Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

New designed protein assemblies

Sabina Božič[1], Tibor Doles[1,2], Helena Gradišar[1,2] and Roman Jerala[1,2,3]

Self-assembly is an essential concept of all organisms. Polypeptides self-assemble either within a single polypeptide chain or through assembly of protein domains. Recent advances in designed protein assemblies were achieved by genetic or chemical linkage of oligomerization domains and by engineering new interaction interfaces, which resulted in formation of lattices and cage-like protein assemblies. The absence of new experimentally determined protein folds in the last few years underlines the challenge of designing new folds. Recently a new strategy for designing self-assembly of a polypeptide fold, based on the topological arrangement of coiled-coil modules as the protein origami, has been proposed. The polypeptide tetrahedron was designed from a single chain concatenating of coiled-coil forming building modules interspersed with flexible hinges. In this strategy the order of coiled-coil segments defines the fold of the polypeptide nanostructure.

Addresses
[1] Department of Biotechnology, National Institute of Chemistry, Ljubljana, Slovenia
[2] Excellent NMR – Future Innovation for Sustainable Technologies Centre of Excellence, Ljubljana, Slovenia
[3] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

Corresponding author: Jerala, Roman (roman.jerala@ki.si)

Current Opinion in Chemical Biology 2013, **17**:940–945

This review comes from a themed issue **Synthetic biomolecules**

Edited by **Shang-Cheng Hung** and **Derek N Woolfson**

For a complete overview see the Issue and the Editorial

Available online 31st October 2013

1367-5931/$ – see front matter,

http://dx.doi.org/10.1016/j.cbpa.2013.10.014

Introduction

Proteins form the most versatile structures in nature, both in terms of structural definition at the nanoscale and their functionality. Tertiary structures of proteins underlie their versatile functionality, such as catalysis, molecular recognition, assembly of cellular scaffolds and many others. Considerable numbers of natural proteins have evolved to form supramolecular structures based on the oligomerization domains, such as oligomeric enzymes or viral protein capsids. Oligomerization subunits are most often symmetric since this requires the minimal number of different domains and interaction surfaces. Oligomerization saturates all of the available binding sites and can lead to closed structures of precisely defined stoichiometry [1]. A recent study proposed that symmetric structures are common because symmetric interfaces are over-represented among the set of all energetically favourable interactions and represent the evolutionary accessible targets [2]. In comparison to other man-made nanostructures protein assemblies can form much more detailed as well as asymmetric nanoscale structures and are as such of considerable interest for many technological applications. Self-assembly of protein domains can lead to the formation of lattices or assemblies of discrete size, such as various cages, comprising from few to tens and hundreds of subunits. In the last two years excellent reviews covered the topic of designed protein assemblies [3•,4].

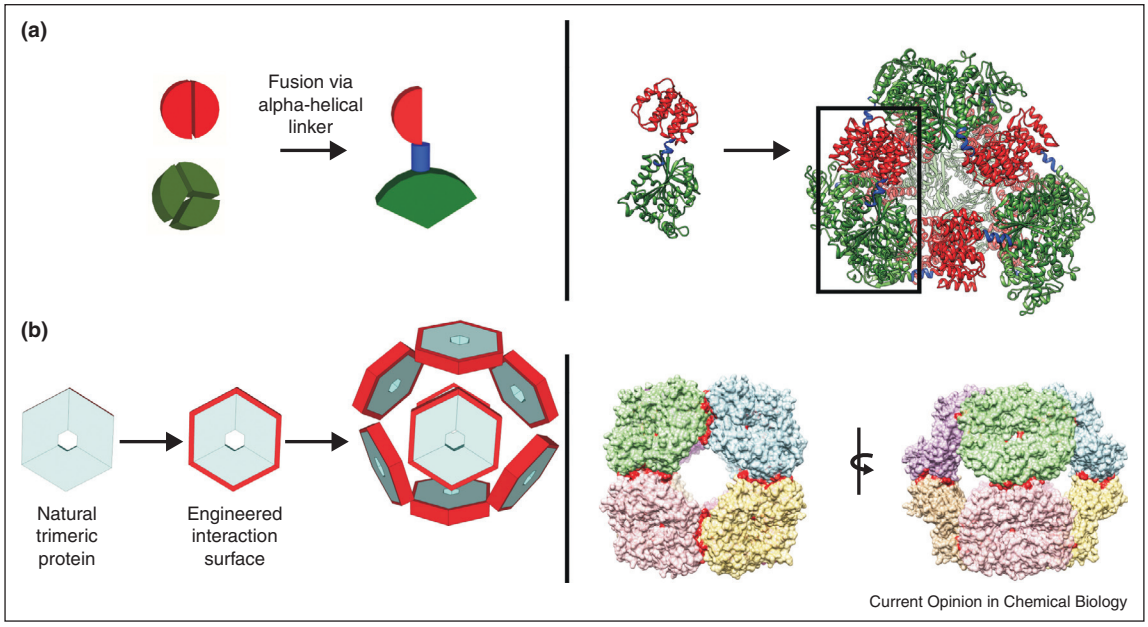
Engineering of natural proteins has been used to modify their functionality but design of new protein folds, not based on the natural template still represents a challenge. Native protein domains are typically composed of packed secondary structure elements stabilized by a large number of weak interactions between the non-consecutive amino acid residues that define the fold, with notable contribution of hydrophobic interactions defining the protein core. Currently the folds of only few small domains can be accurately predicted [5–8], while design of completely new folds is even more challenging [9]. There is a consensus that natural proteins self-assemble into a limited number of folds, estimated to comprise at most few thousand members [10–12]. Currently we can count approximately 1200 protein domain folds among the experimentally determined protein structures. However the absence of any new protein folds, deposited in the PDB since 2008, despite the determination of almost 9000 new structures each year, indicates that we have probably already sampled the large majority of the existing natural protein folds. Therefore *de novo* formation of new protein folds represents an additional challenge.

Design of symmetric intermolecular protein assemblies

Assemblies based on linked protein oligomerization domains

Designed protein oligomerization based on the symmetry was aimed to create a tetrahedral symmetric cage, composed of 12 subunits. Each subunit was composed of one dimerization and one trimerization domain (Figure 1a). These two domains were connected by a continuous semi rigid α-helix, selected to define the relative orientation of the neighbouring domains in the resulting assembly. The designed protein formed cage-like assemblies whose sizes were too heterogeneous to confirm the anticipated structure [13]. Recent introduction of a small number of amino acid modifications into the original protein

Figure 1



Design of symmetric intermolecular protein assemblies. **(a)** Different oligomerization domains are linked together to obtain two or more interacting interfaces within a single polypeptide building block. Fusion of a dimerization domain (red) to a trimerization domain (green), where the predetermined geometry of their symmetry axes is held in place by a semi rigid α-helical linker (blue), resulted in the formation of a tetrahedral cage-like structure (4d9j) (right), composed of 12 subunits [13]. **(b)** Engineering of new protein-protein interfaces into the naturally oligomerizing building blocks leads to the formation of desired symmetric assemblies [22]. By computational docking and interface design new interaction surfaces (marked red) have been introduced into a naturally oligomeric protein, leading to the formation of assemblies with tetrahedral and octahedral symmetries (4ddf) (right).

sequence made it possible to remove the potential steric conflicts and deviations on the designed orientation of the symmetry axes. After those modifications the X-ray crystal structure of the homogenous 12-subunit assembly has been determined. The structure confirmed the tetrahedral geometry of the assembly measuring approximately 16 nm in diameter and deviation of 8 Å from the perfect symmetry [14•,15].

Several other strategies utilizing linkage of protein oligomerization domains have been explored: utilizing protein chemical modifications, small molecule interactions or genetic fusions. C_4-symmetric tetrameric aldolase and D_2 tetrameric streptavidin were used to produce a quadrangular lattice, linked together with a pair of tethered biotin molecules. Arrays of limited size were obtained, probably owing to the flexibility of the biotin linker and imperfect control over the relative orientation of the component oligomers [16]. In another study the protein nanorings were prepared by chemically induced self-assembly of dihydrofolate reductase (DHFR) and histidine triad nucleotide binding 1 (Hint1) fusion proteins. The dimensions of nanorings could be modulated by the length and composition of the peptide linking fusion proteins, in the range from 10 to 70 nm [17]. One study prepared gyraseB-based regulated assembly/disassembly of a fusion polypeptide between gyrase and a trimerization protein domain. The addition of a pseudo-dimeric gyrase B ligand coumermycin induced formation of gyrase B dimers and led to the formation of hexagonal assemblies and its dissociation by a monomeric ligand novobiocin [18].

The main challenge of the oligomerization domain-based assembly is connecting two oligomerization domains in a fixed relative orientation. The key advance was introduction of an extended fusion strategy. Fusion protein, composed of the two oligomerization domains, can generate two or more connections between the adjacent oligomers if the two domains are joined along an axis of symmetry that they both share. The symmetry-matching fusion protein strategy successfully generated linear filaments, 2D arrays extending up to 5 μm and large solid aggregates having crystal-like morphology; however this strategy was only used to construct periodic arrays, but not finite structures like molecular cages [19••].

Engineering new protein interaction surfaces

Although definition of the relative orientation of fusion oligomerizing protein domain partners has been addressed in the studies described above, it requires careful selection of the natural protein domains with specific geometries and arrangements of the symmetry axes, thereby limiting its use as a general platform for the