Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

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

# Rational design of new materials using recombinant structural proteins: Current state and future challenges

Tara D. Sutherland*, Mickey G. Huson, Trevor D. Rapson

*CSIRO, Black Mountain, GPO Box 1700, Acton, ACT 2601, Australia*

ARTICLE INFO

*Keywords:*
Silk
Collagen
Resilin
Elastin
Sequence-defined polymer

ABSTRACT

Sequence-definable polymers are seen as a prerequisite for design of future materials, with many polymer scientists regarding such polymers as the holy grail of polymer science. Recombinant proteins are sequence-defined polymers. Proteins are dictated by DNA templates and therefore the sequence of amino acids in a protein is defined, and molecular biology provides tools that allow redesign of the DNA as required. Despite this advantage, proteins are underrepresented in materials science. In this publication we investigate the advantages and limitations of using proteins as templates for rational design of new materials.

## 1. Introduction

Many of the challenges facing the world could be better addressed with stimuli-responsive or multifunctional materials. In the natural world, protein polymeric materials are examples of such materials with type 1 collagen and tropoelastin having more than 50 documented molecular interactions (Sweeney et al., 2008; Werkmeister and Ramshaw, 2012; Mithieux et al., 2013). Fundamental to the rational design of non-natural stimuli-responsive or multifunctional materials is the precise control of a polymer's composition. Methods to generate sequence- and length-defined synthetic polymers are available (Fig. 1) (Lutz et al., 2013). However, gradual accumulation of near-identical polymers during the synthesis process that can be separated only at high cost is predicted to limit their compatibility with large scale production (Gody et al., 2016). In contrast, biological systems have evolved efficient processes for the rapid synthesis of sequence-defined polymers that are compatible with large-scale production (Fig. 1). The sequence and length of a protein is defined by a DNA sequence and appropriate biological recombinant systems can generate identical polymers at 100–1000 kg scale. Under appropriate conditions, these polymers fold into well-defined three-dimensional structures. Critical for the design process, biotechnologists can precisely and readily change the DNA allowing production of modified protein sequences required for design of new materials.

Despite the promise of proteins in the design of new materials, they are underrepresented in materials science. There is a need for proteins that can be produced at economies of scale in recombinant systems, which are tolerant of amino acid modifications, while remaining compatible with the ability to produce the proteins at scale and fabricate them into material forms. Rational design of functional materials from these templates requires an understanding of the relationship between the protein amino acid sequence, the architecture of the protein in the solid-state, and the desired functional properties. In this article, we investigate the current and future factors that need to be addressed before the potential of protein polymers as general templates for rational design of new materials can be recognised, and realised.

## 2. Addition of functionality into structural protein materials

The key advantage in the use of recombinant proteins for rational design of new materials is the ability to manipulate the protein composition with molecular-level reproducibility and then control materials fabrication to obtain the desired three-dimensional molecular structure (DiMarco and Heilshorn, 2012). Existing examples of recombinant protein-based materials are primarily based on rational design of the protein sequence and then experimental characterisation of the final material to find the most appropriate. These examples can be divided into two classes: the most common class derives from a modular approach where structural protein domains are separate from functional peptide domains (Fig. 2A). Less commonly, functional motifs have been incorporated within a structural protein domain without compromising the structural domain (Fig. 2B). Examples of these two classes are summarised in Table 1.

The latter approach has primarily been explored using either collagen or aculeate (bee, ant, hornet) silk protein sequences as the structural component. With collagen, motifs from animal collagens have either been replaced with cell-binding sites from other collagens (Que et al., 2015), or cell-binding sites from animal collagens have been

* Corresponding author.
*E-mail address:* tara.sutherland@csiro.au (T.D. Sutherland).

http://dx.doi.org/10.1016/j.jsb.2017.10.012
Received 26 September 2017; Received in revised form 26 October 2017; Accepted 28 October 2017

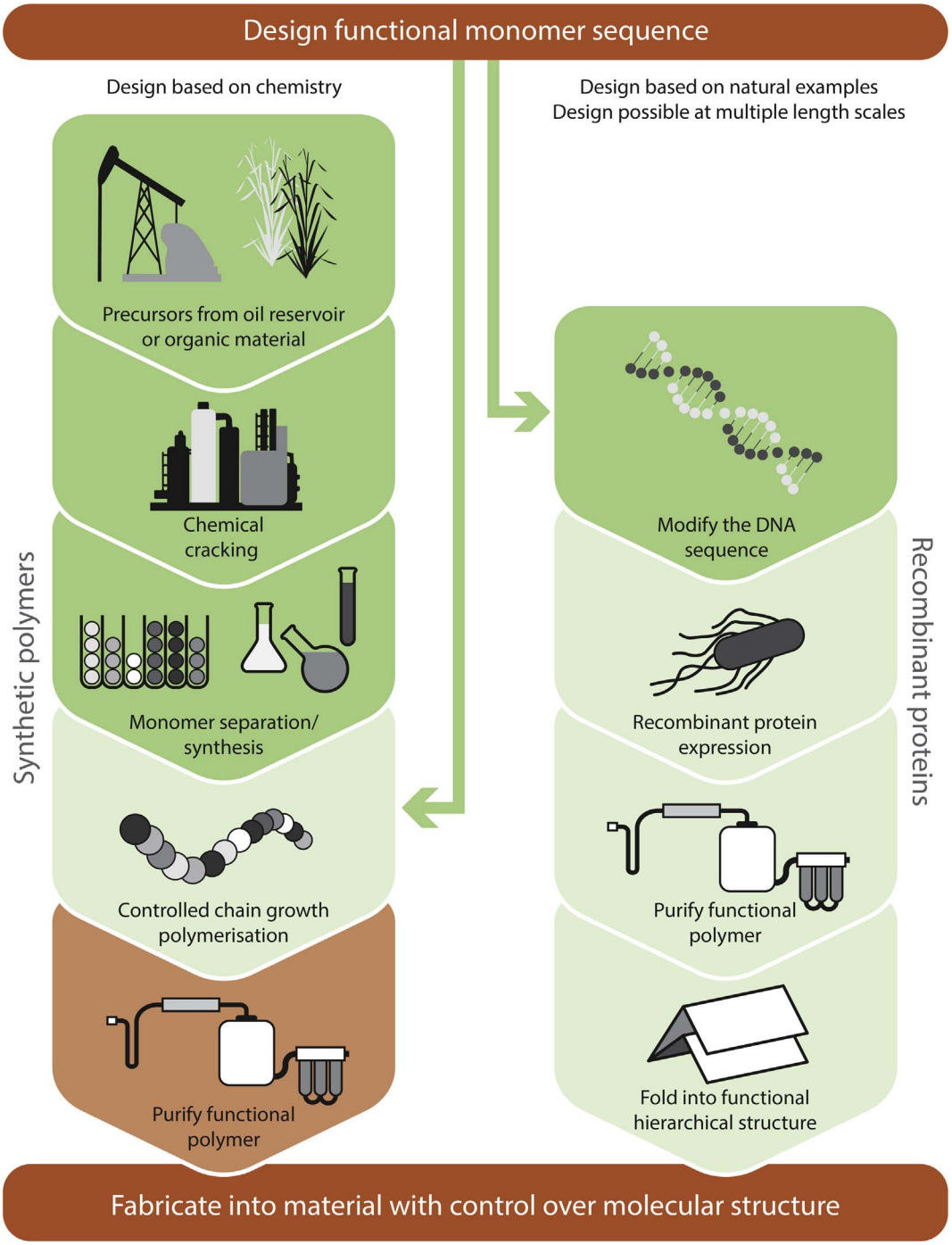


**Fig. 1.** Process of rational design of materials from synthetic polymers or recombinant protein polymers. Mature processes: dark green; processes with some development: light green; processes that are not economical at large-scale: dark brown (from Gody et al., 2016); and those that are not currently possible but should be achievable: light brown.

inserted into bacterial collagen proteins (Cosgriff-Hernandez et al., 2010; Seo et al., 2010; Peng et al., 2014a; An et al., 2014; Que et al., 2015). In honeybee silks, antimicrobial peptides (Trueman et al., 2017) and heme-binding sites have been introduced into the coiled coil domain (Rapson et al., 2017a). Heme-silk materials function as nitric oxide sensors, oxygen sensors, or as materials capable of oxygen reduction or heterologous catalysts (Rapson et al., 2015; Horgan et al., 2016; Rapson et al., 2017b).

### 2.1. Understanding the sequence-structure-function relationship of proteins

Rational design of protein-based materials requires the ability to predict the effect that modifications in the amino acid sequence will have on the final materials properties. This implies that the designer understands, and can control during the fabrication process, the sequence-structure-function relationship of the protein.

Control of material fabrication at the molecular level is currently possible only on a very limited scale. Tertiary structures such as triple helices (as found in collagen) and coiled coils are formed in solution and can be retained in the solid-state under appropriate fabrication conditions. However, methods to control the molecular relationship of these tertiary structures in the solid-state after material fabrication have not been developed. Other naturally occurring structural proteins do not retain their solution structure in the solid-state (i.e. spider and