Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

# Protein cross-linking tools for the construction of nanomaterials

Natalia E Domeradzka[1,2], Marc WT Werten[1], Frits A de Wolf[1] and Renko de Vries[2]

Across bioengineering there is a need to couple proteins to other proteins, or to peptides. Although traditional chemical conjugations have dominated in the past, more and more highly specific coupling strategies are becoming available that are based on protein engineering. Here we review the use of protein modification approaches such as enzymatic and autocatalytic protein–protein coupling, as well as the use of hetero-dimerizing (or hetero-oligomerizing) modules, applied to the specific case of linking together *de novo* designed recombinant polypeptides into precisely structured nanomaterials. Such polypeptides are increasingly being investigated for biomedical and other applications. In this review, we describe the protein-engineering based cross-linking strategies that dramatically expand the repertoire of possible molecular structures and, hence, the range of materials that can be produced from them.

**Addresses**
[1] Wageningen UR Food and Biobased Research, 6708 WG Wageningen, The Netherlands
[2] Physical Chemistry and Soft Matter, Wageningen University and Research Centre, Dreijenplein 6, 6703 HB Wageningen, The Netherlands

Corresponding author: de Vries, Renko (renko.devries@wur.nl)

**Current Opinion in Biotechnology** 2016, **39**:61–67

This review comes from a themed issue on **Nanobiotechnology**

Edited by **Michael Nash** and **Oded Shoseyov**

http://dx.doi.org/10.1016/j.copbio.2016.01.003



## Introduction

Self-assembling proteins are intensively being investigated as building blocks for nanostructured soft materials [1,2••]. Sequence motifs from natural structural proteins can be used to design novel protein materials consisting of polypeptides with simple repetitive sequences ('protein-based polymers', or simply 'protein polymers') [3,4], or one can design entirely novel peptide or polypeptide sequences with predictable self-assembly behaviour [5,6]. Self-assembling (poly)peptides can be either chemically synthesized or genetically engineered [7,8]. The latter route, which is the focus of this review, allows the production of long precisely defined amino acid sequences, thus offering a level of control that is highly attractive in a range of biomedical applications [9–14].

These applications depend on precise control over the way in which the polypeptides are assembled into their final supramolecular structure. It is here that novel protein-engineering techniques for cross-linking can play a crucial role. Although chemical cross-linking reactions can, and have been used extensively [15•], they typically lack the specificity needed to create precisely defined polypeptide nanomaterials. Various protein-engineering tools such as enzymatic coupling [16–19,20••] and the use of bio-recognition modules [21] do offer such specificity. By bio-recognition modules, we here specifically refer to the use of peptide- or small protein domains that can be incorporated in polypeptide chains in order to mediate self-assembly of polypeptides [22] via hetero-dimerization (or hetero-oligomerization). Although these engineering tools can be used to link specific bioactive groups to proteins [23–25], this review mainly focusses on their use in linking designed polypeptides into higher order nanostructures. An overview of the tools discussed is given in Table 1.

## Covalent protein cross-linking

For permanently linking different protein subunits into larger assemblies, a range of cross-linking enzymes and autocatalytic protein modules are available. Protein cross-linking enzymes differ substantially in their degree of substrate specificity. First we consider enzymes with a rather broad substrate specificity. They are especially useful for the creation of irreversibly cross-linked (hydrogel) networks. Key examples are: transglutaminase (*TG*), horse-radish peroxidase (*HRP*), lysyl oxidase (*LOX*) and plasma amine oxidase (*PAO*), which couple one specific type of amino acid to another, as illustrated in Figure 1 (e.g. *TG* couples lysines to glutamines). Their substrate specificity is broad in the sense that these enzymes typically provide the engineer with considerable freedom to choose the residues flanking those to be cross-linked. A second group of enzymes and autocatalytic modules have a much more narrow substrate specificity: Sortase A (*SrtA*), Inteins, SpyTag/SpyCatcher, see Figure 2. These tools are particularly useful for creating molecular conjugates with precisely defined architectures.

## Transglutaminase

*TG*s catalyse the formation of isopeptide bonds between glutamine and lysine residues [26] and are found in both

**Table 1**

**Cross-linking tools and their application to designer polypeptides**

| Strategy | Cross-linking mechanism | Polypeptide | Application | Ref. |
|---|---|---|---|---|
| Transglutaminase | Formation of an isopeptide bond between Lys and Gln. | • Q11 peptide<br>• Collagen-like dendrimers<br>• Elastin-like proteins<br>•Designer protein polymers $K_n$ and $(BQ)_n$ | • Network functionalization<br>• Two-component network<br>• Cross-linking of fibers | [32,34–37] |
| Lysyl oxidase and plasma amine oxidase | Spontaneous aldol condensation or formation of a Schiff base between two Lys residues | • $K_2(SL)_6K_2$ | • Cross-linking of self-assemblies | [40•] |
| Horse radish peroxidase | Spontaneous radical couplings between two Tyr residues | • Silk proteins<br>• Resilin-like polypeptides | • Cross-linking of self-assemblies | [45–46] |
| Sortase A | Formation of threonine-glycine peptide bond between C-terminal LPETGG and N-terminal poly-G motifs | • P1 peptide<br>• Micelles made of elastin-like proteins | • Network functionalization | [51,52] |
| Split–inteins | Formation of a covalent bond by (1) non-covalent association of the halves of a split–intein pair, and (2) joining of the flanking protein fragments and concomitant self-excision by the reconstituted intein | • CutA trimer-forming peptide | • Two-component network | [57] |
| SpyTag–SpyCatcher | SpyTag and SpyCatcher undergo autocatalytic isopeptide bond formation between Asp117 on SpyTag and Lys31 on SpyCatcher | • Elastin-like proteins<br>• CsgA protein | • New self-assembling topologies, two-component network<br>• Enzyme immobilization | [60••,61,62] |
| Heterodimer-forming coiled coils | Hydrophobic forces and electrostatic interactions between complementary units | • Hydrophilic random coli<br>• Self-assembling polypeptides | • New self-assembling topologies<br>• Two-component network; network functionalization | [72,73•,74] |
| WW domain and its proline rich ligand | Selective binding of WW domain to specific tyrosine-containing sequences, based on a combination of interaction types | • Hydrophilic random coli | • New self-assembling topologies; two-component network; network functionalization | [80–82,83••] |
| PDZ domain and its ligand | Strong protein–protein interaction based on hydrophobicity, reinforced by an engineered disulphide linkage | • CutA trimer-forming peptide | • Two-component network; network functionalization | [86] |

microbial [27] and mammalian cells [26]. They are quite promiscuous with respect to the residues flanking the lysine substrate in the polypeptide chain [28,29], but less so with respect to the preferred context of the glutamine [30,31]. For *de novo* designed polypeptides one can often exploit some freedom of design, such that *TG* can be a very useful tool to create well-defined polypeptide architectures. Collier and Messersmith showed that *TG* can be used to couple the glutamines in the self-assembling peptide Q11 (QQKFQFQFEQQ) to lysines in other peptides and proteins [32]. The same group proposed more *TG* substrate peptides, including an amine donor sequence K (EDGFFKI) that can be targeted by *TG* when placed at the C-terminus [33], and which can be coupled to the amine acceptor substrate peptide Q (APQQEA) [34]. The Chilkoti group studied elastin-like polypeptide (ELP) sequences, tandem repeats of the motif $VPGXG(VPAGVG)_6$, and showed that variants for which X is either lysine (K) or glutamine (Q), were efficiently cross-linked by tissue-derived *TG*s [35]. Bozzini *et al.* designed protein polymers with an $(A\text{-}B)_8$ structure, where A is an ELP sequence and B is the motif AAAAAAKAAAKAAQGFL. These protein polymers were efficiently cross-linked by microbial *TG* [36]. Finally, Davis *et al.* developed a protein hydrogel system that was cross-linkable by *TG* [37]. This system consists of two *de novo* designed protein polymers $K_n$ and $(BQ)_n$, where the $K_n$ block contains lysines, B is a random coil hydrophilic block, and the Q block serves as the glutamine substrate.

## Lysyl oxidase and plasma amine oxidase

*LOX* are a family of oxido-deaminases that cross-link components of the extracellular matrix, in particular collagen and elastin [38]. *LOX* cross-linking is based on the oxidation of the primary amine of lysine to an aldehyde. This aldehyde can spontaneously react with another amine and form a Schiff base or undergo an aldol condensation with another aldehyde [39]. In this way a covalent bond is established between two lysines. *PAO* is a commercially available alternative, which also functions by oxidation of primary amines. Bakota *et al.* showed that *LOX* and *PAO* can cross-link the self-assembled structures of the *de novo* designed peptide $K_2(SL)_6K_2$ [40•]. It appears that the activity of *LOX* is less affected by sequence context than that of *TG* [19], although some