

Chapter 2

Characterization of Protein–Protein Interfaces, Considering Surface-Roughness and Local Shape

Abstract This chapter attempts to provide an account of works that have attempted to characterize protein–protein interaction interfaces, with fractal dimension. However, such characterization of interfaces is not solely dependent upon interface roughness. Without involving the biophysical factors, we will concentrate only on geometric characterization of these interfaces. To be specific, we will attempt to talk about a possible algorithm to quantify the changes in two parameters describing any protein–protein interaction interface; namely, the curvature of shape of protein–protein interaction interface and the surface roughness of it. One can connect these two parameters through a novel methodology, ‘extended unit iterated shuffle transformation’. Results show that although the interface patch for enzyme-inhibitor interaction is flatter and smoother than the non-interfacial surface patches, absolute magnitudes of shape curvatures and surface roughness of bound interfaces are greater than what they were in unbound states of concerned entities. Trends observed on antigen–antibody interfaces are found to be somewhat contradictory to the trends observed in case of enzyme-inhibitor interfaces. Antigen–antibody interfaces, like the enzyme-inhibitor interfaces, are found to be flatter and smoother than the non-interfacial surface patches. However, unlike the enzyme-inhibitor interfaces, absolute magnitudes of shape curvatures and surface roughness of bound antigen–antibody interfaces are observed to be less than what they were in unbound states of concerned entities. Algorithm proposed in the present work could quantify the effects due to changes in two extremely important interfacial parameters, through a unified scheme.

2.1 Introduction to Protein–Protein Interaction Interfaces

Most proteins interact, at least transiently, with other protein molecules. Given that many fundamental biological processes such as antigen–antibody recognition, hormone-receptor binding, and signal transduction are regulated through association and dissociation of proteins, we need to study protein–protein interactions in details. Characterizing the physico-chemical properties of interfaces through which protein–protein interactions take place, naturally, has always been a

primary aim of molecular biology. Several studies have addressed protein–protein interactions and their applications in varied paradigms, ranging from rational drug design to structure prediction of multimeric complexes. Among these, some (Katchalski-Katzir et al. 1992; Todd et al. 2002; Arkin et al. 2003; Nooren and Thornton 2003) have focused on the complementarity of chemical and structural features (shape, hydrophobic patterns, distribution of electrical charges, etc.) at the binding interface as major contributors to intersubunit interactions. These studies were useful, but they were banking on an implicit assumption that protein–protein interaction interfaces can be suitably characterized by using a single reference configuration for the interacting subunits; that is, by treating them as if they were rigid molecules. Not all studies, however, had resorted to such ‘rigid molecule’ framework. Due to the increasing availability of computational resources and more refined theoretical models, many other studies were made possible, which investigated the role of a complementary physical effect, namely the internal dynamics of the interacting subunits. A number of studies (Rajamani et al. 2004; Li et al. 2004; Smith et al. 2005; Yogurtcu et al. 2008) have shown that even in the absence of the partner subunits, the dynamical properties of amino acids at the known interface region can differ from those of other amino acids at the protein surface.

2.2 Why One Needs Detailed Geometric Characterization of Protein–Protein Interaction Interfaces?

As mentioned earlier, many essential cellular processes such as signal transduction, transport, cellular motion, and most regulatory mechanisms are mediated by protein–protein interaction (PPI)s. Owing to such overwhelming biological significance, PPIs have been the object of much attention; especially as they relate to interactions and associations across the entire proteome. PPIs are optimized locally (Keskin et al. 2005) and they are mediated through protein interfaces.

A huge body of knowledge about PPI interfaces has been amassed. We know that PPI sites are (mostly) hydrophobic, (mostly) planar but at times globular and protruding (Chothia and Janin 1975; Argos 1988; Jones and Thornton 1997), they are composed of relatively large surfaces with shape and electrostatic complementarity (Jones and Thornton 1996; Janin 1995, 1997)—These facts separate them clearly from enzyme catalytic sites. The catalytic site of an enzyme is a cleft, often buried, sometimes deeply. Catalytic site cleft is enzyme’s workshop; it is here that the catalytic reaction occurs. In comparison to protein–protein interaction interfaces, the catalytic site house a relatively small number of amino acids that are involved in binding the substrate (and/or cofactor). Interestingly, an even smaller subset of these residues is found to be vital to the enzyme’s catalytic function (Bartlett et al. 2002)—Such simple patterns in either the residue composition or in geometry are not observed for protein–protein interaction interfaces. Furthermore, it is known that these interfaces can be dynamic, facilitating binding to different proteins with diverse compositions and shapes (DeLano et al. 2000;

Ma et al. 2002; DeLano 2002; Kuhlmann et al. 2000). Indeed it has been noted that simple rules to identify protein recognition sites and prediction of energetic “hot spots” (Bogan and Thorn 1998) in protein–protein interfaces fail consistently, primarily because of the extreme diversity in shape, chemical character, and plasticity of protein–protein interfaces (Joachimiak et al. 2006). Hence, no single parameter is found to differentiate protein–protein interfaces from other surface patches with absolute confidence (Bradford and Westhead 2005).

It is observed though that, geometry is an important determinant of interfaces. Hence, shape complementarity stands out as one of the principal ingredients for all scoring functions for docking methods (Chen and Weng 2003). In varying extent, almost all the docking algorithms tend to rely on the assumption that interacting proteins have a certain degree of shape complementarity. But the exact algorithm to quantify shape complementarity varies among docking algorithms, which can either be functions based on surface curvatures or functions based on features of surface areas (Chen and Weng 2003). Probing the (elusive) connection between these, viz. functions based on interface curvatures and functions of interface surface features, therefore, assumes immense importance in the paradigm of PPI studies. An attempt to connect them for a system that is dynamic and context-dependent, however, demands an algorithm that is robust, yet sensitive.

2.3 FD-Based Characterizations of Protein–Protein Interaction Interfaces

As it was stated in the last chapter, it was known that protein surfaces are fractal in nature (Lewis and Rees 1985); furthermore, more than two decades ago it was proposed that PPI interfaces can be studied with fractal dimension (FD) as a reliable marker to quantify protein surface roughness (Aqvist and Tapia 1987). But, although promising, such a scheme (Aqvist and Tapia 1987) can be observed to have found little application in future. Possible reason behind the lack of application of this method might well be due to the absence of a suitable mathematical structure that, first, respects the fractal nature of protein surfaces and second, constructs a model to describe the change in local curvature of protein exterior (PPI interfaces in particular). Applying a recently proposed mathematical structure, ‘extended unit-iterated shuffle transformation’ (EU-IST) (Fujimoto and Chiba 2004), a case for this methodology can be submitted where a marker is constructed. Such a marker can connect the pattern of changes of surface roughness in enzyme-inhibitor and antigen–antibody interfaces, with the pattern of changes in the local curvature of the shapes of these two types of interface.

Neither the ‘Grid’—based shape complementarity methods (Katchalski-Katzir et al. 1992), nor the ‘pairwise shape complementarity’ (Chen and Weng 2003) methods, take into account the fractal nature of protein surface roughness. Interestingly, the FD-based approach to describe shape changes (Bowman 1995) is radically different than those studying shape deformations in a continuous way (Sederberg

and Parry 1986). While the later deforms the whole shape continuously, the former deforms the shape in a way that each subpart of the shape in all scales is deformed recursively. The EUIST-based approach (Fujimoto and Chiba 2004), owes its origin to an iterated function system studies (basically, a finite set of contraction mappings on a complete metric space) (Hutchinson 1981). With the unit-IST approach, a fractal-type repetitive structure is constructed on ‘between-the-points’ in the domain (atoms on interface of protein PR₁, before interacting with protein B) and those in the range (same set of atoms, after PPI). Thus, this approach ensures that local resemblance in space (viz. the scale directions) is maintained. [For further details on this, please refer to (Fujimoto and Chiba 2004)]. The EUIST-based approach elaborated in this chapter, therefore, could describe the continuous process of change of interface shape during the entire course of PPI, measuring the change in surface roughness with FD at every step.

2.4 A Small Study to Probe Antigen–Antibody and Enzyme-Inhibitor Interfaces with Surface FD

To conduct such a ‘pilot–study’, 10 pairs of enzyme-inhibitor complex and 10 pairs of antigen–antibody complex, as provided in the dataset of a 2003 paper (Gray et al. 2003), were taken. To obtain the most biologically relevant idea of the process, biological unit information (and not that of the asymmetric unit) for the uncomplexed and complexed units [as provided by the protein data bank (PDB) (Berman et al. 2000)], were considered. [Multifaceted significance of working with the biological units of proteins, especially while studying PPI, has been described in a recent work (Jefferson et al. 2006)].

Each of these protein complex is composed of a pair of proteins, one of which was dubbed as “receptor”, while the other was called “ligand”, in the paper (Gray et al. 2003) from which the dataset of the current work was taken. Ideally, to study the change in comparative surface-roughness profile between interface and non-interface regions, before and after the PPI, one should have studied the roughness of receptors and ligands both—to quantify the parameters in unbound state. However, the ligand protein molecules presented in the dataset turned out to be small sized proteins, in many (if not most) of the cases. Smallness of ligand proteins posed a problem for the calculation of surface-FD for a focused solid-angle zone (discussed later); because to calculate the surface-FD one requires having statistically significant number of atoms, which were not present in most of the cases. Furthermore, since the present work resorted to mean magnitude comparisons to assess the patterns in changes of surface roughness, a significant number of surface patches with statistically significant number of atoms were required. Ligand protein molecules, as presented in the aforementioned dataset, failed to match this constraint, consistently.

Hence the current study focused only on the surface patches of the receptor protein molecule (interface, non-interface alike), while studying the features for

unbound and bound state. But in the bound state, the receptor molecule resides in the state of a complex. Hence, for the studies of ‘before the PPI’ states, the receptor protein molecules were considered; while, for the studies of ‘after the PPI’ studies, the complex protein molecules were considered. Furthermore, since a nomenclature that accommodated the receptor protein’s name and PDB id., alongside that of the complex protein molecule—would have been cumbersome to make sense of, in the current study, these informations were not enlisted. The receptor protein molecule is referred to as the complex molecule in unbound state—for all the ‘before the PPI’ investigations. Details of the receptor and ligand protein molecules can be obtained from the reference (Gray et al. 2003).

While the implementation strategy of solid angle-based interface patch characterization is discussed later, it is important at this point to talk about the biophysical importance of resorting to such a scheme. Despite the fact that the respective geometries of PPI interfaces and catalytic sites play important roles, the geometries merely provide a helpful scaffold for the actual entities to interact, which are the amino acids present therein. Now, how to identify the residues that are participating in PPI? Though PPI interfaces are quite different than enzyme catalytic sites, in this context, it helps to recall how the catalytic site residues are characterized.

To investigate and understand the role and function of the catalytic amino acids in enzymes, the first necessary step was to define a ‘catalytic residue’ unambiguously. One finds such a clear-cut definition in a recent work (Holliday et al. 2007), where the definition first proposed by (Bartlett et al. 2002) was made more pointed. The (Holliday et al. 2007) scheme to define catalytic residues have two distinct parts; viz.: a catalytic residue is any residue involved in the reaction that either has direct involvement in the reaction mechanism (that is, the ‘reactant residues’ whose chemical structure is modified during the course of the reaction, for example, the residue is involved in covalent catalysis, electron shuttling, proton shuttling, etc.); or, has indirect but essential, involvement in the reaction mechanism [that is, the so-called spectator residues, whose chemical structure does not change during the course of the reaction—these are the residues that polarize or alter the pKa of a residue, a water molecule or part of the substrate directly involved in the reaction, affect the stereospecificity or regiospecificity of the reaction, or stabilize the reactive intermediates (either by stabilizing the transition states or the intermediates themselves, or destabilizing the ground states of the substrates)]. Thus, in order to characterize the catalytic mechanism, one requires to study both the ‘reactant residues’ and ‘spectator residues’. Whether the residues participating in PPI can be classified in the same manner or not—needs investigation, but the idea that some of the residues can actually participate in PPI and the other adjoining ones can help the PPI without actually participating in it—seems logical. Hence, one needs a construct that present the information about the spatial location of both types of residues participating in PPI; moreover, the construct should be such that varying number of (purported) ‘spectator residues’ can be studied (to probe their effects on PPI), without altering the scheme significantly—The solid angle-based construct is an ideal candidate that satisfies all these requirements, hence it was resorted to.

2.4.1 Algorithms, Theoretical Basis, and Implementation Strategies

The algorithm to conduct this study comprises of two sections.

2.4.1.1 Algorithm 1st Part: Calculation Scheme for the Surface Fractals

If a given PPI is known to be involving protein PR₁ and protein PR₂, the methodology described below is presented with respect to arbitrarily chosen protein PR₁. Exactly the same methodology was implemented on protein PR₂.

Placing the center of mass of the protein at the origin of spherical polar coordinate system, solid angles of 30°, 45°, and 60° were constructed to (suitably) map the surface atoms for the entire protein. To calculate the FD of a patch of protein surface constituted by these atoms, the following scheme was implemented. (Details of implementation of this scheme can be found in the last chapter of this book.)

$$D_i = 2 - \left\{ \frac{(\log A_{vdW})_i / (\log A_{vdW})_{i-1}}{(\log R_p)_i / (\log R_p)_{i-1}} \right\} \text{ and } FD = 1/n \sum_i^n D_i$$

where $(A_{vdW})_i$ represents the Van der Waals surface area as measured by probe sphere with radius $(R_p)_i$; and FD is calculated as the mean magnitude of several sets of D_i s; n being the number of such sets. Probe radius (R_p) range between 1.4 and 4.4 Å was considered, with an interval of 0.3 Å; thus n was 11. Formula stated above is the implementable form of the original (Lewis and Rees 1985) formula:

$$FD = 2 - \frac{d(\log A_{vdW})}{d(\log R_p)}$$

FD magnitude was evaluated from linear profile of A_{vdW} (ordinate) versus R_p (abscissa) graph on a log–log scale. Applying this basic scheme, contact-FD (CFD), reentrant-FD (RFD), and total-FD (TFD) were calculated for contact, reentrant and total surface description respectively, for each of the protein molecules in both unbound and complex form. In the complexed form the interface atoms reside in ‘buried’ state, which makes it difficult to obtain information about them directly. But by systematically studying the surface roughness of the peripheral zone of the interface patch, before and after the interaction, valuable insights about the trend of change of surface roughness could be gathered—This is exactly what has been pursued throughout the present study.

Solid angles, drawn from center-of-mass of a protein, can map the protein exterior in a unique way that helps in rigorous analyses of subtended surface patches in a focused manner. In this work, protein surfaces were mapped through a series of non-overlapping solid angles. A small solid angle, viz., a solid angle with small magnitude

of both $\varphi - \theta$, can ‘zoom in’ on focused patches of protein surface; whereas a large solid angle (that is, solid angle with large values of $\varphi - \theta$ grid) are helpful to map large vistas of surface with minimal expenditure of time. Since the objective of the current work was to extract and quantify minuscule features of protein surfaces and interfaces, the smaller grids of solid angles were predominantly used.

However, since smaller solid angles often tend to map extremely focused areas on protein surfaces, more often than not, statistically significant number of atoms (necessary requirement for calculating the fractal dimension of the patch) could not be found in those small patches. During such cases, progressively higher grids of solid angles were employed to meet the requirement. Thus, although for every protein, a $\varphi - \theta$ grid of 15° – 15° was chosen as the initial probing scheme; often, the grids 15° – 30° , 30° – 30° , 30° – 45° , etc.,—had to be employed. In some extreme cases, the large grid, viz. 60° – 60° —had to be employed also.

The grid interval for solid-angle zones for a particular protein was kept constant. That means, for any protein, if the solid angle grid $\varphi - \theta$ was observed to ensure statistically significant number of atoms for most of the surface patches, it was kept constant while mapping the surface of that protein.

2.4.1.2 Algorithm 2nd Part: Scheme to Quantify Change of Interface Shapes

Next, to identify the points on the exterior of any arbitrarily chosen protein PR_1 , each point p on its shape S ($p \in S$) was assigned a unique address of the form ($a \in A_L$) (address space with L symbols). This could easily be achieved with one-to-one and onto mapping scheme $M : S \rightarrow A_L$; implying an address mapping $M(p) = a$. The surface roughness depends essentially upon positions of the atoms P ($P = \{p_i\}_{i=1}^n$ $n \geq 32$, to ensure statistical significance) that constitute the local patch of protein surface. The very same set of atoms holds the local shape of the interface for any one of the proteins (PR_1). Hence, if surface roughness is denoted by fd , it can be (trivially) noted that $fd = f(P)$. One may note that the address space defined here might not have to be merely limited to the crystallographic coordinate information; instead, it will change in its composition whenever a new property is considered (for example, the state of atomic polarizability values (Nagle 1990; Noorizadeh and Parhizgara 2005)—before, during and after the interaction might constitute three different address spaces; each of whom are dependent on atomic positional coordinates, but are not atomic positional coordinates themselves). To formalize, an attribute function φ_α was defined, such that $\varphi_\alpha : S \rightarrow Attr$ —that uniquely assigns the magnitude of relevant attribute ($\alpha \in Attr$) to a point ($p \in S$) (in other words, $\varphi_\alpha(p) = a$).

If position of every interface atom is defined in a space X , one can describe the change in the local shape of the interface of PR_1 by defining a shift vector \vec{v} ($\vec{v} \in X$). A systematic analysis of \vec{v} was found to be useful in monitoring every intermediate shape that the PR_1 interface may assume during PPI. This could be

expressed in compact form as $\varphi_{X,\vec{v}} : S \rightarrow X$ or (from an equivalent microscopic perspective) as $\varphi_{\vec{v}}(P) \rightarrow \vec{v}$. However, since it is reasonable to view PPIs as time-dependent and context-dependent phenomena, it will be realistic to assert $\varphi_{X,\vec{v}} : S, t \rightarrow X$, or equivalently, $\varphi_{X,\vec{v}}(P, t) \rightarrow \vec{v}$; where t denoted the time duration of PPI. One notes that $\varphi_{\vec{v}}$ is continuous in S and t . (Although any fractal surface is inherently non-differentiable, evolution of the shift vector, capturing the change of local shape of PR_1 is continuous). Furthermore, evolution of $\varphi_{\vec{v}}$ is self-similar in nature. Hence, one could sum up the interface shape change process as:

$$\varphi_{X,\vec{v}}(P, t) = \varphi_{\vec{v}}(P) + \varphi_{\vec{v}}(P, t),$$

or else,

$$\varphi_{X,\vec{v}}(fd, t) = \varphi_{\vec{v}}(fd) + \varphi_{\vec{v}}(fd, t).$$

For FD calculation of the surface patch(described before), our algorithm depended (implicitly) on the solid angle-based algorithm (Connolly 1986) to calculate local shape of protein. Thus in the present scheme the range of solid angle was considered to be $(0, 2\pi)$; solid angle $(<\pi)$ denoted (subtended) surface of a locally concave shape, solid angle $(>\pi)$ denoted the (subtended) surface of a locally convex shape and solid angle $(=\pi)$ denoted a flat surface. However, during PPI, the local shape of PR_1 was found to change. While minuscule at times, the change itself was found to be a consistent trend. This implied that the cardinality of the set of atoms representing the surface patch subtended by the solid angles, might not be constant. In such a case the FD, describing the roughness of the surface patch, was found to have undergone subtle change in magnitude too. The subtle change in the magnitude of FD, in fact, could be identified as a consistent trend; because even in cases where the aforementioned cardinality was found to be invariant, FD magnitude was still found to be changing. So to describe the entire situation from an atom-centric perspective, by denoting less or more number of atoms with ‘ l ’ and ‘ m ’ respectively, the present scheme could define an EUIST for $32 \leq l \leq m$, such that:

1. $\mu(l, m, a) = a$, when number of atoms counted in the solid angle did not change, viz. $l = m$
2. $\mu(l, m, a) = \mu(l, m, \mu(l, m - 1, a))$, number of atoms counted in the solid angle did change.

Mean values of l and m drawn from representative subset of enzyme-inhibitor interactions or antigen–antibody interactions were considered as constant value for l and m for that subset of interactions.

With the help of basic address mapping $M(p) = a$, one could connect the top-down scheme with the bottom-up (‘exact number of atom’-centric) view with a one-to-one onto mapping: $\psi(l, m, p) = M^{-1}(\mu'(l, m, M(p)))$. Thus the shift vector to describe shape transition could be expressed (more appropriately) as:

$$\varphi_{\vec{v}}(l, m, P, t) = \varphi_{\vec{v}}(P) + \varphi_{\vec{v}}(\psi^{-1}(l, m, P), t) = \varphi_{\vec{v}}(P) + \psi_{\vec{v}}^{-1}(l, m, P, t),$$

which is essentially the analogous bottom-up view of $\varphi_{X,\vec{v}}(P, t) = \varphi_{\vec{v}}(P) + \varphi_{\vec{v}}(P, t)$.

Due to sensitive nature of each of these parameters on context dependence, derivation of an analytic expression that connects P, t, fd, v, l, m — could not be achieved. In absence of analytic evaluation of aforementioned mapping schemes, to ascertain the shift vector, the present work resorted to an empirical scheme. Here, the mean values of the pertinent parameters were calculated. With a more focused dataset (that contained, for example, enzyme-inhibitor interactions for class specific enzymes in statistically significant number), better representative mean values could have been assigned to corresponding entities and better information about shift vector transition function could be obtained.

2.4.2 Results

Results for comparative investigation between surface roughness of enzyme-inhibitor interfaces and that of the non-interfacial parts are obtained for each solid angle zone of the proteins. Similarly results for comparative investigation between surface roughness of antigen–antibody interfaces and that of the non-interfacial parts are obtained for each solid angle zone of the proteins. Data were obtained for Connolly-defined CFD and TFD. (RFD could not be calculated for many solid angles, because of the paucity of number of atoms in them. These are therefore not included in the final analyses.) While the huge bulk of (solid angle specific) raw data obtained thereby is difficult to make sense of, the mean magnitudes of surface roughness before and after the interaction (presented in a consummate way in Table 1.1) helps to decipher certain broad trends.

2.4.2.1 Comments on the General Trends

Obtained results confirmed the long held view that surface roughness magnitudes of interfaces for PPI are low. There was not a single case from the entire spectrum of enzyme-inhibitor and antigen–antibody interactions when the interface TFD was found to be greater than 2.25. This magnitude is less than what some old studies (Pfeifer et al. 1985; Dewey 1994) had suggested. Such low magnitude

Table 1.1 Summary of trends in change of interface roughness, during enzyme-inhibitor and antigen–antibody interactions

Comparative profile of surface roughness by FD	Interface regions of enzymes and inhibitors	Non-interface regions of enzymes and inhibitors	Interface regions of antigens and antibodies	Non-interface regions of antigens and antibodies
Before the interaction	2.119 ± 0.029	2.111 ± 0.033	2.114 ± 0.043	2.131 ± 0.041
After the interaction	2.150 ± 0.029	2.179 ± 0.030	2.108 ± 0.044	2.142 ± 0.041

of interface-FD could be explained [in the line of a previous study (Pfeifer et al. 1985)] as continuous compromise between two contrasting tendencies for optimization over the course of evolution. Higher FD magnitude of interfaces would have helped in maximizing the capture of substrate molecules from the bulk phase, whereas increase of speed of migration of substrate molecules along the protein surfaces would have been more probable if FD of interfaces assumed low magnitudes. Aqvist and Tapia (1987), while commenting upon the causality behind protein surface smoothness, had also emphasized on the contradiction between capture rate and diffusion. The fact that regularity of FD magnitude for protein surfaces as obtained from the present investigation is not a protein-specific feature but of statistical nature, pointed to the fact that aforementioned explanation may well be true.

The CFD magnitudes, ranging between 2.30 and 2.92, on the other hand, depicted a completely different picture. But this trend should not be considered as a contradiction to the trend observed for the TFDs. CFD magnitudes are calculated only from the contact surface roughness; therefore, they suffer from drastic changes with slightest change of probe radius (Islam and Weaver 1991). More importantly, the CFD magnitudes are merely artifacts of computational methodology. Roughness of protein surfaces owes its origin to both CFD and RFD. The RFD magnitude could not always be calculated because of the unavailability of statistically significant number of atoms in the reentrant surface of interfaces. While the CFD and TFD trends do not contradict each other, no relationship could be established among them, to predict one from another. Though a broad trend, viz., increment (or decrement) in CFD, almost always accounted for the increment (or decrement) of TFD—could be observed, some exceptions to this trend could be observed also. In other words, no proportionality could be found between CFD and TFD, at least within the scope of the current study. To minimize the scope of confusion, by ‘FD’ from here onwards, only the TFD will be implied.

Since FD values ranged only within 2.04–2.23, attention was frequently drawn to the magnitude of the second decimal point to identify patterns in the changing profile of FD magnitudes, both in enzyme-inhibitor and antigen–antibody studies. The words “more” or “less”, “increase” and “decrease” etc., in the present report should be understood in this context. To what extent such small changes in FD values can be decisive in molecular recognition, is a debatable question. An objective quantification of the same, nevertheless, was necessary.

2.4.2.2 Surface Roughness Comparison Between Interfacial and Non-interfacial Patch Before and After PPI

This set of results was derived to identify any possible distinguishing trend in the magnitude of surface roughness that the interfacial surface patch might possess over the non-interfacial ones. As a predominant trend, FD of the interface patch for enzyme-inhibitor interactions could be observed to be more than that in non-interfacial patches in uncomplexed state. Although the difference was small, it was hardly ignorable. That is, before PPI, $FD^{\text{enz-inh}}_{\text{interface}} = 2.119 \pm 0.029$

and $FD^{enz-inh}_{non-interface} = 2.111 \pm 0.033$ —was observed. Studies on uncomplexed units of antigens, and antibodies, however, revealed the trend just opposite to what was observed in enzymes and inhibitors. For antigens and antibodies, the leading trend (eight out of ten cases) revealed that interfacial patches of these proteins, even in their uncomplexed forms, have smooth surfaces in comparison to non-interfacial surface patches of the same proteins. This is why, before the PPI, the magnitudes $FD^{antg-antbd}_{interface} = 2.114 \pm 0.043$ and $FD^{antg-antbd}_{non-interface} = 2.131 \pm 0.041$ —could be recorded. The margin of smoothness possessed by the interfacial patches of antigens and antibodies, could be noted to be less than the margin of roughness that characterized the interfacial surface patches for enzymes and inhibitors.

Principal trend for the enzyme-inhibitor interactions after PPI suggested an increase of surface roughness of the interface (and adjoining) surface patch from its magnitude in the uncomplexed state. This could be observed in the comparison of mean magnitudes of FD for interfacial surface patch before and after PPI. While the results for surface roughness of enzyme-inhibitor interfaces before PPI has been presented above; the same for after PPI, states were recorded as: $FD^{enz-inh}_{interface} = 2.150 \pm 0.029$ and $FD^{enz-inh}_{non-interface} = 2.179 \pm 0.030$. However, for the antigen–antibody interactions, the trend was less clear-cut. Here, in four out of ten cases, FD magnitude of the interfacial surface patch could be observed to have become less in the interface (and adjoining surface patch) in the complexed form, than what it was in the uncomplexed state. For another four cases though, the interfacial FD magnitude was found to be slightly higher in the interface than what it was in the uncomplexed state. For two other cases, the interface-FD magnitudes (almost) equalled the non-interface-FD. Because of these contradictory trends, the global result of comparative roughness profile between antigen–antibody interfaces (along with their adjoining surface patches) and non-interface regions, after the PPI, viz. $FD^{antg-antbd}_{interface} = 2.108 \pm 0.044$ and $FD^{antg-antbd}_{non-interface} = 2.142 \pm 0.041$ —could not show a distinct pattern. Although the mean magnitude of FD values tends to suggest an unambiguous reduction of surface roughness of antigen and antibody interfaces, one has to admit that the trend of reduction of interface-FD is not as general as the trend of increase of interface-FD in enzyme-inhibitor interactions.

2.4.2.3 Patterns in Change of Surface Roughness (FD) in Interfacial Patches

Findings presented in the last section for before and after PPI, provide a composite picture of the process. In case of the enzyme-inhibitor interactions, the portion of the surface patch in the uncomplexed form that serves as interface during PPI, for both enzymes and inhibitors, could be observed to have (slightly) higher surface roughness than the other (non-interfacial) portions of the same molecules. Surface roughness of interfaces increased by a little margin in the course of PPI for enzyme-inhibitor interactions; however, during PPI, the roughness of the

non-interfacial parts of both enzymes and inhibitors, increased by slightly bigger margin. Hence, the interfacial parts of the enzyme-inhibitor complexes appear to be flat, when observed with respect to the non-interfacial patches of the same proteins.

For the antigen–antibody interactions, such clear (and dramatic) trends could not be observed. The portion of the surface patch in the uncomplexed form that serves as interface during PPI, for both antigens and antibodies, were found to have (slightly) less surface roughness than the non-interfacial portions of the same molecules. Aforementioned roughness was observed to decrease further in the course of PPI; whereby in four out of ten cases, the interfacial roughness assumed less magnitude than that of the (non)-interfacial roughness. In another four cases interfacial roughness was found to assume higher magnitude than that of the (non)-interfacial roughness; and finally, in two cases they were found to show exactly the same magnitude, presenting ultimately an inconclusive result bereft of any tangible trend.

Since it is believed that enzymes and their inhibitors co-evolve to form interfaces with high degree of surface complementarity, reason behind the consistent trend of FD increase in interface regions of both of them, could easily be understood. Higher magnitude of surface roughness could ensure an increase in the probability of better surface complementarity for the enzyme-inhibitor interactions. However, since it is known that enzyme-inhibitor and antigen–antibody complexes represent two different classes of binding (Lawrence and Colman 1993), the same logic could not be applied on the later. The immune system produces different types of antibodies in response to an antigen, some of which bind their respective epitopes quite well while others bind poorly. Thus a particular antigen–antibody complex does not necessarily possess the best possible binding interface. Such an interpretation could help in understanding the (rather) inconclusive nature of the obtained results, when it came to antigen–antibody interactions.

Obtained results vindicated the age-old assertions that PPI interfaces are notoriously difficult to be identified keeping surface roughness as a probe. Work presented in this chapter presented the entire anatomy of the process by quantifying surface roughness of the interfacial and non-interfacial surface patches for all the involved monomers before and after the PPI. Trends observed from obtained results tend to suggest that, for all practical purposes, the magnitude of surface roughness alone, account for little help in identifying the interfacial regions. However, this difficulty could be overcome, to somewhat extent, by incorporating the shape-vector based information about trends in the change of local shape on the interfaces.

2.4.2.4 Patterns in the Magnitude of Shift Vector

From an empirical perspective, magnitude assumed by the function describing the shift vector [microscopic description: $\psi_v^{-1}(l, m, P, t) (= \varphi_v(l, m, P, t) - \varphi_v(P))$, macroscopic description: $\varphi_v(P, t) (= \varphi_{X,\vec{v}}(P, t) - \varphi_v(P))$] was calculated from the mean magnitudes of $\varphi_v(l, m, P, t)$, $\varphi_v(P)$, $\varphi_{X,\vec{v}}(P, t)$ and $\varphi_v(P)$ —as observed in enzyme-inhibitor and antigen–antibody cases. Positive magnitudes of shift vectors in most cases for enzyme-inhibitor interactions suggested

an increment of convexity and complementary concavity in the local shape of interfacial area, associated with complex formation. The $\Delta FD_{enz-inh} = 0.031$ —indicates that for uncomplexed enzymes and uncomplexed inhibitors, the curvature of the local shape that holds their interfacial surface patches, increases during interaction, albeit by a little margin. Whereas, demonstrating just the opposite trend, consistent decrease (four out of five cases) of shift vector magnitudes in case of antigen–antibody interactions symbolized a decrease of convexity and concavity that characterized the interfacial shape of the uncomplexed antigens and antibodies. As a result of such decrease (the $\Delta FD_{antgn-antbd} = -0.006$), albeit small in margin, the antigen–antibody complexes could be observed to possess flat shaped interfacial surface patches in them. The fact that shape correlations might not be an important parameter while studying antigen–antibody interactions, was pointed out long ago (Lawrence and Colman 1993; Novotný et al. 1986). Current study provided a comprehensive theoretical characterization behind such claim.

The utility of resorting to shift vector calculation became apparent with these observations; that is, even without knowing the exact (non-linear, time-dependent, context-dependent) mappings between P, t, fd, \vec{v}, l, m ; one could easily form an idea about the nature of change in the curvature of interface from the aforementioned formulae. Although the current work considered only the final magnitudes of these parameters, one may easily implement this algorithm in molecular dynamics centric studies to obtain series of magnitudes of these parameters, which might help in obtaining an elaborate observation of the evolution of shape change during the course of an interaction. Furthermore, one notes that changes in FD magnitudes are extremely small in the current work, involving mostly the second or third decimal places. The fact that shift vectors could still reveal a change in local shape building upon these minuscule differences—suggested affirmatively that they are sensitive.

An increase or decrease of curvature (either in convexity or concavity) does not necessarily imply a change in the roughness of the surface that is holding it. However, interesting geometrical aspects of such surface patch could be detected with attempts that concentrated on particular local regions of protein exterior by mapping it with solid angles drawn from any suitably chosen reference (say, centre of mass of the protein with any suitable grid of inclination and azimuth). During the course of PPI, change in the number of atoms in the surface patch described by solid angles, becomes a non-trivial possibility. In such a case, when local curvature is changing alongside the number of atoms in surface patch subtended by the solid angle, a measure of surface roughness or a measure of local curvature alone, would have merely accounted for incomplete information. The current methodology presented an approach to connect these two. One notes the existence of another algorithm that attempts to connect local shape with the roughness of the surface that holds this shape (Banerji 2011),—however, this method fails to throw much light on the nature of change in local shape (viz. the increase or decrease of curvature of local shape etc.), which can readily be obtained from an implementation of the present one.

This work can at best be viewed as a ‘pilot-project’ to investigate the paradigm of protein–protein interaction interfaces from a new and rigorous theoretical and

computational standpoint. While some of its findings merely vindicated the old assertions, the new computational structure proposed in this work could unite the information about surface and shape deformation study successfully. The inclusive approach of the employed methodology could ensure that protein–protein interaction interfaces are not only investigated from studying the monomeric surface roughness, but also from the (tiny but hardly ignorable) evidences, collected from peripheral surface patches of the (buried) interface. Since the shape changes during PPI were considered and connected with the surface roughness change profile, the current methodology could present a reliable template for dissection of geometric features of PPI interfaces.

References

- Aqvist J, Tapia O (1987) Surface fractality as a guide for studying protein–protein interactions. *J Mol Graph* 5:30–34
- Argos P (1988) An investigation of protein subunit and domain interfaces. *Protein Eng* 2:101–113
- Arkin M, Randal M, DeLano W, Hyde J, Luong T, Oslob J, Raphael D, Taylor L, Wang J, McDowell R, Wells J, Braisted A (2003) Binding of small molecules to an adaptive protein–protein interface. *Proc Natl Acad Sci USA* 100:1603–1608
- Bartlett GJ, Porter CT, Borkakoti N, Thornton JM (2002) Analysis of catalytic residues in enzyme active sites. *J Mol Biol* 324:105–121
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H (2000) Shindyalov i.n., bourne p.e., the protein data bank. *Nucleic Acids Res* 28(1):235–242
- Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. *J Mol Biol* 280:1–9
- Bowman RL (1995) Fractal metamorphosis: a brief student tutorial. *Comput Graph* 19(1):157–164
- Bradford JR, Westhead DR (2005) Improved prediction of protein–protein binding sites using a support vector machines approach. *Bioinformatics* 21(8):1487–1494
- Chen R, Weng Z (2003) A novel shape complementarity scoring function for protein–protein docking. *PROTEINS: structure. Funct Genet* 51:397–408
- Chothia C, Janin J (1975) Principles of protein–protein recognition. *Nature* 256:705–708
- Connolly ML (1986) Measurement of protein surface shape by solid angles. *J Mol Graph* 4:3–6
- DeLano WL, Ultsch MH, de Vos AM, Wells JA (2000) Convergent solution to binding at a protein–protein interface. *Science* 287:1279–1283
- DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. *Curr Opin Struct Biol* 12:14–20
- Dewey TG (1994) Fractal analysis of proton exchange kinetics in lysozyme. *Proc Natl Acad Sci USA* 91:12101–12104
- Fujimoto T, Chiba N (2004) Fractal deformation using displacement vectors and their increasing rates based on extended unit iterated shuffle transformation, “Thinking in patterns: fractals and related phenomena in nature”. World Scientific, Ed. Miroslav M Novak, pp 57–68
- Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D (2003) Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol* 331:281–299
- Holliday GL, Almonacid DE, Bartlett GJ, O’Boyle NM, Torrance JW, Murray-Rust P, Mitchell JBO, Thornton JM (2007) MACiE (mechanism, annotation and classification in enzymes): novel tools for searching catalytic mechanisms. *Nucleic Acids Res* 35:D515–D520
- Hutchinson JE (1981) Fractals and self similarity. *Indiana Univ Math J* 30:713–747
- Islam SA, Weaver DL (1991) Variation of folded polypeptide surface area with probe size. *Proteins Struct Funct Genet* 10(4):300–314

- Janin J (1995) Principles of protein–protein recognition from structure to thermodynamics. *Biochimie* 77:497–505
- Janin J (1997) Specific versus non-specific contacts in protein crystals. *Nat Struct Biol* 4:973–974
- Jefferson E, Walsh T, Barton G (2006) Biological units and their effect upon the properties and prediction of protein–protein interactions. *J Mol Biol* 364:1118–1129
- Joachimiak LA, Kortemme T, Stoddard BL, Baker D (2006) Computational design of a new hydrogen bond network and at least a 300-fold specificity switch at a protein–protein interface. *J Mol Biol* 361:195–208
- Jones S, Thornton JM (1996) Principles of protein–protein interactions. *Proc Natl Acad Sci USA* 93:13–20
- Jones S, Thornton JM (1997) Analysis of protein–protein interaction sites using surface patches. *J Mol Biol* 272:121–132
- Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci USA* 89:2195–2199
- Keskin O, Mab B, Nussinov R (2005) Hot Regions in Protein–Protein Interactions: The Organization and Contribution of Structurally Conserved Hot Spot Residues. *J Mol Biol* 345(5):1281–1294
- Kuhlmann UC, Pommer AJ, Moore GR, James R, Kleanthous C (2000) Specificity in protein–protein interactions: the structural basis for dual recognition in endonuclease colicin–immunity protein complexes. *J Mol Biol* 301:1163–1178
- Lawrence MC, Colman PM (1993) Shape complementarity at protein–protein interfaces. *J Mol Biol* 234:946–950
- Lewis M, Rees DC (1985) Fractal surfaces of proteins. *Science* 230:1163–1165
- Li X, Keskin O, Ma B, Nussinov R, Liang J (2004) Protein–protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *J Mol Biol* 344:781–795
- Ma B, Shatsky M, Wolfson HJ, Nussinov R (2002) Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci* 11:184–197
- Nagle JK (1990) atomic polarizability and electronegativity. *J Am Chem Soc* 112(12):4741–4747
- Nooren IMA, Thornton JM (2003) Structural characterisation and functional significance of transient protein–protein interactions. *J Mol Biol* 325:991–1018
- Noorizadeh S, Parhizgara M (2005) The atomic and group compressibility. *J Mol Str Theochem* 725(1–3):23–26
- Novotný J, Handschumacher M, Haber E, Bruccoleri RE, Carlson WB, Fanning DW, Smith JA, Rose GD (1986) Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). *Proc Natl Acad Sci USA* 83(2):226–230
- Pfeifer P, Welz U, Wippermann H (1985) Fractal surface dimension of proteins: lysozyme. *Chem Phys Lett* 113:535–540
- Rajamani D, Thiel S, Vajda S, Camacho CJ (2004) Anchor residues in protein–protein interactions. *Proc Natl Acad Sci USA* 101:11287–11292
- Sederberg TW, Parry SR (1986) Free-form deformation of solid geometric models. *ACM Comput Graph* 20(4):151–160
- Smith GR, Sternberg MJE, Bates PA (2005) The relationship between the flexibility of proteins and their conformational states on forming protein–protein complexes with an application to protein–protein docking. *J Mol Biol* 347:1077–1101
- Todd A, Orengo C, Thornton J (2002) Sequence and structural differences between enzyme and nonenzyme homologs. *Structure* 10:1435–1451
- Yogurtcu ON, Erdemli SB, Nussinov R, Turkay M, Keskin O (2008) Restricted mobility of conserved residues in protein–protein interfaces in molecular simulations. *Biophys J* 94:3475–3485

Fractal Symmetry of Protein Exterior

Banerji, A.

2013, X, 55 p. 5 illus., 2 illus. in color., Softcover

ISBN: 978-3-0348-0653-4