Con A. However, it is interesting that all the indices favor, substantially or marginally, the observed arrangements.

DISCUSSION

An attempt has been made to systematically describe and to rationalize the observed patterns of quaternary association in legume lectins in terms of hydrophobic surface area buried on oligomerization, interaction energy, and shape complementarity. This attempt is also of considerable general interest in terms of identifying the important determinants of quaternary association. In as much as all the legume lectins have essentially the same tertiary structure, they constitute an ideal system to study the effect of small alterations in the tertiary structure arising from variations in the sequence. As the alterations are only in the details of the structure, the system could be used to check the sensitivity of the major factors that have been proposed to be responsible for quaternary association and, indeed, protein-protein interactions.

The results of the different calculations indicate that the observed variability in the quaternary association in legume lectins can be described in terms of simple geometrical parameters and explained on the basis of the indices mentioned above. All the three indices favor the observed arrangements. However, the discrimination provided by buried hydrophobic surface area is marginal in a few instances. It is true to a lesser extent with shape complementarity. Perhaps surprisingly, the discrimination pro-

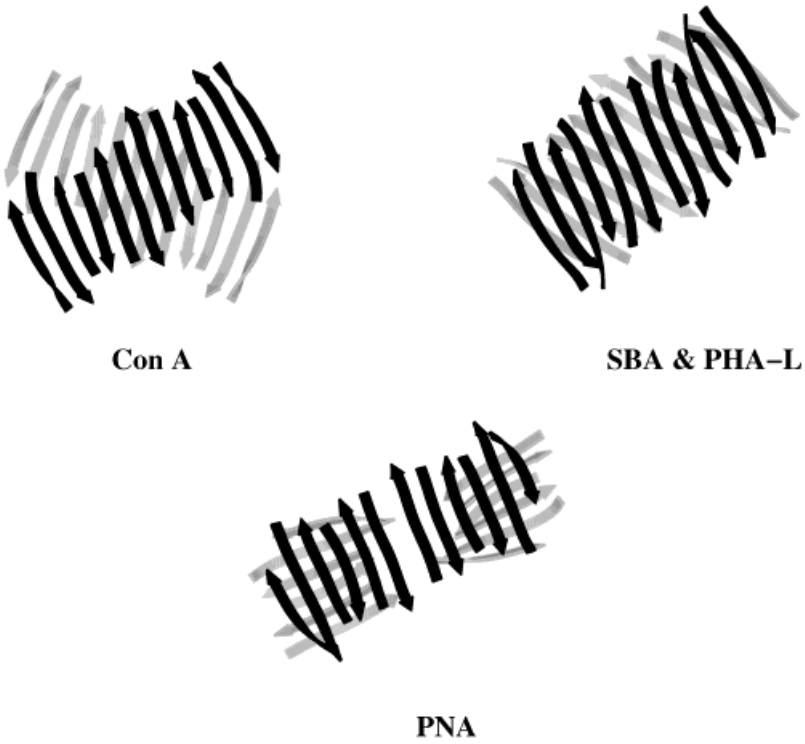


Fig. 7. Back β-sheets of Con A, SBA (PHA-L), and PNA. Dimers that form side-by-side interface are shown in black while subunits 3 and 4 are in grey.

TABLE IV. Buried Hydrophobic Surface Area (A in Å²), Interaction Energies (E in kcal/mol), and Shape Complementarity (S) in the Different Dimeric Models

Protein model	Interface type		
	WBAI	GSIV	PSL
WBAI			
A	1,271	1,253	899
E	−72	−47	−42
S	0.708	0.395	0.631
GSIV			
A	799	1,471	817
E	−29	−110	−66
S	0.536	0.623	0.494
PSL			
A	1,123	922	1,339
E	−21	−61	−99
S	0.352	0.379	0.552

TABLE V. Buried Hydrophobic Surface Area (A in Å²), Interaction Energies (E in kcal/mol), and Shape Complementarity (S) in the Tetrameric Models Involving Con A, SBA, and PSL Dimers

Protein model	Type of interface	
	Con A	SBA
Con A		
A	8,655	7,201
E	−255	−59
S	0.593	0.527
SBA		
A	7,729	7,825
E	−26	−101
S	0.545	0.745
PSL		
A	6,294	6,091
E	−13	−55
S	0.560	0.477

vided by interaction energy appears to be unambiguous in all except one case. The energy functions and, therefore, the absolute values of the interaction energies, are recognized to be very approximate.[59] Yet, the results presented here demonstrate that they could be extremely valuable, and indeed reliable, when comparing related systems.

There are other protein families as well in which substantial variability in quaternary structure results from small alterations in tertiary structure, although structural data pertaining to such variability are most extensive among legume lectins. Galectins provide another recent example of such variability.[60] The different modes of dimerization of alpha and beta chemokines, despite their having similar tertiary and related primary structures, have been explored earlier on the basis of hydrophopbic surface clusters.[61] Cystine-knot growth factors exhibit variability in quaternary association resulting from larger differences in primary and tertiary structures.[62] In a recent example, difference in quaternary association in highly homologous bulb (garlic and snowdrop[63]) lectins has been implicated in the generation of different sugar specificities.[64] Thus, the present study deals with a widely observed phonomenon

TABLE VI. Buried Hydrophobic Surface Area (A in Å^2), Interaction Energies (E in kcal/mol), and Shape Complementarity (S) in the Tetrameric Models Involving PNA, Con A, and SBA Monomers

Protein model	Type of interface		
	PNA	Con A	SBA
PNA			
A	9,495	9,277	7,956
E	−229	−222	−180
S	0.445	0.281	0.221
Con A			
A	11,207	11,230	10,444
E	−325	−404	−226
S	0.505	0.513	0.420
SBA			
A	10,364	11,163	11,874
E	−152	−188	−230
S	0.426	0.542	0.560

and is relevant to the general problem of protein folding as the large variability in quaternary association is ultimately determined by variations in the sequence.

ACKNOWLEDGMENTS

We thank Michael C. Lawrence for providing the program to compute shape complementarity and A. Surolia and Nagasuma R. Chandra for useful suggestions. Computations were done at the Super Computer Education and Research Centre and the Interactive Graphics Based Molecular Modelling Facility supported by the Department of Biotechnology, India.

APPENDIX

Calculation of ϕ and θ

Mean planes were constructed using the C^α positions of strands 1 to 4 from the back β-sheets in each subunit. A perpendicular line was dropped to the plane from the centroid of each monomer determined using the C^α positions. Cross products of the vector corresponding to this line and those of the least squares lines through the C^α positions in the four strands, were then computed. The average of the four resulting vectors would then be normal to a plane perpendicular to the plane of the β-sheet as well as the strands comprising it. The normal to the mean plane of the sheet in subunit 2 was then projected on this plane in subunit 1. The angle between this projected line and the plane normal of the sheet in subunit 1 is defined as θ. The propeller angle ϕ is the torsion angle of the two plane normals about the line joining the two centroids.

REFERENCES

1. Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N. What should be called a lectin? Nature 1980;285:66.
2. Barondes SH. Lectins: their multiple endogenous cellular functions. Annu Rev Biochem 1981;50:207–231.
3. Sharon N, Lis H. Lectins as cell recognition molecules. Science 1989;246:227–234.
4. Sharon N. Lectin-carbohydrate complexes of plants and animals: an atomic view. Trends Biochem Sci 1993;18:221–226.
5. Peumans WJ, van Damme EJM. Lectins as plant defense proteins. Plant Physiol 1995;109:347–352.
6. Drickamer K. Multiplicity of lectin-carbohydrate interactions. Nature Struct Biol 1995;2:437–439.
7. Drickamer K. Making a fitting choice: common aspects of sugar-binding sites in plant and animal lectins. Structure 1997;5:465–468.
8. Liao D-I, Kapadia G, Ahmed H, Vasta GR, Herzberg O. Structure of S-Lectin, a developmentally-regulated vertebrate beta-galactoside binding protein. Proc Natl Acad Sci USA 1994;91:1428–1432.
9. Crennel S, Garman E, Laver G, Vimr E, Taylor G. Crystal structure of Vibrio cholerae neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. Structure 1994;2:535–544.
10. Lobsanov YD, Gitt MA, Leffler H, Barondes SH, Rini JM. X-ray crystal structure of the human dimeric S-lac lectin, L-14-II, in complex with lactose at 2.9 Å resolution. J Biol Chem 1993;268: 27034–27038.
11. Emsley J, White HE, O'Hara BP, Oliva G, Srinivasan N, Tickle IJ, Blundell TL, Pepys MB, Wood SP. Structure of pentameric human serum amyloid P component. Nature 1994;367:338–345.
12. Young NM, Oomen RP. Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate-binding sites. J Mol Biol 1992;228:924–934.
13. Sharma V, Surolia A. Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. J Mol Biol 1997;267:433–445.
14. Derewenda Z, Yariv J, Helliwell JR, Kalb (Gilboa) AJ, Dodson EJ, Papiz MZ, Wan T, Campbell J. The structure of the saccharide-binding site of concanavalin A. EMBO J 1989;8:2189–2193.
15. Bourne Y, Roussel A, Frey M, Rouge P, Fontecilla-Camps J-C, Cambillau C. Three-dimensional structures of complexes of Lathyrus ochrus isolectin I with glucose and mannose. Proteins: Struct Funct Genet 1990;8:365–376.
16. Bourne Y, Rouge P, Cambillau C. X-ray structure of a (alpha-Man(1–3)beta-Man(1–4)GlcNAc)-lectin complex at 2.1 Å resolution. J Biol Chem 1990;265:18161–18165.
17. Bourne Y, Rouge P, Cambillau C. X-ray structure of biantennary octasaccharide-lectin complex refined at 2.3 Å resolution. J Biol Chem 1992;267:197–203.
18. Shaanan B, Lis H, Sharon N. Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. Science 1991;254:862–866.
19. Delbaere LTJ, Vandonselaar M, Prasad L, Quail JW, Wilson KS, Dauter Z. Structure of the lectin IV of Griffonia simplicifolia and its complex with the Lewis b human blood group determinant at 2.0 Å resolution. J Mol Biol 1993;230:950–965.
20. Rini JM, Hardman KD, Einspahr H, Suddath FL, Carver JP. X-ray crystal structure of a pea lectin-trimannoside complex at 2.6 Å resolution. J Biol Chem1993; 268:10126–10132.
21. Naismith JH, Emmerich C, Habash J, Harrop S, Helliwell JR, Hunter WN, Raftery J, Kalb (Gilboa) AJ, Yariv J. Refined structure of concanavalin A complexed with methyl-alpha- D-mannopyranoside at 2.0 Å resolution and comparison with the saccharide-free state. Acta Cryst 1994;D50:847–858.
22. Banerjee R, Das K, Ravishankar R, Suguna K, Surolia A, Vijayan M. Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex. J Mol Biol 1996;259:281–296.
23. Ravishankar R, Ravindran M, Suguna K, Surolia A, Vijayan M. Crystal structure of peanut lectin—T-antigen complex. Carbohydrate specificity generated by water bridges. Curr Sci 1997;72:855–861.
24. Prabu MM, Sankaranarayanan R, Puri KD, Sharma V, Surolia A, Vijayan M, Suguna K. Carbohydrate specificity and quaternary association in basic winged bean lectin. X-ray analysis of the lectin at 2.5 Å resolution. J Mol Biol 1998;276:787–796.
25. Hardman KD, Ainsworth CF. Structure of concanavalin A at 2.4 Å resolution. Biochemistry 1972;11:4910–4919.
26. Becker JW, Reeke GN, Wang JL, Cunningham BA, Edelman GM. The covalent and three-dimensional structure of concanavalin A. J Biol Chem 1975;250:1513–1524.
27. Einspahr H, Parks EH, Suguna K, Subramanian E, Suddath FL. The crystal structure of pea lectin at 3.0 Å resolution. J Biol Chem 1986;261:16518–16527.

28. Reeke GN Jr, Becker JW. Three-dimensional structure of favin: Saccharide binding-cyclic permutation in leguminous lectins. Science 1986;234:1108–1111.
29. Bourne Y, Abergel C, Cambillau C, Frey M, Rouge P, Fontecilla-Camps J-C. X-ray crystal structure determination and refinement at 1.9 Å resolution of isolectin I from the seeds of *Lathyrus ochrus*. J Mol Biol 1990;214:571–584.
30. Delbaere LTJ, Vandonselaar M, Prasad L, Quail JW, Pearlstone JW, Carpenter MR, Smillie LB, Nikrad PV, Spohr U, Lemieux RU. Molecular recognition of a human blood group determinant by a plant lectin. Can J Chem 1990;68:1116–1121.
31. Banerjee R, Mande SC, Ganesh V, Das K, Dhanaraj V, Mahanta SK, Suguna K, Surolia A, Vijayan M. Crystal structure of peanut lectin, a protein with an unusual quaternary structure. Proc Natl Acad Sci USA 1994;91:227–231.
32. Cornish-Bowden AJ, Koshland DE Jr. The quaternary structure of proteins composed of identical subunits. J Biol Chem 1971;246: 3092–3102.
33. Chothia C, Janin J. Principles of protein-protein recognition. Nature (London) 1975;256:705–708.
34. Janin J, Wodak SJ. Structural domains in proteins and their role in the dynamics of protein function. Prog Biophys Mol Biol 1983;42:21–78.
35. Janin J, Miller S, Chothia C. Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol 1988;204:155–164.
36. Argos P. An investigation of protein subunit and domain interfaces. Protein Eng 1988;2:101–113.
37. Korn AP, Burnett RM. Distribution and complementarity of hydropathy in multisubunit proteins. Proteins: Struct Funct Genet 1991;9:37–55.
38. Thornton JM, MacArthur MW, McDonald IK, Jones DT, Mitchell JBO, Nandi CL, Price SL, Zvelebil MJJ. Protein structures and complexes: what they reveal about the interactions that stabilize them. Phil Trans R Soc Lond 1993;345:113–129.
39. Jones S, Thornton JM. Protein-protein interactions: a review of protein dimer structures. Prog Biophys Mol Biol 1995;63:31–65.
40. Tsai C-J, Lin SL, Wolfson HJ, Nussinov R. Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect. Protein Sci 1997;6:53–64.
41. Larsen TA, Olson AJ, Goodsell DS. Morphology of protein-protein interfaces. Structure 1998;6:421–427.
42. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The protein data bank: A computer based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
43. Hardman KD, Agarwal RC, Frieser MJ. Manganese and calcium binding sites of concanavalin A. J Mol Biol 1982;157:69–86.
44. Loris R, Steyaert J, Maes D, Lisgarten J, Pickersgill R, Wyns L. Crystal structure determination and refinement at 2.3 Å resolution of the lentil lectin. Biochemistry 1993;32:8772–8781.
45. Dessen A, Gupta D, Sabesan S, Brewer CF, Sacchettini JC. X-ray crystal structure of soybean agglutinin cross-linked with a biantennary analog of blood group I carbohydrate antigen. Biochemistry 1995;34:4933–4942.
46. Hamelryck TW, Dao-Thi M, Poortmans F, Chrispeels MJ, Wyns L, Loris R. The crystallographic structure of phytohemagglutinin-L. J Biol Chem 1996;271:20479–20485.
47. Carrington DM, Auffret A, Hanke DE. Polypeptide ligation occurs during post-translational modification of concanavalin A. Nature 1985;313:64–67.
48. Chrispeels MJ, Hartl P M, Sturm A, Faye L. Characterization of endoplasmic reticulum-associated precursor of concanavalin A. Partial amino acid sequence and lectin activity. J Biol Chem 1986;261:10021–10024.
49. Bowles DJ, Marcus SE, Pappin DJ, Findlay JB, Eliopoulos E, Maycox PR, Burgess J. Post-translational processing of concanavalin A precursors in jackbean cotyledons. J Cell Biol 1986;102:1284–1297.
50. Brunger AT. X-PLOR manual version 3.1.New Haven: Yale University. 1992.
51. Brunger AT. Simulated annealing in crystallography. Annu Rev Phys Chem 1991;42:197–223.
52. Schoichet BK, Kuntz ID. Protein docking and complementarity. J Mol Biol 1991;221:327–346.
53. Walls PH, Sternberg JE. New algorithm to model protein-protein recognition based on surface complementarity. J Mol Biol 1992;228: 277–297.
54. Lawrence MC, Colman PM. Shape complementarity at protein-protein interfaces. J Mol Biol 1993;234:946–950.
55. Helmer-Citterich M, Tramontano A. PUZZLE: A new method for automated protein docking based on surface shape complementarity. J Mol Biol 1994;235:1021–1031.
56. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 1997;272:106–120.
57. Padlan EA. X-ray crystallography of antibodies. Adv Prot Chem 1996;49:57–133.
58. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 1971;55:379–400.
59. Adman ET, Mather MW, Fee JA. Molecular modeling studies on the proposed NaCl-induced dimerization of Chromatium vinosum high-potential iron protein. Chem Biophys Acta 1993;1142:93–98.
60. Seetharaman J, Kanigsberg A, Slaaby R, Leffler H, Barondes SH, Rini JM. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1 Å resolution. J Biol Chem 1998;273:13047–13052.
61. Covell DG, Smythers GW, Gronenborn AM, Clore GM. Analysis of hydrophobicity in the alpha and beta chemokine families and its relevance to dimerization. Protein Sci 1994;3:2064–2072.
62. Sun PD, Davies DR. The cystine-knot growth-factor superfamily. Annu Rev Biophys Biomol Struct 1995;24:269–291.
63. Hester G, Wright CS. The mannose-specific bulb lectin from Galanthus nivalis (snowdrop) binds mono- and dimannosides at distinct sites. Structure analysis of refined complexes at 2.3 Å and 3.0 Å resolution. J Mol Biol 1996;262:516–531.
64. Chandra NR, Ramachandraiah G , Bachhawat K, Dam TK, Surolia A, Vijayan M. Crystal structure of a dimeric mannose-specific agglutinin from garlic: quaternary association and carbohydrate specificity. J Mol Biol 1999;285:1157–1168.
65. Kraulis PJ. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Cryst 1991;24:946–950.