PROTEINS: Structure, Function, and Bioinformatics 56:130–142 (2004)

Hydrophobic Complementarity in Protein–Protein Docking

Alexander Berchanski,[1] Boaz Shapira,[1] and Miriam Eisenstein[2]*
[1]*Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel*
[2]*Department of Chemical Services, Weizmann Institute of Science, Rehovot, Israel*

ABSTRACT **Formation of hydrophobic contacts across a newly formed interface is energetically favorable. Based on this observation we developed a geometric–hydrophobic docking algorithm that estimates quantitatively the hydrophobic complementarity at protein–protein interfaces. Each molecule to be docked is represented as a grid of complex numbers, storing information regarding the shape of the molecule in the real part and information regarding the hydropathy of the surface in the imaginary part. The grid representations are correlated using fast Fourier transformations. The algorithm is used to compare the extent of hydrophobic complementarity in oligomers (represented by D_2 tetramers) and in hetero-dimers of soluble proteins (complexes). We also test the implication of hydrophobic complementarity in distinguishing correct from false docking solutions. We find that hydrophobic complementarity at the interface exists in oligomers and in complexes, and in both groups the extent of such complementarity depends on the size of the interface. Thus, the non-polar portions of large interfaces are more often juxtaposed than non-polar portions of small interfaces. Next we find that hydrophobic complementarity helps to point out correct docking solutions. In oligomers it significantly improves the ranks of nearly correct reassembled and modeled tetramers. Combining geometric, electrostatic and hydrophobic complementarity for complexes gives excellent results, ranking a nearly correct solution < 10 for 5 of 23 tested systems, < 100 for 8 systems and < 1000 for 19 systems. Proteins 2004; 56:130–142.** © **2004 Wiley-Liss, Inc.**

Key words: molecular recognition; D_2 tetramers; surface complementarity; protein–protein interfaces; oligomers; fast Fourier transform; grid representation by complex numbers

INTRODUCTION

The formation of a protein–protein complex in aqueous solution at room temperature is an entropically driven process.[1] Non-bonded van der Waals and electrostatic interactions between the solute and the solvent, that are lost in the process of complex formation, are mostly compensated for by newly formed interactions at the protein–protein interface, leading to a small change in the enthalpy of the system. In contrast, entropy changes due to rearrangement of water molecules, immobilization of protein side chains at the interface and the assembly of two or more proteins into a single complex, have a considerable contribution to the free energy change. In particular, the rearrangement of water molecules, which were bound at the surface of the protein and are now free to mix with other water molecules in the bulk solvent, increases the entropy thus lowering the free energy of the system. This stabilizing contribution depends on the size of the newly formed protein–protein interface and on its chemical nature. Water molecules located next to non-polar (hydrophobic) portions of the protein surface satisfy their hydrogen bonding potential by interacting with other water molecules, leading to the formation of an ordered layer ("iceberg"), which is entropically unfavorable.[1] Therefore, burying of hydrophobic portions of the surface at the new protein–protein interface brings about a larger entropy gain than burying of polar portions of the surface. The interactions at the interface can be divided in a simplified manner to polar-polar, polar-hydrophobic and hydrophobic-hydrophobic. Interactions between polar portions of the surfaces include ion pairing and hydrogen bonding, which are very important for specificity. The contribution of these electrostatic interactions in protein–protein docking was examined previously by others and us[2] (and see references therein). The contribution of polar–hydrophobic interactions is not obvious. Free energy is gained by burying a hydrophobic portion of the surface but free energy is also lost because the polar residues cannot satisfy their hydrogen bonding potential. In contrast, interactions between hydrophobic portions of the two surfaces are clearly entropically favorable, as described above, and therefore we concentrate in this study on hydrophobic complementarity.

The thermodynamic considerations outlined above lead researchers to study hydrophobic interactions at protein–protein interfaces. Many studies analyzed the amino acid composition and the hydropathy of interfaces in comparison to the surface and the interior of the protein. Some of these studies analyzed oligomers[3,4] and others examined complexes of soluble proteins.[5,6] They concluded that the

A. Berchanski and B. Shapira contributed equally to this work.

*Correspondence to: Miriam Eisenstein, Department of Chemical Services, The Weizmann Institute of Science, Rehovot 76100, Israel. E-mail address: miriam.eisenstein@weizmann.ac.il

Received 4 August 2003; Accepted 4 December 2003

Published online 7 May 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20145

© 2004 WILEY-LISS, INC.

abundance of hydrophobic residues at the interface is generally larger than on the exterior of proteins and smaller than in their interior. Moreover, interfaces in oligomers were found to be more similar to protein interiors than interfaces in complexes of soluble proteins. Most of these studies analyzed statistically the whole interface and did not consider the spatial distribution of hydrophobic residues. More detailed studies of interfaces identified clusters of hydrophobic residues on the surface of one of the component molecules in complexes of soluble proteins[7] and in dimeric and tetrameric oligomers.[8] In a more recent analysis[9] 136 interfaces in dimers were studied qualitatively. Patches of hydrophobic residues were found near the center of the interface for only one-third of the cases, whereas in most dimers hydrophobic interactions were interspersed with hydrogen bonds and water molecules were found scattered over the entire interface area. None of the above studies examined quantitatively the extent of hydrophobic complementarity at the interface. A recent study examined pair-wise interactions between amino acids across the interface.[10] They found a clear preference for hydrophobic–hydrophobic contacts and a considerable contribution of aromatic residues. Their database of interfaces included both, oligomers and hetero-dimers of soluble proteins, and their analyses did not distinguish between the two groups.

Several studies considered the contribution of hydrophobic residues in docking. Vakser and Aflalo[11] tried "hydrophobic docking," in which shape complementarity of only the hydrophobic portions of the surface was examined. However, they unrealistically categorized about 80% of the surface as hydrophobic. Ackermann et al. considered hydrophobic complementarity as an additional measure to shape complementarity and rewarded contacts of surface points with hydrophobicities exceeding a given value.[12] Both approaches were not tested for unbound docking (i.e., docking that starts from the structures of uncomplexed component molecules). More recently Chen and Weng combined geometric complementarity with desolvation and electrostatics in docking and found the desolvation term to be most dominant. In their study all the contacts that involve hydrophobic surface residues were rewarded, (according to their non-pairwise ACE scores), irrespective of which residue they contact.[13] Hence, their hydrophobic term reflects the buried non-polar surface area. In addition, many groups use desolvation terms in post-scan filters.[14,15]

In this study we examine quantitatively the extent of hydrophobic complementarity at the interface. We favor only interactions that juxtapose hydrophobic portions of the surfaces and compare the extent of such complementarity for interfaces in tetramers with D_2 symmetry, which represent oligomers, and in hetero-dimers of soluble proteins (referred to hereafter as complexes). We then ask if hydrophobic complementarity is unique to the correct interface or does it also occur in false interfaces, formed by other parts of the surface (the non-interacting surface). This is done by examining the effect of hydrophobic complementarity in prediction of the structures of oligomers and complexes.

We find that there is hydrophobic complementarity at the interface in both groups of systems, complexes and oligomers. The extent of this complementarity in oligomers is related to the non-polar hydrophobic interface area, which in turn appears to be an approximately constant proportion of the interface. Thus, the hydrophobic complementarity is extensive in large oligomeric interfaces, whereas in smaller interfaces it is partial. Similarly, partial complementarity is observed for interfaces in complexes.

Geometric–hydrophobic docking gives better results than geometric docking in reassembly and modeling of tetramers. It also improves the result of unbound docking for complexes. A combination of geometric, electrostatic and hydrophobic docking is particularly effective.

METHODS

Geometric–Hydrophobic Docking

To study the geometric–hydrophobic complementarity at protein–protein interfaces we use a modified version of our docking program MolFit.[16,17] As before, the molecules are projected onto a 3D grid of complex numbers, storing the geometric descriptor of the molecule in the real part.[2] The imaginary part of each grid point is 0 in the unmodified geometric docking algorithm. In the current study it is changed to H^a or H^b, for molecules **a** and **b**, respectively, where H is the hydrophobic descriptor (see below). Thus, the grid representations of molecules **a** and **b** are as follows:

$$\bar{a}_{l,m,n} = \left\{ \begin{array}{cc} 0 & \text{outside the molecule} \\ 1 + i\mathrm{H}^a & \text{on the surface of the molecule} \\ -15 & \text{in the interior of the molecule} \end{array} \right\} = (G^a + \mathrm{i}H^a)_{l,m,n}$$

$$\bar{b}_{l,m,n} = \left\{ \begin{array}{cc} 0 & \text{outside the molecules} \\ 1 - \mathrm{i}H^b & \text{on the surface of the molecule} \\ & \text{and in the interior} \end{array} \right\} = (G^b - \mathrm{i}H^b)_{l,m,n}.$$

The quality of the match between the molecular surfaces for a given shift of molecule **b** by α, β, γ grid points with respect to molecule **a** is measured by the real part of the correlation value $C_{\alpha,\beta,\gamma}$, where

$$\mathrm{real}(C_{\alpha,\beta,\gamma}) = \sum_{l=1}^{N}\sum_{m=1}^{N}\sum_{n=1}^{N} G^a_{l,m,n} \cdot G^b_{l+\alpha,m+\beta,n+\gamma} + H^a_{l,m,n} \cdot H^b_{l+\alpha,m+\beta,n+\gamma}.$$

The parameter H^a is positive for a selection of hydrophobic side chains or fragments (as described below) and 0 otherwise, whereas H^b is 1 for the hydrophobic moieties and 0 otherwise. Therefore, the correlation value, real($C_{\alpha,\beta,\gamma}$), denoted hereafter as the complementarity score, is a linear combination of the geometric and hydrophobic complementarities at the given relative shift. The correlation matrix is calculated via fast Fourier transformations, as described before.[16]

We tried three alternative hydrophobic representations in which the assignment of the hydrophobic descriptor H was as follows: In the first representation (denoted "*residues*") grid points in the surface layer of molecule **a**, which are situated within the volume of atoms belonging to the most hydrophobic side chains according to the Kyte and Doolittle hydropathy scale[18] (valine, leucine, isoleucine, methionine, alanine, proline, and phenylalanine), were given a positive H^a. Notably, the radius of an atom in MolFit is its van der Waal radius plus the radius of a water molecule.[17] In molecule **b** grid points within the volume of the hydrophobic side chain atoms listed above were assigned the value $H^b = 1$. H was 0 for all the other grid points in both molecules. If a given grid point was situated within the volume of a hydrophobic side chain atom and also within a neighboring hydrophilic atom, it was classified hydrophobic. The second representation (denoted "*residues+*") was similar to the *residues* representation, but it also included the slightly less hydrophobic cysteine, tyrosine, and tryptophane[18] in the list of hydrophobic residues. In the third representation (denoted "*fragments*") the CH, CH_2, CH_3, SH, and S moieties (except CαH) in all the side chains were considered hydrophobic. However, in this case grid points situated within the volume of a hydrophobic atom and a neighboring hydrophilic atom were classified hydrophilic, because otherwise most of the surface was categorized as hydrophobic.

Determination of the Hydrophobic Complementarity Coefficient H^a

The parameter H^a adjusts the relative contribution of the geometric and hydrophobic complementarity to the score. It was determined empirically for oligomers and complexes and for the *residues, residues+* and *fragments* representations separately, using virtual scans.[2] The geometric–hydrophobic score for each solution in the scan is a linear combination of the geometric and hydrophobic complementarity: $S_i = S_i^G + S_i^H \cdot H^a$. Hence, the score for solution i, S_i depends on H^a and on the character of the interface reflected in the geometric and hydrophobic complementarities, S_i^G and S_i^H respectively. Therefore, for several systems, we performed two geometric–hydrophobic scans with different but close values of H^a, and used the scores to obtain linear equations for all the solutions. Next, S_i values were calculated for different values of H^a (virtual scans) and the rank of the nearly correct solution (see definition below) was plotted as a function of H^a. These graphs were used to identify adequate values of H^a for oligomers and complexes and for the different hydrophobic representations, such that produced the lowest ranks for nearly correct solutions. They were also used to examine the sensitivity of the geometric–hydrophobic docking results to the value of H^a.

Selection of Test Systems and Application of the Geometric–Hydrophobic Docking Algorithm

In a previous study we presented a method that combined comparative modeling and docking in order to predict the structures of D_2 tetramers.[19] Tetramers were constructed in two steps: A symmetry restricted docking step, saving five best C_2 homodimers per orientation, and a tetramer forming step, in which the homodimers were combined using vector algebra considerations. Two tetramer forming algorithms were described, the **ab/ac** and **ab/cd** algorithms, which differ in that the first one requires identification of two different C_2 homodimers in the docking step (**ab** and **ac**), whereas the second one forms tetramers from one C_2 homodimer (**ab** = **cd**). In this study we use the same selection of 16 D_2 tetramers as in the previous study, adopting the same translation and rotation intervals in the scans. This allows a detailed comparison of the results of geometric and geometric–hydrophobic docking. The selected tetramers are listed in Table I. These tetramers crystallize either as a monomer, a dimer, or a tetramer in the asymmetric unit, producing perfect or approximate tetramers. In the latter cases we reconstruct the same tetramer two or four times, starting each time from one of the independent subunits in the asymmetric unit. As these subunits differ slightly from one another the predictions correspond to some extent to unbound docking in complexes. We also use the geometric–hydrophobic docking algorithm to form tetramers from model-structures of single subunits, constructed by comparative modeling from templates with different oligomerization states.[19]

Next, we apply the geometric–hydrophobic docking algorithm to 23 hetero-dimers of soluble proteins, using the structures of the unbound proteins where possible. Fourteen (14) of these systems were previously studied by us via geometric and geometric–electrostatic docking,[2] and we added eight unbound/unbound and one bound/unbound systems from the benchmark of Chen et al.[85] (antibody/antigen and "others"). These systems too are listed in Table I. Full rotation/translation scans were performed for the complexes, saving one solution per orientation. The same translation and rotation grid intervals were used in all the scans. Also, the molecules were treated as previously described; hence water molecules and ions were omitted, incomplete side chains were modeled and docking to the surface of the Fc fragments of antibodies was prevented.[2]

The results of our docking searches are presented below in terms of the geometric–hydrophobic complementarity scores for nearly correct predictions, i.e., predictions that deviate by less than 3 Å (RMSD for all the common Cα atoms) from the experimental structures for complexes, and less than 3.5 Å for the considerably larger tetramers. We also list the ranks of the nearly correct solutions, which are their positions in the lists of predicted solutions, sorted by the scores.

RESULTS AND DISCUSSION

The Hydrophobic Complementarity Parameter H^a

We used virtual scans to determine the parameter H^a, as described in the Methods section. Real and virtual scans with different H^a values were calculated for seven independent subunits of D_2 tetramers (1gic/a, 1gic/b, 1ftr/a, 1ftr/b, 1ftr/c, 1a2z/d, 1rhp/b, and 1hdc/c) and for five complexes (2pts, 1cho, 1fss, 1dvf, and 1vfb). These cases were chosen because geometric docking did not rank a nearly correct

TABLE I. List of D_2 Tetramers and Complexes of Soluble Proteins Used to Examine the Contribution of Hydrophobic Complementarity at the Interface and in Docking

System	PDB[20] code[a]	Reference
4-Hydroxybenzoyl CoA thioesterase	1bvq	21
Dialkylglycine decaboxylase	1dge	22
Enol acyl carrier protein reductase	1enp	23
Fructose-1,6-bisphosphatase	1bk4	24
Concanavalin-A	1gic	25
Glycine N-methyl transferase	1xva	26
Tetrahydromethanopterin formyltransferase	1ftr	27
Phosphoglycerate mutase 1	1bq4	28
D-glyceraldehyde-3-phosphate dehydrogenase	1a7k	29
Pyrrolidone carboxyl aminopeptidase	1a2z	30
Fructose 1,6-bisphosphate aldolase	1ado	31
Phosphofructokinase	6pfk	32
Platelet factor 4	1rhp	33
NADP-dependent alcohol dehydrogenase	1ykf	34
3α, 20β-Hydroxysteroid dehydrogenase	1hdc	35
L-Asparagine amidohydrilase	4eca	36
β-Trypsin/BPTI	2ptc:2ptn/4pti	37, 38
Barnase/barstar	2brs:1bni[b]/1bta	39–41
Subtilisin/eglin-C	2sec:1scd/1tec[c]	42–44
α-Chymotripsin/HPTI	1cho:5cha/1ovo	45–47
Tripsin/soy-bean inhibitor	1avw:1ept/1avu[b]	48, 49
Thermitase/eglin-C	1tec:1thm/2sec[c]	42, 44, 50
Thrombin/BPTI	1bth:4htc[b]/4pti	37, 51, 52
Acetypcholinesterase/fasciculin-II	1fss:2ace/1fsc	53–55
β-Lactamase TEM1/BLIP[d]		56–58
Ab Hyhel-10/lysozyme	3hfm:3hfm[c]/1lza	59, 60
Ab Hyhel-5/lysozyme	3hfl:3hfl[e]/lza	60, 61
Ab D1.3/lysozyme	1vfb:1vfa/1lza	60, 62
Ab Je142/HPR	2jel:2jel[e]/1poh	63, 64
Ab D1.3/Ab E5.2	1dvf:1vfa/1dvf[e]	62, 65
Ab 5G9 (Fv)/tissue factor	1ahw:1fgn/1boy	66, 67
Ab Hulysll (Fv)/lysozyme	1bvk:1bvl/3lzt	68–70
Ab Hyhel-63 (Fv)/lysozyme	1dqj:1dqq/3lzt	71, 70
Ab D44.1 (Fv)/lysozyme	1mlc:1mlb/3lzt	72, 70
Ab E8 (Fv)/cytochrome C	1wej:1qbl/1hrc	73, 74
Fyn tyrosine kinase/SH_3 domain	1avz:1avv/1shf	75, 76
Ras activating domain/Ras	1wql:1wer/5p21	77–79
Methylamine dehydrogenase/amicyanin	2mta:2bbk/1aan	80–82
Che A/Che Y	1a0o:1chn/1a0o[c]	83–84

[a]For complexes the PDB codes of the complex is given first followed by the PDB codes of the component molecules.
[b]Several side chains on the surface of the molecule are not complete. These side chains were completed automatically.[2]
[c]The structure of the unbound molecule is not known. We chose a structure from another complex to represent a situation in which the conformation of the docked proteins is not that of the bound molecules.
[d]We thank N. Strynadka and M. James for giving us the coordinates of the β-lactamase, BLIP, and their complex.
[e]The structure of the unbound molecule is not known; therefore, the structure of the bound molecule was used.

structure first. For tetramers, the most prominent result was the wide minima in the graphs of predicted rank of the nearly correct solution versus H^{α} (data not shown). Thus, $1.5 > H^{\alpha} > 0.8$ was adequate for the *residues* representation, $1.1 > H^{\alpha} > 0.8$ was adequate for the *residues*+ representation, and $1.0 > H^{\alpha} > 0.7$ was acceptable for the *fragments* representation. We also noted that when several subunits from the same system were tested the dependence of the rank on H^{α} was very similar. All the computations for tetramers were performed with $H^{\alpha} = 1.3$ for the *residues* representation, $H^{\alpha} = 1.0$ for the *residues*+ representation and $H^{\alpha} = 0.95$ for the *fragments* representation.

The optimization of H^{α} for complexes gave somewhat different results. The dependence of the rank on H^{α} was weak and very shallow minima in the graphs of predicted

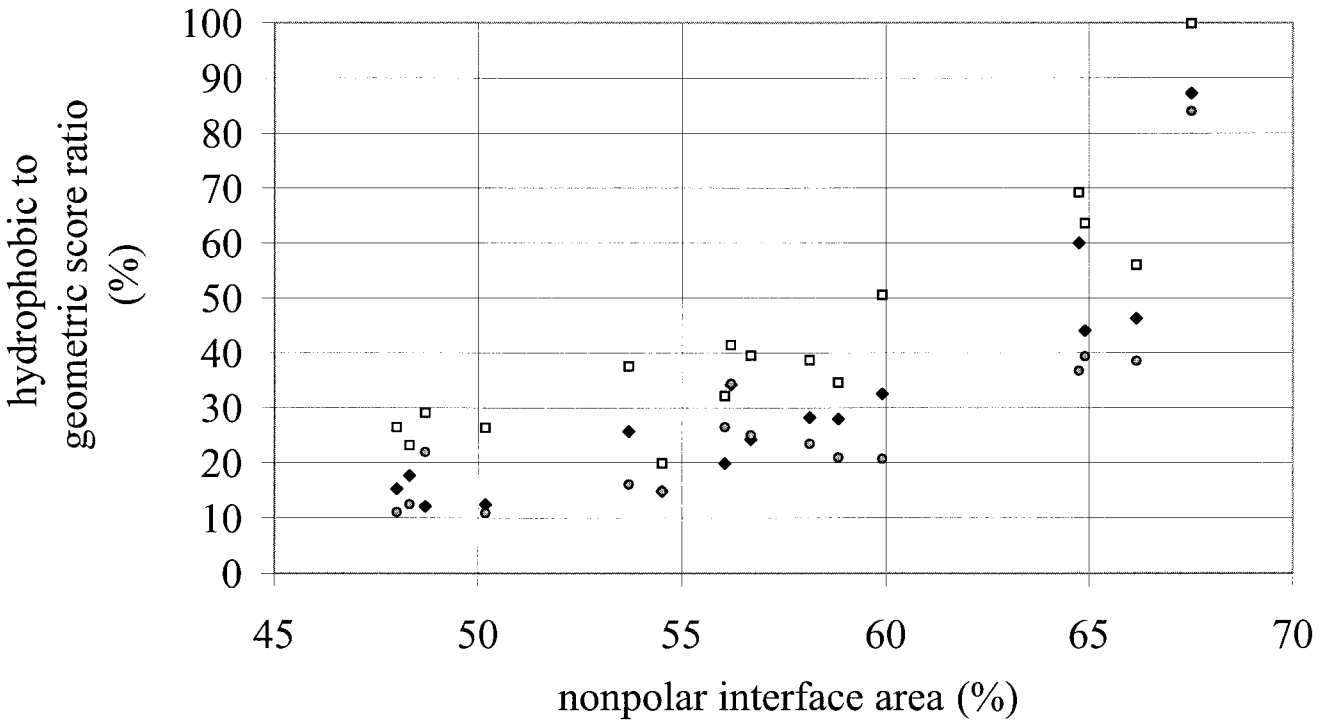

Fig. 1. The relation between the hydrophobic complementarity and the non-polar interface area for 16 D_2 tetramers. The black diamonds, empty squares and circles represent scores obtained with the *residues, residues+* and *fragments* representations, respectively. We consider the ratio between the hydrophobic and geometric scores and not the scores themselves in order to eliminate the bias that may arise from small differences in the grid intervals employed in the docking calculations for the various systems.

rank versus H^a appeared for much smaller values of H^a: $0.6 > H^a > 0.4$ for the *residues* and *residues+* representations and $0.4 > H^a > 0.3$ for the *fragments* representation. For complexes we used $H^a = 0.5$ with the *residues* and *residues+* representations and $H^a = 0.4$ with the *fragments* representation. As mentioned in the Methods section, the H^a values are chosen to produce the lowest ranks of the nearly correct solutions by limiting the formation of false positive solutions. The much smaller optimal H^a values for complexes compared to oligomers indicate that for complexes false solutions with appreciable hydrophobic complementarity are more readily formed than for oligomers. Hence, in complexes the hydrophobic descriptors distinguish less efficiently the interface from the non-interacting surface. This is in line with the smaller average difference between the percent of non-polar area at the interface and on the non-interacting surface in complexes compared to oligomers (see Tables II and III).

Hydrophobic Complementarity in D_2 Tetramers

D_2 tetramers were formed from 48 independent subunits using the **ab/ac** tetramer-forming algorithm.[19] Table II lists the scores and ranks of the nearly correct tetramers in the different predictions. The increase in score upon inclusion of hydrophobic complementarity is referred to hereafter as the "hydrophobic score." Dividing the hydrophobic score by H^a provides a measure for the hydrophobic complementarity. The geometric–hydrophobic scores for the tetramers are consistently higher than the geometric scores, indicating that the hydrophobic complementarity obtained with the *residues* representation is 12–83% of the geometric complementarity; with the *residues+* representation the ratio is 20–100% and with the *fragments* representation it is 11–77% (averaging over the several subunits in each system and taking into consideration that $H^a \neq 1$).

The ratio of the hydrophobic score to the geometric score for tetramers (combining the **ab** and **ac** interfaces) appears to be correlated with the percent of non-polar interface area, as seen in Figure 1. The correlation coefficient between the two measures is 0.85 for the *residues* representation, 0.83 for the *residues+* representation and 0.76 for the *fragments* representation. Notably, the dependence in Figure 1 is not linear and the percent of hydrophobic score is smaller than the percent of non-polar interface area for most of the range. Hence, only when the interface is highly non-polar the hydrophobic complementarity is extensive, and in most cases only partial hydrophobic complementarity is observed. We also note that the non-polar interface area is linearly related to the overall interface area (correlation coefficient of 0.97; data not shown). Considering the two relationships together we conclude that the non-polar portions of large interfaces are more often juxtaposed than non-polar portions of smaller interfaces.

Interestingly, the conclusion above does not reflect similarity of the amino-acid composition or spatial distribution in large or small interfaces. For example, the very large

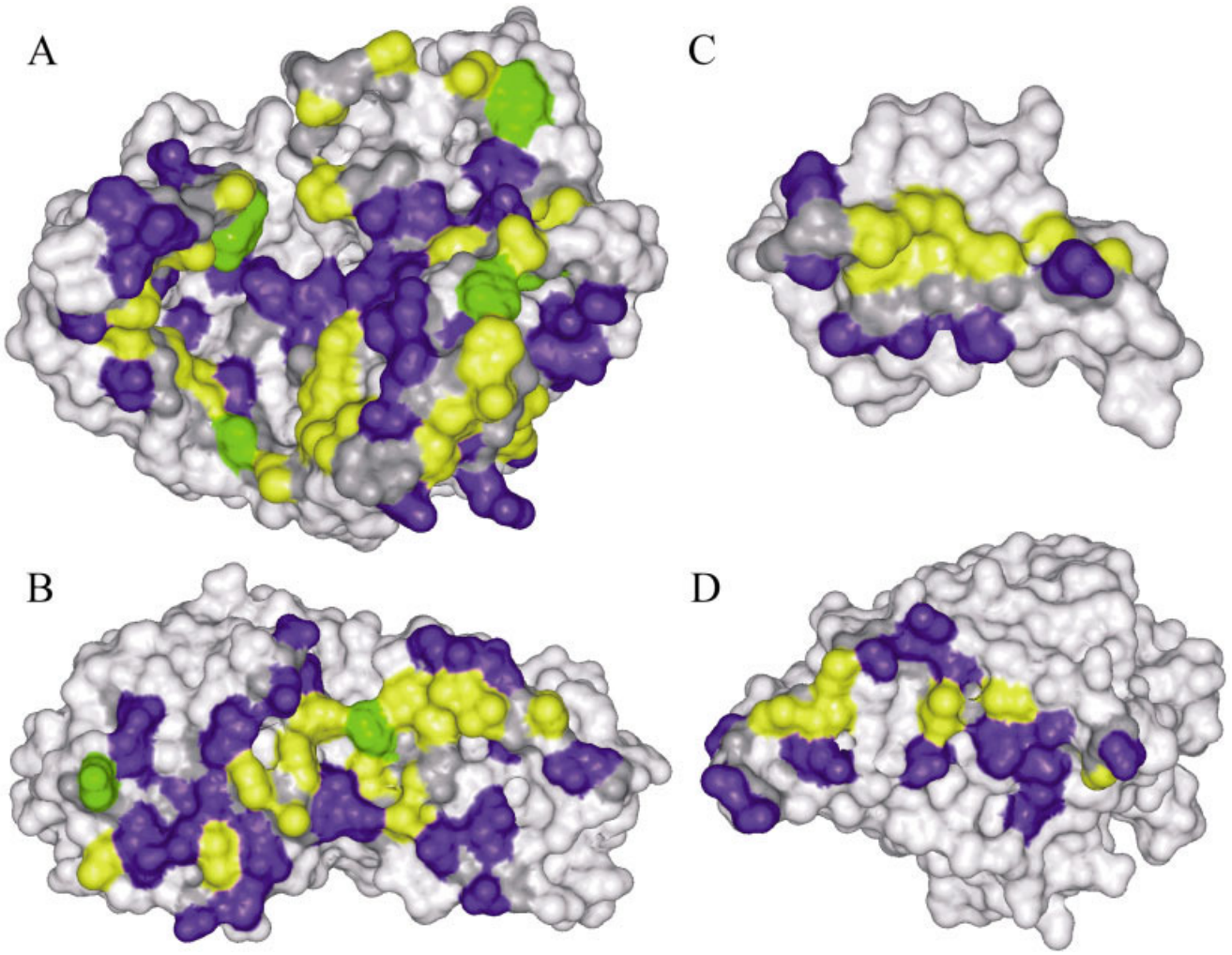

Fig. 2. Examples of different spatial distributions of hydrophobic and hydrophilic residues in large (A, B) and small interfaces (C, D) in D_2 tetramers. The side chains of residues whose fractional accessibility changes by more than 0.1 are colored as follows: Hydrophobic side chains in yellow, tyrosine tryptophane and cysteine in green and hydrophilic residues in purple. The backbone atoms of all the interface residues are colored in dark gray. **A**: The A-B interface in 1dge. **B**: The A-B interface in 6pfk. **C**: The A-B interface in 1rhp. **D**: The A-C interface in 1gic.

A-B interface in 1dge (8524 Å^2) is 61% non-polar and the hydrophobic score for this interface is large. However, looking at Figure 2(A) we note a large hydrophilic patch near the center of the interface, surrounded by smaller hydrophobic patches, which are further surrounded by hydrophilic residues. In contrast, the large A-B interface in 6pfk (4296 Å^2) is characterized by a central hydrophobic patch surrounded by hydrophilic residues [see Fig. 2(B)]. Similarly, different small interfaces have different spatial distributions of hydrophobic residues. In the A-B interface of 1rhp (1596 Å^2) a large hydrophobic patch can be seen with hydrophilic residues around it [see Fig. 2(C)], whereas the A-C interface in 1gic (1904 Å^2) is characterized by a mixture of hydrophobic and hydrophilic residues [see Fig. 2(D)].

The observation that there is disproportionately less hydrophobic complementarity for smaller interfaces is further demonstrated in the comparison of the scores for the **ab** and **ac** interfaces. Figure 3 presents the ratio of scores for the **ac** and **ab** dimers in the nearly correct tetramers. In geometric docking this ratio is always smaller than 1 because the **ab** score is by definition larger than the **ac** score. In geometric–hydrophobic docking the ratio is smaller than in geometric docking for systems in which there is a significant difference between the **ab** and **ac** interface areas (area(**ab**)/area(**ac**) $>$ 1.2 for 1ftr, 1dge, 1xva, 4eca, 6pfk, 1ykf, 1bvq, 1a7k, 1bk4, and 1rhp). For most of the systems the difference in the **ac** to **ab** score ratios, with and without hydrophobic complementarity, is more pronounced for the *residues* and *residues+* representations than for the *fragments* representation, and in general the score ratios obtained with the *fragments* representation are similar to the corresponding values in geometric docking (see Fig. 3). We conclude that for this selection of systems the *fragments* representation is a weaker hydrophobic descriptor than the *residues* or *residues+* representations. Therefore, geometric–hydrophobic docking with the *fragments* representation is less likely to improve the results of geometric docking than the other hydrophobic representations.

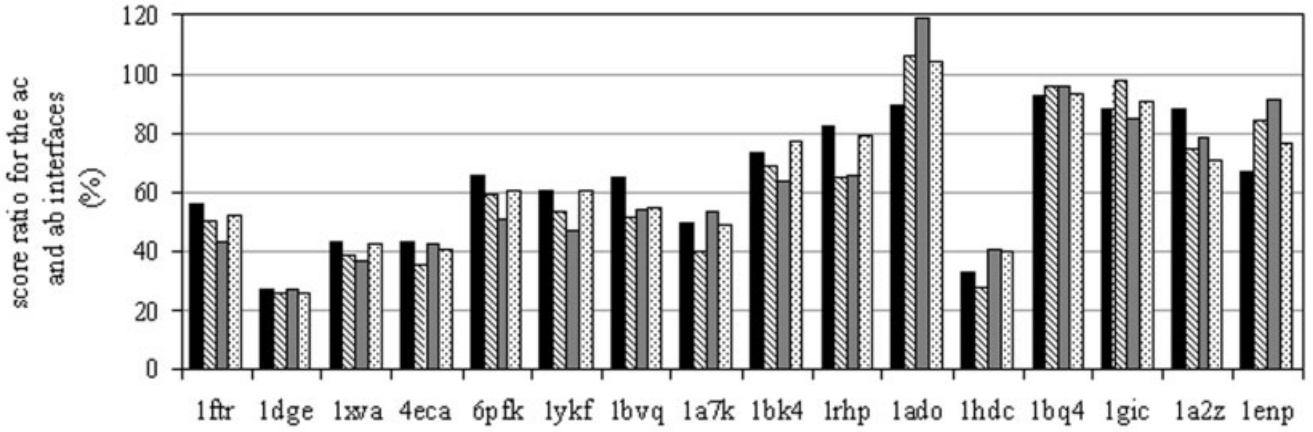

Fig. 3. The ratio of the scores for the **ac** and **ab** interfaces in geometric and geometric–hydrophobic docking. The columns are ordered as follows: Black for geometric docking, hatched for the *residues* representation, solid gray for the *residues*+ representation and dotted for the *fragments* representation. The average difference between the ratios due to introduction of hydrophobic complementarity is 0.09, 0.11, and 0.04 for the *residues, residues*+ and *fragments* representations, respectively.

Reassembling D_2 Tetramers via Geometric–Hydrophobic Docking

Table II also lists the ranks of the nearly correct tetramers obtained by applying the **ab/ac** tetramer-forming algorithm[19] to lists of C_2 homodimers obtained via geometric or geometric–hydrophobic docking. Four of the tetramers crystallize with a single subunit in the asymmetric unit (1bvq, 1dge, 1enp, 1bk4). These are perfect tetramers and the predictions correspond to bound docking. Not surprisingly the ranks of the nearly correct predictions for the perfect tetramers are 1 (2 in one case) with either geometric or geometric–hydrophobic docking. The other 12 tetramers crystallize as a dimer or a tetramer in the asymmetric unit. Reassembly of tetramers was performed starting from each of the 44 independent subunits in these 12 systems, resembling unbound docking. When only geometric complementarity is considered, we find a nearly correct tetramer ranked 1 in 32 out of these 44 cases. Introduction of hydrophobic complementarity significantly improves the results. Thus, with the *residues* and *residues*+ representations the rank improves in 11 of the 12 possible cases and with the *fragments* representation it improves in 10 cases. Particularly interesting is the case of 1hdc/d, for which a nearly correct tetramer was not formed from dimers obtained via geometric docking. The introduction of hydrophobic complementarity in this case produced high-ranking nearly correct solutions. It is also important to note that in only a few cases new false positive tetramers are formed when hydrophobic complementarity is considered, leading to a minor deterioration of the rank in 1 case with the *residues* or *residues*+ representations and in three cases with the *fragments* representation. We conclude that hydrophobic complementarity is important in subunit–subunit recognition in tetramers and therefore useful in docking. The *fragments* representation is slightly less successful in pointing out the nearly correct solution than the other hydrophobic representations, as expected from the hydrophobic complementarity analysis above.

The different levels of hydrophobic complementarity in large and small interfaces lead us to test the possibility of combining geometric and geometric–hydrophobic docking in a two-stage reassembly of D_2 tetramers. This can be readily done with our **ab/cd** tetramer-forming algorithm, which forms C_2 dimers in a first docking step and uses them to form tetramers in a second docking step.[19] We assumed that geometric–hydrophobic docking in the first step would identify better the larger **ab** interface, whereas geometric docking in the second step would identify the smaller **ac** interface. This approach worked well for 1ftr, in which the **ab** and **ac** interfaces differ considerably in size (5670 and 1482 Å^2, respectively) and hydropathy (61.2 and 54.8% non-polar, respectively). Thus, for 1ftr/a the combination of geometric–hydrophobic docking in the first step with either geometric or geometric–hydrophobic docking in the second step improved the rank of the nearly correct prediction to 1. No such improvement was observed when geometric docking in the first step was combined with geometric–hydrophobic docking in the second step. We also tested the combined geometric and geometric–hydrophobic prediction procedure on 1bq4, in which the two interfaces are small compared to other D_2 tetramers, and similar in size (1816 and 1657Å^2) and hydropathy (48.1 and 48.6% non-polar). As expected, the rank obtained for 1bq4/d with the **ab/cd** algorithm employing geometric docking in both steps was better than the rank obtained with any combination. The results for 1ftr/a and 1bq4/d follow the expected trend: when geometric–hydrophobic docking is used to predict a large interface, it improves the results. However, this is not general and for 1rhp/b any combination of geometric and geometric–hydrophobic docking is helpful despite the small **ab** and **ac** interfaces (1596 and 1283 Å^2). This is probably due to the high non-polar proportion of the A-B and A-C interfaces in 1rhp [59 and 52%, respectively and see Fig. 2(C)].

Another parameter that may affect the rank of the nearly correct solution is the hydropathy of the non-interacting surface that takes part in the formation of false solutions. Table II lists the percent of non-polar area at the **ab** plus **ac** interfaces and on the non-interacting surface of each tetramer. The latter value is

TABLE II. Comparison of the Nearly Correct Tetramers Obtained with the ab/ac Tetramer-Forming Algorithm Combined With Geometric or Geometric–Hydrophobic Docking With the *residues*, *residues+* or *fragments* Representations

	Geometric		Geometric–hydrophobic						Non-polar area (%)
			residues		*residues+*		*fragments*		
System	Score[b]	Rank	Score	Rank	Score	Rank	Score	Rank	I; S[a]
1bvq	2868	1	3320	2	3702	1	3466	1	49 ; 51
1dge	8592	1	11746	1	11912	1	10504	1	59 ; 50
1enp	2786	1	4960	1	4714	1	3758	1	65 ; 52
1bk4	4652	1	6342	1	6260	1	5578	1	59 ; 52
1gic/a	1960	9	2474	2	2536	5	2452	1	56 ; 46
1gic/b	2216	2	2782	1	2792	1	2774	1	
1xva/a	3156	1	3818	1	3976	1	3478	1	48 ; 52
1xva/b	3188	1	3786	1	4046	1	3532	1	
1ftr/a	2416	10	3658	2	3872	2	3074	1	60 ; 45
1ftr/b	2832	2	3796	2	4022	2	3374	3	
1ftr/c	2502	11	3932	1	4216	1	2918	10	
1ftr/d	2844	1	3684	1	3838	1	3314	1	
1bq4/a	2440	1	2894	1	3004	1	2704	1	48 ; 49
1bq4/b	2144	1	2592	1	2594	1	2350	2	
1bq4/c	2508	1	3178	1	3098	1	2856	1	
1bq4/d	1924	85	2424	32	2410	28	2174	87	
1a7k/a	3690	1	5524	1	5798	1	4776	1	65 ; 53
1a7k/b	3568	1	5616	1	5704	1	4688	1	
1a7k/c	3868	1	5730	1	5944	1	5160	1	
1a7k/d	3800	1	5656	1	5986	1	5060	1	
1a2z/a	2386	1	3752	1	3616	1	3184	1	66 ; 49
1a2z/b	2396	1	3632	1	3580	1	3096	1	
1a2z/c	2214	1	3490	1	3374	1	2912	1	
1a2z/d	2102	2	3254	1	3198	1	2860	1	
1ado/a	2848	1	3416	1	3500	1	3264	1	55 ; 53
1ado/b	3428	1	3982	1	4064	1	3834	1	
1ado/c	2970	1	3592	1	3566	1	3424	1	
1ado/d	3042	1	3666	1	3602	1	3498	1	
6pfk/a	3422	1	4532	1	4820	1	4252	1	57 ; 51
6pfk/b	3838	1	4966	1	5238	1	4726	1	
6pfk/c	3516	1	4720	1	4954	1	4396	1	
6pfk/d	3344	1	4346	1	4686	1	4100	1	
1rhp/a	2110	1	2802	1	2908	1	2772	1	56 ; 53
1rhp/b	1564	35	2552	2	2392	2	2238	6	
1rhp/c	1854	1	2550	1	2510	1	2370	1	
1rhp/d	2004	1	2974	1	2838	2	2612	1	
1ykf/a	3776	1	5052	1	5202	1	4428	1	54 ; 50
1ykf/b	3474	1	4768	1	4948	1	4028	1	
1ykf/c	4170	1	5628	1	5798	1	4788	1	
1ykf/d	4376	1	5618	1	5772	1	4964	1	
1hdc/a	4078	1	6330	1	6114	1	5190	1	68 ; 51
1hdc/b	3926	1	6058	1	5818	1	5062	1	
1hdc/c	2290	28	5630	4	5492	1	4778	1	
1hdc/d	–	>100	5442	3	4544	3	3032	5	
4eca/a	5384	1	6266	1	6896	1	5952	1	50 ; 48
4eca/b	5884	1	7040	1	7500	1	6468	1	
4eca/c	5906	1	6738	1	7352	1	6526	1	
4eca/d	6112	1	6994	1	7672	1	6750	1	

[a]I stands for the interface (ab + ac) and S stands for the non-interacting surface of the tetramer.
[b]The complementarity score for a tetramer is twice the sum of the scores for the ab and ac dimers.[19]

significantly smaller than the former (difference of 2% or more) for 11 of the 16 systems. For these systems the improvement in the rank of the nearly correct solution obtained with geometric–hydrophobic docking is somewhat larger than for the four systems in which the percent of non-polar surface is similar at the interface and on the non-interacting surface (see for example the results for 1hdc and 1bq4 in Table II). In only one case the percent of non-polar area on the non-interacting surface is slightly larger than at the interface (1xva). This, however, does not lead to formation of false positive predictions with either hydrophobic representation and the rank of the nearly correct solution remains 1.

TABLE III. Prediction of the Structure of D_2 Tetramers From Modeled Subunits

Modeled tetramer	Modeling template	Non-polar area (%) I ; S[a]	Score and rank of the nearly correct prediction	
			Geometric docking	Geometric–hydrophobic docking[b]
1bdb: cis-biphenyl-2,3-dihydrodiol-2,3-dehydrogenase from *Pseudomonas Sp.*[86]	1ae1: tropinone reductase-I from *Datura stramonium*[87] tetramer	68 ; 52	2012 ; 23	3148 ; 21
1ar5: Fe-Mn superoxide dismutase from *Propionibacterium shermanii*[88]	1vew: Mn superoxide dismutase from *Escherichia coli*[89] dimer	58 ; 54	1726 ; 132	2568 ; 67
1nhk: nucleoside diphosphate kinase from *Myxococcus xanthus*[90]	1b99: nucleoside diphosphate kinase from *Dictyostelium discoideum*[91] hexamer	58 ; 53	1670 ; 104	NF

[a]See footnote a in Table II.
[b]The *residues+* representation was used in these computations.

Modeling D_2 Tetramers via Geometric–Hydrophobic Docking

In a previous study we successfully applied the tetramer-forming procedures based on geometric docking to model-structures of single subunits. Hence, we combined comparative modeling with docking to form D_2 tetramers that cannot be modeled by comparative modeling because the modeling template is not a tetramer and new subunit–subunit interfaces must be predicted.[19] In this study we test if introduction of hydrophobic complementarity improves the predictions. Table III lists the modeled subunits, the modeling templates and the scores and ranks of the predicted nearly correct tetramers obtained via geometric or geometric–hydrophobic docking with the *residues+* representation, which appeared most successful in the reassembly of tetramers.

For the first system, 1bdb, the contribution of hydrophobic complementarity to the score is considerable but it does not affect the prediction considerably. The best nearly correct tetramer obtained via geometric–hydrophobic docking is more accurate than the tetramer obtained via geometric docking (3.0 versus 3.9 Å RMSD), but its rank is barely changed (21 versus 23). In the second case, 1ar5, hydrophobic complementarity improves the results considerably. The nearly correct tetramer obtained with geometric–hydrophobic docking is more accurate than the tetramer obtained with geometric docking (4.0 versus 4.5 Å) and it is ranked higher (67 versus 132). In the third case, 1nhk, a nearly correct tetramer was not formed via geometric–hydrophobic docking, whereas with geometric docking such a tetramer was ranked 104.

The success or failure of geometric–hydrophobic docking for predicting the structures of the three tetramers above appears to be related mostly to the relative sizes of the interfaces. In 1ar5 the A-D interface is large, 3725 Å^2 and 2.1 times larger than the A-B interface (1789 Å^2). In 1bdb the A-D interface is also large, 3306 Å^2, and it is 1.4 times larger than the A-C interface (2233 Å^2). In 1nhk both the R-L and R-A interfaces are small, 1839 and 760 Å^2, respectively. The correct R-A dimer is not formed in the geometric–hydrophobic docking scan possibly because the non-polar fraction of this interface is similar to that of the non-interacting surface (53%).

Hydrophobic Complementarity in Complexes of Soluble Systems

Table I lists 23 protein–protein complexes that were used for testing hydrophobic complementarity in complexes of soluble proteins, and Table IV compares the geometric, geometric–electrostatic and geometric–hydrophobic docking results for these systems (unbound–unbound or bound–unbound docking). One notes that with either hydrophobic representation the score increases compared to the geometric score or remains unchanged (when the *residues* representation is used and either one or both interacting molecules do not display hydrophobic side chains at the interface). These results indicate that there is hydrophobic complementarity at the interface in complexes of soluble proteins. However, the ratio of hydrophobic complementarity to geometric complementarity is considerably smaller than the corresponding values for tetramers: 0–38% with the *residues* representation, 4–42% with the *residues+* representation and 4–33% with the *fragments* representation (taking into consideration the H^a values, which are smaller than 1). Notably, for these 23 complexes the percent of hydrophobic complementarity is smaller than the percent of non-polar interface area (see Fig. 4), indicating that the hydrophobic complementarity is partial, as described above for interfaces of similar size in D_2 tetramers. The dependence in Figure 4 is considerably smoother for the *fragments* and *residues* representations than for the *residues+* representation, which distinguishes antibody–antigen systems (see below).

The next question concerning hydrophobic complementarity in complexes is whether the correct interface is significantly different from the non-interacting surface. We note a distinction between the group of 9 enzyme-inhibitor systems, for which the percent of non-polar area is generally larger at the interface than on the non-interacting surface and the group of 10 antibody-antigen systems, in which a weak tendency in the opposite direction is observed (a similar observation was mentioned

TABLE IV Geometric–Hydrophobic Docking for Complexes of Soluble Molecules Using the Unbound Structures

System	Non-polar area (%) I ; S[a]	Geometric docking (Score; rank[b])	Geometric–electrostatic (Score; rank)	Geometric–hydrophobic			Geometric–electrostatic–hydrophobic (Score; rank)
				residues (Score; rank)	*residues+* (Score; rank)	*fragments* (Score; rank)	
2ptn/4pti	52 ; 47	425 ; 271–278	516 ; 63–64	441 ; 259–264	451 ; 316–331	444 ; 269–275	544 ; 50–52
1bni/1bta	49 ; 49	462 ; 10–11	659 ; 1	462 ; 16	480 ; 13	471 ; 15	677 ; 1
1scd/1tec	57 ; 49	423 ; 918–949	427 ; 961–992	452 ; 724–744	468 ; 894–915	454 ; 667–691	472 ; 755–778
5cha/1ovo	60 ; 47	423 ; 131–136	483 ; 27–29	450 ; 124–127	505 ; 30–32	460 ; 78–79	565 ; 5
1ept/1avu	55 ; 48	410 ; 2288–2348	425 ; 1688–1729	424 ; 2594–2669	447 ; 2183–2246	432 ; 1995–2047	456 ; 1178–1194
1thm/2sec	58 ; 49	480 ; 180–186	482 ; 170–176	529 ; 65–69	524 ; 175–181	505 ; 148–157	526 ; 156–161
4htc/4pti	51 ; 50	434 ; 508–531	559 ; 99–102	447 ; 686–707	646 ; 676–690	452 ; 595–613	585 ; 137–138
2ace/1fsc	53 ; 50	435 ; 689–705	978 ; 3	470 ; 395–404	526 ; 111–115	456 ; 640–655	1066 ; 1
TEM1/BLIP	59 ; 54	688 ; 1	610 ; 98–101	708 ; 1	777 ; 1	726 ; 1	699 ; 18
3hfm/1lza	44 ; 47	488 ; 119–128	636 ; 1096–1103	488 ; 168	527 ; 75–76	504 ; 119–122	675 ; 156
3hfl/1lza	47 ; 48	580 ; 1	594 ; 521–525	589 ; 1	589 ; 1	599 ; 1	808 ; 1
1vfa/1lza	43 ; 47	397 ; 932–961	400 ; 894–920	397 ; 1125–1160	445 ; 296–305	407 ; 1062–1103	448 ; 279–291
2jel/1poh	48 ; 48	425 ; 286–292	422 ; 381–389	440 ; 195	475 ; 55–59	440 ; 307–318	471 ; 85–92
1vfa/1dvf	42 ; 46	425 ; 2186–2262	437 ; 2390–2435	425 ; 2590–2670	464 ; 1132–1169	425 ; 2594–2670	470 ; 1215–1242
1fgn/1boy	44 ; 48	514 ; 61–62	514 ; 126–129	514 ; 77–79	535 ; 120–125	532 ; 65	535 ; 122–127
1bvl/3lzt	44 ; 45	467 ; 1236–1261	469 ; 2137–2189	467 ; 1306–1339	519 ; 533–549	484 ; 1153–1189	521 ; 490–497
1dqq/3lzt	44 ; 47	451 ; 296–303	449 ; 317–327	451 ; 321–331	492 ; 159–169	464 ; 288–299	490 ; 170–177
1mlb/3lzt	52 ; 45	427 ; 327–340	427 ; 327–339	440 ; 217–225	450 ; 283–294	439 ; 340–353	450 ; 282–292
1qbl/1hrc	46 ; 49	494 ; 267–272	495 ; 265–270	494 ; 306–314	508 ; 368–378	526 ; 176–179	509 ; 357–367
1avv/1shf	58 ; 53	267 ; 7556–7642[c]	267 ; 7562–7652	301 ; 5195–5332	310 ; 6081–6193	290 ; 6886–6994	310 ; 4230–4290
1wer/5p21	52 ; 50	478 ; 844–861	479 ; 727–854	468 ; 1443–1484	488 ; 1000–1024	525 ; 405–413	493 ; 841–873
2bbk/1aan	69 ; 49	459 ; 3550–3617	458 ; 3638–3708	547 ; 918–935	547 ; 1047–1069	520 ; 1602–1642	515 ; 1939–1978
1chn/1a0o	57 ; 51	510 ; 59–66	510 ; 58–64	581 ; 11	594 ; 4–5	568 ; 15–16	594 ; 4–5

[a]See footnote a in Table II.
[b]The rank is given as a range of numbers when the score of the nearly correct solution is the same as for other, false solutions. The range encompasses all these solutions.
[c]For this system no solution was found with RMSD < 3 Å. The numbers refer to the best solution with RMSD = 3.4 Å.

before[92]). Based on this observation one expects that for enzyme–inhibitor systems inclusion of hydrophobic complementarity will improve the ranks of nearly correct solutions more than for antibody–antigen systems. Indeed, for the enzyme–inhibitor systems the three hydrophobic representations have a similar effect, improving the rank for 5 of the 9 systems and lowering it for 2 or 3 systems. The results obtained with the *residues* and *fragments* representations for the antibody–antigen systems are also as expected, deterioration of the rank of the nearly correct solution for 7 or 5 of the 10 systems and a small improvement in 2 cases. In contrast, the *residues+* representation is particularly effective for antibody-antigen systems, improving the rank for 7 of the 10 systems, and the rank of the 8th system remains 1. This representation appears to be much better for distinguishing correct from incorrect solutions for antibody–antigen systems, probably due to the abundance of tyrosine and tryptophane residues in the antigen binding loops of antibodies. The systems 1avz, 1wq1, 2mta, and 1a0o (which are not enzyme–inhibitor neither antibody–antigen) have a higher ratio of non-polar area at the interface than on the non-interacting surface as the group of enzyme–inhibitor systems. Introduction of hydrophobic complementarity improves significantly the ranking of the nearly correct solutions for 3 of the 4 systems in this group.

Combining Geometric, Geometric–Electrostatic, and Geometric–Hydrophobic Docking

Currently our docking procedure tests either geometric, or geometric–electrostatic or geometric–hydrophobic complementarity, producing separate lists of solutions. However, nearly correct solutions are likely to be included in each of the three lists, that is in a subset of common solutions. For this subset of solutions it is possible to combine the geometric, electrostatic and hydrophobic scores to obtain a combined geometric-electrostatic-hydrophobic score, $S_{geo+elec+hydro} = S_{geo} + e \cdot S_{elec} + h \cdot S_{hydro}$, where S_{elec} is the difference between the geometric–electrostatic and the geometric scores and S_{hydro} is the difference between the geometric–hydrophobic and geometric scores for a given solution.

Combined geometric–electrostatic–hydrophobic (*residues+* representation) scores were calculated for the subset of common solutions for each system (3427–7046 common solutions out of 8760 solutions in each scan), with different values of e and h, and used to rerank the solutions. The best results were obtained with $e = h = 1$. This is not unexpected as the electrostatic and hydrophobic contributions were already scaled to match the geometric complementarity score. The combined scores and the new ranks of the nearly correct solutions are listed in Table IV. The latter are much better than the ranks in the

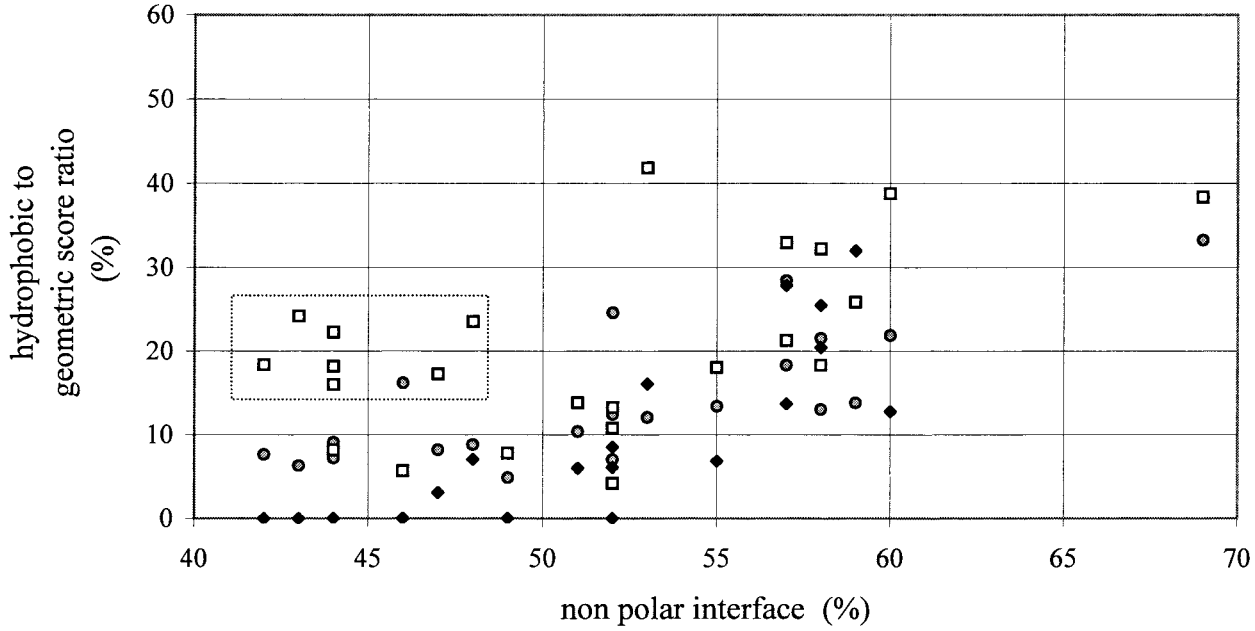

Fig. 4. The relation between the hydrophobic scores and the non-polar interface areas for 23 hetero-dimers of soluble proteins. The black diamonds, empty squares and circles represent scores obtained with the *residues*, *residues*+, and *fragments* representations, respectively. Note the large ratio obtained with the *residues*+ representation for most antibody-antigen complexes (seven points enclosed in a rectangle).

geometric scans for 17 of the 23 systems and remain unchanged for 2 other systems. Small deterioration of the rank is observed for 4 systems (tem1/blip, 3hfm/1lza, 1fgn/1boy and 1qbl/1hrc) for 3 of which geometric-electrostatic docking produced many false positive solutions (see the "good electrostatic docking rules" in Heifetz et al.[2]). Clearly the combination of electrostatic and hydrophobic complementarity (with the *residues*+ representation) with geometric complementarity improves the docking results for complexes of soluble proteins. This improvement is due, in some cases, to the higher combined score for the nearly correct solution. For example, in the system 5cha/1ovo both the electrostatic and the hydrophobic scores for the nearly correct solution are positive and large and help to point out this solution. It also reflects the weak anti-correlation between the effect of electrostatic and hydrophobic complementarity on the rank of the nearly correct solution. Some of the improvement in rank is also due to the elimination of false solutions, which are found in only one of the 3 lists (geometric, geometric–electrostatic or geometric–hydrophobic) and are eliminated in the intersection process.

CONCLUSIONS

We present in this study a geometric–hydrophobic docking algorithm and apply it to test the hydrophobic complementarity at the interface in oligomers (represented by D_2 tetramers) and in complexes of soluble proteins. We then test the effect of hydrophobic complementarity in docking. Thus, we compare the ability of the geometric–hydrophobic algorithm to distinguish the correct solution from false ones to the performance of geometric docking.

We find several interesting results: First, there is hydrophobic complementarity at the interface in oligomers and in complexes of soluble proteins. The level of such complementarity is considerably larger for the large interfaces in oligomers for which the portion of non-polar area is also large. For smaller interfaces in oligomers and for the comparable size interfaces in complexes, the hydrophobic complementarity is only partial. These observations suggest that the contribution of the hydrophobic effect in binding is expected to be variable, as previously noted by Tsai et al.[93]

Next, we find that the hydrophobic complementarity for the correct interface is generally more extensive than for false interfaces, leading to a significant improvement in the ranks of nearly correct reassembled and modeled tetramers. The *residues* and *residues*+ representations give better results for these systems than the *fragments* representation. For complexes we favor the *residues*+ representation, which improves the docking results for all the systems, including antibody–antigen systems. Combining geometric, electrostatic, and hydrophobic scores in unbound docking for complexes gives excellent results, with a nearly correct solution ranked < 10 for 5 of 23 systems, < 100 for 8 systems and < 1000 for 19 systems. In the future we shall combine geometric, electrostatic, and hydrophobic complementarity in a single scoring function.

ACKNOWLEDGMENTS

We thank Professor Ephraim Katchalski-Katzir for his support and advice throughout this study and the members of the group for stimulating discussion. This study was supported in part by the Daat Consortium for Developing Generic Technologies for Design and Development of Drugs and Diagnostics. The docking program MolFit is available in http://www.weizmann.ac.il/Chemical_Services/molfit

REFERENCES

1. Creighton TE. Protein structures and molecular properties. New York: W.H. Freeman, 1997.
2. Heifetz A, Katchalski-Katzir E, Eisenstein M. Electrostatics in protein–protein docking. Protein Sci 2002;11:571–587.
3. Argos P. An investigation of protein subunit and domain interfaces. Protein Eng 1988;2:101–113.
4. Janin J, Miller S, Chothia C. Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol 1988;204:155–164.
5. Jones S, Thornton JM. Principles of protein–protein interactions. Proc Natl Acad Sci USA 1996;93:13–20.
6. Lo Conte L, Chothia C, Janin J. The atomic structure of protein–protein recognition sites. J Mol Biol 1999;285:2177–2198.
7. Young L, Jernigan RL, Covell DG. A role for surface hydrophobicity in protein–protein recognition. Protein Sci 1994;3:717–729.
8. Miller S. The structure of interfaces between subunits of dimeric and tetrameric proteins. Protein Eng 1989;3:77–83.
9. Larsen TA, Olson AJ, Goodsell DS. Morphology of protein–protein interfaces. Structure 1998;6:421–427.
10. Glaser F, Steinberg DM, Vakser IA, Ben-Tal N. Residue frequencies and pairing preferences at protein–protein interfaces. Proteins 2001;43:89–102.
11. Vakser IA, Aflalo C. Hydrophobic docking: a proposed enhancement to molecular recognition techniques. Proteins 1994;20:320–329.
12. Ackermann F, Herrmann G, Posch S, Sagerer G. Estimation and filtering of potential protein–protein docking positions. Bioinformatics 1998;14:196–205.
13. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47:281–294.
14. Jackson RM, Gabb HA, Sternberg MJ. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 1998;276:265–285.
15. Palma PN, Krippahl L, Wampler JE, Moura JJ. BiGGER: a new (soft) docking algorithm for predicting protein interactions. Proteins 2000;39:372–384.
16. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 1992;89:2195–2199.
17. Eisenstein M, Shariv I, Koren G, Friesem AA, Katchalski-Katzir E. Modeling supra-molecular helices: extension of the molecular surface recognition algorithm and application to the protein coat of the tobacco mosaic virus. J Mol Biol 1997;266:135–143.
18. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–132.
19. Berchanski A, Eisenstein M. Construction of molecular assemblies via docking: modeling of tetramers with D2 Symmetry. Proteins 2003;53:817–829.
20. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
21. Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D, Holden HM. The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from Pseudomonas sp. Strain CBS-3. J Biol Chem 1998;273:33572–33579.
22. Hohenester E, Keller JW, Jansonius JN. An alkali metal ion size-dependent switch in the active site structure of dialkylglycine decarboxylase. Biochemistry 1994;33:13561–13570.
23. Rafferty JB, Simon JW, Baldock C, et al. Common themes in redox chemistry emerge from the X-ray structure of oilseed rape (Brassica napus) enoyl acyl carrier protein reductase. Structure 1995;3:927–938.
24. Weeks CM, Roszak AW, Erman M, Kaiser R, Jornvall H, Ghosh D. Structure of rabbit liver fructose 1,6-bisphosphatase at 2.3 A resolution. Acta Crystallogr D Biol Crystallogr 1999;55:93–102.
25. Harrop SJ, Helliwell JR, Wan TCM, Kalb AJ, Tong L, Yariv J. Structure solution of a cubic crystal of concanavalin a complexed with methyl alpha-D-glucopyranoside. Acta Crystallogr D Biol Crystallogr 1996;52:143.
26. Fu Z, Hu Y, Konishi K, et al. Crystal structure of glycine N-methyltransferase from rat liver. Biochemistry 1996;35:11985–11993.
27. Ermler U, Merckel M, Thauer R, Shima S. Formylmethanofuran: tetrahydromethanopterin formyltransferase from Methanopyrus kandleri - new insights into salt-dependence and thermostability. Structure 1997;5:635–646.
28. Rigden DJ, Walter RA, Phillips SE, Fothergill-Gilmore LA. Polyanionic inhibitors of phosphoglycerate mutase: combined structural and biochemical analysis. J Mol Biol 1999;289:691–699.
29. Kim H, Hol WG. Crystal structure of Leishmania mexicana glycosomal glyceraldehyde-3- phosphate dehydrogenase in a new crystal form confirms the putative physiological active site structure. J Mol Biol 1998;278:5–11.
30. Singleton M, Isupov M, Littlechild J. X-ray structure of pyrrolidone carboxyl peptidase from the hyperthermophilic archaeon Thermococcus litoralis. Structure Fold Des 1999;7:237–244.
31. Blom N, Sygusch J. Product binding and role of the C-terminal region in class I D-fructose 1,6-bisphosphate aldolase. Nat Struct Biol 1997;4:36–39.
32. Schirmer T, Evans PR. Structural basis of the allosteric behaviour of phosphofructokinase. Nature 1990;343:140–145.
33. Zhang X, Chen L, Bancroft DP, Lai CK, Maione TE. Crystal structure of recombinant human platelet factor 4. Biochemistry 1994;33:8361–8366.
34. Korkhin Y, Kalb AJ, Peretz M, Bogin O, Burstein Y, Frolow F. NADP-dependent bacterial alcohol dehydrogenases: crystal structure, cofactor-binding and cofactor specificity of the ADHs of Clostridium beijerinckii and Thermoanaerobacter brockii. J Mol Biol 1998;278:967–981.
35. Ghosh D, Erman M, Wawrzak Z, Duax WL, Pangborn W. Mechanism of inhibition of 3 alpha, 20 beta-hydroxysteroid dehydrogenase by a licorice-derived steroidal inhibitor. Structure 1994;2:973–980.
36. Palm GJ, Lubkowski J, Derst C, Schleper S, Rohm KH, Wlodawer A. A covalently bound catalytic intermediate in Escherichia coli asparaginase: crystal structure of a Thr-89-Val mutant. FEBS Lett 1996;390:211–216.
37. Marquart M, Walter J, Deisenhofer J, Bode W, Huber R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr B 1983;39:480.
38. Walter J, Steigemann W, Singh TP, Bartunik H, Bode W, Huber A. New algorithm to model protein–protein recognition based on surface complementarity. Applications to antibody-antigen docking. J Mol Biol 1992;228:277–297.
39. Buckle AM, Henrick K, Fersht AR. Crystal structural analysis of mutations in the hydrophobic cores of barnase. J Mol Biol 1993;234:847–860.
40. Buckle AM, Schreiber G, Fersht AR. Protein–protein recognition: crystal structural analysis of a barnase- barstar complex at 2.0-Å resolution. Biochemistry 1994;33:8878–8889.
41. Lubienski MJ, Bycroft M, Freund SM, Fersht AR. Three-dimensional solution structure and 13C assignments of barstar using nuclear magnetic resonance spectroscopy. Biochemistry 1994;33:8866–8877.
42. McPhalen CA, James MN. Structural comparison of two serine proteinase-protein inhibitor complexes: eglin-c-subtilisin Carlsberg and CI-2-subtilisin Novo. Biochemistry 1988;27:6582–6598.
43. Fitzpatrick PA, Ringe D, Klibanov AM. X-ray crystal structure of cross-linked subtilisin Carlsberg in water vs. acetonitrile. Biochem Biophys Res Commun 1994;198:675–681.
44. Gros P, Fujinaga M, Dijkstra BW, Kalk KH, Hol WGJ. Crystallographic refinement by incorporation of molecular dynamics. The thermostable serine protease thermitase complexed with eglin-c. Acta Crystallogr B 1994;45:488–499.
45. Fujinaga M, Sielecki AR, Read RJ, Ardelt W, Laskowski M, James MN. Crystal and molecular structures of the complex of alpha-chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution. J Mol Biol 1987;195:397–418.
46. Papamokos E, Weber E, Bode W, et al. Crystallographic refinement of Japanese quail ovomucoid, a Kazal-type inhibitor, and model building studies of complexes with serine proteases. J Mol Biol 1982;158:515–537.
47. Blevins RA, Tulinsky A. The refinement and the structure of the dimer of alpha-chymotrypsin at 1.67-Å resolution. J Biol Chem 1985;260:4264–4275.
48. Huang Q, Wang Z, Li Y, Liu S, Tang Y. Refined 1.8 Å resolution crystal structure of the porcine epsilon- trypsin. Biochim Biophys Acta 1994;1209:77–82.
49. Song HK, Suh SW. Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from Erythrina caffra and tissue-type plasminogen activator. J Mol Biol 1998;275:347–363.
50. Teplyakov AV, Kuranova IP, Harutyunyan EH, et al. Crystal

structure of thermitase at 1.4 Å resolution. J Mol Biol 1990;214: 261–279.
51. van de Locht A, Bode W, Huber R, et al. The thrombin E192Q-BPTI complex reveals gross structural rearrangements: implications for the interaction with antithrombin and thrombomodulin. Embo J 1997;16:2977–2984.
52. Rydel TJ, Tulinsky A, Bode W, Huber R. Refined structure of the hirudin-thrombin complex. J Mol Biol 1991;221:583–601.
53. Harel M, Kleywegt GJ, Ravelli RB, Silman I, Sussman JL. Crystal structure of an acetylcholinesterase-fasciculin complex: interaction of a three-fingered toxin from snake venom with its target. Structure 1995;3:1355–1366.
54. Raves ML, Harel M, Pang YP, Silman I, Kozikowski AP, Sussman JL. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (-)-huperzine A. Nat Struct Biol 1997;4:57–63.
55. LeDu MH, Housset D, Marchot P, Bougis PE, Navaza J, Fontecilla-Camps C. Structure of fasciculin 2 from green mamba snake venom: Evidence for unusual loop flexibility. Acta Crystallogr D Biol Crystallogr 1992;52:87.
56. Strynadka NC, Jensen SE, Alzari PM, James MN. A potent new mode of beta-lactamase inhibition revealed by the 1.7 Å X- ray crystallographic structure of the TEM-1-BLIP complex. Nat Struct Biol 1996;3:290–297.
57. Strynadka NC, Adachi H, Jensen SE, et al. Molecular structure of the acyl-enzyme intermediate in beta-lactam hydrolysis at 1.7 Å resolution. Nature 1992;359:700–705.
58. Strynadka NC, Jensen SE, Johns K, et al. Structural and kinetic characterization of a beta-lactamase-inhibitor protein. Nature 1994;368:657–660.
59. Padlan EA, Silverton EW, Sheriff S, Cohen GH, Smith-Gill SJ, Davies DR. Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab-lysozyme complex. Proc Natl Acad Sci USA 1989;86:5938–5942.
60. Maenaka K, Matsushima M, Song H, Sunada F, Watanabe K, Kumagai I. Dissection of protein-carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity. J Mol Biol 1995;247:281–293.
61. Cohen GH, Sheriff, S., Davies, D. R. Refined structure of the monoclonal antibody HyHEL-5 with its antigen hen egg-white lysozyme. Acta Crystallogr D Biol Crystallogr 1996;52:315.
62. Bhat TN, Bentley GA, Boulot G, et al. Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci USA 1994;91:1089–1093.
63. Prasad L, Waygood EB, Lee JS, Delbaere LT. The 2.5 Å resolution structure of the jel42 Fab fragment/HPr complex. J Mol Biol 1998;280:829–845.
64. Jia Z, Quail JW, Waygood EB, Delbaere LT. The 2.0-Å resolution structure of Escherichia coli histidine-containing phosphocarrier protein HPr. A redetermination. J Biol Chem 1993;268:22490–22501.
65. Braden BC, Fields BA, Ysern X, et al. Crystal structure of an Fv-Fv idiotope-anti-idiotope complex at 1.9 Å resolution. J Mol Biol 1996;264:137–151.
66. Huang M, Syed R, Stura EA, Stone MJ, Stefanko RS, Ruf W, Edgington TS, Wilson IA: The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab 5G9 and TF.G9 complex. J Mol Biol 1998;275:873–894.
67. Harlos K, Martin DM, O'Brien DP, Jones EY, Stuart DI, Polikarpov I, Miller A, Tuddenham EG, Boys CW: Crystal structure of the extracellular region of human tissue factor. Nature 1994;370:662–666.
68. Holmes MA, Buss TN, Foote J: Conformational correction mechanisms aiding antigen recognition by a humanized antibody. J Exp Med 1998;187:479–45.
69. Holmes MA, Foote J. Structural consequences of humanizing an antibody. J Immunol 1997;158:2192–2201.
70. Walsh MA, Schneider TR, Sieker LC, Dauter Z, Lamzin VS, Wilson KS: Refinement of triclinic hen egg-white lysozyme at atomic resolution. Acta Crystallogr D Biol Crystallogr 1998;54:522–546.
71. Li Y, Li H, Smith-Gill SJ, Mariuzza RA: Three-dimensional structures of the free and antigen-bound Fab from monoclonal antilysozyme antibody HyHEL-63(,). Biochemistry 2000;39:6296–6309.
72. Braden BC, Souchon H, Eisele JL, Bentley GA, Bhat TN, Navaza J, Poljak RJ: Three-dimensional structures of the free and the antigen-complexed Fab from monoclonal anti-lysozyme antibody D44.1. J Mol Biol 1994;243:767–781.
73. Mylvaganam SE, Paterson Y, Getzoff ED: Structural basis for the binding of an anti-cytochrome c antibody to its antigen: crystal structures of FabE8-cytochrome c complex to 1.8 Å resolution and FabE8 to 2.26 A resolution. J Mol Biol 1998; 281: 301–22.
74. Bushnell GW, Louie GV, Brayer GD: High-resolution three-dimensional structure of horse heart cytochrome c. J Mol Biol 1990;214:585–595
75. Arold S, Franken P, Strub MP, Hoh F, Benichou S, Benarous R, Dumas C: The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling. Structure 1997;5:1361–1372.
76. Noble ME, Musacchio A, Saraste M, Courtneidge SA, Wierenga RK: Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. Embo J 1993;12:2617–2624.
77. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, Schmitz F, Wittinghofer A: The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. Science 1997;277:333–338.
78. Scheffzek K, Lautwein A, Kabsch W, Ahmadian MR, Wittinghofer A: Crystal structure of the GTPase-activating domain of human p120GAP and implications for the interaction with Ras. Nature 1996;384:591–596.
79. Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A: Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. Embo J 1990;9:2351–2359.
80. Chen L, Durley RC, Mathews FS, Davidson VL: Structure of an electron transfer complex: methylamine dehydrogenase, amicyanin, and cytochrome c551i. Science 1994;264:86–90.
81. Chen L, Doi M, Durley RC, Chistoserdov AY, Lidstrom ME, Davidson VL, Mathews FS: Refined crystal structure of methylamine dehydrogenase from Paracoccus denitrificans at 1.75 Å resolution. J Mol Biol 1998;276:131–149.
82. Durley R, Chen L, Lim LW, Mathews FS, Davidson VL: Crystal structure analysis of amicyanin and apoamicyanin from Paracoccus denitrificans at 2.0 Å and 1.8 Å resolution. Protein Sci 1993;2:739–752.
83. Welch M, Chinardet N, Mourey L, Birck C, Samama JP: Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. Nat Struct Biol 1998;5:25–29.
84. Bellsolell L, Prieto J, Serrano L, Coll M: Magnesium binding to the bacterial chemotaxis protein CheY results in large conformational changes involving its functional surface. J Mol Biol 1994;238:489–495.
85. Chen R, Mintseris J, Janin J, Weng Z: A protein–protein docking benchmark. Proteins 2003;52:88–91.
86. Hulsmeyer M, Hecht HJ, Niefind K, et al. Crystal structure of cis-biphenyl-2,3-dihydrodiol-2,3-dehydrogenase from a PCB degrader at 2.0 Å resolution. Protein Sci 1998;7:1286–1293.
87. Nakajima K, Yamashita A, Akama H, et al. Crystal structures of two tropinone reductases: different reaction stereospecificities in the same protein fold. Proc Natl Acad Sci USA 1998;95:4876–4881.
88. Schmidt M, Meier B, Parak F. X-ray structure of the cambialistic superoxide dismutase from Propionibacterium Shermanii active with Fe or Mn. J Biol Inorg Chem 1996;1:532.
89. Edwards RA, Baker HM, Whittaker MM, Whittaker JW, Jameson GB, Baker EN. Crystal structure of Escherichia coli manganese superoxide dismutase at 2.1-angstrom resolution. J Biol Inorg Chem 1998;3:161–165.
90. Munoz Dorado J, Arias JM. Protein kinases and phosphatases during the developmental cycle of myxobacteria. Microbiologia 1995;11:376–378.
91. Gonin P, Xu Y, Milon L, et al. Catalytic mechanism of nucleoside diphosphate kinase investigated using nucleotide analogues, viscosity effects, and X-ray crystallography. Biochemistry 1999;38: 7265–7272.
92. Wodak SJ, Janin J. Structural basis of macromolecular recognition. Advances in Protein Chemistry 2003;61:9–73.
93. Tsai CJ, Xu D, Nussinov R: Structural motifs at protein–protein interfaces: protein cores versus two-state and three-state model complexes. Protein Sci 1997;6:1793–805.