Journal of Molecular Graphics and Modelling 42 (2013) 92–103

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

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

An *in silico* method for designing thermostable variant of a dimeric mesophilic protein based on its 3D structure

Sohini Basu, Srikanta Sen*

Molecular Modeling Group, Biolab, Chembiotek, TCG Lifesciences Ltd., Bengal Intelligent Park, Tower-B 2nd Floor, Block-EP & GP, Sector-V, Salt Lake Electronic Complex, Calcutta 700091, India

ARTICLE INFO

Article history:
Received 1 December 2012
Received in revised form 25 February 2013
Accepted 27 February 2013
Available online 21 March 2013

Keywords:
Protein design
Protein dimer
Mutation
Conformer library
Self-energy
Stability

ABSTRACT

Designing proteins with enhanced thermostability has been a major interest of protein engineering because of its potential industrial applications. Here, we have presented a computational method for designing dimeric thermostable protein based on rational mutations on a mesophilic protein. Experimental and structural data indicate that the surface stability of a protein is a major factor controlling denaturation of a protein and ion-pairs are most efficient in enhancing the stability of the surfaces of the monomers and the interface between them. Our mutation based strategy is to first identify several polar or charged residues on the protein surface, interacting weakly with the rest of the protein and then replacing the side-chains of suitable neighboring residues to increase the interaction between these two residues. In stabilizing the interface, mutation is done in the interface for forming an ion pairs between the monomers. Application of this design strategy to a homo-dimeric protein and a hetero-dimeric protein as examples has produced excellent results. In both the cases the designed mutated proteins including the individual monomers and the interfaces were found to be considerably more stable than the respective mesophilic proteins as judged by self-energies and residue-wise interaction patterns. This method is easily applicable to any multi-meric proteins.

1. Introduction

Organisms whose optimal growth temperatures are in the range from 20 to 50 °C are known as mesophilic organisms. Mesophilic organisms are most common in nature. Over the last decade, new organisms have been found in hot places such as hot springs and hot vents at the bottom of the sea. Such organisms grow and function at high temperatures and are classified as thermophiles [1,2]. If the optimal growth temperature (T_{opt}) of the thermophilic organism is in the range 50 °C < T_{opt} < 80 °C, the organism is called a moderate thermophile while, organisms with T_{opt} > 80 °C, are called hyperthermophiles. Proteins isolated from thermophilic and hyperthermophilic organisms remain structurally stable and functionally active at much higher temperatures at which their mesophilic counterparts are structurally unstable and hence inactive [3,4]. Thermophilic and hyperthermophilic proteins and their mesophilic counterparts generally possess high degree of sequence similarity and similar three-dimensional structures [5,6].

Designing proteins with enhanced thermostability is one of the major focuses of modern protein engineering. Thermophilic proteins are of great interest not only for their extraordinary stability but also because of their potential industrial uses [2,7–9]. Thus, it is important to understand the physical basis of the enhanced stabilities of thermophilic proteins at elevated temperatures and, also to use that acquired knowledge in designing tailor-made thermostable proteins for industrial applications. There are a few approaches for making a given protein a thermophilic one [10–14]. However, these methods depend mostly on the sequence comparison, and it may happen that a specific amino acid bias in thermostable proteins is more related to the evolutionary changes than a direct relevance to its thermostability. Thus, designing methods based on the 3D structure of the protein appears to be more straight forward and reliable. In the present work we have proposed a structure-based effective and robust computational approach for converting a mesophilic multimeric protein into a thermophilic one. This method is an extension of our earlier design approach for turning a single chain single globule mesophilic protein into a thermophilic variant [15,16].

The physical origin of the enhanced stability of thermophilic proteins has been investigated extensively [16–25]. Our current understanding indicates that there are a number of factors that individually or in combination can cause the extra stability. Such factors include optimized electrostatic interactions such as increased number of salt-bridges, improved packing, networks of hydrogen bonds, increased hydrophobic interactions, and decreased number and volume of internal cavities, *etc.* [17–33]. Till now, no general rules have been identified to determine thermostability, and different thermophilic proteins appear to achieve their increased

* Corresponding author. Tel.: +91 33 23572744/45; fax: +91 33 23574113.
E-mail addresses: srikantasen@ymail.com, srikanta@chembiotek.com (S. Sen).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.02.011

thermostability due to different combinations of the above mentioned factors. In the natural systems the enhancement of stability due to multiple factors is more likely. As a consequence, in designing thermophilic variants of mesophilic proteins one can prefer a single factor or a few factors of choice. However, it is now well accepted that electrostatic interactions like ion-pairs are the most common factors observed in hyper-thermophilic proteins. Comparative studies have revealed that many thermophilic proteins have increased number of salt bridges [1,17,22,23,25] compared to the homologous mesophilic proteins. Moreover, a large number of salt bridges are generally found to be present on the surface of thermophilic proteins, and these salt bridges may participate not only in intra-protein interactions but also in inter-domain and inter-subunit interfaces of proteins for the stabilization of the whole protein. In a recent paper, it has been reported that a protein CutA1 isolated from *Pyrococcus Horikoshii*, has a very high melting temperature (150 °C) and its crystal structure indicates that the surface of this protein is almost covered by ion-pairs [25]. Based on these facts, we have developed a computational method for generating thermophilic proteins by introducing mutations that improve the electrostatic energy of the 3D structure of the respective mesophilic counterpart. We have published earlier the main concept and its validation along with the demonstration of designing thermostable variants of single chain single globule mesophilic proteins [15,16]. The purpose of the present work is to demonstrate that the same approach can be extended to convert a dimeric mesophilic protein into a dimeric thermophilic protein. It is important to emphasize that in converting a multimeric mesophilic protein into a multimeric thermostable protein, it is not enough to simply enhance the stabilities of the individual monomers only, but also to stabilize further the interfaces between the monomers such that the quaternary complex is stabilized as a whole. For the sake of simplicity, here we have demonstrated two examples of dimeric proteins of which, one is a homo-dimer and the other is a hetero-dimer but the approach can be applied to any multimeric protein as well. In both the examples it has been demonstrated explicitly that the designed mutations have improved the stability of the individual monomers as well as the stability of the interface. As experimental verification for the enhanced stability of the designed protein is not amenable to us, we have demonstrated the enhanced stability in terms of computational properties. We have chosen the modulation of electrostatic interactions to achieve enhanced stability because, so far, it is known to be the most common stabilizing factor in many cases. Consideration of other factors in designing thermophilic proteins may be addressed in the future. The present method is found to be a versatile and straight forward design approach in converting a multimeric mesophilic protein into a multimeric thermostable one. Validation of the general approach of our design method based on rational mutations in converting a single chain single globule mesophilic protein into a thermophilic protein has already been described in our earlier work where, it was demonstrated by considering several naturally occurring homologous thermophilic and mesophilic protein pairs that such natural protein pairs satisfy all the criteria we have used in converting a mesophilic protein into a thermophilic one [15,16]. In the present work the design method has been extended to cover dimeric proteins through additional stabilization of the interface.

2. Materials and methods

2.1. Rational for stabilizing the interface at high temperature

It is known that hydrophobic interaction plays an important role in forming dimeric proteins. However, as hydrophobic interaction is weakened with increase in temperatures, it is unlikely that hydrophobic interactions provide enhanced stability of the interface for a thermophilic protein. On the other hand, there are evidences that the interfaces for dimeric thermostable proteins are stabilized by ion pairs formed between the monomers [34–36]. Based on this fact, we have developed a method for identifying mutations on the two monomers at the interface allowing the formation of inter-monomer ion pairs and thus, stabilizing the dimer.

2.2. Choice and preparation of the dimeric mesophilic proteins of interest

The RCSB database was searched to find the crystal structures of dimeric proteins of moderate sizes and a homo-dimeric protein (*Fibroblast growth factor receptor 2*, PDB ID: 3EUU) and another hetero-dimeric protein (*Sugar binding protein Lectin 1*, PDB ID: 1GGP) were selected as the 'test systems'. We prepared each protein as follows. The homo-dimeric crystal structure 3EUU contains 100 residues for each monomer. First of all H-atoms were assigned to the crystal structure of the selected protein using CHARMM. Then the 3D structure was refined by performing energy minimization by 5000 steepest descent steps in vacuum. This removed bad steric contacts from the structure if any. The hetero-dimeric protein 1GGP contains 234 residues in chain A and 254 residues in chain B and was prepared in the same way. These two protein structures were then used separately to design their thermostable variants.

2.3. Generation of side-chain conformers libraries

We generated the conformer libraries for the 19 different kinds of amino acid residues directly from 49 PDB files randomly selected from the RCSB database and using our in-house tool [15,16]. We have not considered Gly in this library as it contains only one H-atom as its side-chain. We visually inspected the generated conformers and found it to represent the conformational space available to the side-chain reasonably well. The conformer models also cover their diverse orientations satisfactorily. The advantage of this method is that the conformers of the side-chains in all the cases are experimentally found ones.

2.4. Designing of more stable protein dimer

The entire designing process has two major steps (1) designing of more stable individual monomers and (2) designing of more stable monomer–monomer interface.

2.4.1. Designing of more stable individual monomers

The stability of an individual monomer is enhanced by introducing rational based mutations in it. The rational and the detailed method have been discussed earlier [15,16]. Here we are giving the essential summary of the approach. Based on our current understanding of the physical origin of the enhanced stability of thermophilic proteins we have selected to modulate the electrostatic component of the self-energy of the protein by mutating a set of rationally selected residues on or close to the surface of the protein. We have chosen mutations on or close to surface as such mutations are reported to enhance thermostability [25–33]. Moreover, it is expected that stabilization of the structural fluctuation of the surface will reduce the possibility of the onset of denaturation. Further more, mutations at the surface are less likely to alter the folded structure. The details of the original approach of converting a single chain mesophilic protein into a designed thermostable one are available in our earlier publications [15,16]. Fig. 1 shows the summary of the schematic diagram of the work flow. The steps of the present design approach are summarized below.

Step-1: First, the energy-minimized 3D structure of the original protein was considered and the interaction energies of the