Progress in Biophysics and Molecular Biology 128 (2017) 113–120

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

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

Understanding protein domain-swapping using structure-based models of protein folding

Nahren Manuel Mascarenhas [1], Shachi Gosavi*

Simons Centre for the Study of Living Machines, National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India

ARTICLE INFO

Article history:
Received 23 July 2016
Received in revised form
5 September 2016
Accepted 26 September 2016
Available online 17 November 2016

Keywords:
Mechanism of domain swapping
Protein topology
Symmetrized structure-based models
Molecular dynamics simulations

ABSTRACT

In domain-swapping, two or more identical protein monomers exchange structural elements and fold into dimers or multimers whose units are structurally similar to the original monomer. Domain-swapping is of biotechnological interest because inhibiting domain-swapping can reduce disease-causing fibrillar protein aggregation. To achieve such inhibition, it is important to understand both the energetics that stabilize the domain-swapped structure and the protein dynamics that enable the swapping. Structure-based models (SBMs) encode the folded structure of the protein in their potential energy functions. SBMs have been successfully used to understand diverse aspects of monomer folding. Symmetrized SBMs model interactions between two identical protein chains using only intra-monomer interactions. Molecular dynamics simulations of such symmetrized SBMs have been used to correctly predict the domain-swapped structure and to understand the mechanism of domain-swapping. Here, we review such models and illustrate that monomer topology determines key aspects of domain-swapping. However, in some proteins, specifics of local energetic interactions modulate domain-swapping and these need to be added to the symmetrized SBMs. We then summarize some general principles of the mechanism of domain-swapping that emerge from the symmetrized SBM simulations. Finally, using our own results, we explore how symmetrized SBMs could be used to design domain-swapping in proteins.

Contents

* Corresponding author.
E-mail address: shachi@ncbs.res.in (S. Gosavi).
[1] Present address: Department of Chemistry, Sacred Heart College (Autonomous), Tirupattur 635601, Tamilnadu, India.

http://dx.doi.org/10.1016/j.pbiomolbio.2016.09.013
0079-6107/

1. Introduction

Domain-swapping is a process in which two or more identical protein monomers exchange structural elements and fold into dimers or multimers (Bennett et al., 1994a; Liu and Eisenberg, 2002). The individual units of such oligomers are structurally similar to the original monomer (Fig. 1). The earliest evidence for domain-swapping was seen in RNase A (Crestfield et al., 1962) and the first determination of a domain-swapped structure was that of the dimeric diphtheria toxin (Bennett et al., 1994a,b). Since then, it has become clear that domain-swapping is common, and several domain-swapped proteins have been crystallized (Shameer et al., 2011). Recent data indicates that RNA structures may also domain-swap (Suslov et al., 2015). In this section, we first define several terms used in the study of domain-swapping (Liu and Eisenberg, 2002; Gronenborn, 2009; Rousseau et al., 2012) and then summarize the biological consequences of domain-swapping (Rousseau et al., 2012; Wodak et al., 2015).

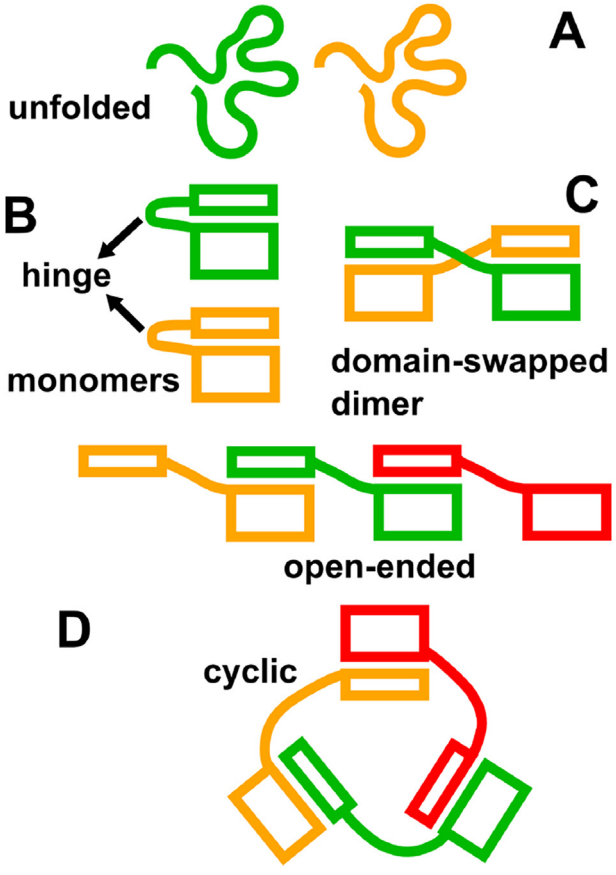

Fig. 1. The elements of domain swapping. (A) Two unfolded monomer chains shown in green and orange. These can either fold to two monomers (B) or exchange structural elements and fold to a domain-swapped dimer (C). The hinge is the peptide connector between the two domains in a single protein and is shown in (B). The hinge is in different conformations in the monomer (B) and the swapped structures (C, D). (D) Alternate outcomes of domain swapping which could lead to larger aggregates.

1.1. Domain-swapping nomenclature

In order for a protein to be considered domain-swapped the structures of both the monomer and the domain-swapped oligomer need to be present. If one of the structures is that of a homologous protein then the protein is said to be quasi-domain swapped. If only the domain-swapped structure is available then the protein is considered a candidate for domain-swapping. Several inter-protein interfaces are formed upon domain-swapping. The interface whose inter-protein interactions mimic those present in the monomer is called the primary interface. However, new interfaces which are not present in the monomer can also be formed and such interfaces are called secondary interfaces. Many proteins swap only a single secondary structural element and in such cases this "domain" can be thought of as being swapped. However, there are proteins where the size of the "swapped" region is similar to that of the other region. To account for both cases, in this review, we call both parts of the monomer swapped domains. So, a domain-swapping protein usually consists of two swapping-domains connected by a peptide segment (of about 4–5 residues) called the hinge (Fig. 1B and C). The hinge undergoes a conformational change upon swapping, usually being a loop or a turn in the monomer (Fig. 1B) and adopting an extended conformation after domain-swapping (Fig. 1C). There are also examples where two different hinges induce two different modes of domain swapping in the same protein (Liu et al., 1998, 2001; Chen et al., 2010). In some proteins (Nilsson et al., 2004), one swapped domain is inserted into the sequence of the other domain and there are two peptide segments connecting the two domains. Upon swapping, both segments undergo a conformational change. Given the diversity of structural hinge types, there have been efforts to understand the sequence determinants of hinges. Early efforts concentrated on the proline composition of hinges but although prolines contribute to domain-swapping in some proteins (Bergdoll et al., 1997; Rousseau et al., 2001; Miller et al., 2010) this is not a universal effect (Barrientos et al., 2002; Cho et al., 2005). A recent study analyses the distribution of all amino acids in the hinge regions of domain swapping proteins and finds that the amino acid distribution of hinge residues is similar to that of other loop regions. However, there are some amino acids such as valine which are more likely to be found in hinges (Shingate and Sowdhamini, 2012).

1.2. Biological consequences of domain swapping

1.2.1. Aggregation

Open-ended domain-swapping (Fig. 1D) can lead to the formation of large protein aggregates (Esposito et al., 2010). Such aggregation can lead to the loss of protein function and cause disease (Law et al., 2006). Domain-swapping has also been implicated in the formation of disease-causing fibrillar aggregates (Chiti and Dobson, 2006; Herczenik and Gebbink, 2008; Bennett et al., 1995, 2006; Janowski et al., 2005). For instance, mutants of cystatin-C with reduced domain-swapping also have reduced fibrillar aggregation (Nilsson et al., 2004). The question that arises, then, is, "Why are the amino acids that cause increased domain-swapping retained through evolution?" In a recent study, we showed that in stefin-B, residues that increase the propensity of domain-swapping are part of the protease binding-site and thus, domain-swapping can be a by-product of the need to conserve protein