RECOGNITION AND INTERACTIONS CONTROLLING THE ASSEMBLIES OF β BARREL DOMAINS

E. D. GETZOFF, J. A. TAINER, AND A. J. OLSON
Research Institute of Scripps Clinic, Department of Molecular Biology, La Jolla, California 92037

ABSTRACT We present a qualitative computer graphics approach to the characterization of forces important to the assembly of β domains that should have general utility for examining protein interactions and assembly. In our approach, the nature of the molecular surface buried by the domain contacts, the specificity of the residue-to-residue interactions, and the identity of electrostatic, hydrophobic, and hydrophilic interactions are elucidated. These techniques are applied to the β barrel domains of Cu, Zn superoxide dismutase (SOD), immunoglobulin Fab, and tomato bushy stunt virus coat protein (TBSV), a plant viral capsid protein. By looking at a set of proteins having different numbers of interacting β domains, we have been able to see some of the variety and also some of the patterns common to these assembled domains. Strong β domain interactions (identified by their biochemical integrity) are apparently due to chemical, electrostatic, and shape complementarity of the molecular surfaces buried from interaction with solvent molecules. Although the amount of hydrophobic buried surface area appears to correlate with the strength of the interaction, electrostatic forces appear to be important in both stabilizing and destabilizing specific contacts. In TBSV, analysis of electrostatic interactions may help explain mechanisms of subunit accommodation to different environments, particle expansion, and pathways of assembly. The possible molecular basis for observed differences in the stability and flexibility of the domain complexes is discussed.

INTRODUCTION

We are developing methodology to examine different stages involved in macromolecular assembly, including precollision orientation by electrostatic forces; local recognition involving shape and chemical complementarity between specific interacting residues; and induced fit accomplished by local rearrangement of the flexible regions of one or both molecules. Computer graphic analyses are used to examine and document modeling of macromolecular associations too prolonged in time for analysis by current dynamics calculations. To evaluate the role of electrostatic forces, we calculate the electrostatic field surrounding each domain, the electrostatic potential on the buried molecular surfaces (Getzoff et al., 1983), and the electrostatic stabilization energies of each interface in the assembled complex. Shape and chemical complementarity are examined by using molecular surface calculations (Connolly, 1983*a*, *b*) together with the atomic geometry.

Where possible, aspects of flexibility are also considered. For example, the existence of both conserved and variable packing contacts among the 180 subunits of tomato bushy stunt virus (Harrison et al., 1978) and southern bean mosaic virus (Abad-Zapatero et al., 1980) implies a degree of flexibility in assembly. Divalent metal cations bound to carboxylic acid side chains by ionic interactions are important in stabilizing intersubunit contacts in the protein shell; removal of these ions, coupled with increased pH, leads to electrostatic repulsion between the now negatively charged acids and transformation of the virus structure to an expanded state that is also well characterized (Robinson and Harrison, 1982). One of our specific interests is in defining the possible roles of electrostatic forces in transformations between the compact and expanded states of the icosahedral viruses.

The purpose of this exploration is one of pattern discovery—the identification and classification of characteristic patterns for a variety of protein interfaces—with the goal of gaining the ability to recognize similar patterns in a number of protein systems. Although our analyses are designed for general application to many systems and to suggest specific, experimentally answerable questions, our current focus is the interactions and assembly of β barrel domains including Cu, Zn superoxide dismutase, the immunoglobulins, and the plant viral capsids. The techniques described below can give information about the packing interactions in far more detail than we can reasonably include in a short paper, especially for complex systems such as TBSV. Here we present some details from a simpler system, the SOD dimer contact, and summarize our results for the immunoglobulins and TBSV.

METHODS

Calculation and Analysis of Buried Molecular Surfaces

Molecular surfaces buried from solvent accessibility by interdomain or intersubunit contacts were calculated using the program MS (Connolly, 1983*a*, *b*). The molecular surface calculation mathematically rolls a

water-sized probe sphere (1.4 Å radius) representing a solvent water molecule over the van der Waals surface of the protein. Molecular surface calculations used individual van der Waals radii (McCammon et al., 1979; Bondi, 1964) including either explicit (for SOD) or implicit (for all other molecules) hydrogen atoms. Surface points of one domain blocked from solvent accessibility by the second domain of an interdomain interface are defined as "buried". Holes and discontinuities in buried surfaces indicate gaps in the interface that are large enough to accommodate a water molecule.

Computer graphics analysis of the buried molecular surfaces and their underlying stereochemistry was done using the interactive graphics language GRAMPS (O'Donnell and Olson, 1981) and the molecular modeling program GRANNY (Connolly and Olson, 1985). Two surfaces were considered to have global shape complementarity when their curvature and twist matched. Assessment of local shape complementarity was based on the size and number of gaps between the two interacting surfaces (Tainer et al., 1985). The degree of chemical complementarity was judged using color computer graphics to determine the atom types contributing to the interacting buried surfaces. Chemical complementarity is thus assessed from the interaction of appropriate hydrogen bonding atom pairs and salt links as well as from the match of hydrophobic-to-hydrophobic and hydrophilic-to-hydrophilic surfaces.

Assignment of Partial Charges

Using the atomic positions from the Brookhaven Protein Databank (Bernstein et al., 1977), we examined the role of electrostatic forces in the assembly of β barrel domains. The electrostatic potential was calculated for the arrangement of partial charges in each domain and the color-coded display of the patterns of potential mapped onto the molecular surfaces buried in each interface. Partial charges were assigned to the atomic coordinates using the updated charges (Weiner et al., 1984) of the AMBER data base (Weiner and Kollman, 1981). This dictionary of Mulliken net atomic charges was developed from ab initio LCAO-SCF-MO (linear combination of atomic orbitals—self consistent field—molecular orbital) calculations on individual amino acids using the STO-3G basis set (three gaussian functions per Slater-type orbital) (Hayes and Kollman, 1976; Weiner et al., 1984). All nonhydrogen atoms and all potentially hydrogen-bonding atoms were included explicitly, while a united-atom representation was used for aliphatic and aromatic hydrogen atoms. The amino acid residues were assumed to be in their predominant state of ionization in aqueous solution at physiological pH. Histidine residues were considered neutral; in each histidine the imidazole proton was placed to allow maximal hydrogen bonding. The COOH-termini were assigned a net charge of -1 and the NH_2-termini a net charge of $+1$ (except where acetylated). Artificial termini created by the division of the polypeptide chain into domains or by the absence of disordered residues were left uncharged. No solvent water molecules were included in the calculation.

Calculation of the Electrostatic Potential

The electrostatic potential V at a given point r_o, resulting from the partial charges q_i assigned to the atomic positions r_i is given by:

$$V = \Sigma_i \, q_i / D|r_i - r_0|.$$

Dielectric effects arise from the shielding of charges by solvent atoms. In dilute aqueous solution, the dielectric constant surrounding a protein varies from a low value at distances less than a water diameter to that of bulk water at large distances (Hopfinger, 1973; Dean, 1981). Distance-dependent dielectric constants modeling this variation have been used for electrostatic calculations on macromolecules (Dean, 1981; Dearing et al., 1981; Hingerty et al., 1985;) and predicted by self-consistent calculations modeling the water near lysozyme by Langevin orientation of point dipoles (Warshel and Levitt, 1976). We have used a linearly distance-dependent dielectric model, replacing the dielectric constant (D) with a constant value (ϵ) multiplied by the distance in Ångstroms between each partial charge and the calculation point ($|r_i - r_o|$). The contributions of each partial charge within 20 Å of the calculation point were summed. To correct for dipoles split by the finite cutoff radius, we used the method of neutral spheres (Weiner and Kollman, 1981): a compensatory charge placed at the cutoff radius ensures that the net charge inside the sphere is zero. The electrostatic potential assigned to each molecular surface point was calculated 1.4 Å above that point along a surface normal vector; this calculation point (r_o) represents the center of the solvent probe sphere at the position of its closest approach to the molecule. The electrostatic potentials were mapped onto points of the molecular surface using a five-color code where red is most negative (less than -21 kcal/mol for $\epsilon = 1$, less than -5.25 kcal/mol for $\epsilon = 4$); yellow is negative, (between -21 and -7 kcal/mol for $\epsilon = 1$ and between -5.25 and -1.75 for $\epsilon = 4$); green is approximately neutral (between -7 and $+7$ kcal/mol for $\epsilon = 1$, between -1.75 and $+1.75$ kcal/mol for $\epsilon = 4$); cyan is positive (between $+7$ and $+21$ kcal/mol for $\epsilon = 1$, between $+1.75$ and $+5.25$ for $\epsilon = 4$); and blue is most positive (more than $+21$ kcal/mol for $\epsilon = 1$ and more than $+5.25$ kcal/mol for $\epsilon = 4$). Electrostatic stabilization energies were calculated by summing electrostatic interactions between all pairs of partial charges located in opposing subunits of each interface using a constant dielectric value of 4.

Calculation and Display of the Electrostatic Field

The electrostatic field is the gradient (the magnitude and direction of maximum change) of the electrostatic potential. By convention, electrostatic field vectors indicate the direction in which a positive charge would move when placed in the field: the direction of maximum decrease in potential energy. Electrostatic field vectors were calculated by evaluating the partial derivative of the electrostatic potential with respect to each of the three coordinate axes. For the x component with the linearly distance-dependent dielectric model described above:

$$dV/dx = \Sigma_i - 2q_i|x_i - x_0|/\epsilon|r_i - r_0|^4.$$

For computer graphics studies of the field direction, a computer program to display each vector as an arrow was developed. The arrows were color-coded by the electrostatic potential, because color-coding by the magnitude of electrostatic field showed simply that the field was largest near the molecular surface and decreased as it extended into solution. These arrows, defining the force vectors on charged molecules, were used to evaluate long-range orienting effects between molecules prior to their collision.

RESULTS AND DISCUSSION

Structural Features of β Barrels

Before discussing our results, we summarize some structural information on β barrel domains as it relates to their assembly. β barrels can contain 5–13 strands. The interiors are closely packed with hydrophobic side chains. All β barrels, regardless of strand number, show a similar slight flattening in one direction (Richardson, 1981). Twist is one of the most conspicuous features of β-structure and it is always of the same handedness. If defined in terms of the angle at which neighboring β strands cross each other, then the twist is left-handed (Quiocho et al., 1977). If defined in terms of the twist of the hydrogen bonding direction or of the peptide planes viewed along a strand, then the twist is right-handed (Chothia et al., 1977). The overall twist of a β barrel can also be defined as the angle at which strands on opposite sides of the barrel cross one another. The β

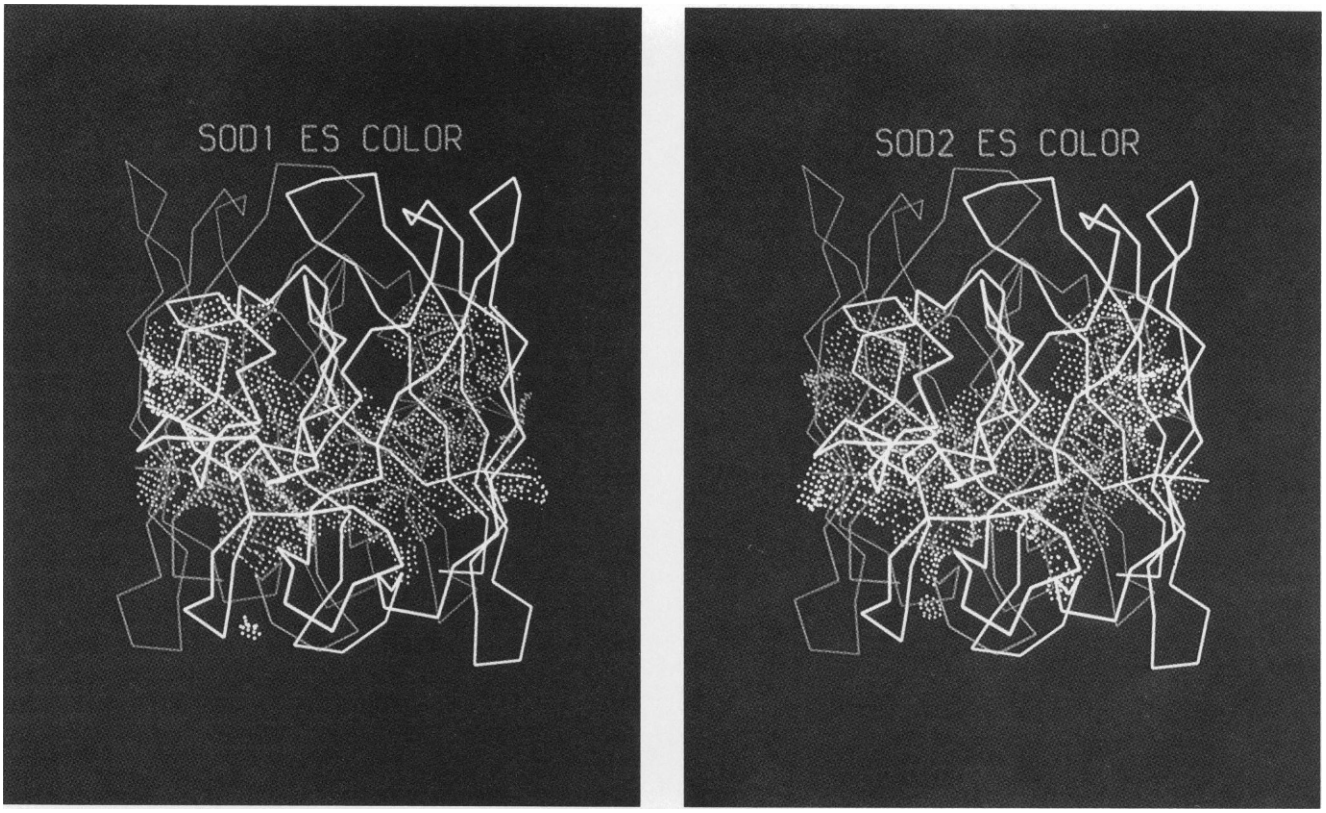

FIGURE 1 Solvent buried surface for the dimer interface of Cu, Zn superoxide dismutase (SOD). Surfaces are color-coded by electrostatic potential (*red* = most negative; *yellow* = negative; *green* = approximately neutral; *cyan* = positive; and *blue* = most positive; see methods for details). (*a*) Buried surface on SOD1 subunit viewed looking toward the surface from SOD1. (*b*) Corresponding interface on SOD2 from the same point of view as (*a*). Please refer to the color figure section at the back of this book.

barrel twist angle varies depending on the number of strands: ~95° for 5–6, 40° for 7–8, and 30° for 9–13 strands in the β barrel. We also use twist to describe the overall shape of surface buried by a domain interaction; twist is then measured by the angle between the front and back edges of the buried surface when viewed perpendicular to the interface.

Based on the pattern of temperature factors in β barrel structures that have been refined (Getzoff, Tainer, and Olson, unpublished results), the relative local mobility is lowest in the middle of the strands and increases toward the ends. Thus, one would expect more conformational flexibility for β barrel interactions involving the ends rather than the middle of β strands.

Copper, Zinc Superoxide Dismutase Dimer Contact

Molecular Surface Buried by the Dimer Contact. The SOD dimer contact is representative of a strong domain interaction, and the patterns of interaction for this interface should provide clues about the forces stabilizing such strong domain interactions. For the two SOD subunits to be separated, their interaction must be weakened by conditions which denature or change the tertiary fold of the subunit, such as denaturing agents, extremes of pH or temperature, reduction of the intrasubunit disulfide bond, or removal of one or more of the active-site metals. Bovine SOD remains dimeric in 8 M urea (Malinowski and Fridovich, 1979), and in 4% sodium dodecylsulfate (Keele et al., 1971). The x-ray structure of bovine SOD has been refined at 2-Å resolution (Tainer et al., 1982). Because the four subunits in the crystallographic asymmetric unit were fit and refined independently (Tainer et al., 1983), there is a redundancy of information that allows an internal check for the accuracy of the residue positions. This level of information allows a detailed analysis not practical for the TBSV domain interactions.

The molecular surfaces buried on each subunit of SOD by the dimer contact are continuous and complementary in both shape and chemical composition (Fig. 1): the buried surfaces are interdigitated without any gaps the size of a water molecule and the areas of contact match in terms of hydrophobic and hydrophilic atom types. Twenty-two residues contribute to the average of 519 Å^2 of molecular surface buried on each subunit by the contact, which represents ~8% of the total external molecular surface. When viewed along the contact surface perpendicular to the twofold axis, the two closely fitted surfaces in the contact have a right-handed twist of ~60° (from the edge

of the active site face to the edge of the β barrel face of the subunit). This twist results in a larger interacting buried surface area on the two globular subunits than would occur by the interaction of other comparable parts of the subunit surface. The twisted surfaces are also effectively wrapped around each other to provide an overall specificity of fit with regard to possible subunit rotations and translations.

Although the contact is primarily hydrophobic (Richardson et al., 1975), a significant portion (30%) of the buried surface is hydrophilic. The pattern of hydrophilic atoms is distributed throughout the buried surface; in contrast, the interior regions are electrostatically neutral while the exterior edges each show both positive and negative electrostatic potentials (Fig. 1). About 36% of the buried contact area is formed by main chain atoms, which are located in patches across the surface. Ten β strand residues and 12 loop residues are buried by the dimeric interaction. In addition to these residues in direct contact, there is apparently a surrounding region where a number of hydrophilic residues may be interacting through bridging water molecules, such that the effective contact area is increased. Three residues from each subunit form four main chain hydrogen bonds between the two subunits, and there are no intermolecular hydrogen bonds involving side chain atoms. All of these hydrogen bonds are between a β strand residue and a loop residue, and they form a hydrophilic patch near the center of the contact. The majority (77%) of side-chain to side-chain contacts occur between a β strand residue and a loop residue with only three β-residue to β-residue contacts, and no contacts between two loop residues.

There is no overwhelming preference manifest in the 14 types of residues involved in the dimer interaction. The residues Gly, Val, and Ile occur three times, and Ala and Thr occur twice. Two of the three Gly are involved in the main-chain hydrogen bonds between the subunits, whereas the third (Gly 148) packs against the other two (Gly 49 and Gly 112). These three sequence-invariant Gly come together in an area of closest approach between the polypeptide chains of the two subunits. Substitution of a different amino acid type in this region would disrupt the contact.

Possible Determinants for the Dimer Contact. A stripe of largely sequence-invariant residues runs approximately perpendicular to the twofold axis and accounts for about one-half of the surface buried in the dimer contact. Because this strip is also associated with almost all of the buried main chain area, it forms about two-thirds of the entire buried surface. The neighborhood of the stripe includes most of the specific residue side chain contacts, and the known sequence variations are relatively conservative with regard to residue type. These contact residues have very small deviations in position among the four subunits in the crystallographic asymmetric unit. Moreover, computer graphics analysis of the model, color-coded by temperature factors, indicates that this stripe has unusually low thermal factors comparable to those around the active site Cu ion. The structural position, high degree of order, and sequence conservation of this set of contact residues suggest that they are important determinants of the dimeric interaction and stability.

Domain Interactions in Immunoglobulin Fab

β domain interactions have been calculated for four immunoglobulin G Fab structures available from the Brookhaven Protein Data Bank (Bernstein et al., 1977) and of sufficient quality to examine their detailed interactions: NEWM at 2-Å resolution (Saul et al., 1978), MCP at 2.7-Å resolution (Segal et al., 1974), KOL at 1.9-Å resolution (Marquart et al., 1980), and RHE at 1.6 Å resolution (Furey et al., 1983). Our current results suggest that the domain interactions in these structures are quite similar. The most complete analysis to date has been done on Fab NEWM, and we will use this structure as an example.

The Fab V_L–V_H complex is formed by the interaction of the two flattened β barrels at ~60° angle. The surface buried by this contact has strong shape complementarity, being interlocked by interdigitating surface bumps and by an overall twist that wraps the buried surfaces. The depth and diameter of the local surface undulations are on the scale of 3–5 Å from the average surface plane. A single large hydrophobic patch accounts for most of the hydrophobic surface area buried on the V_L domain by the V_H. In contrast to the electrostatic potential of the SOD dimer contact, the electrostatic potential of the V_L is strikingly positive along one side with a small area of negative potential along the other edge (Fig. 2). Using clipping planes to slice through the two contact surfaces reveals that these areas of high potential are quite complementary to the potential on the mostly negative V_H surface. Detailed studies are in progress using the electrostatic fields calculated independently from these two domains to determine whether or not these electrostatic features may function to orient their assembly.

The C_L–C_H contact surface formed by the interaction of the two highly flattened β barrels at an almost 90° angle is more broken than the surface buried by the Fab V_L–V_H complex (Fig. 3). This contact has a reasonably large area of β–β interaction that is largely hydrophobic. This area is surrounded by surface from loop regions that bend around and cup the surface of the opposing domain. While similar to the SOD dimer contact previously described in terms of roughness and pattern of buried hydrophobic-hydrophilic atoms, the C_L–C_H contact surface has only about half the twist. The C_L domain buried surface is more negative, the C_H surface more positive.

In sum, both the Fab V_L–V_H and Fab C_L–C_H buried surfaces show good shape, chemical, and electrostatic

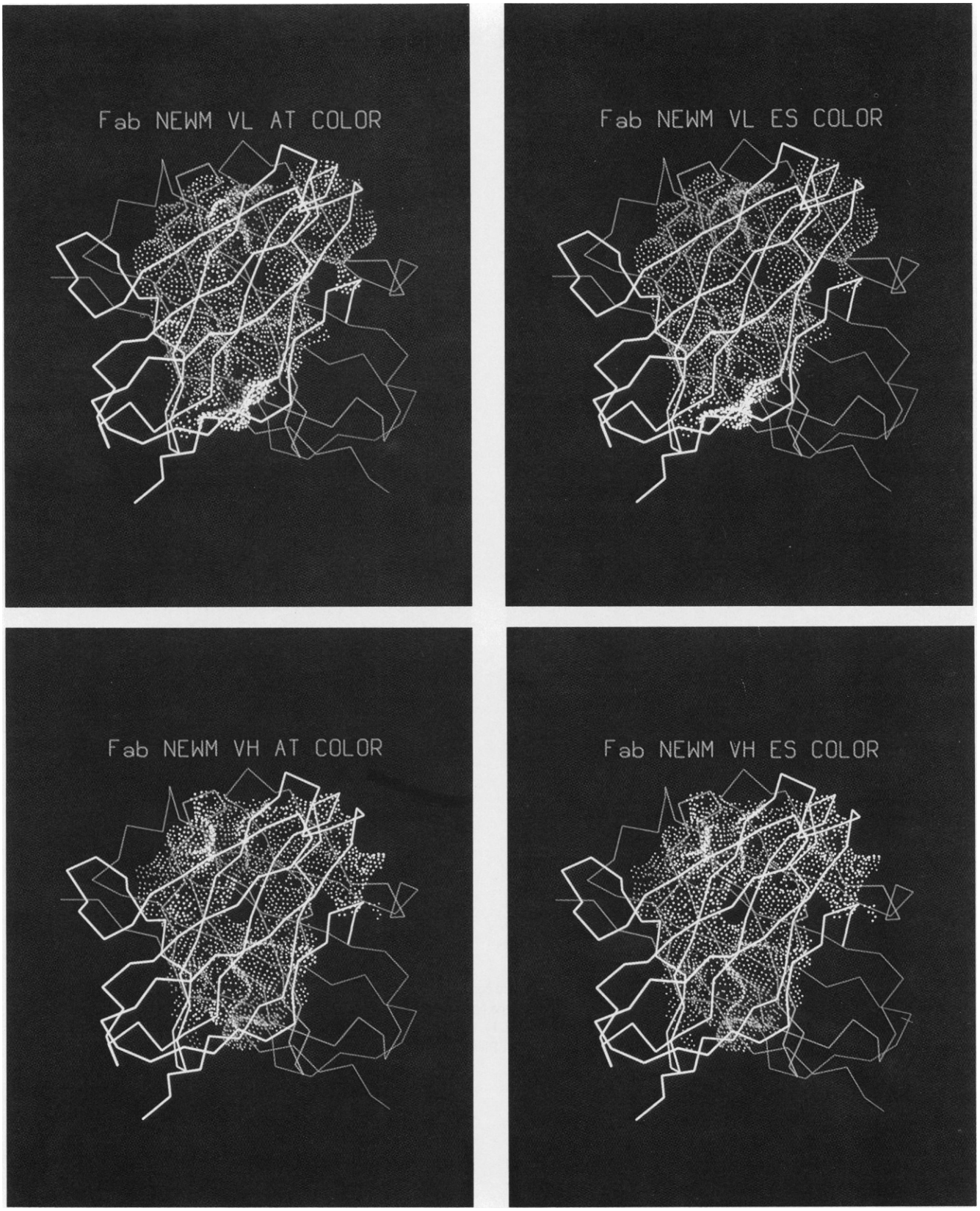

FIGURE 2 Interface for Fab NEWM between the variable light (red α-carbon trace) and variable heavy (cyan α-carbon trace) domains. Viewing conventions are the same as in Fig. 1. (*a*) Buried surface on variable light domain color-coded by atom type (*green* = carbon; *blue* = nitrogen; *red* = oxygen; *yellow* = sulfur). (*b*) Same surface as in (*a*) color-coded by electrostatic potential (see Fig. 1). (*c*) Buried surface on variable heavy domain color-coded by atom type. (*d*) Same surface as in (*c*) color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

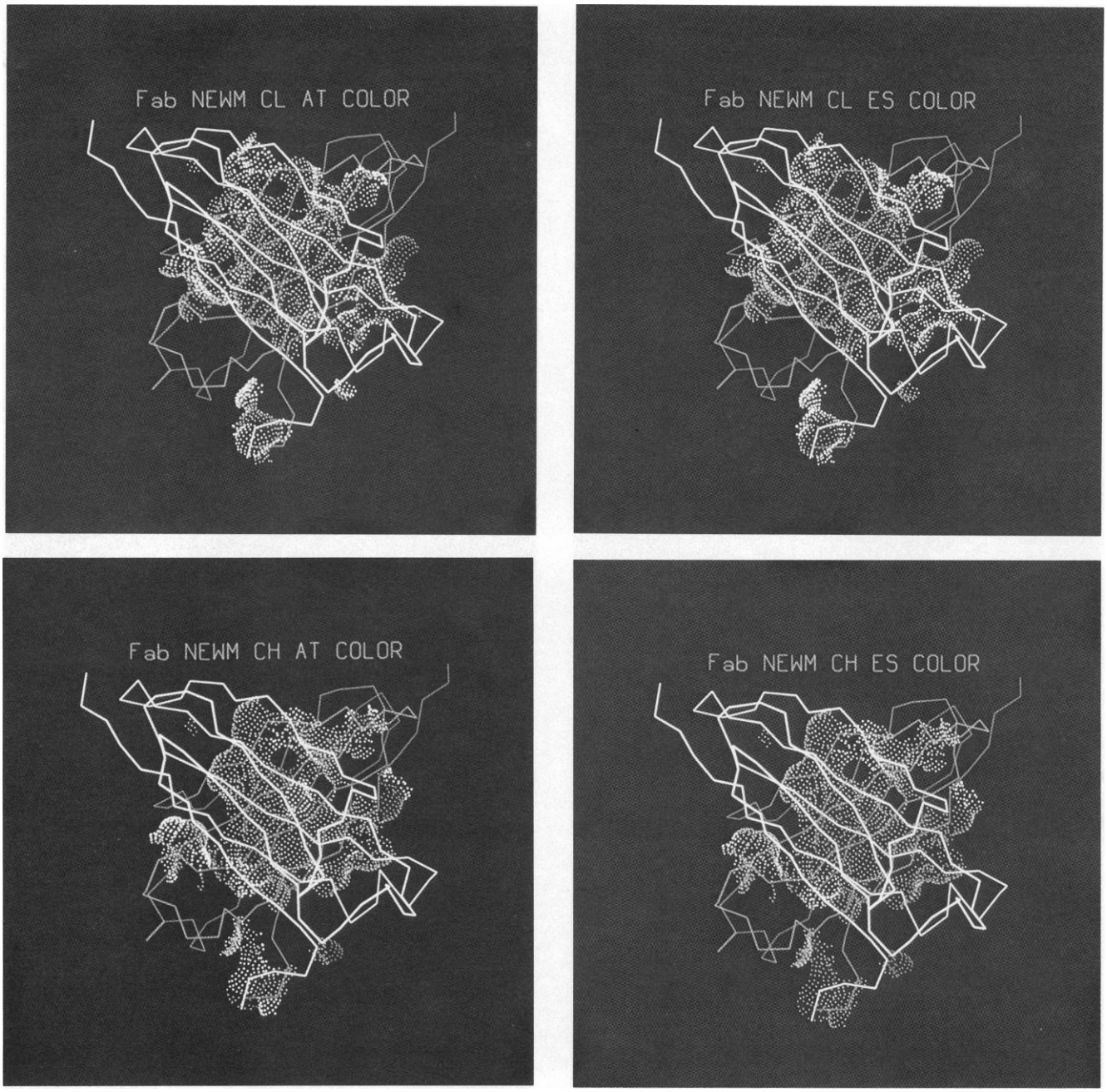

FIGURE 3 Interface for Fab NEWM between the constant light and constant heavy domains. Viewing conventions are the same as in Fig. 1. (*a*) Buried surface on constant light domain color-coded by atom type (*green* = carbon; *blue* = nitrogen; *red* = oxygen; *yellow* = sulfur). (*b*) Same surface as in (*a*) color-coded by electrostatic potential (see Fig. 1). (*c*) Buried surface on the constant heavy domain color-coded by atom type. (*d*) Same surface as in *c* color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

complementarity. Both include large areas of significant β–β interaction, but the Fab V_L–V_H is more hydrophilic and has electrostatic potentials of larger magnitude than the Fab C_L–C_H (Table I). The observed buried surface features appear to be consistent with the known biochemistry. For example, the higher electrostatic potentials associated with the variable domain interaction may account for the 10-fold faster rate of association for the variable domains relative to the constant domains (Klein et al., 1979). The calculated electrostatic stabilization energy for the variable domain interaction (−43.2 kcal/mol) is also somewhat higher than for the constant domain (−40.9 kcal/mol). The significantly lower ratio of hydrophobic-to-hydrophilic buried surface area for the variable domains (V_L = 1.8, V_H = 1.7) as compared with the constant domains (C_L = 2.3, C_H = 3.2) is consistant with the lower affinity of the variable domain interactions compared with the constant domains (Klein et al., 1979).

Domain Interactions in TBSV

The crystal structures of the intact spherical plant viruses provide an opportunity to study several aspects of recognition and interactions among β-barrel domains. Since the assembled states of these viruses are known to better than 3-Å resolution (Olson et al., 1983; Abad-Zapetero, et al., 1980; Liljas, et al., 1982), detailed analysis of domain and subunit interfaces can present a quantitative picture of shared and unique modes of interaction, and may even suggest models for pathways of assembly.

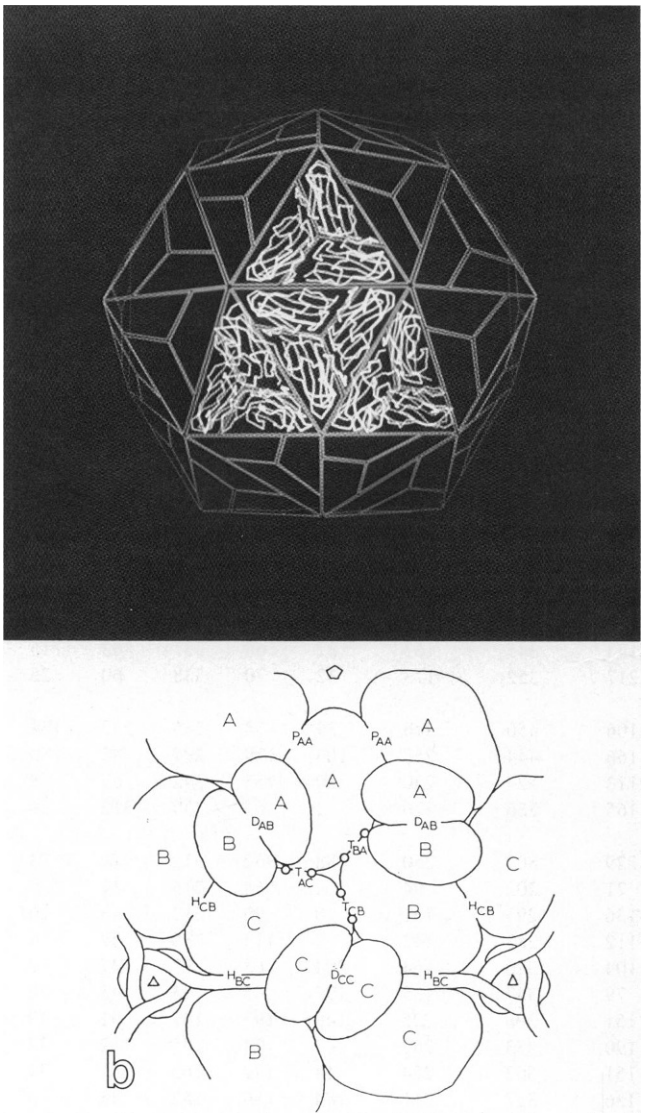

FIGURE 4 TBSV packing. (*a*) Packing of TBSV shell (S) domains on the surface of the virus particle. Each quadrilateral represents the position of a viral protein subunit. Twelve positions are filled with a schematic of the S domains, showing the directionality of the axis of the β barrel. Please refer to the color figure section at the back of this book. (*b*) Diagram showing packing of subunits in the icosahedral asymmetric unit of TBSV and notation for interfaces. The different kinds of subunit contact are labeled *D* (dimer), *T* (trimer), *P* (pentamer) and *H* (hexamer), with a subscript showing the types of subunit interacting across the contact in question (from Olson et al., 1983). This naming convention is used to indicate the interfaces in Table I.

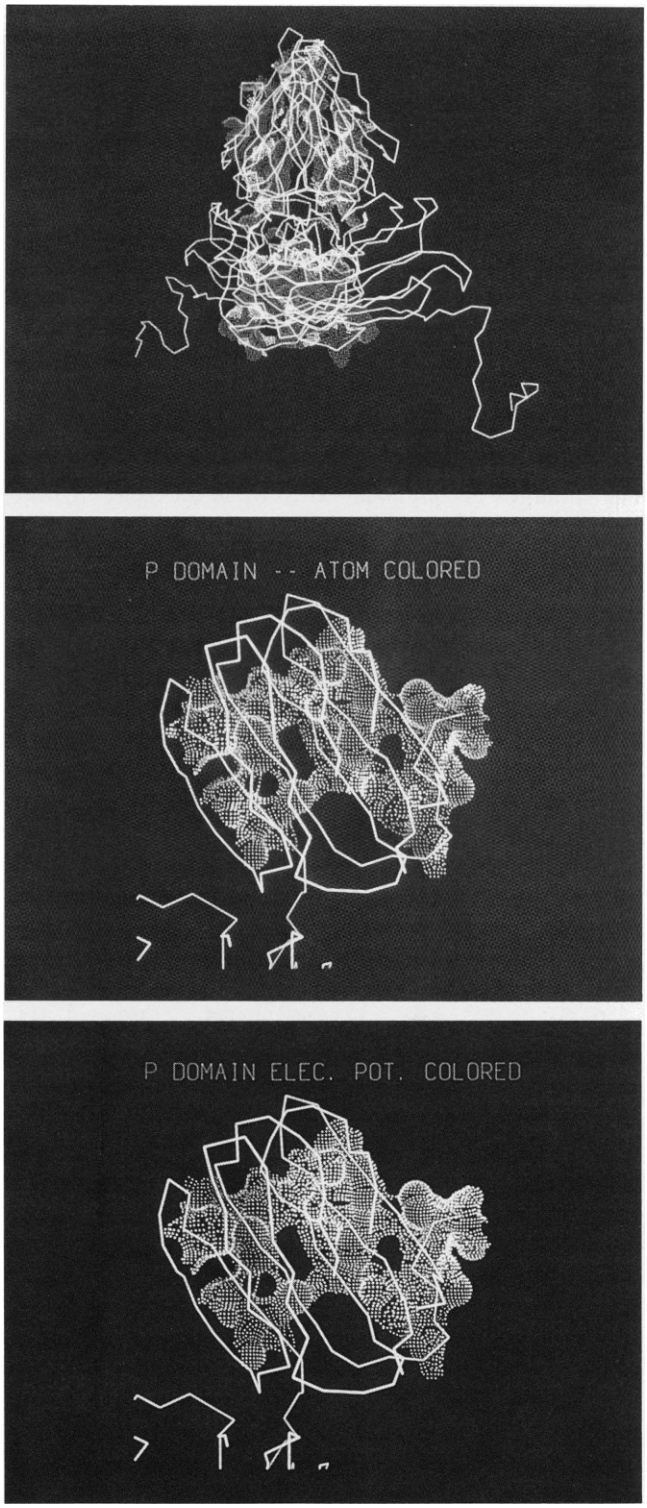

FIGURE 5 Dimer contacts for TBSV. (*a*) The dimer interface between C subunits. Two C subunits are seen as α-carbon traces. Only the buried surface of one of the subunits is shown, color-coded by atom type (see Fig. 2). The projecting domains are at the top center, pointing upward. The shell domains are below. Two NH_2-terminal arms can be seen at the bottom. All subsequent figures of TBSV subunits retain an orientation in which the P domain is on top and the S domain is on the bottom. (*b*) Face-on view of P domain buried surface color-coded by atom type. (*c*) Same surface as in (*b*) color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

Of the currently solved multisubunit β-barrel structures, TBSV offers the largest variety of interdomain contacts. Each of the 180 protein subunits in the T = 3 icosahedral structure has an identical amino acid sequence and folds into a number of distinct domains, two of which are clearly seen for all subunits in the crystal structure. These two domains are both of β-barrel Greek key topology and are designated *S* for shell and *P* for projecting. The three distinct symmetry environments of the T = 3 particle differentiate three subunits by their positions in the icosahedral asymmetric unit, designated *A*, *B*, and *C*. The A subunits pack against each other around the fivefold axes; the C subunits pack against each other around the twofold axes; and the B subunits alternate with the C's in a hexameric packing around the threefold axes (see Fig. 4). The subunit's accommodation to these different environments provides the variety of contacts in the system. The major modes of accommodation involve relative movement of P and S domains and the ordering of ~40 additional residues toward the NH_2-terminal end of the S domain of the C subunits. This ordered stretch, designated the NH_2-terminal arm, makes interactions both with adjacent subunits (C and B) across a subunit interface and with its nonadjacent threefold related mates, creating a mini β-domain termed a β annulus. Thus one can see both the variety of interfaces in the intact TBSV structure, and the sequence-of-event specific nature of some of them.

Some general remarks can be made about the relative orientations of the axes of the interacting β barrels in the icosahedral viral structures, and the general shape of the interfaces made. In the observed cases, in the shell domains these axes run essentially tangent to the particle surface; thus the orientations of interacting β barrel domains are dictated by the local symmetry axes (see Fig. 4 *a*). The projecting domains, on the other hand, are oriented with the barrel axes radial to the particle surface, and are packed with their axes ~45° to each other (see Fig. 5 *a*). Unlike the dimer contacts observed in SOD and the immunoglobulins, no significant twist is seen in any of the interfaces examined for TBSV. On the other hand, interdigitation of surfaces due to undulations of ~3–5 Å in diameter and depth is the norm, as with the SOD and immunoglobulin interfaces.

Fig. 4 *b* diagrams the subunit contacts in TBSV. The contacts are classified according to the nature of the subunit coordination forming each specific interface. Our

TABLE I
SURFACES BURIED IN β DOMAIN INTERFACES

Interface	Domain	Surface Area (Å²)												
		Total	β	Non-β	Side Chain	Main Chain	Hydrophobic	Hydrophilic	Electrostatic Potential					
									−−	−	0	+	++	
SOD														
Dimer	One	508	176	332	315	193	345	163	32	68	331	63	15	
	Two	525	176	350	308	217	352	173	32	70	338	60	25	
Fab NEWM														
Variable	L	695	469	227	589	106	450	246	39	34	245	237	141	
	H	702	662	40	536	166	444	258	103	179	297	43	80	
Constant	L	753	318	435	641	113	524	230	80	255	342	67	8	
	H	742	557	185	577	165	566	176	1	66	330	312	34	
TBSV														
D_{CP}		1063	611	452	833	229	803	260	188	352	312	179	34	
D_{CS}		296	0	296	225	71	202	94	4	64	216	11	0	
D_{ASBS}	AS	446	125	321	311	136	295	151	0	90	312	43	0	
	BS	426	124	302	314	112	283	142	0	114	280	32	0	
T_{BSCS}	BS	490	20	469	386	104	326	164	204	68	97	47	72	
	CS	473	10	462	393	79	337	135	197	65	95	45	70	
P_{AS}	AS	620	431	189	469	151	396	225	149	193	147	91	40	
	AS5	595	230	365	405	190	333	262	27	194	185	152	37	
H_{CSBS}	CS	526	213	313	375	151	303	224	14	172	165	141	34	
	BS	568	385	183	412	156	357	211	184	156	182	46	0	
H_{BSCS}	BS	442	210	232	320	122	287	155	16	139	178	106	4	
	CS	454	367	87	344	110	284	170	148	122	109	48	27	
H_{BSCSA}	BS	747	367	381	515	232	467	280						
	CS	1024	728	296	759	265	665	359						
	Arm	856	113	743	518	338	576	280						
I_{ASP}	AS	271	44	227	153	118	177	93	11	133	126	1	0	
	AP	353	146	208	251	103	240	114	0	13	144	137	60	
I_{CSP}	CS	242	2	240	126	116	164	78	36	109	98	0	0	
	CP	331	165	166	240	91	225	106	0	9	132	121	69	

Fab abbreviations: L = light chain, H = heavy chain TBSV abbreviations: D = dimer, T = trimer, P = pentamer, H = hexamer, I = intramolecular

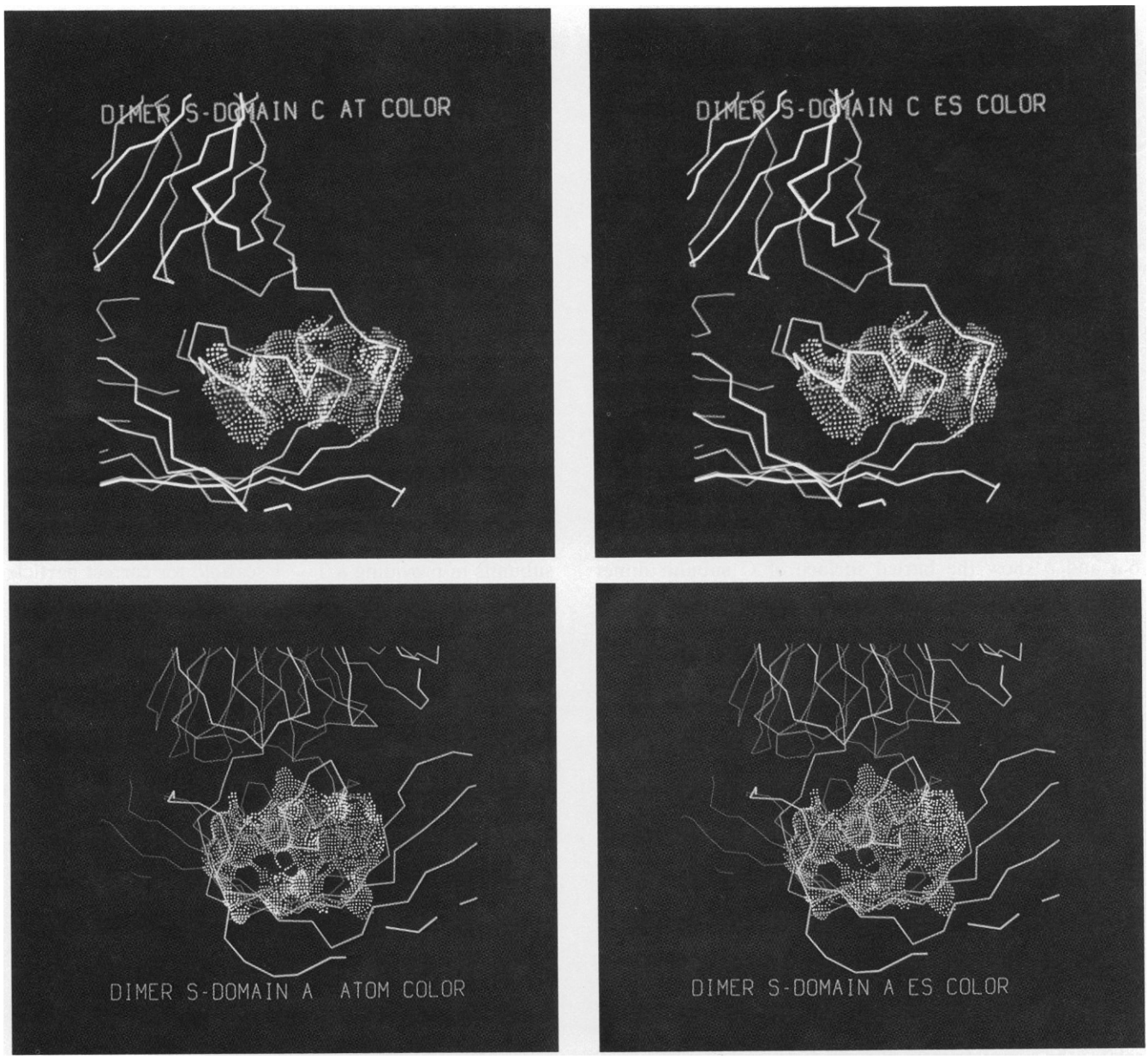

FIGURE 6 Dimer contacts for the shell (S) domains of TBSV. The backbone α-carbon trace shows A subunit in blue, B subunit in cyan, and C subunit in orange. (*a*) Twofold interface for the S domains of C subunits, removing the arm from the calculation. Buried surface is color-coded by atom type. (*b*) Same surface as in *a* color-coded by electrostatic potential. (*c*) Dimer contact for the S domains of A and B subunits, color-coded by atom type. (*d*) Same surface as in *c* color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

approach to analyzing this array of interactions has been to calculate and display the surfaces buried by each pairwise contact, mapping and/or calculating a variety of surface characteristics, and looking for patterns of interest in these representations. Because we cannot, in this presentation, go into a detailed analysis of each contribution to every interdomain contact, we list exhaustively particular surface characteristics of each interface (see Table I), illustrate patterns observed in particular classes of contacts, and discuss some of the implications of these characteristics and patterns.

TBSV Dimer Contacts. Fig. 5 *a* depicts the interaction around the twofold axis of both the P and the S domains of the C subunit. The color coding of the buried surface by atom type is useful both for seeing the extent and distribution of hydrophilic and hydrophobic surface, and for detecting patterns of hydrogen bond networks. The P and S domains differ quite noticeably in their distribution of oxygen and nitrogen atoms over the interacting surface. Although no striking pattern of atom type can be seen in the P domain, quite noticeable parallel stripes of alternating oxygen and nitrogen atoms can be seen running

along the strand directions at the top and bottom of the shell domain. Such a pattern is suggestive of sets of alternative hydrogen bonds at this interface that could be broken and remade by opposing motion of the two surfaces in a direction along the stripes.

A closer look at the P domain interface offers additional information. The ratio of hydrophobic-to-hydrophilic surface buried in the contact is the highest of any interface, as are both the total and hydrophobic surface area (see Table I). Fig. 5 *b* and *c* show atom and electrostatic coloring of the P domain interface. It is noteworthy that, even though the surface is predominantly hydrophobic, most of it exhibits very high electrostatic potential, and that the distribution of potential is in large part complementary to its twofold related interface, particularly at the two edges. Thus, the observed surface characteristics and distributions indicate that the P domain dimer interface should rank as the most stable in the particle.

In contrast, further examination of the S domain dimer contact indicates much less stability to this interface. Figs. 6 *a* and *b* show the buried surface of C subunit dimer contact without considering the arm. Figs. 6 *c* and *d* show the equivalent dimer contact between the A and B subunits around a local twofold axis. In both cases we see the alternating oxygen/nitrogen pattern running along the strand direction, and a low ratio of buried hydrophobic to hydrophilic surface area. The electrostatically color-coded surfaces show overall weak potentials. The one area that shows electrostatic complementarity to the twofold related surface is located in the part of the chain between the S and P domains, i.e., the flexible hinge region. Thus, the combination of alternate hydrogen bonds, low hydrophobic buried area, and chain flexibility suggests a moveable interface. Diffraction studies of the expanded form of TBSV (Robinson and Harrison, 1982) indicate that this interface is disrupted in the particle expansion.

TBSV Trimer Contacts. A local threefold axis relates the symmetrically distinct A, B, and C subunits of TBSV. The three interfaces created are for all purposes identical. This class of interface presents several unique aspects. In this study the trimer interface between the B and C subunits is chosen for examination. In addition to having a low ratio of hydrophobic-to-hydrophilic surface area, there is practically no participation from residues in β strands, and the side-chain to main-chain surface area ratio is very high. Furthermore, the buried surface appears less continuous than in any of the other interfaces, indicating patches of solvent accessibility. These characteristics point to a flexible interface. The most striking feature, however, can be seen in Fig. 7 and in Table I. The figure shows both sides of the interface, that is, the buried surface for the C and B subunits, color coded by both atom type and electrostatic potential. The concentration of oxygen atoms (carbonyls) and of high negative potential at both interfaces indicate an explosive contact. There is no complementarity of potential across the interface. This is the proposed site of two divalent cations (calciums) that stabilize the interface, and provide the mechanism for a pH dependent switch to disrupt the contact and expand the particle.

TBSV Pentamer/Hexamer Contacts. Because the curvature of the T = 3 particle is dictated by the nature of the five-coordinate and six-coordinate interfaces, the notion of quasiequivalence (Caspar and Klug, 1962) was proposed to explain such icosahedral designs built from a single type of subunit. Examination of these interfaces shows both the truth and the limitations of the original notions. Certainly, the order/disorder switch of the NH_2-terminal arm was not anticipated in the original theory of quasiequivalence. The arm acts to provide alternative bonding for a change of dihedral angle between particle "facets," thus alternating hexamer interactions between direct and divided (arm intervening). Even more surprisingly, the arm serves as a link between nonadjacent subunits in providing a T = 1 scaffold for correct particle sizing (Olson et al., 1983). Here, however, we look at direct hexamer and pentamer contacts and the divided contact in the absence of the arm. This is done to explore the notion that the arm folds into and stabilizes an already existing contact.

Fig. 8 illustrates both sides of each of the hexamer and pentamer buried surfaces, color coded by electrostatic potential. The striking feature of these interfaces is the cross-strand extent of the strong potential surfaces, and their complementarity on the opposite face. This pattern suggests the possibility of alternative electrostatic links across the interface, consistent with the concept of a fulcrum between divided and direct contacts. Additional features of these interfaces point to electrostatic orientational functions. The strong adjacent positive and negative potentials seen across the divided C subunit buried surface suggests that electrostatic forces may facilitate alignment with a nearby subunit. The fact that one face in each of these contacts has an overall stronger negative potential than the other may also help in speeding proper assembly. Work in progress uses electrostatic fields calculated for an isolated subunit to determine if these or other interfaces may function to orient assembly.

Other TBSV Contacts. The remaining contacts in the assembled TBSV particle seen in the x-ray structure are of two types. First are the S to P domain interactions, both intra and inter-subunit. These contacts involve the smallest buried surface areas of all of the interfaces. As Table I indicates, the two alternative orientations of intra-domain contact (I_{ASP} and I_{CSP}) are comparable in all characteristics listed. There are indications from the structure of the expanded particle that these are not the only alternatives for SP interaction.

The last "interdomain" contact is by far the most

interesting, because it forms a new domain on assembly of the particle, the NH_2-terminal arm/β-annulus. Study of this interface involves examination of more than pair-wise contacts. Table I (H_{BSCSA}) lists some characteristics of the three-way interface created when the arm folds up against its own C shell domain, and the interfacing B shell domain (interactions from the other two C domains in the β-annulus are not considered). The arm forms an additional β-strand with both its own S domain and that of the adjacent BS domain. Studies are continuing with electrostatic energy and field calculations to characterize sequential stabilization of the divided contact and formation of the β-annulus. This holds out for us the possibility of examining a preformed environment (the hexamer S contacts) and its relationship to the formation of a β domain from disordered but tethered chains.

SUMMARY

We have been examining β-barrel interdomain contacts in a number of assembled protein systems by using computer

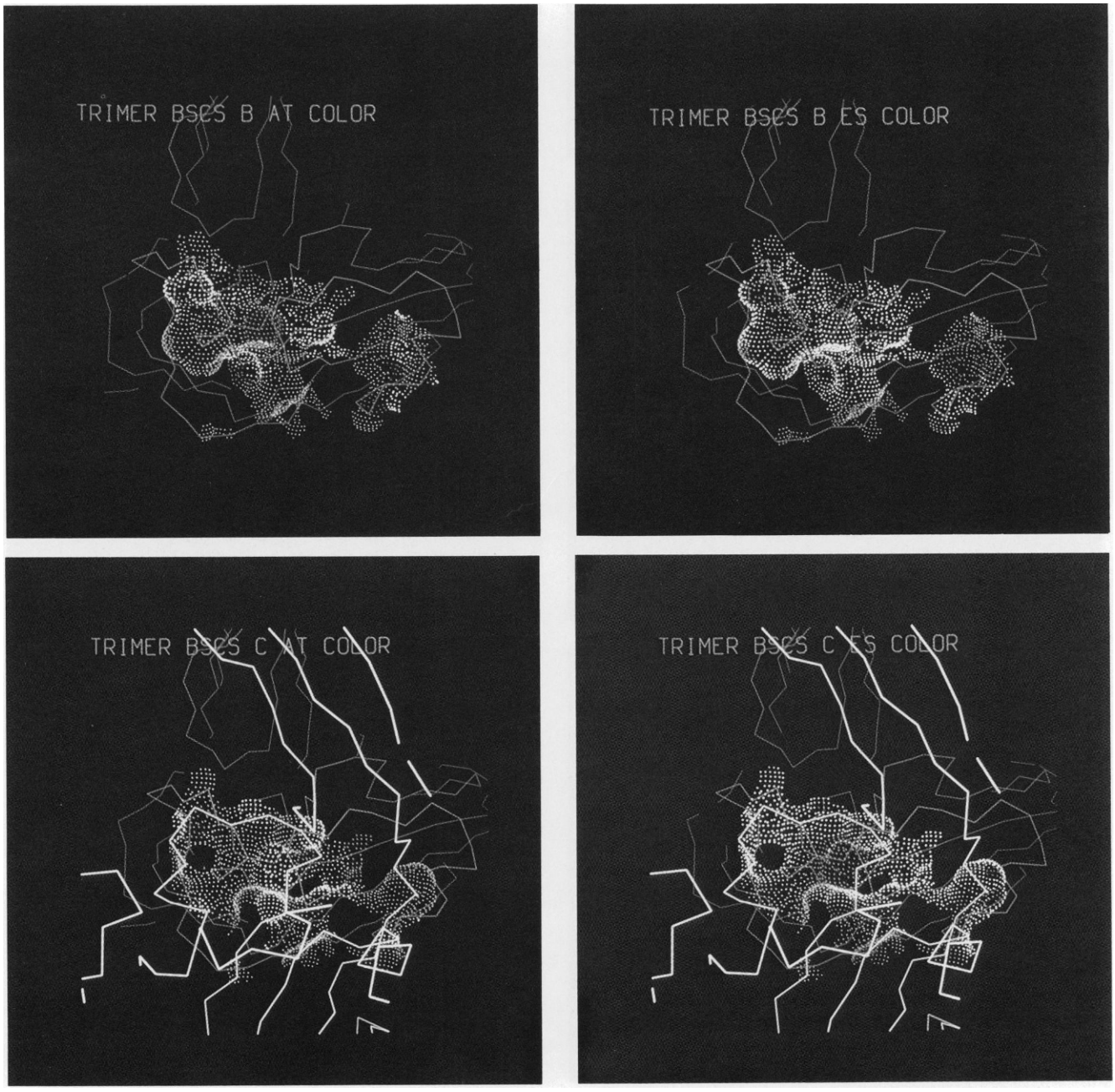

FIGURE 7 Trimer contacts in TBSV. (*a*) Buried surface for the B subunit at the trimer contact with the C subunit, color-coded by atom color. (*b*) Same surface as in (*a*) color-coded by electrostatic potential. (*c*) Buried surface for the C subunit at the trimer contact with the C subunit, color-coded by atom color. (*d*) Same surface as in (*c*) color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

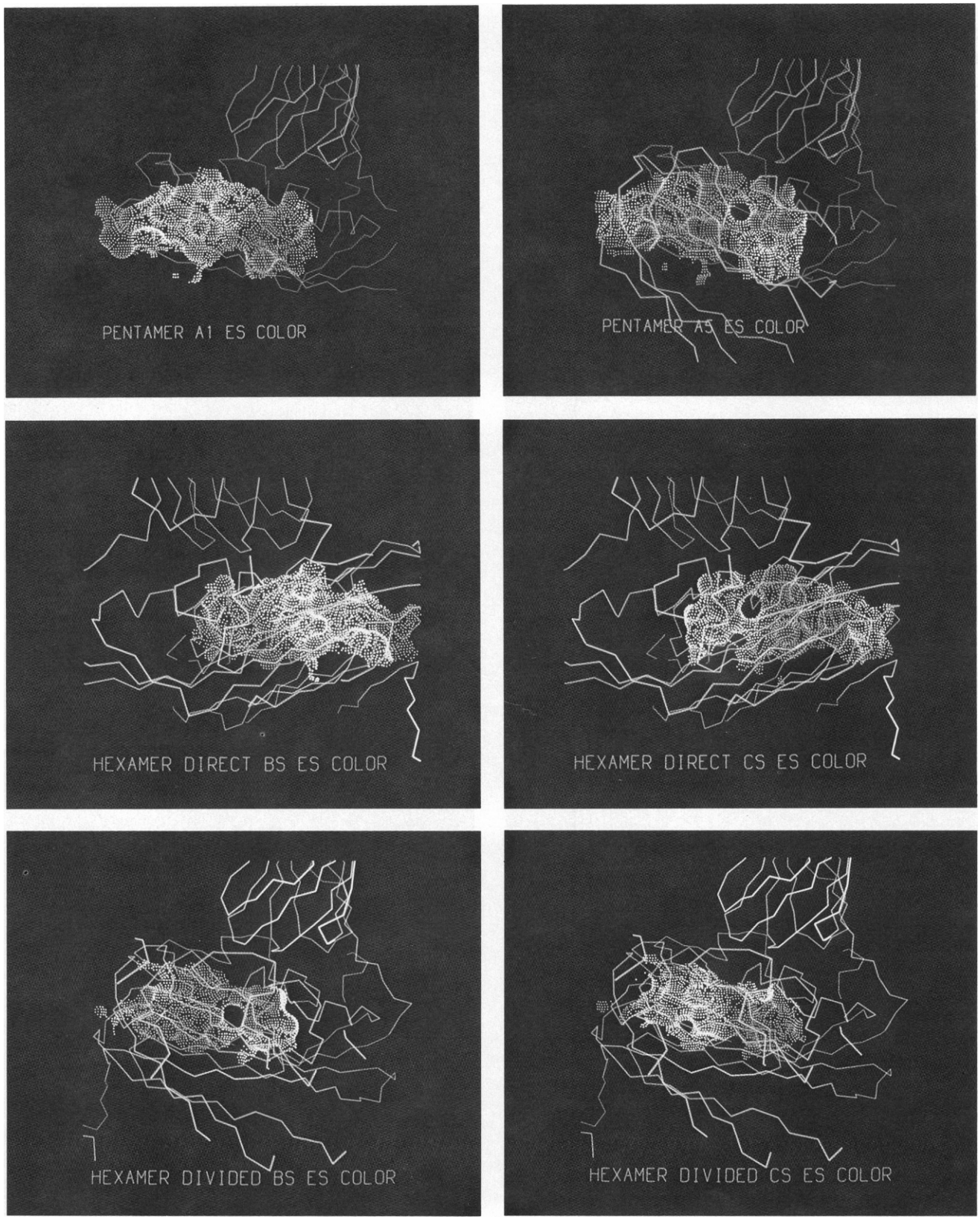

FIGURE 8 Pentamer and hexamer contacts in TBSV. (*a*) Buried surface for A subunit at pentamer interface color-coded by electrostatic potential. (*b*) Companion surface to that in (*a*) on fivefold-related A subunit with electrostatic potential coloring. (*c*) Buried surface for B subunit at the hexamer direct interface, color-coded by electrostatic potential. (*d*) Companion surface to that in (*c*) on adjacent C subunit, color-coded by electrostatic potential. (*e*) Buried surface for B subunit at hexamer divided interface. The surface was calculated without the coordinates for residues in the NH_2-terminal arm and is color-coded by electrostatic potential. (*f*) Companion surface to that in (*e*) on the adjacent C subunit. This surface is also calculated without the NH_2-terminal arm and color-coded by electrostatic potential. Please refer to the color figure section at the back of this book.

graphics to reveal patterns of atomic, steric, and electrostatic distribution. Our objective is to understand the nature of macromolecular recognition, stabilization, and assembly. We have found a series of function specific patterns in the contacts examined. The stable dimer interfaces in SOD, the IgG heavy-chain/light-chain contacts, and the projecting domain of TBSV all have very large complementary buried surfaces characterized by extensive hydrophobic contacts and by stabilizing electrostatic interactions at the solvent interface.

One of the most variable parameters identified is the relative amount of buried surface contributed by the β barrel residues vs. other parts of the domain. Among the TBSV domain interactions, the ratio of β to non-β surface area varies from 0 for the D_{CS} and T_{BSCS} contacts to over four for the H_{CS} interaction. Although the ratio of hydrophobic-to-hydrophilic buried surface is greater than one for all contacts examined, it varies from a high of 3.2 for the Fab constant heavy chain domain to a low of 1.3 for the AS5 domain in the TBSV pentameric interface. The SOD dimer interface, which is known to be highly stable, has an unusually high ratio of 1.5 for main chain-to-side chain buried surface area, but its 2.1 ratio of hydrophobic-to-hydrophilic buried surface is about average. The differences in the stability and assembly rate of the variable and constant Fab domain interfaces can be explained by the greater involvement of electrostatic interactions in the variable interfaces and hydrophobic interactions in the constant domain interfaces.

Three types of variable contacts are seen in TBSV: (*a*) the parallel stripes of alternating oxygen and nitrogen atom type, indicating alternative hydrogen bonding networks, coupled with weak electrostatic interactions seen at the S domain dimer contact, which allows the "sliding contact" inferred from the expanded particle structure; (*b*) The "explosive" electrostatically energetic interface, stabilized by counter ions, seen at the trimer contacts; and (*c*) the cross-strand bands of strong electrostatically complementary interactions seen at the hexamer and pentamer contacts, allowing some notion of a "fulcrum" for quasi-equivalence between the direct and divided (intervening arm) interfaces. In addition, several observations of electrostatic potential distribution suggest that the electrostatic fields surrounding these surfaces may act to align particular interfaces at a distance to facilitate productive collisions.

Received for publication 22 May 1985 and in revised form 28 June 1985.

REFERENCES

Abad-Zapatero, C., S. S. Abdel-Mequid, J. E. Johnson, A. G. W. Leslie, I. Rayment, M. G. Rossmann, D. Suck, and T. Tsukihara. 1980. Structure of southern bean mosaic virus at 2.8 Å resolution. *Nature* (*Lond.*). 286:33–39.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rogers, O. Kennard, O., T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Bondi, A. 1964. van der Waals volumes and radii. *J. Phys. Chem.* 68:3 441–451.

Caspar, D. L. D., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27:1–24.

Chothia, C., M. Levitt, and D. Richardson. 1977. Structure of proteins: packing of α-helices and pleated sheets. *Proc. Natl. Acad. Sci. USA.* 74:4130–4134.

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

Connolly, M. L. 1983*b*. Analytical molecular surface calculation. *J. Appl. Crystallogr.* 16:548–558.

Connolly, M. L., and A. J. Olson. 1985. GRANNY, A companion to GRAMPS for the real-time manipulation of macromolecular models. *Comput. and Chem.* 9:1–6.

Dean, P. M. 1981. Drug-receptor recognition: Electrostatic field lines at the receptor and dielectric effects. *Br. J. Pharmacol.* 74:39–46.

Dearing, A., P. Weiner, and P. A. Kollman. 1981. Molecular mechanical studies of proflavine and acridine orange intercalation. Nucleic Acids. 9:1483–1497.

Furey, W. Jr., B. C. Wang, C. S. Yoo, and M. Sax. 1983. Structure of a novel Bence-Jones protein (Rhe) fragment at 1.6 Å resolution. *J. Mol. Biol.* 167:661–692.

Getzoff, E. D., J. A. Tainer, P. K. Weiner, P. A. Kollman, J. S. Richardson, and D. C. Richardson. 1983. Electrostatic recognition between superoxide and Cu, Zn superoxide dismutase. *Nature* (*Lond.*). 306:287–288.

Harrison, S. C., A. J. Olson, C. E. Schutt, F. K. Winkler, and G. Bricogne. 1978. Tomato bushy stunt virus at 2.9 Å resolution. *Nature* (*Lond.*). 276:368–373.

Hayes, D. M., and P. A. Kollman. 1976. Electrostatic potentials of proteins. 2. Role of electrostatics in a possible catalytic mechanism for carboxypeptidase A. *J. Am. Chem. Soc.* 98:7811–7816.

Hingerty, B. E., R. H. Ritchie, T. L. Ferrell, and J. E. Turner. 1985. Dielectric effects in biopolymers: The theory of ionic saturation revisited. *Biopolymers.* 24:427–439.

Hopfinger, A. J. 1973. Conformational properties of macromolecules. Academic Press, Inc., New York.

Keele, B. B., Jr., J. M. McCord, and I. Fridovich. 1971. Further characterization of bovine superoxide dismutase and its isolation from bovine heart. *J. Biol. Chem.* 246:2875–2880.

Klein, M., C. Kortan, D. I. Kells, and K. J. Dorrington. 1979. Equilibrium and kinetic aspects of the interaction of isolated variable and constant domains of light chain with the Fd' fragment of immunoglobulin G. *J. Am. Chem. Soc.* 18:1473–1481.

Liljas, L., T. Unge, T. A. Jones, K. Fridborg, S. Lovgren, U. Skoglund, and B. Strandberg. 1982. Structure of satellite tobacco necrosis virus at 3.0 Å resolution. *J. Mol. Biol.* 159:93–108.

Malinowski, D. P., and I. Fridovich. 1979. Subunit association and side-chain reactivities of bovine erythrocyte superoxide dismutase in denaturing solvents. *Biochemistry.* 18:5055–5060.

Marquart, J., J. Deisenhofer, R. Huber, and W. Palm. 1980. Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 and 1.9 Å resolution. *J. Mol. Biol.* 141:369–391.

McCammon, J. A., P. G. Wolynes, and M. Karplus. 1979. Pico-second dynamics of tyrosine side chains in proteins. *Biochemistry.* 18:927–942.

McCammon, J. A., and M. Karplus. 1983. The dynamic picture of protein structure. *Acc. Chem. Res.* 16:187–193.

O'Donnell, T. J., and A. J. Olson. 1981. GRAMPS–A graphics language interpreter for real-time, interactive, three-dimensional picture editing and animation. *Computer Graphics.* 15:133–142.

Olson, A. J., G. Bricogne, and S. C. Harrison. 1983. Structure of tomato

bushy stunt virus IV: The virus particle at 2.9-Å resolution. *J. Mol. Biol.* 171:61–93.

Quiocho, F. A., G. L. Gilliland, and G. N. Phillips, Jr. 1977. The 2.8-Å resolution structure of the L-arabinose-binding protein from *Escherichia coli*. *J. Biol. Chem.* 252:5142–5149.

Richardson, J. S. 1981. The anatomy and taxonomy of protein structure. *In* Advances in Protein Chemistry. C. B. Anfinson, J. T. Edsal, and I. M. Richards, editors. Academic Press, Inc., New York. 168–339.

Richardson, J. S., K. A. Thomas, B. H. Rubin, and D. C. Richardson. 1975. *Proc. Natl. Acad. Sci. USA.* 72:1349–1353.

Robinson, I. K., and S. C. Harrison. 1982. Structure of the expanded state of tomato bushy stunt virus. *Nature* (*Lond.*). 297:563–568.

Saul, F.A., L. M. Amzel, and R. J. Poljak. 1978. Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin New at 2.0-Å resolution. *J. Biol. Chem.* 253:585–595.

Segal, D. M., E. A. Padlan, G. H. Cohen, S. Rudikoff, M. Potter, and D. R. Davies. 1974. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA.* 71:4298–4302.

Tainer, J. A., E. D. Getzoff, K. M. Beem, J. S. Richardson, and D. C. Richardson. 1982. Determination and analysis of the 2 Å structure of Cu, Zn superoxide dismutase. *J. Mol. Biol.* 160:181–217.

Tainer, J. A., E. D. Getzoff, J. S. Richardson, and D. C. Richardson. 1983. Structure and mechanism of Cu, Zn superoxide dismutase. *Nature* (*Lond.*). 306:284–287.

Tainer, J. A., E. D. Getzoff, J. Sayre, and A. J. Olson. 1985. Modeling intermolecular interactions: topography, mobility, and electrostatic recognition. *J. Mol. Graphics.* In press.

Warshel, A., and M. Levitt. 1976. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103:227–249.

Weiner, P. K., and Kollman P. A. 1981. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comp. Chem.* 2:3 287–303.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner. 1984. A new force field for molecular-mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.

DISCUSSION

Chairman: Adrian Parsegian
Scribes: Marc J. Glucksman and Joe D. J. O'Neil

BURNETT: It is possible to look at TBSV and SBMV not in the usual way as polypeptides, but as pentamers and hexamers. I favor this approach since when you remove Ca^{++} ions, the particle swells and the contacts within the pentamers and hexamers remain intact. It appears that what you are saying is that the hydrophobic contacts are mainly within the pentamers and hexamers. Is this true? Additionally, there are two types of contacts in the hexamer, one of which is similar to that within the pentamers. Are both types of these hexamer contacts equally hydrophobic?

GETZOFF: Table I shows that the pentameric and hexameric faces are comparable in hydrophobicity, but the most hydrophobic contact in TBSV is between the projecting domains across the twofold axis.

BURNETT: Let's suppose you have a pentamer of dimers, so that there is a fivefold axis relating them. Then one dimer can link to a further dimer and start off another closed set either around a pseudo sixfold or a fivefold axis. This then establishes a geometry for either a T = 1 or a T = 3 capsid.

BLUM: Do you really need a specific description of the model with a resolution up to 1 Å? How close are the points on your surface model?

GETZOFF: The density of the surface is about four dots per $Å^2$.

BLUM: What is the total area that a probe can see on the residue surface?

GETZOFF: That depends on the size of the residue.

BLUM: Let's assume the available residue area on a protein surface is ~ 20 $Å^2$. Imagine a charge is approaching together with hydrogen-bonded polarized water molecules. Then the electrostatic potential of the incoming group will see an area ~ 4–5 $Å^2$ on the surface of the residue.

GETZOFF: Do you mean that the incoming molecule is shielded by water molecules?

BLUM: Yes, especially for the charge. My question is, if you attempt to look at a less-detailed picture of the surface, using two or three $Å^2$ for a grid interval, will you see a similar picture of the charges?

GETZOFF: The surface points are for purposes of calculation only. They are distributed on the surface of the protein, but they each represent the accessibility to a probe representing one water molecule. If I change the surface density of the dots so that the dots are further apart, they would still each represent a calculation at a discrete point.

TAINER: The reason to look at such detail is to look at complementarity. The complementarity is certainly there in the assembled complexes, in terms of polar, electrostatic, and van der Waals interactions. In the case of superoxide dismutase (SOD) we have explicit hydrogen atoms and when these are included, the cavities and imperfections in and between the subunits, disappear. We also have four independent subunits in the crystallographic asymmetric unit of SOD and these packing interactions are reliably conserved. I believe that the packing is real and is telling us about the specificity of subunit contacts. It is true that you may want to look at different resolutions, but first you would like to see the atomic detail present in the structure. This is the first step to building models that explain the structure and allow us to use site-directed mutagenesis to test the models. We would like to see cases where the models fail, because those tell us where to improve the model. We need the fine detail to make these predictions.

BLUM: I agree with you totally as far as the complementarity of the surfaces is concerned. What I want to point out is that if you look at

electrostatic interactions and potentials, you are better off being a little less specific, because electrostatic potential is always transferred through the hydrogen bonds of the water molecule that projects as a cone from whatever position the incoming charge looks upon the surface. For the purpose of complementarity and looking at the crevices you are completely correct.

GLAZER: One interesting point in what you said with respect to SOD intersubunit contact is that it is stable to SDS and urea. This is a very strong contact. This suggests that when these two monomers come together their surfaces form a fairly strong set of interactions that requires some type of reorganization of both monomer faces.

GETZOFF: I agree with you completely. This is a difficult problem because in this case you can't look at the structure of the individual subunits. We have been pursuing this area, looking at antibodies raised against the protein myohemerythrin, and our initial results indicate that the antigenic sites of the molecule correspond to the sites of flexibility. Flexibility is important; the missing piece in the story is what movements occur in the molecules.

GLUCKSMAN: Would you say that there is a hierarchy of importance with regard to multimer contacts comparing the ratio of hydrophilic to hydrophobic areas in the buried surface to high electrostatic potential for fruitful interactions?

GETZOFF: I think both are important for different reasons. If you look for instance at the immunoglobulin Fab structures, the interface between the constant domains is more hydrophobic than that between the variable domains. The constant regions have been shown to have a higher affinity for each other experimentally. In contrast, the variable domains show a greater contribution of electrostatic interactions, and experimentally assemble 10-fold more rapidly in solution. I think the electrostatic forces play a kinetic role in orienting and setting up the specificity for the interaction, so it can occur rapidly, whereas the hydrophobic forces have a more dominant role in defining the final stability.

GLUCKSMAN: So the electrostatic interaction first recognizes the event and then the hydrophobic confers stability?

GETZOFF: Yes, to some extent.

GLUCKSMAN: Concerning the recent elucidation of virus structures with dominant β structures such as rhinovirus by Michael Rossman's group, polio virus by Jim Hogle's group, and adenovirus by Roger Burnett's group, does this work, taken together, suggest a general importance of the β-barrel motif?

GETZOFF: I think it is clear that β-barrel domains make nice protein building blocks for assembly. There are a number of multimeric enzymes formed from β barrels. The interesting aspect from my point of view is that the interactions between the β barrels are very different with respect to the β-barrel regions in the interface and the angles between them.

PARSEGIAN: One of the referees asked how you justify your decision to use a dielectric constant that is proportional to distance. On what basis you think it is possible to avoid solving the real electrostatic problem, which includes dielectric interfaces?

Last summer I read a paper by Warshel, who is apparently the father of the $\epsilon = r$ you use (Warshel. 1984. *Quart. Rev. Biophys.* 17:283–522). He said, "A frequently asked practical question is: what type of distance-dependent dielectric constant should be used in calculating electrostatic interactions in proteins? The relevant issue is not simple, and frequently misunderstood. The function $\epsilon = r$ used in some calculations is attributed to the d(r) if Warshel and Levitt (1976), which represents the screening parameter of the dipole and not the proper $\epsilon(r)$. In general, we feel that calculations of electrostatic interactions of proteins should be based on detailed microscopic models, and not on an ill-defined effective dielectric constant."

My own impression is that questions asked on such detailed scale overload the calculation. There are clear rules, but when you have so much detail, it is harder to think about what is going on. You have to back off and think on a 3–4 Å scale, instead of a 0.1–0.2 Å scale, perhaps as B. K. Lee has been doing for years. Otherwise your computations have the appearance of a logical procedure but not the accuracy that must go with it.

GETZOFF: We are looking for patterns in the interactions. I don't mean to argue that the electrostatic potential at any one surface point is quantitatively exact. I agree that electrostatic interactions still have a way to go before being completely correct. We are working on some newer models for calculating electrostatic forces in superoxide dismutase where there is more specific information. We now know the crystallographic positions of bound water molecules, and we would like to include them explicitly. The first step includes orienting the hydrogen atoms reasonably, then using partial charges for the water molecules as well as the protein. The second aspect of improving our calculations involves distinguishing the dielectric constant within the protein from that of the solvent which is currently under investigation. Unfortunately this requires a great deal of computer time.

PARSEGIAN: That assumes the molecular interactions are sufficiently well known so that it is only a matter of crunching numbers.

GETZOFF: No, I don't think that is the case, but we are looking at the structures of known complexes. You should build better models and then devise experiments to test them. With respect to the correct choice of a dielectric model, we have been focussing on what happens if you change the model. We have used a dielectric constant equal to the distance in ängströms between the calculation point and a partial charge (see Fig. 2*b* and *d*). We recalculated these electrostatic potentials using a dielectric constant of four. The differences in the dielectric model did not make a significant difference in the patterns we see. Any constant dielectric from 1–80 will give the same pattern, but the magnitude of the electrostatic potentials will be different.

PARSEGIAN: The qualitative lesson doesn't really seem to depend on the values of the numbers even to within an order of magnitude.

LEE: I think you are not more guilty than other people in using the distance dependent dielectric constant. In coming up with a quantitative measure of shape complementarity, if you just deduce the gaps, I would like to know what results you get. Have you come up with numbers for comparing the number of gaps in the interface vs. the gaps in the protein?

GETZOFF: We haven't examined that quantitatively. Qualitatively, there are indications of gaps in interfaces. For interfaces with gaps where water molecules could intervene between the two surfaces, you see a hole instead of a continuous surface. If you make thin slices, you can see protruding and depressed pieces fitting into each other, but this

shape complementarity is hard to quantitate. Another aspect of shape complementarity is the twist of the interacting surfaces, as for example, in the Fab fragment. This twist adds to both the specificity of orientation and the amount of surface that contributes to the interaction.

LEE: I have a question about hydrophobicity/hydrophilicity ratios. In a number of systems the dissociation of subunits involves both enthalpic and entropic components. It is not clear to me whether the strength of the interactions, however you define them, will correlate with the "hydrophobic interaction." It may have a lot to do with hydrogen bonding. Do you have some feelings about this in this study and whether you can say that the hydrophobic interactions are more important than the hydrophilic?

GETZOFF: The data that address this question most directly were collected by our collaborator at the Research Institute of Scripps Clinic, James Sayer. He was working on defining the contributions of given types of atoms to the binding energy. He started with Joel Janin's values, calculated energy contributions for each type of atom, and examined the energy as a function of hydrophilic and hydrophobic surface area. The binding energy roughly correlates with the magnitude of the hydrophobic area. I don't think this excludes hydrogen bonds from being important. The biggest concern with hydrogen bonds is that the presence of a hydrogen bond in a complex is not necessarily a change from zero to one hydrogen bond, but may represent a change from a hydrogen bond with solvent to a hydrogen bond in the complex.

LEE: But if you look at effects of entropy and enthalpy the correlation tends to break down. There are many more enthalpic components than you expect, so you must be careful.

GETZOFF: That is why I didn't stress numbers in the study. I agree, they are tentative.

NOVOTNY: We have studied the nature of domain-domain contacts in immunoglobulin variable domains in some detail (Novotny and Haber, 1985. *Proc. Natl. Acad. Sci. USA.*, vol. 82). We also estimated the electrostatic contribution to the domain-domain stabilization energy. To compute the electrostatic energy of VL, VH domains and their dimers, we used two models: the Coulomb equation, with dielectric constant = 50 (Churg, Russell, and Warshel), and the solvent-screening model (Northrup). The results of the two computations were similar to your results. What seemed to make a big difference was evaluation of the potential to infinity, as opposed to using a cutoff distance. We found in the four crystallographic structures studied, that the electrostatic contribution to dimer formation varies greatly, from stabilizing to destabilizing. It thus seems that electrostatics is not the true structural invariant of dimer formation, in these cases, and I would like to ask you, whether you have any information about how the electrostatic stabilization varies in different structures such as SOD or viruses.

TAINER: That's a very interesting comment. If you look at these contacts mapped with electrostatic potentials, many values are close to zero. The final outcome of an energy calculation comes from the sum of large numbers of small terms that are not known very accurately. This problem forced us to go to pattern analysis. From the work we have done on other immunoglobulins, the patterns look similar to those for Fab NEWM. There may be electrostatic complementarity in these other cases, and the difference in energy you observe may result from a large number of small terms that we would consider neutral when we map them onto the surface. Have you checked which terms dominate the differences you see in electrostatic energy?

NOVOTNY: The immunoglobulin variable domains have very different net charges, so differences are mainly due to unit charges.

POTSCHKA: I would like to make a general note of caution to this sort of approach even though I believe in it in principle and use it myself. The quantitative amount of hydrophobicity of accessible surface area depends crucially on how you define the interface and the surfaces. The different algorithms that have been used by Chothia and others to define hydrophobicity give widely different results. It depends on how you program your computer. This is as crucial as the choice of dielectric constant in electrostatics. One thing you can say qualitatively is that hydrophobicity should give some contribution to stability but no angular preference. Even though contribution of the charges may be small in terms of total energy, charges should have an important contribution in detailed positioning. I don't think one should make inferences of hierarchy of electrostatics. That is beyond this work.

GETZOFF: I don't advocate a specific hierarchy, except to point out that electrostatic forces work at a greater distance than hydrophobic forces. As Ray Salemme said earlier, there is a significant amount of information suggesting the importance of electrostatic orientation at a distance.

GLUCKSMAN: In terms of computation, is a 1.4 A radius probe the most appropriate value to use in Michael Connolly's Molecular Surface program?

TAINER: This question also bears on the previous one. What we are really doing is comparing a number of surfaces. It doesn't matter *per se* what size probe radius you use, as long as you are consistent in comparing among a number of examples. This is one of our goals.

GETZOFF: We use a 1.4 A radius probe representing a water molecule. There are two reasonable bounds for such a probe. The radius of an oxygen atom and the radius which would also include the hydrogen atoms. For solvent accessible surfaces a water molecule can approach in many orientations, so the radius of an oxygen atom seems appropriate.