Biochimica et Biophysica Acta 1834 (2013) 404–414

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

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

Understanding the importance of the aromatic amino-acid residues as hot-spots

I.S. Moreira *, J.M. Martins, R.M. Ramos, P.A. Fernandes, M.J. Ramos

REQUIMTE/Departamento de Química, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

ARTICLE INFO

Article history:
Received 27 April 2012
Received in revised form 26 June 2012
Accepted 17 July 2012
Available online 25 July 2012

Keywords:
Alanine scanning mutagenesis
Hot-spot
Aromatic residue
Aromatic cluster
Cooperativity
SASA

ABSTRACT

Protein–protein interactions (PPI) are crucial for the establishment of life. However, its basic principles are still elusive and the recognition process is yet to be understood. It is important to look at the biomolecular structural space as a whole, in order to understand the principles behind conformation–function relationships. Since the application of an alanine scanning mutagenesis (ASM) study to the growth hormone it was demonstrated that only a small subset of residues at a protein–protein interface is essential for binding — the hot-spots (HS). Aromatic residues are some of the most typical HS at a protein–protein interface. To investigate the structural role of the interfacial aromatic residues in protein–protein interactions, we performed Molecular Dynamic (MD) simulations of protein–protein complexes in a water environment and calculated a variety of physical–chemical characteristics. ASM studies of single residues and of dimers or high-order clusters were performed to check for cooperativity within aromatic residues. Major differences were found between the behavior of non-HS aromatic residues and HS aromatic residues that can be used to design drugs to block the critical interactions or to predict major interactions at protein–protein complexes.
© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Protein–protein interactions (PPI) are essential for all biological processes. Since the application of an alanine scanning mutagenesis study to the growth hormone by Wells et al. the world become aware that only a small subset of residues at a protein–protein interface is essential for binding — the hot-spots (HS) [1]. These residues were defined as amino-acids that upon alanine mutation present a $\Delta\Delta G_{binding}$ (binding free energy difference) value higher than 2.0 kcal/mol [2]. Statistical studies have shown that the aromatic residues, especially, tyrosine (Tyr) and tryptophan (Trp) can be found more frequently at the core of interactions as they establish energetically favorable interactions with the surrounding environment [3]. For example, tyrosine residues are located in the antigen-binding sites for antibodies and are responsible for their high specificity. The two remaining aromatic residues, phenylalanine (Phe) and histidine (His), can also occur at the binding interface [4]. The reason why aromatic residues are preferred in PPI can be understood by their physicochemical properties. Tyrosines have both a hydrophobic ring and a hydrophilic hydroxyl group in one side chain contributing to binding by hydrogen bond via the side chain 4-hydroxyl group, as well as hydrophobic interactions, van der Waals interactions and amino-aromatic (cation–π) interactions via the aromatic ring. In comparison, although phenylalanine has an aromatic ring, its frequency in PPI is a third of the one of Tyr residues. Tryptophan can also contribute with π-interactions, hydrophobic interactions and hydrogen bonding [4]. Its large size facilitates binding with a large number of residues at the same time, allowing it to be a central node in protein–protein interfacial contact network [5]. Histidine is an aromatic residue that has an imidazole ring. As a heteroaromatic moiety it can interact by its rings with nonpolar and aromatic groups and by its heteroatoms participate in hydrogen bonds. Depending on the protonation state it can be also involved in salt-bridges with acidic groups [6].

π–π interactions are of utmost importance in drug chemistry as 61% of Phe residues, 54% of Tyr residues and 59% of Trp residues in a protein–protein interface are involved in this kind of interactions [7,8]. They generally depend upon charge distribution and the shape of the molecule. Although individually they are weak, being highly numerous, they have a great influence over key chemical and biological processes. Cation–π interactions constitute a strong specific driving force that plays a key role in molecular recognition. It involves short distance interactions and it is among the strongest non-covalent binding forces [7]. The geometries most commonly observed for these interactions are not necessarily the ones that have the highest interaction energy between the intervening pair, but the ones that can provide the maximum overall stability to the protein structure by the optimum use of all hydrogen bonding sites [9]. A complete review about aromatic interactions can be found in [10,11]. These aromatic interactions could be essential not only by the formation of pair interactions but instead by the formation of higher order clusters, beyond the dimer that can adopt specific conformations. It has been shown that interactions between these residues can contribute to the stability of the native fold [8] and it has been suggested that the aromatic clusters could be a determinant of

* Corresponding author. Tel.: +351 220 402 000.
E-mail address: irina.moreira@fc.up.pt (I.S. Moreira).

1570-9639/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbapap.2012.07.005

thermal stability in thermophilic proteins [12]. Statistical studies have shown that aromatic clusters can be found in protein structures and that in about half of the times form high order clusters that bring nonlocal primary structure distant sites together [13]. In order to understand the role and importance of the interfacial aromatic residues in protein–protein interactions, we performed Molecular Dynamic (MD) simulations of various protein–protein complexes in a water environment and calculated a variety of physical–chemical characteristics. Our main objective was to differentiate the structural pattern involving aromatic residues that behaved as HS from the ones that act as NS (Null-Spots: residues that upon alanine mutation generate a binding free energy difference lower than 2.0 kcal/mol). Hence, four different characteristics were studied: energetic character, micro-environment, water's role and the formation of high order clusters.

2. Methodology

2.1. System setup

Seven different complexes for a total of 50 aromatic residues were studied: between Barnase and Barnstar (PBDID: 1BRS [14]); Igg1 Kappa D1.3 FV and Igg1 Kappa E5.2 FV (PDBID:1DVF [15]); Ribonuclease A and Ribonuclease Inhibitor (PDBID: 1DFJ [16]); Bacterial cell division ZipA and Ftsz (PDBID: 1F47 [17]); Vascular Endothelial Growth Factor and FLT-1 Receptor (PDBID: 1FLT [18]); Fibroblast Growth Factor 2 and Fibroblast Growth Factor Receptor 1 (PDBID: 1FQ9 [19]); and Igg1 Kappa D1.3 FV and Hen Egg white lysozyme (PDBID: 1VFB [20]). These systems were selected based on the existence of experimental binding free energy ($\Delta\Delta G_{binding}$) values for the interfacial residues upon alanine mutations and the existence of a high number of aromatic residues at their interfaces. The protonation state of the different residues of the various proteins was determined using the PDB2PQR server at http://kryptonite.nbcr.net/pdb2pqr/ [21] by the PROPKA methodology [22–24] that computes the pKa values of the ionizable residues in a protein by determining a perturbation to the model pKa value due to the protein environment.

2.2. Molecular Dynamic simulations

The MD simulations were performed using the AMBER9 package [25] with the Cornell force field ff03 [26]. Two different 8 ns simulations were made, one in an implicit solvent using the Generalized Born (GB) solvent [27], and another using TIP3P explicit water molecules. The GB simulations (GB^{OBC}) [28] were used in the alanine scanning mutagenesis protocol as it was shown in previous work their capacity to reproduce more accurately the experimental binding free energy values [29,30]. The explicit solvent simulations were used to study the dynamical behavior and crucial interactions of the aromatic residues between each other or with other molecules at the system. The complexes were solvated by explicit waters that extended 10 Å from any edge of the cubic box to the protein atoms. Counter ions were added to the boxes to neutralize the system. In the GB simulations, the ionic strength was set to 0. In each of the simulations, the system was initially minimized to remove bad contacts by steepest descent followed by conjugated gradient. The systems were then subjected to 2 ns of heating procedure (in NVT ensemble) in which the temperature was gradually raised to 300 K, followed by 6 ns runs in NPT ensemble. The Langevin [31,32] thermostat was used and the electrostatic interactions were calculated by using the particle mesh Ewald (PME) method [33]. Bond lengths involving hydrogens were constrained using the SHAKE algorithm [34]. The equations of motion were integrated with a 2 fs time-step and the non-bonded interactions were truncated with a 16 Å and a 10 Å cutoff, in the GB and in explicit solvent simulations respectively.

2.3. Alanine scanning mutagenesis (ASM) protocol

The MM-PBSA (Molecular Mechanics Poisson Boltzmann Surface Area) script [35] integrated into the AMBER9 package [25] was used to calculate the binding free energy difference upon alanine mutation. It combines a continuum approach to model solvent interactions with a MM-based approach to atomistically model protein–protein interactions. This provides speed and accuracy and has been quite used in the last years [29,30,35–45]. The MM-PBSA approach first developed by Massova et al. [35] was improved by Moreira et al. [30] and can now be applied with an accuracy of 1 kcal/mol. The mutant complexes are generated by a single truncation of the mutated side chain, replacing Cγ with a hydrogen atom and setting the Cβ-H direction to that of the former Cβ-Cγ. For the binding energy calculations, a total of 26 snapshots of the complexes were extracted in the last 1 ns of the run. The ΔΔG is defined as the difference between the mutant and wild type complexes defined as:

$$\Delta\Delta G = \Delta G_{cpx-mutant} - \Delta G_{cpx-wild\ type} \quad (1)$$

Typical contributions to the free energy include the internal energy (bond, dihedral, and angle), the electrostatic and the van der Waals interactions, the free energy of polar solvation, the free energy of nonpolar solvation, and the entropic contribution:

$$G_{molecule} = E_{internal} + E_{electrostatic} + E_{vdW} + G_{polar\ solvation} + G_{non-polar\ solvation} - TS \quad (2)$$

For the calculations of relative free energies between closely related complexes, it is assumed that the total entropic term in Eq. (2) is negligible as the partial contributions essentially cancel each other [43]. The first three terms of Eq. (2) were calculated with no cutoff. The $G_{polar\ solvation}$ was calculated by solving the Poisson–Boltzmann equation with the software DELPHI [46,47]. In this continuum method, the protein is modeled as a dielectric continuum of low polarizability embedded in a dielectric medium of high polarizability. We used a set of values for the DELPHI parameters that proven in a previous study to constitute a good compromise between accuracy and computing speed [48]. We used a value of 2.5 grids/Å for scale (the reciprocal of the grid spacing); a value of 0.001 kT/c for the convergence criterion; a 90% for the fill of the grid box; and the Coulombic method to set the potentials at the boundaries of the finite-difference grid. The dielectric boundary was taken as the molecular surface defined by a 1.4 Å probe sphere and by spheres centered on each atom with radii taken from the Parse [49] vdW radii parameter set. The key aspect of the new improved approach is the use of a three dielectric constant set of values ($\varepsilon = 2$ for nonpolar residues, $\varepsilon = 3$ for polar residues and $\varepsilon = 4$ for charged residues plus histidine) to mimic the expected rearrangement upon alanine mutation (the method is described in [29,30]). It is important to highlight that we used only one trajectory for the computational energy analysis as it has been proven to give the best results [30]. Side-chain reorientation was implicitly included in the formalism by raising the internal dielectric constant. The nonpolar contribution to the solvation free energy due to van der Waals interactions between the solute and the solvent was modeled as a term dependent of the solvent accessible surface area (SASA) of the molecule. It was estimated by $0.00542 \times SASA + 0.92$ using the molsurf program developed by Mike Connolly [50]. As a systematic mutation of residues on PPI is a fastidious and time consuming methodological approach we have recently developed a VMD [51] plugin (http://compbiochem.org/Software/compasm/Home.html) [52]. This plugin has an easy-to-use graphical interface to prepare the input files, run the calculations and analyze the final results and was used in this work. The performance of the ASM method can be assessed by the use of the F1 score (Eq. (3)), which is defined as a function of Precision (P, also called specificity,

Download English Version:

https://daneshyari.com/en/article/10537181

Download Persian Version:

https://daneshyari.com/article/10537181

Daneshyari.com