Journal of Molecular Graphics and Modelling 27 (2009) 706–711

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Development of new indices to evaluate protein–protein interfaces: Assembling space volume, assembling space distance, and global shape descriptor

Miki H. Maeda [a,*], Kengo Kinoshita [b]

[a] Bioinformatics Research Unit, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
[b] Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minatoku, Tokyo 108-8639, Japan

ARTICLE INFO

Article history:
Received 14 July 2008
Received in revised form 30 October 2008
Accepted 3 November 2008
Available online 8 November 2008

Keywords:
Assembling space volume
Delaunay tessellation
Gap volume
Shape complementarity
Subunit assembly

ABSTRACT

Protein–protein interaction is fundamental to initiate the cellular functions of proteins, and thus structural analyses of protein interfaces are the first step to understand their functions at the molecular level. Although shape complementarities have been used to evaluate the fitness of interfaces, the conventional method did not distinguish between two main components of complementarity, the well-fitting of the surface shapes and the size of the gap regions between the pair of molecules, and could not evaluate the global shape of the interfaces. Therefore, we now propose three new indices to describe protein interfaces: assembling space volume (ASV), assembling space distance (AS-distance), and global shape descriptor (GS). The ASV is calculated using Delaunay tessellation, and the AS-distance is calculated as the ratio of ASV to delta-ASA (accessible surface area). The GS is calculated as the ratio of the volume of Delaunay tetrahedra with a long edge to that of all tetrahedra in the assembling space. To evaluate the feasibility of the three indices, we applied our method to homo-dimer proteins, and performed systematic comparisons with SURFNET by Laskowski and shape complementarity by Lawrence and Colman. As a result, our indices behave differently from the existing ones, and shed light on new features of protein–protein interfaces, as general trends of AS-distances for all protein interfaces.

1. Introduction

Protein functions are often initiated by interactions with partner molecules. Thus, for better understanding of protein structure–function relationships, the first step would be to analyze the molecular recognition in the vast amount of structural data accumulated by structural genomics projects [1]. The number of hetero-oligomer structures is not enough to massive analysis yet, but many homo-oligomer structures have been experimentally determined. In the process of molecular recognition of proteins, the concept of "shape complementarity" between a protein and a partner molecule has been a key to evaluate the fitness of the interaction sites.

The study of shape complementarity was begun by Lawrence and Colman [2]. They considered the direction and the distance between pairs of normal vectors on the molecular surfaces of proteins [3], one from protein A and the other from protein B. They calculated the complementarity score for each pair of the vectors, and then took the median value as an indicator to evaluate the complementarity of interfaces. In more recent methods, the local shape of molecular surfaces is also considered by using quadratic approximations [4], by using minimum and maximum curvature of the surfaces [5], by using Fourier correlation [6], or by the number of atoms in the interface [7]. Although shape complementarities can be divided into two main components conceptually, the degree of surface shape correspondence between convex and concave and the sizes of the gap regions between the pair of molecules, all of the above methods treated the two components at the same time or focused only on fitness. However, as we will show, the two components are not tightly dependent on each other, and the size of the gap regions can also be an important index of the complementarity of the interfaces. Therefore, in this paper, we address the calculation of the size of the gap regions in protein–protein interactions, and applied it to homo-dimer interfaces.

Laskowski was the first to attempt to calculate the volume in the interface, known as the gap volume, which can be obtained by the program SURFNET [8]. The gap volume is similar concept to our approach (assembling space volume, ASV) but completely different in calculation method, as shown later, and the results are quite different (for example, see Fig. 1). The most important difference is that SURFNET is based on the probe, while our method uses

* Corresponding author. Tel.: +81 29 838 7065; fax: +81 29 838 7065.
E-mail address: mmaeda@nias.affrc.go.jp (M.H. Maeda).

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.11.002

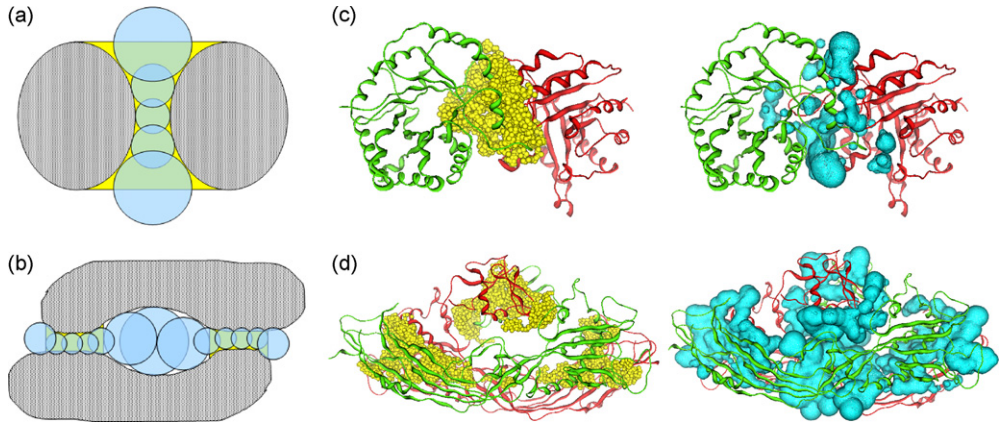

Fig. 1. Comparison between ASV and gap volume. Schematic comparisons are shown in (a) and (b), where the light blue spheres are the spheres used to define the gap volume in SURFNET, and the yellow-shaded portion indicates the assembling space defined by our method. The actual cases are shown in (c) for 1amk (triose phosphate isomerase, ASV = 1835 Å^3, gap volume = 2603 Å^3) and (d) for 1pre (proaerolysin, ASV = 4120 Å^3, gap volume = 21,361 Å^3). Yellow dots and blue surfaces represent the assembling spaces and gap regions, respectively. The figures were prepared with MOE^{TM} (Chemical Computing Group Inc., http://www.chemcomp.com/).

Delaunay tessellation. The probe-based approach is good for the detection of small pockets [9], but may not be suitable when considering larger interfaces such as in protein–protein interaction sites. The difference between two detections is the important point in our analyses. Detection of the surface pockets is often done to find the potential ligand binding sites, and thus larger estimation of the pocket space can be more favorable than smaller estimation. On the other hand, however, in protein–protein interactions, the main purpose is to estimate how and what kinds of interactions exist in the interacting space. Therefore, overestimation in size is not acceptable. As we will show later, the space detected by a probe-based approach is much larger than that by a Delaunay tessellation approach.

Delaunay tessellation [10] is a widely used technique to describe the geometry of three-dimensional objects. It defines a unique convex hull made by tetrahedra from a set of vertex. For protein structures, it is also used, for example, to detect structural motifs [11], to classify folding topology [12], to carry out structural alignments [13], to identify the protein–protein interaction sites [14], to assign secondary structural elements [15], and to find cavities and pockets [16,17]. It may be noteworthy to emphasize the difference between Delaunay and Voronoi tessellations [18]. The two tessellation-techniques are usually applied to define the *space* belonging to the set of points. The Delaunay tessellation to discrete points corresponds to the dual graph of the Voronoi tessellation for the set. Thus, when the spatial distribution of some points is considered in the three-dimensional space, Delaunay tessellation gives a set of tetrahedra with the points as vertex, while Voronoi tessellation gives a set of polyhedra around the points. Therefore, Voronoi tessellation may be used to define the space belonging to each point, and Delaunay tessellation can be used to define the space among the points. Actually, Voronoi tessellations have been used to calculate the volume or packing density in each subunit [19], and here we used Delaunay tessellation to calculate the size of the interface regions.

In this paper, we propose three new descriptors to evaluate the complementarity in molecular recognition of proteins: (1) ASV, the volume of the *assembling space* or gap space in protein–protein interfaces, (2) AS-distance, the ratio of ASV to interface area, and (3) global shape descriptor (GS) to describe the global shape of the interfaces. Systematic comparisons between gap volume and ASV, and between shape complementarity and AS-distance are also discussed. Our method is available as a web server at http://asvcalc.dna.affrc.go.jp/. It should be noted that the term of assembling space is a new word made in this study to discriminate more popular term such as gap space or interface region. The precise definition will be given as a space we calculate the volume, but it is equivalent to the usual gap space or interface region conceptually.

2. Materials and methods

2.1. Calculation methods

The calculation of assembling space volume, ASV, is composed of three independent steps, which are (1) definition of the dimer interface, (2) Delaunay tessellation and trimming of tetrahedra with long edges, and (3) calculation of the volume of the Delaunay tetrahedra and elimination of the space occupied by atoms.

First, the solvent accessible surface area (ASA) of each atom is calculated for the dimer structure and for each subunit by using our own program which implements Lee and Richards' algorithm [20]. Then, the interfacial atoms of the subunits are defined as the atoms with difference in ASA (ΔASA) between the ASA in the dimer state and that in each separate subunit. The atoms with $\Delta\text{ASA} > 0.1$ Å^2 were considered as the atoms in the interface regions. In these calculations, we did not consider all hetero atoms in the PDB file, because we focus on the correspondence of proteins in their interface regions, by assuming that the contribution of the hetero atoms other than water molecules in the interface will be small. It is true that this assumption is a naïve one but large compounds were rarely found in our data set used in this study with an exception of a tryptophan in 5csm. Our observation that the small molecules are not often found in the interface regions may suggest the small contribution of the small molecule for the stabilization of protein complexes structures. It may be noteworthy that effects of water molecules for the stability of complex structure is not small, but all the water molecules are not appeared in every PDB file, thus, in this study, we did not treat water molecules in the interface.

Second, Delaunay tessellation is performed on the coordinates of all atoms in the interfaces to define the assembling space, by using Qhull [21]. Then, Delaunay tetrahedra are classified by their location as follows when we do not consider the order of vertexes in each tetrahedron. When an interface between subunit A and subunit B is considered, each tetrahedron is classified as one of five types, A4, A3B1, A2B2, A1B3 and B4, where A or B means a vertex in the subunit A or B, respectively, and the number after the letter indicates the number of vertex in the subunit. For example, the A3B1 tetrahedron means the three vertexes in subunit A and one

vertex is in subunit B. Thus, every tetrahedron can be classified into one of two categories: either a tetrahedron with vertices that belong to only one subunit (A4 and B4) or to both subunits (others). We call the latter tetrahedron an interfacial tetrahedron (or I-tetrahedron), hereafter.

The interfaces of proteins are sometimes very twisted (for example, Fig. 4a: 1fip). In the twisted interfaces, the above procedures were not sufficient to define the interfacial regions precisely, because of our definition of I-tetrahedron, where only one vertex from a different subunit is required for it to be judged as I-tetrahedron. Thus, we added another step to trim the tetrahedra with non-contacting pairs of vertices. To include the indirect, water mediated interactions, we set the threshold of the long edge for trimming to be more than 10.0 Å in this study. (The threshold value will be further discussed later.) We call the trimmed tetrahedron a T-tetrahedron, hereafter. It should be noted that the trimming step was introduced to eliminate the unusual tetrahedra, but the ratio of the volume of all T-tetrahedra to that of all I-tetrahedra might be used as an indicator to describe the global shape of the interfaces, as discussed later.

Finally, we calculate ASV by using the grid approximations, because it is one of the easiest ways to extract the volume occupied by atoms from the sum of the volume of each tetrahedron. All of the grid points in each tetrahedron are generated with 0.2 Å separations in each *XYZ*-direction, and then the grid points with distances to the nearest atom that are less than the van der Waals distance of the atom are eliminated. The van der Waals radii of the atoms are from Lee and Richards' values [20]. The volume of the assembling space is calculated from the total number of grid points with a volume allocated to each point; that is, a cubic of an interval (=0.2 $Å^3$ in this study).

3. Results and discussion

3.1. Data set of homo-dimer proteins

The data of homo-dimer structures were collected based on the data set of Ponstingl et al. [22] from the PDB site on September 2006. Their original data set contained 76 entries. Because some of the PDB files had no definitions about oligomerization, the dimer coordinates were given by the generated crystal lattice. In several pairs of the possible dimer forms, the interface of the clearly largest ΔASA was selected. As the result, seven PDB structures (1a3c, 1czj, 1jsg, 1otp, 1puc, 1vlb, and 2tgi) were removed because their native interface could not be determined due to the lack of the information of oligomer state, and one old entry (1hlr) was replaced by a new one (1vlb). Thus, 69 entries were used for the following analyses.

3.2. Comparison between ASV and gap volume

Interfacial spaces of proteins have often been evaluated by the gap volume obtained by SURFNET, where the gap regions were detected by using probe-spheres and the volume was calculated as the sum of the van der Waals volumes of the spheres. The probe-spheres were shrunk until they contacted a pair of atoms in different subunits (see Fig. 1a and b). This approach is suitable for detecting a cavity or cleft in small molecule binding sites, but it will have some problems in large and complex protein–protein interaction sites. For example, if the interface of the proteins is rounded, then the space in the edge regions can be included as the gap regions (Fig. 1a), and if there is a large cleft in the interface without any contacts, then the cleft can be treated as part of the gap regions (Fig. 1b). These situations were actually observed in our data set (Fig. 1c and d). The first example can be observed in 1amk (triose phosphate isomerase), where the ASV is 1835 $Å^3$ and the gap volume is 2603 $Å^3$. The other example is for 1pre (proaerolysin), where the interface consists of two separate contacting interfaces but SURFNET recognized it as one single, large interface, where the ASV is 4120 $Å^3$ and the gap volume is 21,361 $Å^3$.

To observe the general trends of the differences between ASV and the gap volume, we created a plot between ASV and gap volume (Fig. 2b). As seen in the figure, ASV is weakly correlated with the gap volume, qualitatively, but the absolute value of the gap-volume is far larger than ASV, on average. The case of 1pre

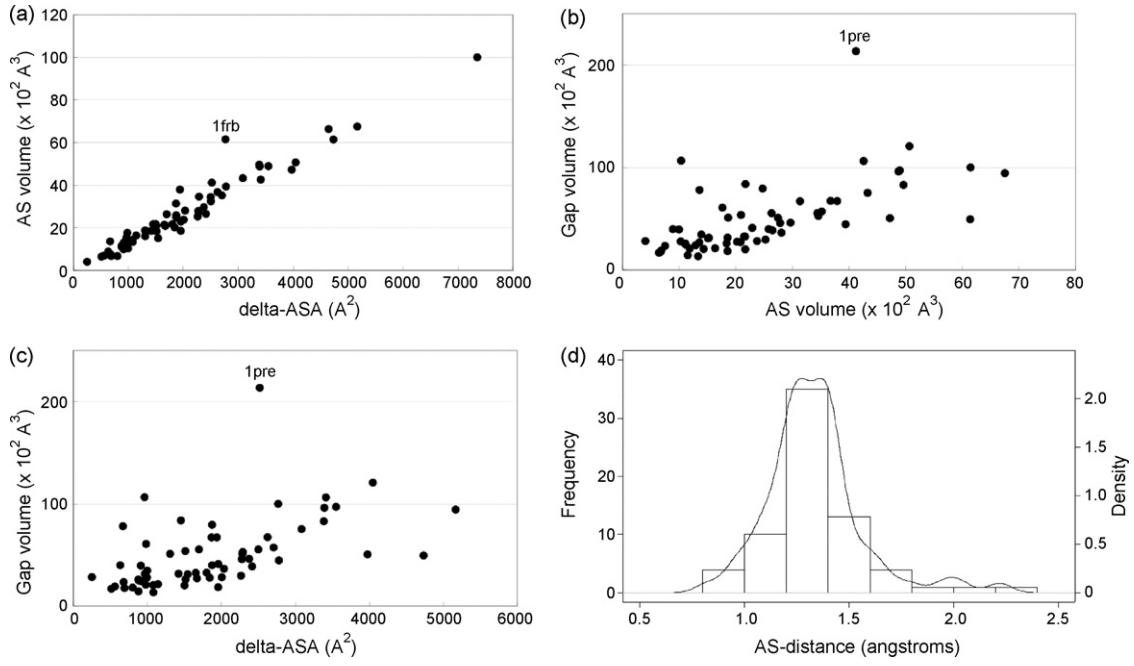

Fig. 2. Quantitative comparison between gap volume and ASV. (a) Scatter plot between gap volume and ASV, (b) the same plot between ASV and dASA, and (c) the same plot between gap volume and dASA. (d) Distribution of AS-distance.

(proaerolysin, Fig. 1d) represents an extreme case where the gap volume has a much larger value than ASV.

The concept of ASV is similar to the gap volume of SURFNET, but the results are quite different as described. The gap volume is determined by filling the interface with probe spheres, and the filling is performed for all clefts near interface regions. On the other hand, the ASV measures only the interfacial volume between two subunits, and does *not* include any intra-molecular clefts. In addition, ASV defines the border as planes/lines connecting three/two atom centers, while the border in the gap volume is determined by spheres. The projection of the spheres in the border will enlarge the gap volume value, as compared with ASV. As a result, it is quite reasonable that the gap volume tends to be larger than the ASV, as shown in Fig. 2a. In short, the main differences between ASV and gap volume exist in two points, i.e., (1) edge definition and (2) intra-molecular cleft treatment.

3.3. AS-distance, as an index to evaluate the average distance of interfaces

Interfacial volumes will depend on the areas of interface regions to some extent, because each interfacial volume is the product of interface area and gap distance. To evaluate the dependence, we plotted the ASV and the gap volume against ΔASA (Fig. 2b and c, respectively). The correlation between ASV and ΔASA was high (Pearson's correlation coefficient (PCC) = 0.97), while surprisingly, the correlation between gap volume and ΔASA was rather weak (PCC = 0.46). As we discussed above, the differences between ASV and gap volume are edge definition and intra-cleft treatment. Both will cause a low correlation between gap volume and ΔASA, but the latter effect may be critical, because the intra-cleft volume will not depend on the ΔASA at all. Therefore, ASV/ΔASA will be a new indicator to evaluate an average distance between protein–protein interfaces as opposed to the gap index (=gap volume per interface area). Hereafter, we refer to the ratio, ASV/ΔASA, as the AS-distance.

As shown in Fig. 2d, the AS-distances are sharply distributed around 1.3 Å. In other words, the average distances between subunit in homo-dimers are not large. It may be noted that this observation is not consistent with those of Jones and Thornton [23], where some differences in gap indexes among homo-dimers were observed. The largest AS-distance, 2.22 Å, was observed for 1rfb (bovine interferon γ). The protein is like a swapped dimer with very unique interfaces, and it contains 10 water molecules in the interface region, which should enlarge the distance between subunits [24]. The second largest AS-distance was 2.03 Å for diphtheria toxin (PDB: 1tox). The interfacial area (ΔASA) of the toxin is relatively small (671 Å^2), as compared with those of the other proteins. Carroll et al. previously reported that dimeric toxins slowly dissociated to native monomers in solution at neutral pH [25], indicating the weak interactions among toxins. Such weak interactions may be understood in terms of the unusually high AS-distance and the relatively small contact area. The smallest AS-distance, 0.826, was obtained for 1bam (endonuclease BamH1), where the well-folded globular subunits form a tight interface. As in these examples, the AS-distance may be able to differentiate among the types of shape complementarity of protein–protein interfaces, that is, tight complex or weak ones.

3.4. Comparison between shape complementarity (SC) and AS-distance

To clarify the features of AS-distance as an index to describe the shape complementarity of interfaces, we compared AS-distance with SC by Lawrence and Colman. The calculation of SC was done by the CCP4 package [26], and the results are shown in Fig. 3. As seen in the figure, the correlation between SC and AS-distance is rather weak. For example, the SC scores of 1ad3, 1ajs, 1amk, 1fip, 1iso, 1nsy, 1rfb, 1smt, 1tox, and 5csm are almost the same (around 0.71), but the AS-distance values are quite different (1.25, 1.38, 1.21, 1.24, 1.25, 1.30, 2.22, 1.18, 2.03 and 1.96, respectively; Table 1).

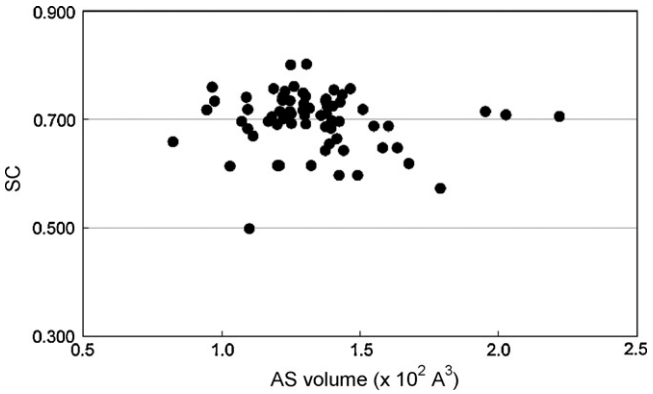

Fig. 3. Scatter plot between SCs and AS-distances.

The structures 1rfb (bovine interferon γ), 1tox (diphtheria toxin) and 5scm (chorismate mutase) have the three highest AS-distances among the ten example proteins, which have AS-distances that are much higher than the average value of the AS-distance. A possible reason for such large values is extra molecules, such as water molecules and ligands, in the interface. For example, 1rfb, which has the largest AS-distance (2.22 Å), contains ten water molecules in the assembling region, and one tryptophan molecule was found at the interface of 5csm (AS-distance = 1.95 Å). However, 1afw (peroxisomal thiolase) has fifty waters in the interface, but its AS-distance is only 1.37. Therefore, a large AS-index is not always caused by indirect interactions through water molecules or the insertion of small molecules. It should be noted that all of these examples have virtually the same SC values.

3.5. Global shape descriptor

As described in Section 2, we introduced a trimming step to eliminate unusual tetrahedra with long edges, due to some twisted or bent interfaces, as in the cases of 1fip (Fig. 4a) and 1ajs (Fig. 4b). In these proteins, the volume of trimmed tetrahedra (T-tetrahedra) will be larger than that of the proteins with flat interfaces (Fig. 4c). Therefore, the ratio of the volume of I-tetrahedra to that of T-tetrahedra will be a good indicator to describe the shape of the interface.

Table 1
Structural data for ten proteins with SC scores around 0.7: values of SC, ΔASA, ASV, AS-distance, and GS_{10}.

PDB ID	SC	ΔASA (Å^2)	ASV (Å^3)	AS-distance (Å)	GS_{10}
1ad3	0.711	4048	5067	1.252	0.453
1ajS	0.711	3550	4896	1.379	0.488
1amk	0.715	1515	1835	1.211	0.055
1fip	0.710	1673	2080	1.244	0.272
1iso	0.715	3412	4256	1.247	0.354
1nsy	0.708	2706	3518	1.300	0.421
1rfb	0.706	2770	6147	2.219	0.366
1smt	0.705	2008	2374	1.182	0.331
1tox	0.709	671	1359	2.027	0.529
5csm	0.715	1941	3790	1.953	0.277

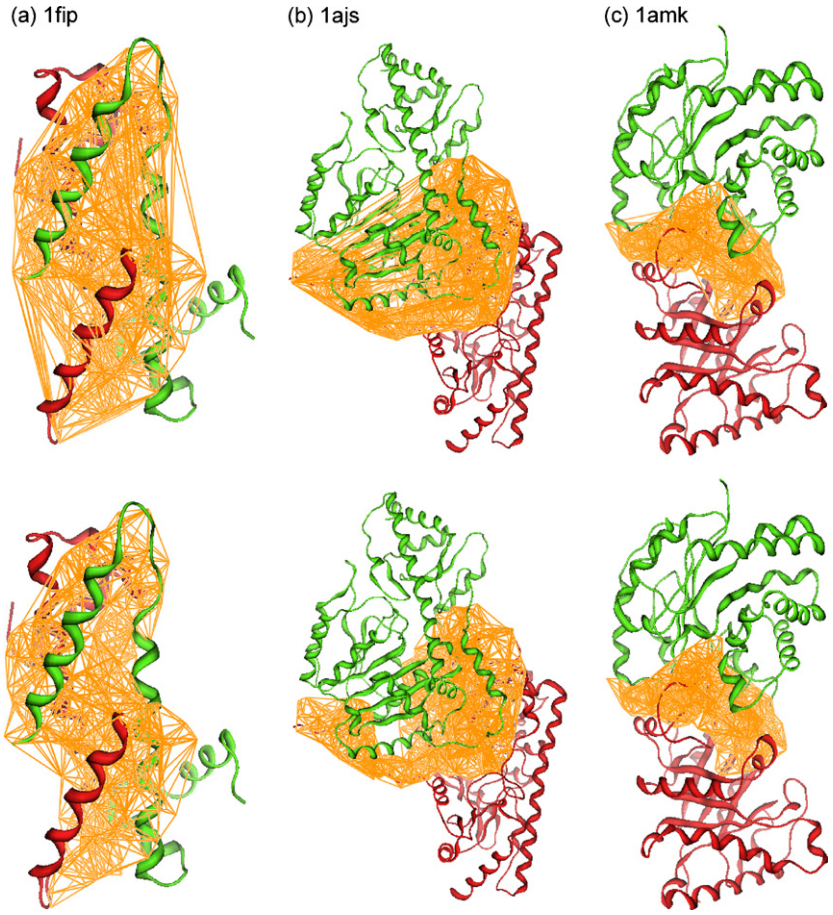

Fig. 4. Examples of long edge trimming. The upper figures show the I-tetrahedra in orange color, and the lower figures indicate the tetrahedra after removing the long edges for (a) 1fip (factor for inversion stimulation), (b) 1ajs (aspartate aminotransferase), and (c) 1amk (triose phosphate isomerase).

In this study, we define the *GS* as (volume of T-tetrahedra)/(volume of I-tetrahedra). Then, the shape descriptor will vary from 0.0 to 1.0, according to the degree of twisting or bending in the interface. For the ideal flat interface, the value will be 0.0, and it will become larger as the interface becomes more twisted or bent. Among all of the structures in our data set, the smallest value of the GS was 0.034 for 1ctt (cytidine deaminase), and the structure has a rather flat interface. The largest value was 0.74 for 1pre (proaerolysin), and the structure has a complicated interface with multiple contacts. It should be noted that GS is largely depend on the threshold for trimming. In other words, GS is a descriptor with a resolution of the threshold distance for trimming. To clarify this point, we will call GS as GS_{10}, where 10 is the threshold distance (in Å unit) used in this study.

To determine the relationship between SC and GS_{10}, we again chose ten proteins with similar SC scores around 0.71 (1ad3, 1ajs, 1amk, 1fip, 1iso, 1nsy, 1rfb, 1smt, 1tox and 5csm). Their GS_{10}s were 0.45, 0.49, 0.055, 0.27, 0.35, 0.42, 0.37, 0.33, 0.53, and 0.28, respectively. The interfacial structures were visually inspected, according to the GS_{10} values. As a result, a structure with a GS_{10} value less than 0.20 has a rather flat interface, whereas a higher GS_{10} corresponds to a twisted or bent interface. In addition, some structures with significantly large GS_{10} values (more than 0.40) have multi-contact interfaces. For example, the smallest GS_{10} in the ten structures was found for 1amk (triose phosphate isomerase, $GS_{10} = 0.055$) (Fig. 4c). As shown in the figure, the interface is flat and compact. On the other hand, the interface of 1fip (FIS protein, $GS_{10} = 0.23$) has a twisted shape (Fig. 4a), and the interface of 1ajs (aspartate aminotransferase, $GS_{10} = 0.49$) is surrounded by the other subunit (Fig. 4b).

3.6. Trimming threshold dependency

To calculate the assembling space, one parameter has to be determined, that is the threshold in the trimming step. We evaluated the threshold dependency of the GS_{10} by changing the value from 3.0 Å to 20.0 Å with 1.0 Å intervals for the 69 structures (shown in Fig. 5). The ratios of the number of T-tetrahedra to that of I-tetrahedra in each threshold value have large values when the threshold value is small, and the ratios gradually decrease as the threshold becomes large. Most of the proteins have the same tendency, with some exceptions such as 1rpo and 1tox. For the other 64 structures, their ratios became constant in the range above 10 Å (the average ratio is 0.069 and the standard deviation is 0.022 at 10.0 Å). We also visually inspected the three-dimensional shapes of the tetrahedra in 1ajs, 1amk, and 1fip, using cutoffs of 8.0 Å, 10.0 Å, and 12.0 Å. The tetrahedra at 8.0 Å and 10.0 Å were almost the same, but the tetrahedra at the 12.0 Å cutoff still included inappropriate space. Thus, the threshold will be suitable between 10 Å and 12 Å. The average rates at 8 Å, 10 Å, and 12 Å are 0.132, 0.083, and 0.056, respectively. On the other hand, the difference of the ratios between the 12 Å and 10 Å threshold values

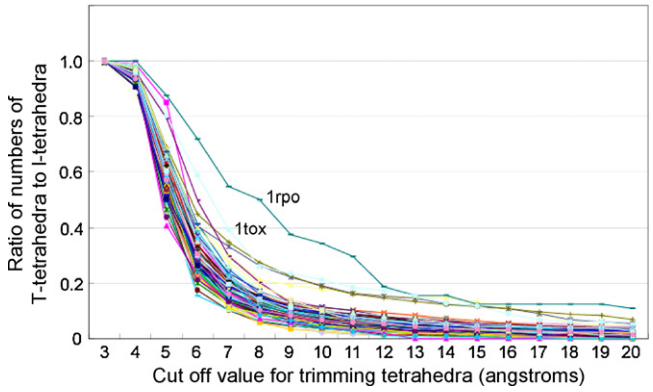

Fig. 5. Relationship between cutoff value and the ratio of the numbers of T-tetrahedra to I-tetrahedra.

are rather smaller than that between 10 Å and 8 Å (0.027 for 10–12 Å, and 0.049 for 8–10 Å). This means that the smaller threshold removes more tetrahedra. From the above results, a larger value will be good in the range of 8–10 Å. Therefore, we use 10 Å for our analysis in this study. However, our results would not change much with a different threshold in most cases, and the differences are small.

4. Conclusion

We reported a new method to calculate the space between two protein surfaces, named the ASV. Two descriptors for protein–protein assembly derived from ASV were also proposed. The AS-distance, the ratio of ASV to ΔASA of dimer, was almost constant among homo-dimer interfaces. This observation may be used to develop a method to identify the biological protein–protein interfaces in the protein crystals. The other one, named global shape descriptor (GS), was the ratio of the volumes of the E-polyhedron and I-polyhedron. These polyhedra were defined during calculation of the interfacial polyhedron. The interface bending, the GS value is larger. The descriptor could be applied to the classification of the interfacial shape.

Acknowledgments

The authors thank Dr. Roman Laskowski (European Bioinformatics Institute) for supplying the SURFNET program and helping to apply it to our dataset. We also acknowledge Mr. Kinya Toda (Ryoka Systems Inc.), Dr. Zui Fujimoto (National Institute of Agrobiological Sciences) and Mr. Hiroya Nobori (Mitsubishi Space Software Co. LTD.) for technical support. The visualization by MOE was performed with the assistance of the Computer Center for Agriculture, Forestry and Fisheries Research, MAFF, Japan. This research was supported by a Grant-in-Aid for Scientific Research (Ministry of Education, Culture, Sports, Science and Technology of Japan).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2008.11.002.

References

[1] J.M. Chandonia, S.E. Brenner, The impact of structural genomics: expectations and outcomes, Science 311 (2006) 347–351.
[2] M.C. Lawrence, P.M. Colman, Shape complementarity at protein/protein interfaces, J. Mol. Biol. 234 (1993) 946–950.
[3] M.L. Connolly, Solvent-accessible surfaces of proteins and nucleic acids, Science 221 (1983) 709–713.
[4] B.B. Goldman, W.T. Wipke, QSD Quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock), Proteins: Struct. Funct. Genet. 38 (2000) 79–94.
[5] Y. Tsuchiya, K. Kinoshita, H. Nakamura, Structure-based prediction of DNA-binding sites on proteins using the empirical preference of electrostatic potential and the shape of molecular surfaces, Proteins: Struct. Funct. Genet. 55 (2004) 885–894.
[6] H.A. Gabb, R.M. Jackson, M.J.E. Sternberg, Modelling protein docking using shape complementarity, electrostatics and biochemical information, J. Mol. Biol. 272 (1997) 106–120.
[7] R. Chen, Wing, Docking unbound proteins using shape complementarity, desolvation, and electrostatics, Proteins: Struct. Funct. Genet. 51 (2003) 397–408.
[8] R.A. Laskowski, SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions, J. Mol. Graph. 13 (1995) 323–330.
[9] T. Kawabata, N. Go, Detection of pockets on protein surfaces using small and large probe spheres to find putative ligand binding sites, Proteins: Struct. Funct. Bioinform. 68 (2007) 516–529.
[10] B. Delaunay, Sur la sphère vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk 7 (1934) 793–800.
[11] H. Wako, T. Yamato, Novel method to detect a motif of local structures in different protein conformations, Protein Eng. 11 (1998) 981–990.
[12] D.L. Bostick, M. Shen, I.I. Vaisman, A simple topological representation of protein structure: implications for new, fast, and robust structural classification, Proteins: Struct. Funct. Bioinform. 56 (2004) 487–501.
[13] J. Roach, S. Sharma, M. Kapustina, C.W. Carter, Structure alignment via Delaunay tetrahedralization, Proteins: Struct. Funct. Bioinform. 60 (2005) 66–81.
[14] J. Bernauer, J. Azé, J. Janin, A. Poupon, A new protein–protein docking scoring function based on interface residue properties, Bioinfomatics 23 (2007) 555–562.
[15] T. Taylor, M. Rivera, G. Wilson, I.I. Vaisman, New method for protein secondary structure assignment based on a simple topological descriptor, Proteins: Struct. Funct. Bioinform. 60 (2005) 513–524.
[16] J. Liang, H. Edelsbrunner, P. Fu, V. Sudhakar, S. Subramaniam, Analytical shape computation of macromolecules. II. Inaccessible cavities in proteins, Proteins: Struct. Funct. Genet. 33 (1998) 18–19.
[17] J. Liang, H. Edelsbrunner, C. Woodward, Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design, Protein Sci. 7 (1998) 1884–1897.
[18] G. Voronoi, Nouvelles applications des paramètres continus à la théorie des formes quadratiques, Journal für die Reine und Angewandte Mathematik 133 (1907) 97–178.
[19] A. Poupon, Voronoi and Voronoi-related tessellations in studies of protein structure and interaction, Curr. Opin. Struct. Biol. 14 (2004) 233–241.
[20] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (1971) 379–400.
[21] C.B. Barber, D.P. Dobkin, H.T. Huhdanpa, The Quickhull algorithm for convex hulls, ACM Trans. Math. Software 22 (1996) 469–483.
[22] H. Ponstingl, K. Henrick, M. Thornton, Discriminating between homodimeric and monomeric proteins in the crystalline state, Proteins: Struct. Funct. Genet. 41 (2000) 47–57.
[23] S. Jones, J.M. Thornton, Principles of protein–protein interactions, Proc. Natl. Acad. Sci. U.S.A. 93 (1995) 13–20.
[24] C.T. Samudzi, J.R. Rubin, Structure of recombinant bovine interferon-gamma at 3.0 Å resolution, Acta Cryst. D 49 (1993) 513–521.
[25] S.F. Carroll, J.T. Barbieri, R.J. Collier, Dimeric form of diphtheria toxin: purification and characterization, Biochemistry 25 (1986) 2425–2430.
[26] Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. Acta Cryst. D 50 (1994), 760–763.