Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

The design of symmetric protein nanomaterials comes of age in theory and practice

Todd O Yeates[1,2,3], Yuxi Liu[1] and Joshua Laniado[3]

In nature, protein molecules have evolved as building blocks for the assembly of diverse and complex structures, many of which exhibit a high degree of symmetry. This observation has motivated a number of recent engineering efforts in which the advantages of symmetry have been exploited to design novel self-assembling protein structures of great size. Materials ranging from cages to extended two and three-dimensional arrays have been demonstrated. Especially for extended arrays, a vast number of geometrically different design types are possible. A table of geometric rules is provided for designing a universe of novel materials by combining two component symmetries.

Addresses
[1] UCLA Department of Chemistry and Biochemistry, United States
[2] UCLA-DOE Institute for Genomics and Proteomics, United States
[3] UCLA-Molecular Biology Institute, United States

Corresponding author: Yeates, Todd O (yeates@mbi.ucla.edu)

Current Opinion in Structural Biology 2016, **39**:134–143

This review comes from a themed issue on **Engineering and Design**

Edited by **Den Tawfik** and **Raghavan Varadarajan**

http://dx.doi.org/10.1016/j.sbi.2016.07.003

Introduction

Building blocks that have self-complimentary interfaces can self-assemble into elaborate structures. Nature serves as a rich source of inspiring specimens. At the macromolecular scale, viral capsids are quintessential examples, but other equally extraordinary macromolecular assemblies abound in nature (reviewed in [1–3]). The beauty and functional utility of these assemblies have long-motivated engineering efforts to create comparable structures in the laboratory. Beginning in the 1980s Ned Seeman pioneered ideas for using DNA molecules as building blocks for nanostructures [4]. Over the years, those ideas and various strategic variations led to the creation of elaborate supramolecular architectures and design patterns built from nucleic acids (reviewed in [5]). In nature, protein molecules have been the choice for the evolution of large assemblies with diverse form and function. But the engineering path to following Nature's lead has been challenged by the complexity of the rules that govern protein folding and assembly. To overcome those challenges, special strategies are needed.

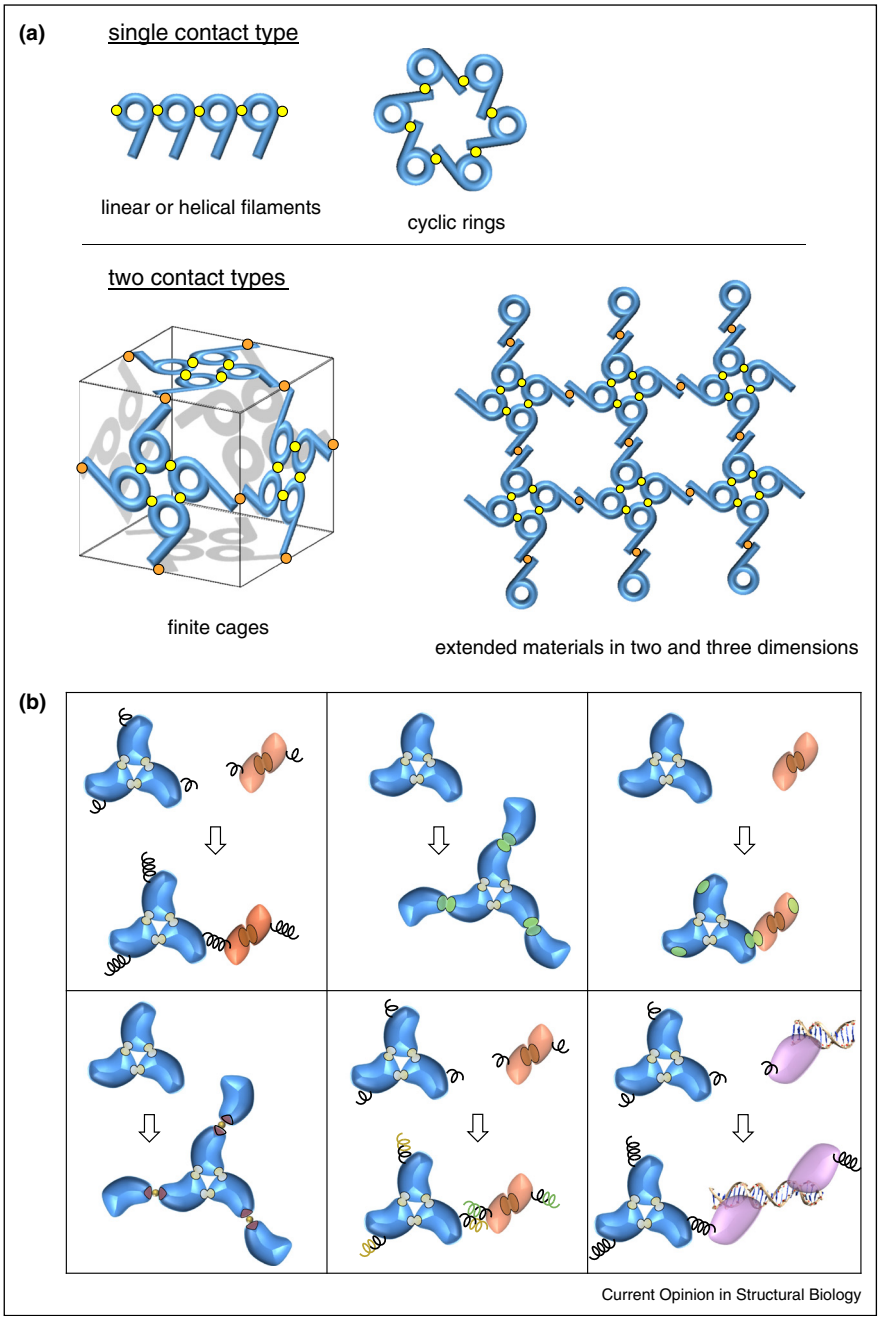
In developing a strategic approach for building with protein molecules, Nature provides a major clue. Symmetry prevails in naturally evolved protein assembles. This is an empirical fact evident in the vast database of known macromolecular structures [6,7], but the prevalence of high symmetry in large protein assemblies was anticipated at least as far back as 1956 when Crick and Watson emphasized that viral capsids were likely able to evolve more easily in symmetric forms because symmetric assemblies require the fewest number of distinct interfacial contacts between individual subunits [8]. That key observation applies as well to designed structures, and indeed the early history of designing protein assemblies is rich with cases of relatively simple symmetric structures such as dimers and helical filaments [9,10,11,12]. The push in recent years to create very large protein assemblies has been guided even more strongly by principles of symmetry.

Symmetry-based design strategies

The symmetry of an object is fully described by the set of spatial operations (*e.g.* rotations) that leave the entire object unchanged except for an undetectable exchange of identical subunits. Because the symmetry of an object obeys the properties of a mathematical group, each specific type of symmetry is often referred to as a symmetry group. The symmetry group of a structure can be used to understand how many structurally distinct contact types are required to hold all the subunits together in one connected object. Certain simple types of architecture can be created from a building block that touches itself in just one way; i.e. using a single contact type. The possible outcomes are limited to structures like cyclic rings of subunits, or head-to-tail filaments (Figure 1a). More complex architectures require building blocks with more than one distinct interface.

A relatively simple group theory analysis explains the minimum number of distinct contact types required to achieve a given target symmetry. This was articulated first in the context of three-dimensional crystals [13] and then in the context of designed protein assemblies by Padilla *et al.* [14]. For example, if all the elements of a symmetry group can be generated by repeated application of a single element of the group (i.e. a rotational operation), then one contact type is sufficient. The cyclic or head-to-tail filament architectures noted above are examples of this type.

Figure 1



Assembly consequences and strategies for introducing multiple contact types into protein building blocks. **(a)** Illustration of varied symmetric architectural forms and the number of distinct contact types required for connectivity between molecular building blocks. Two contact types are sufficient to create diverse assemblies. **(b)** Different molecular strategies for creating a building block having two distinct contact types in a defined orientation. Left to right (top): alpha helical fusion; 1-component interface design; 2-component interface design. Left to right (bottom): metal or ligand bridging; coiled-coil helical fusions; designed symmetrization of DNA binding proteins.

If two elements from the symmetry group must be chosen in order to obtain the full symmetry group by repeated operations, then two contact types are required, and so on. Surprisingly, it turns out that a great many types of symmetry — including finite cages and many extended two and three-dimensional arrays — can be generated using just two properly chosen symmetry elements in combination (Figure 1a). This key point frames the problem of designing novel protein assemblies by prescribing the number of distinct contact types that must be