[Frontiers in Bioscience, Landmark, 18, 289-304, January 1, 2013]

Recombinant protein polymers in biomaterials

Wookhyun Kim[1]

[1]Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, and the Wyss Institute of Biologically Inspired Engineering of Harvard University, Boston, MA, USA

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Synthesis of repetitive genes
- *3.1. Concatermerization*
- *3.2. Recursive directional ligation*
- *3.3. Convergent assembly*

4. Natural protein-based materials
- *4.1. Recombinant protein polymers from mammals*
  - *4.1.1. Collagen*
  - *4.1.2. Elastin*
- *4.2. Recombinant protein polymers from insects or spiders*
  - *4.2.1. Silk*
  - *4.2.2. Resilin*
- *4.3. Recombinant protein polymers from marine organisms*
  - *4.3.1. Mussel adhesive protein*
  - *4.3.2. Reflectin*

5. Summary and perspective
6. Acknowledgements
7. References

1. ABSTRACT

Naturally occurring protein-based materials have been found that function as critical components in biomechanical response, fibers and adhesives. A relatively small but growing number of recombinant protein-based materials that mimic the desired features of their natural sources, such as collagens, elastins and silks, are considered as an alternative to conventional synthetic polymers. Advances in genetic engineering have facilitated the synthesis of repetitive protein polymers with precise control of molecular weights which are designed by using synthetic genes encoding tandem repeats of oligopeptide originating from a modular domain of natural proteins. Many repeat sequences as protein polymer building blocks adopt a well-defined secondary structure and undergo self-assembly to result in physically cross-linked networks or with chemical cross-linking so that further form three-dimensional architectures similar to natural counterparts. In this review, recombinant protein polymers currently developed will be presented that have emerged as promising class of next generation biomaterials.

2. INTRODUCTION

In the search for next-generation biomaterials, scientists have been learning valuable lessons from nature's superior products. Of those products, protein-based materials found in nature are attractive sources for the design of biologically-inspired materials. They play critical roles such as maintaining organ structure, desirable biomechanical responses without rupture and adhesives in their native environment. Collagens and elastins from animals, silks and resilins from spiders and insects, mussel byssus threads, abductin and reflectins from marine species are produced to survive and adjust in their respective environment. As a primary component in biological system, protein-based materials have evolved to meet structural and functional requirements through continuous optimization. Inspired from these aspects, a growing number of new protein-based materials have emerged as promising biomaterials in medicine and materials science (1). For example, nature has developed protein glues which incredibly stick to a solid surface in harsh milieu, such as waves, temperature and saline water (2). Another

example is silk protein that has been used as textures and medical sutures (3). However, only relatively limited class of protein-based materials has been identified and developed as useful biomaterials for practical applications when compared to synthetic polymeric materials.

The structural analysis of natural protein-based materials has revealed that the primary sequence deduced from cDNA is largely comprised of structurally and functionally distinct domains. These domains are frequently featured by relatively simple sequence combination and high composition of specific amino acid residues, which means tandem repeats of structurally or functionally distinct peptide unit. For example, the primary sequence of soluble elastin, tropoelastin, is composed of alternating sequence of two distinct domains, elastomeric domain and crosslinking domain (4). The elastomeric domains are rich in glycine, valine, and proline, and consist of high tandem repeats of oligopeptide sequence motif that is responsible for elasticity in the native state. The crosslinking domains are rich in lysine residues, which are enzymatically processed to form covalent crosslinks to stabilize elastin fiber networks. Presence of repetitive oligopeptide motifs within the protein polymer sequence is not only found in elastomeric proteins, but also often found in other proteins including adhesive proteins from marine invertebrates. The modular domain of these proteins is composed of the tandem repeats of oligopeptides of approximately 3-10 amino acid residues that may be repeated up to several hundred times. As illustrated by structural studies, recent research focus of protein-based materials has been directed to these repetitive oligopeptide domains because they are responsible for the regularity of local conformation and underlie the self-assembly of the materials into the supramolecular architectures that define their biological function (1). Thus, to better understand structure-function relationships of these natural protein-based materials, oligopeptide unit or recombinant polypeptide-based polymers (protein polymers) consisting of repetitive oligopeptide were synthesized to access the biological relevance and role of the repetitive peptide sequence in natural protein-based materials.

Over the past two decades, genetic engineering approaches have enabled the design and creation of artificial protein polymers with desirable molecular weights and properties (1,5). The synthetic protein polymers displayed similar behaviors to their native counterpart, which have the ability to form a regularly repeating, well-defined secondary structure and self-assemble into fibrous structures. There are distinct advantages of recombinant protein polymers over synthetic polymers. Because of easy control of the primary sequence with defined length and the potential for large-scale production, the use of biosynthetic machinery is the promising approach to the synthesis of protein-based biomaterials. Various biosynthetic strategies have been employed for the production of protein-based biomaterials having regularly repeating structural motifs in bacteria, yeast, plant, and mammalian cells (6,7). Of the various host cell expression systems, *Escherichia coli* is widely used to express protein polymer genes due to the ease of manipulating gene of interest and large-scale production of target proteins. However, expression of eukaryotic coding sequences and lack of post-translational modification, for example hydroxyproline in collagen and 3,4-dihydroxyphenylalanine (DOPA) in mussel adhesive protein, limit the wide utility of *E. coli* expression systems. Through the recombinant DNA techniques, it was possible to make the large synthetic genes encoding the repetitive amino acid sequences by concatemerization of a monomeric gene and facilely synthesize the repetitive protein polymers in bacterial host (8). Silk-, elastin-, collagen- and mussel adhesive mimetic protein polymers have been synthesized using E. coli expression system and they have shown unique mechanical and biological properties (9,10,11). Repetitive sequence motifs derived from naturally occurring protein-based materials that are extensively studied are provided in Table 1.

The design strategy of previously developed protein polymers was based on the primary sequence and the control of mechanical property of the protein materials (1,5). In addition to design of protein polymer sequence, biologically functional peptide motifs can be added to the protein polymers for the purpose of specific cell binding. The versatility of recombinant protein polymers has been realized a few decades ago and potential applications are now being explored in drug delivery, tissue engineering and other applications.

3. SYNTHESIS OF REPETITIVE GENES

Recent advances in recombinant DNA technology and metabolic engineering enable large-scale synthesis of repetitive protein polymer with precise control of molecular weights, sequence and composition when compared to synthetic polypeptide-based polymer. To prepare the protein polymer, monomer oligonucleotide encoding a defined number of oligopeptide repeats needs to be design and chemically synthesized and then the monomer DNA fragments are multimerized to create larger repetitive genes with desirable length. A few approaches to protein polymer mimetic genes, such as elastin-, silk- and resilin-mimetic genes, have been developed to prepare repetitive polypeptides by using bacterial protein expression system (6,8,9,12). Two strategies are widely employed to generate a variety of repetitive genes for the synthesis of protein polymers, which is classified as concatemerization and recursive directional ligation (RDL).

3.1. Concatemerization

Concatemerization strategy is widely used technique for assembling monomer genes encoding oligopeptide units to multimer genes (Figure 1A) (13). This technique involves annealing of single-stranded DNA to afford double-stranded DNA with internal restriction sites that is not necessarily different. Typical restriction endonucleases (E) were employed to generate the complementary cohesive ends upon cleavage. Use of special restriction endonucleases with non-palindromic sequences (termed as type IIs restriction endonuclease) facilitates the design of monomer DNA sequence without any limitations of enzyme recognition sequence. This type

Table 1. Examples of naturally occurring protein-based materials, corresponding repetitive oligopeptide building blocks and characteristic structure and their properties (65). Repetitive sequence is represented as one-letter code. (Xaa: 20 amino acids)

Protein polymer	Repetitive sequence	Initial modulus (MPa) (65)	Resilience (%) (65)	Features
Collagen	Xaa-Yaa-Gly	1200	90	Hydroxyproline (Yaa), triple helical conformation
Elastin	VPG(Zaa)G, VPAVG	1.1	90	β-turn, (Zaa): 19 amino acids except proline, LCST, temperature responsive
Silk	silkworm: $(AG)_n(Xaa)G$	10000	35	crystalline β-sheet
	dragline spider: $A_nGG(Xaa)$			
Resilin	GGRPSDSYGAPGGGN	2	92	Di-tyrosine crosslink, Resilience, LCST and UCST
Mussel adhesive protein	AKPSYPPTYK	870	28	DOPA (Y), DOPA crosslink

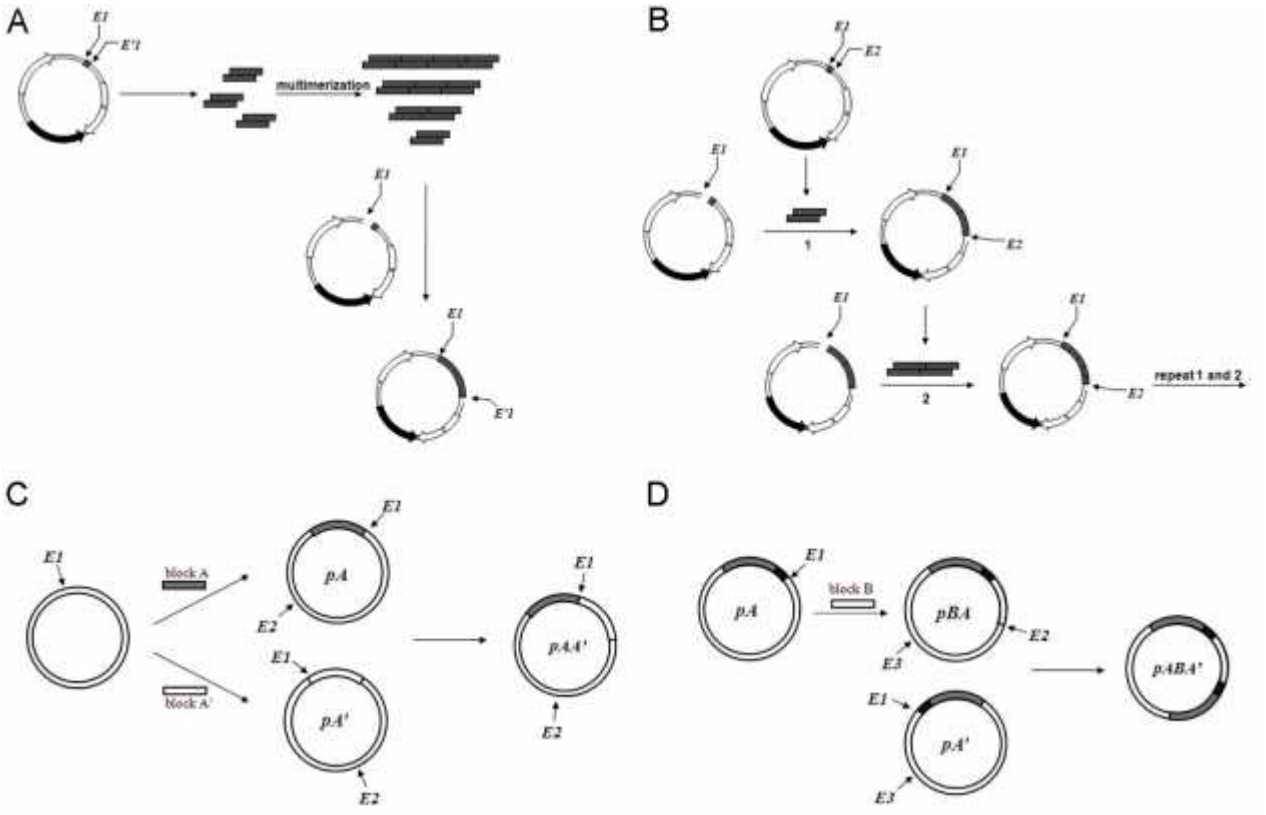

Figure 1. Strategies for construction of synthetic repetitive genes. (A) Concatermerization strategy and (B) Recursive directional liagtion strategy (RDL). (C,D) convergent assembly of protein blocks. E: restriction endonuclease.

of endonucleases recognizes the unique oligonucleotide sequence and cleaves the DNA strand outside their recognition sequences (BbsI, 5'-GAAGAC (+2/+6)-3') so that additional base pairs are not introduced between defined repeats. Other restriction endonucleases can also be utilized but their usage is restricted to specific sequence because they cleave the DNA strand within the recognition sequence involved in monomer design, such as CCN-GGN coding for Pro-Gly motif frequently found in elastomeric protein polymers. Otherwise additional base pairs as a seam need to be inserted between repeats. A set of multimer DNA with increments by one unit length of monomer DNA can be obtained by ligation of a head-to-tail fashion to create a library of multimerized DNA. This multimer DNA library can be directly used for subsequent cloning into a vector and screening multimer inserts.

3.2. RECURSIVE DIRECTIONAL LIGATION

Recursive directional ligation strategy is also commonly employed to generate the repetitive genes with desired size by a step-by-step manner and has been developed by Chilkoti *et al.* to produce different sets of elastin-like polymers from a monomer unit consisting of defined number of pentapeptides (Figure 1B) (14). This method enabled to precisely calculate the molecular weight of recombinant protein polymer by amino acid sequence translated from the number of monomer DNA assembly. Design of monomer DNA in RDL has to meet several requirements as two distinct recognition sequences at each ends to be cut by respective restriction endonucleases, production of complementary overhangs upon cleavage and no interruption of repeat

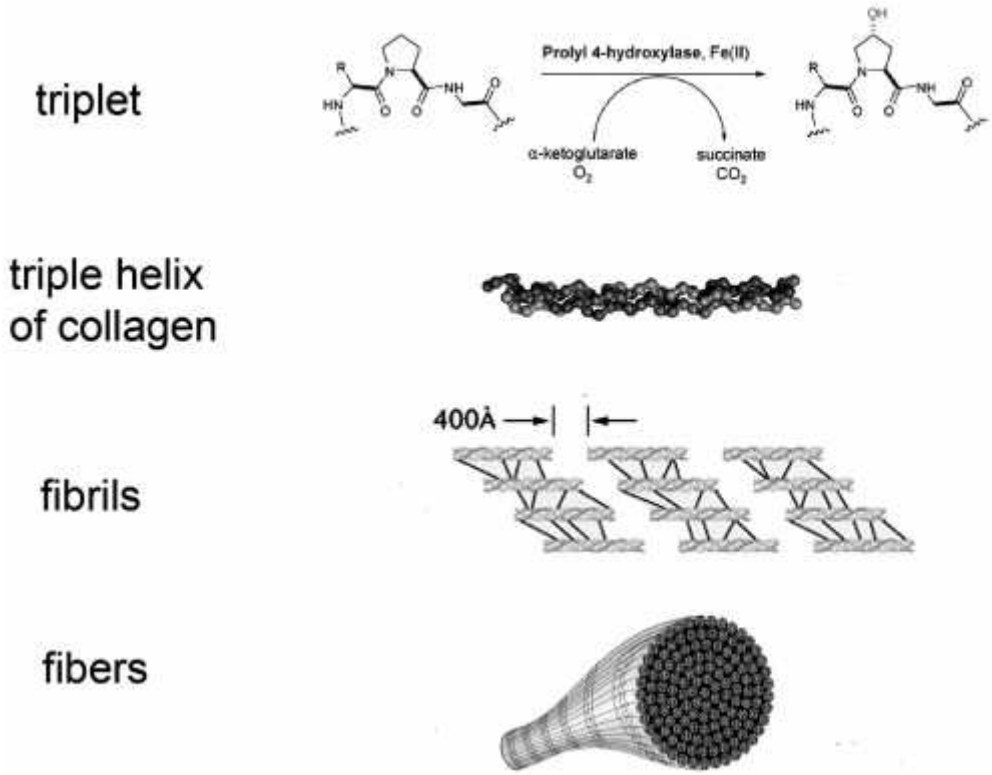

Figure 2. Hierarchical structure of collagen. Hydroxylation of proline residue in collagen triplet ((Xaa)(Yaa)Gly) is catalyzed by prolyl 4-hydroxylase.

sequence by recognition sequences because those recognition sites using RDL are usually located within the coding sequence. Two different restriction endonucleases are used for the purpose of generating complementary cohesive ends. Examples of restriction endonucleases used for the RDL monomer are PflM I (5'-CCANNNN↓NTGG-3') and Bgl I (5'-GCCNNNN↓NGGC-3'). Nhe I coding for Thr-Ser motif and Spe I coding for Ala-Ser motif (5'-(A/G)↓CTAG(T/C)-3') are used a complementary cohesive ends and involved in the coding sequence without interruption of oligopeptide repeat. However, Bgl II and BamH I (5'-(A/G)↓GATC(T/C)-3') was utilized in the synthesis of ELP gene with introduction of additional residues between repeats of monomer. The monomer DNA with a complementary cohesive ends is ligated into a linearized vector cut by one of two restriction endonucleases resulting in two repeats of monomer DNA in the vector. This procedure is performed recursively to grow the number of repeats of monomer DNA. Although RDL needs a step-by-step cloning, desired length of repetitive genes can be achieved after certain number of ligation rounds.

3.3. CONVERGENT ASSEMBLY

Recombinant protein block copolymers can also be synthesized by assembling one multimer DNA to another multimer DNA that each multimer DNA encodes individual protein polymer block with distinct structural and functional properties (15,16). The protein block copolymers have capacity to form physical crosslinks through the self-assembly of one block while the other block remains flexible and hydrated. This self-assembling process by a defined block will result in microphase separation to form the stable protein network structure. The synthetic strategy and method of for assembly of multimer genes have been described in literatures (13,15,16). A convergent assembly from plasmids containing individual blocks is performed to generate di-, tri- or higher block copolymer genes. For diblock construct, the pair of recombinant plasmids pA and pA' encoding each blocks is digested with restriction enzymes, E1 and E2 to afford two fragments. Ligation of the fragment from pA and the fragment from pA' afforded the recombinant plasmid pAA'. For triblock construct, the AA' gene encodes diblock with a single E1 restriction site that can separate into two blocks upon cleavage. Plasmid pAA' is then cleaved by restriction enzyme E1 that enables the insertion of a variety of central blocks. Subsequently another multimer DNA encoding B block is inserted into the compatible E1 site of pAA' to afford a plasmid, pABA', containing a protein triblock copolymer gene.

4. NATURAL PROTEIN-BASED MATERIALS

4.1. Recombinant protein polymers from animals

4.1.1. Collagen

Collagen is the most ubiquitous protein in animals, occurring in extracellular matrix as fibrils. The primary structure of collagen polypeptides consists of tandem repeats of (Gly-Xaa-Yaa) amino acids sequence. Glycine must be in every third position and is arranged such that the glycine NH from one polypeptide forms a

hydrogen-bond to a proline CO from another polypeptide that stabilizes the collagen structure. *Xaa* and *Yaa* in the triplet are often proline and trans-4-hydroxyproline (Hyp) residues, respectively. The hydroxyproline residues of the *Yaa* position play important roles in the thermal stability of the triple helical conformation (17). The structure of collagen molecules in fibrils is characterized by the formation of right-handed triple helix, in which three left-handed polypeptides are assembled in rope-like structure. Triple helices extend to 300 nm in length with 1.5 nm in diameter. These nanometer-sized triple helices self-assemble to form higher-level supramolecular structures with the fibril (Figure 2). Details of the collagen biosynthesis have been understood since the 1970s (18). Unlike syntheses of other structural proteins, biosynthesis of collagen requires a multiple step process of procollagen modification by collagen-specific enzymes to completely fold into triple helix and self-assemble into fibrils. One of the most important post-translational modifications for procollagen is hydroxylation of proline residues in the *Yaa* position. This modification is a prerequisite for thermally stable triple helical formation.

Prolyl 4-hydroxylase (P4H) catalyzes the conversion of proline to trans-4-hydroxyproline residues in -Gly-Xaa-Pro- sequences in collagen and collagen-like domains (Figure 2) (19). Mammalian P4H is a tetrameric enzyme consisting of $\alpha_2\beta_2$ subunits in which the α subunits catalyze hydroxylation of proline residues and the β subunits are multifunctional protein disulfide isomerases (PDI) that contain chaperone-like activity, and stabilize the α subunits by assembling into heterotetramer against inactive aggregates. Since unhydroxylated procollagen synthesized in the absence of this enzymatic activity is unable to form stable triple helix, P4H is the key enzyme in collagen biosynthetic pathway (20). In comparison with prolyl hydroxylation, it is considered that the other post-translational modifications involved in collagen biosynthesis are not essential for triple helical formation.

Currently, most collagen employed in commercial applications is derived from animal sources, specifically bovine. However, there are potential problems such as immunological responses and transmission of animal-related diseases including prions. These problems motivated the development of an alternative biosynthetic methods using genetic and protein engineering. A number of host expression systems have been developed to synthesis large quantities of recombinant collagen. Mammalian, insect cells, and mouse milk have been developed to produce hydroxylated collagen (11,21,22); however, the production of collagen was not high enough for further applications and cost-effective. Among the common expression systems, *E. coli* and yeasts cannot produce fully hydroxylated collagen due to the absence of the gene of P4H. In order to hydroxylate proline residues in these host cells, it is essential to co-express the genes encoding P4H along with recombinant collagen. Significantly, co-expression of recombinant collagen and P4H in *Pichia pastoris* has enabled higher-level production of hydroxylated recombinant collagen (23).

The bacterial expression system from *Escherichia coli* is quick and easy to manipulate. In addition, *E. coli* has been frequently used for expression of multiple genes and enabled high-level production of recombinant proteins. However, there has been only limited success in production of collagen using this expression system. Collagen-mimetic polypeptides consisting of the most frequently occurring tripeptide sequences in natural collagen have been designed and produced by using recombinant DNA techniques (24,25). To stabilize triple helices by protein engineering, short collagen-mimetic peptides $(Pro\text{-}Pro\text{-}Gly)_{10}$ was fused with the timer-forming foldon domain from T4 fibritin (26). One critical problem of these collagen-mimetic peptides is that proline residues in *Yaa* position were not hydroxylated, which leads to thermally unstable triple helices. Studies of the production of hydroxylated collagen in *E. coli* have been reported, in which cotranslational incorporation of trans-4-hydroxyproline (Hyp) was demonstrated using multi-site sense-codon substitution and hyperosmotic shock. This approach had partial success in the production of hydroxyproline-containing collagen in proline auxotrophic *E. coli* strain (27). Although this method is successful for incorporation of hydroxyprolines into recombinant collagen, any proline codons in the recombinant protein sequences could be replaced by Hyp and collagen polypeptides are over-hydroxylated. Another approach to production of hydroxylated collagen in *E. coli* will be coexpression of genes encoding P4H and collagen. However, expression of P4H tetramer in *E. coli* was not successful. The insoluble nature of the α subunits makes it difficult to assemble into an active P4H due to the reduction of two intramolecular disulfide bonds (28). *In vivo* assembly of an active P4H subunit is important for collagen biosynthetic pathway engineering. Recently, genetically engineered *E. coli* host systems facilitated the high-level production of the active P4H tetrameric enzyme (29,30).

Triple helix forming collagens are one of the versatile structural proteins and can provide building block for self-assembling biomaterials. Although collagens have mechanical properties and biological functions suitable for use as an innovative biomaterial, there has not been much effort directed toward collagen engineering in bacterial hosts systems. Several reasons can be suggested. Because of tendency to form insoluble fibrils and need to essentially hydroxylate proline residues for stable triple helix formation, it has been difficult to engineer collagen as an element of protein-based biomaterials. In addition, a large amount of collagens is still available from animal sources despite the potential problems associated with this process.

Synthetic collagen-mimetic peptides derived from chemical synthesis have been extensively investigated for their potential application as biomaterials. Collagen peptide-amphiphile and triblock peptides were designed to facilitate the supramolecular self-assembly of triple helical collagen (31,32). Another approach to synthetic collagen is design of covalently knotted collagen-like peptides, in which cystine-knots mimics C-terminal region of type III collagen (33). In contrast to synthetic collagen-mimetic peptides, most studies of collagen biosynthesis focused solely on expressing recombinant collagens that were identical to native collagens. Few studies of collagen engineering using biosynthetic pathways have been reported. In recent studies, high-level expression strategy for recombinant collagen-mimetic polymers in *E. coli* has been reported and repetitive cell-binding domains

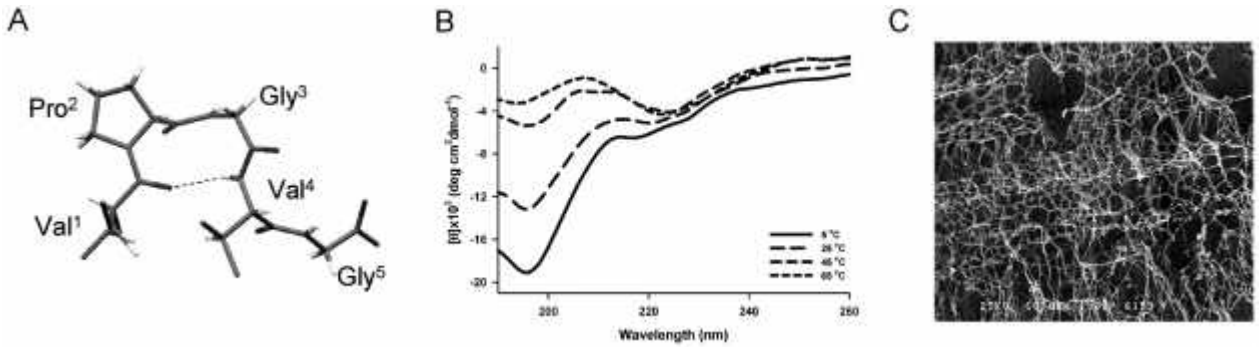

Figure 3. (A) Structural feature of the type II β-turn unit of the elastin pentapeptide repeat characterized by hydrogen bond between Val^4 N-H and Val^1 C=O obtained from the crystal structure of cyclo-$(VPGVG)_3$. (B) Temperature-dependent CD spectra of elastin-mimetic polypeptide depicting the thermally induced conformational transitions. (C) Cryo-etch HRSEM images of coacervates of elastin-mimetic polypeptide in a dilute aqueous solution.

designed from native collagen sequences have showed the ability to promote cell adhesion (34,35). Engineering collagen-mimetic polymers consisting of repeating tripeptides derived from native collagen sequence will be important for development of functionalized collagen. Through rational design strategies, it is possible to create collagen-mimetic biomaterials with enhanced molecular organization and biological properties. Such collagen-mimetic polymers could be used in tissue engineering, wound-healing and drug delivery systems. In addition, recombinant DNA techniques enable the introduction of new functional motifs in the repetitive sequence beyond cell adhesion motifs, while maintaining the triple helical nature. The ability of collagen to self-assemble and organize *in vivo* and *in vitro* is of interest for fabricating nanostructured biomaterials. Considered with its tensile strength and rod-like fibril formation, collagen-mimetic polymers are able to serve as nanowires and nanopatterning materials for nanobiotechnological application (36,37).

4.1.2. Elastin

Elastin is one of major protein components in extracellular matrix (ECM) and often found in skin, lung, artery, ligament and cartilage. The primary sequence of the soluble precursor protein, tropoelastin, is composed of distinct, alternating elastomeric domains and cross-linking domains (4). The elastomeric domains largely consist of repetitive oligopeptides which are rich in valine, proline and glycine. The cross-linking domains are rich in alanine and lysines, which are involved in the enzymatic cross-linking reactions. Native elastin is extremely insoluble and forms filamentous fiber networks due to interaction between hydrophobic structural segments.

Repetitive sequence most commonly found in the elastomeric domain is the pentapeptide (VPGVG) which is a structurally well-defined oligopeptide motif (7). With extensive pioneering works by Urry *et al.*, the thermodynamic and viscoelastic properties derived from the elastomeric domains were emulated by repetitive elastin-mimetic polymers (EMP) created by using recombinant DNA technology (7,10). It is found that EMP based on the pentapeptide repeat sequence undergoes a thermally reversible self-assembly from aqueous solution in analogy to the phase transition behavior of native tropoelastin. Precise determination of the detailed molecular structure of the native elastin is difficult using high-resolution structural analysis, because the molecule is insoluble above the transition temperature (T_t) and conformationally mobile in solution below the temperature. Use of model repetitive oligopeptides has facilitated to better understand the structure-function relationships of the native protein-based materials. Spectroscopic and structural analysis of EMP have suggested that the pentameric sequence units undergo a conformational transition from a random coil to a type II β-turn as the temperature approaches the T_t (38,39). Self-assembling process results from spontaneous phase separation of the protein polymer above a lower critical solution temperature (LCST), which coincides with a conformational rearrangement of the local secondary structure within the pentapeptide motifs (Figure 3).

Amino acid substitution at the specific position within the pentapeptide sequence has been studied to provide insights into detailed understanding of the structure-property relationships of elastin-based biomaterials (10,40). Substitution of valine by isoleucine adding a methylene group, (IPGVG), lowered the T_t of the protein polymer without change of its secondary structure. It has been suggested that the Proline-Glycine (Pro-Gly) sequence motif is frequently found in the repetitive elastomeric proteins and plays a critical role in the structural development and elastomeric properties (5). Studies on proline substitution have indicated an essential role in the periodically occurring Pro-Gly motif within the elastin-mimetic protein polymer with respective to the formation of a type II β-turn structure. Kim *et al.* incorporated structurally diverse set of unnatural imino acids into elastin-mimetic polymer (41). Especially, introduction of fluoroprolines has demonstrated the importance of proline ring in the local conformation and subsequent influence on thermal self-assembly of elastin analogues (42,43). Both (4R)- and (3R)-fluoroprolines preferably adopt the type II β-turn structure and (4R)-fluoroproline reduces the transition temperature in comparison with proline.

Two elastin repeat sequences, elastomeric [(I/V)PGVG] or plastic [(I/V)PAVG] sequence, are

thoroughly studied as they show tunable thermodynamic and mechanical properties through amino acid substitution in a position-specific manner and length of pentapeptide repeat. Of amino acid residues within elastin repeat unit, valine at the forth position within the elastomeric pentapeptide unit can be substituted for other naturally occurring amino acids except for proline without the structural change. Most importantly, the T_t of elastin-mimetic polymers, [VPG-(Xaa)-G], can be modulated depending on the chemical identity and polarity of side chain. Hydrophilic amino acid residues such as aspartic acid and glutamic acid shifted the Tt to higher temperature whereas hydrophobic amino acid residues such as tryptophan and tyrosine shifted Tt to lower temperatures (45). Moreover, glycine residue at the third position is replaced with alanine which has additional methyl group. The plastic polymer (or plastin) consisting of one of the pentapeptide sequences [(I/V)PAVG] or their combination exhibits dramatic changes in the mechanical behavior and secondary structure with respect to the elastomeric polymer. First, mechanical differences were observed between the crosslinked elastin and plastin hydrogels. Young's modulus of the plastin is 2.0X108 dynes/cm2, which is approximately two orders of magnitude greater than that of the elastin (1.0X106 dynes/cm2) (7). In addition, DSC and turbidity studies have indicated that plastin exhibited difference in thermal behavior during heating and cooling process (ΔT = 25 oC) in which restoration to the soluble state required a significant cooling (Tt = 6 oC for cooling) (45). The plastic polymer also exhibited different enthalpy values (ΔH) between heating (34 J g-1) and cooling (66 – 91 J g-1). In contrast, the elastin showed completely reversible thermal behavior in which the transition temperature and the enthalpy values are almost identical during heating and cooling process. The difference in secondary structure is correlated with the observed difference in reversibility of the phase transition. NMR and Raman spectroscopic studies of poly(VPAVG) have indicated that the Pro-Ala sequence motif adopts different secondary structure from poly(VPGVG), but other residues (-VGV-) preferably form a β-sheet structure through intermolecular hydrogen bonding interactions that result in stabilization of the self-assembled plastic polymers upon cooling. Differences in the transition temperature and mechanical response between the two elastin protein polymers can be used as a design concept for generating a variety of elastin-mimetic block copolymers, which exhibit similar properties to synthetic thermoplastic elastomers but improve biocompatibility (13,15,46).

Synthetic polymers can self-assemble into a range of supramolecular structures that facilitate the loading of hydrophobic drugs into local microenvironments. Likewise, elastin-mimetic polymers undergo phase separation through self-assembly to afford the formation of hydrophobic aggregates ranging from the nano to the micron scale. In addition, diverse material platforms, including nano- or micro-particles, fibers, hydrogels and films that incorporate drugs can be formed through appropriate material process (46-49). Elastin-mimetic biomaterials have great potential in drug delivery and implant application. Chilkoti and coworkers have designed and synthesized a diverse set of elastin-like polymers (ELP) that are able to self-assemble above the body temperature to 42 ^{o}C (50,51). These ELPs were conjugated to fluorescent probes or anticancer drug, doxorubicin through an acid-sensitive linker to allow the pH-triggered release of the drug. They have demonstrated the uptake of their ELP conjugates within the tumor during hyperthermia treatment (heating to 42 ^{o}C) which is clinically used in cancer therapy. The ELP diblock copolymers were genetically fused with a variety of bioactive peptides such as Tat peptide, cell penetrating peptide (CPP) and Arg-Gly-Asp (RGD) peptide so that they are able to form multivalent micro-/nano-particles after heating (52). These ELPs were used to target for cancer cells or enhanced the selective uptake of the ELP within the tumor. Herrero-Vanrell et al. generated self-assembled microparticles of the poly(VPAVG) that display good biocompatibility and thermal hysteresis behavior (53,54). Dexamethasone phosphate (DMP) and bone morphogenetic protein (BMP) as model drugs were encapsulated within the plastic microparticle by heating above its transition temperature without influence on self-assembly into the microparticle. The DMP and BMP release profiles showed initial burst release to almost 80 % and 15 %, respectively, and then a sustained release. For the development of protein-based gene delivery system, Chen et al. designed and synthesized oligolysine-containing elastin-mimetic polymer for gene therapy application (Figure 4A) (55). Chaikof and coworkers have employed elastomeric and plastic polymers for the construction of triblock copolymers, which have robust viscoelastic and mechanical responses (15,46,47). A central block comprised of [(VPGAG)2(VPGEG)(VPGAG)2] repeats displays elastomeric behavior and is hydrophilic due to the presence of alanine and glutamic acid that is negatively charged at physiological pH, while identical end blocks consisting of [(I/V)PAVG] repeats display thermoplastic behavior and self-assembly of end blocks acts as physical crosslinks above its transition temperature. The triblock copolymer films containing sphingosine-1-phosphate (S1P) as a model drug were cast from water and trifluoroethanol (TFE) and then the release of S1P from the elastin-mimetic films was measured for fifteen days. The calculated diffusion coefficient of S1P from water-cast films was approximately 30 fold higher than that from TFE-cast films because different process of elastin triblock induced significant changes in microstructure of the protein film. It was found that release rate of drugs was dependent upon film processing conditions that influence microstructure and microphase separation of the material. Temperature-responsiveness of polymeric materials has been utilized for drug delivery, cell encapsulation and tissue engineering in a variety of platforms. Block copolymers such as poly(N-isopropylacrylamide) (pNIPAAm)-based polymers and the poloxamer (Pluronics®) are widely used temperature-sensitive polymers. They are solution below the transition temperature of the polymer and can be injected as solutions mixed with drugs or cells, while forming gels above the transition temperature or at body temperature (56). Likewise, elastin-based polymers can undergo temperature-dependent micro-phase separation and processing as injectable formats mixed with various drugs is possible. Adams *et al.* synthesized ELP, $\{[VPGKG(VPGVG)_{16}]_{102}\}$, containing crosslinkable site by THPP (β-[tris(hydroxymethyl) phosphino] proprionic acid) in which the ELP can form an injectable in-situ gel. They loaded

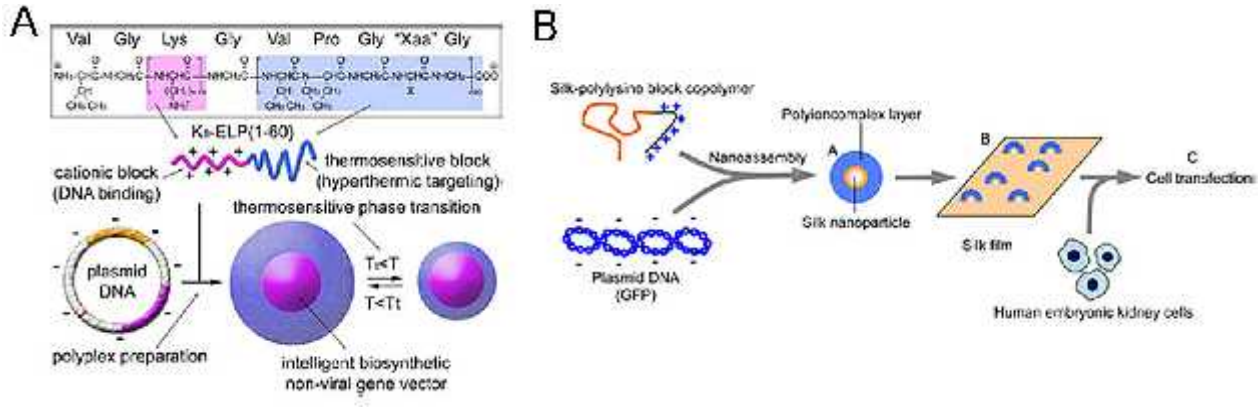

Figure 4. Strategies for non-viral gene delivery using protein-based block polymers. (A) poly(L-lysine)-elastin-mimetic polymer construct and (B) silk-mimetic polymer-poly(L-lysine) construct . Reproduced from ref. 55 and 74, respectively.

antibiotics, cefazolin and vancomycin, for local delivery to infected sites in orthopedic applications, such as bone and joint (57). The drug release profile from crosslinked ELP showed that 80 and 60 % of cefazolin and vancomycin were released for 3 and 20 days, respectively. Shamji *et al.* have reported another in-situ gel forming ELP, $(VPGVG)_{120}$, after perineural injection, which provide the possibility of perineural drug delivery (58). Hart *et al.* have studied the release of model compounds, such as theophylline, vitamin B12 and ovalbumin from elastin triblock copolymer gel (59).

Elastin-mimetic polymers for use in tissue engineering are also of significant interest as 3-D scaffolds for vascular grafts and tissue repair. Successful implantable small diameter vascular graft material requires cell adhesion onto vascular grafts and crosslinks to modulate mechanical response of the material. In this regard, cell adhesion can be obtained by adding cell-binding domains to elastin-mimetic sequences and crosslink by incorporating lysine residues either within the elastin repeat sequence or flanking the elastin sequence. Heilshorn *et al.* have reported that incorporation of CS-5 cell-binding domains enabled the adhesion of human vein endothelial cells (HUVEC) to this material (60,61). It was found that the ability of this material to cell binding and spreading largely depend on the frequency of cell-binding domains and the location of crosslinking lysine residues. Cabello and coworkers employed similar elastin-mimetic polymer construct containing the cell-binding motif and the lysine residues for development of ocular surface extracellular matrix that support cell attachment and growth (62).

Elastin-mimetic polymers are a promising new class of biomaterials for the drug delivery and implant applications. These materials exhibit high level of biocompatibility and a tunable range of biomechanical properties as a result of the precise control of pentapeptide sequence, block sequence, and processing conditions. A variety of opportunities exists in the processing of these protein polymers into nanoparticles, hydrogels, films or fibers for a number of implant applications in orthopedics, as well as in plastics, cardiovascular, and general surgery. These applications can be further extended by incorporating bioactive drugs, proteins, and cells, within or onto the protein polymer network.

4.2. Recombinant protein polymers from insects or spiders

4.2.1. Silk

Natural silks from both spiders and insects (silkworms) have attracted research interest due to their self-assembling behaviors, remarkable mechanical properties, biodegradability and biocompatibility that make silks useful biomaterials (3). Silks, therefore, have been of significant interest for a variety of application in the field of biomedical engineering and materials sciences.

Silk proteins from both insects and spiders are fibrous materials and characterized by a modular sequence of alternating, structurally distinct crystalline and amorphous domains largely consisting of repetitive sequences (63,64). Spiders produce various types of silk proteins with distinct mechanical properties and different roles in nature. The silkworm, *Bombyx mori*, also produces large quantities of silk fibroin that are used as textiles.

Detailed studies on silk fibers have been focused on two silks of *Nephila clavipes* and *A. diadematus*, dragline silk and flagelliform silk. The dragline silks, known as a rigid fiber, of *Nephila clavipes* form a safety line and a framework of the web, and display an initial modulus of 10 GPa, a high tensile strength (1.1 GPa) and about one-third extensibility (65). Spider dragline silk proteins are comprised of alanine-rich segments and glycine-rich segments. The alanine-rich segments adopt beta-sheet secondary structures that form highly oriented crystalline while glycine-rich segments are amorphous. The capture spiral of the web consists of flagelliform silk and non-fibrous aqueous coating of aggregate silk. In contrast to the dragline silk, the flagelliform silk exhibits a high extensibility stretching over twice its length and it has an initial modulus of 0.003 GPa and a tensile strength of 0.45 GPa (65). The flagelliform silk contains tandem repeats of Gly-Pro-Gly-Gly-Xaa where Xaa can be Ala,

Ser, Tyr, and Val. The crystalline domain of *Bombyx mori* silk fibroin is characterized by repetitive sequences of the hexapeptide motif, Ser-Gly-Ala-Gly-Ala-Gly containing high content of Gly-Ala sequence motif.

Inspired by the structure and properties of natural silks, a variety of recombinant silk-mimetic polymers was prepared by using E. coli expression system which is the most widely used protein expression system for large scale production. Cappello and coworkers have first explored the synthesis and expression of synthetic genes encoding GAGAGS repeats derived from the crystalline region of silkworm silk fibroin (66). Krejchi *et al.* used the repetitive sequence, GAGAGAG (E or M), to control assembly of beta sheet and thus thickness in crystalline (67). Introduction of glutamic acid or oxidation-reduction of methionine at the end of repeat unit prevented the β-sheet formation. Prince *et al.* and Winkler *et al.* have designed, cloned and expressed the repetitive synthetic genes encoding [GRGGLGGQGAGAAAAAGGAGQGGYGGLGSQG] repeats derived from the conserved repetitive sequences of spider dragline silk (68,69). Repetitive silk-mimetic polymer consisting of $[(GPGGSGPGGY)_2GPGGK]$ derived from spider flagelliform silk has been synthesized by Zhou *et al.* to investigate the structure-function relationships suggesting that the periodic presence of type II β-turns may contribute to the elastomeric behavior of this material like other elastomeric polypeptides such as elastin-mimetic polymer (70).

Due to high-performance mechanical properties, biocompatibility, biodegradability and stability, natural silk proteins have been used as sutures in surgery for a long time. In a selective solvent, silk-mimetic polymers can be processed into various formats of biomaterials including fibers, hydrogels, films and microspheres. Fusion of silk-mimetic polymers to various functional domains can provide a variety of opportunity in biomedical and tissue engineering applications. Bini *et al.* have made fusion of a cell-binding motif, RGD, to the end of the recombinant dragline silk-mimetic polymer to enhance cell adhesion on the silk-based film (71). Furthermore, silaffin (R5 peptide) and dentin matrix protein 1 (DMP1) domains were fused to the recombinant silk-mimetic polymer to induce a selective nucleation with silica and hydroxyapatite, respectively, on the self-assembled silk-mimetic material through beta-sheet formation (72,73). These mineralized silk-based biomaterials have potential applications in bone tissue engineering due to the combination of mechanical properties derived from the spider silk domain and the nucleating domains.

For drug delivery application, especially as nonviral gene therapy, Numata *et al.* designed and synthesized amphiphilic silk-mimetic diblock copolymers consisting of the repetitive oligopeptide block derived from the dragline silk protein sequence and oligolysine (poly(L-lysine), pLL) block (Figure 4B) (74). The anionic plasmid DNA (pDNA) can form ionic pairs with the cationic block, pLL, via electrostatic interactions. Silk-mimetic protein block self-assembles in aqueous solution and pLL block forms the complexes with pDNA. Depending upon molar ratio of two molecules, size of silk-pLL-pDNA complexes ranged from 100 to 800 nm particles. The silk-pDNA complexes showed no cytotoxicity and low level of cell transfection. The amphiphilic silk-mimetic diblock copolymer was further modified with cell-binding motif to enhance cell binding and improve transfection efficiency (75). The pDNA complexes of silk-pLL-RGD protein block copolymers varying number of RGD motif were employed for gene delivery to several cell types. It was found that transfection efficiency was linearly dependent on the number of RGD sequences.

Silk proteins are of great interest in the development of new biomaterials because of their robust mechanical properties, biocompatibility and biodegradation. With use of recombinant DNA techniques and material processing techniques, creation of silk-mimetic block copolymers offer a variety of opportunity for potential applications in drug/gene delivery, tissue engineering and coatings of medical device.

4.2.2. Resilin

Resilin and elastin are naturally occurring elastomeric proteins found in wing hinges of insects and arteries of animals, respectively, which continuous movements occur during their lifetime. Both proteins possess high resilience (> 90%) and low stiffness (1-2MPa). Resilin protein was first discovered from studies about flight systems of desert locusts and dragonflies and plays significant role in flying and jumping mechanisms of insects (76). Structural studies of this protein suggest that resilin is an unstructured amorphous protein matrix and dynamic (12,77). FTIR and circular dichroism (CD) studies have recently provided further evidence for random coil conformation (77).

Resilin from insect joints and tendons are found to have covalent crosslinks where the crosslinking occurs between tyrosine residues, forming di- and tri-tyrosine catalyzed by an enzyme, peroxidase (78). The crosslinked material exhibits mechanical properties of a high resilience (92%, tendon from the wing region of the dragonfly) and a high fatigue lifetime over 300 million cycles (79). It is reported that tendon reilin in dragonfly could be stretched three times and compressed one third its original length (80). These outstanding properties make resilin a promising new protein polymer material having great potential for biomedical and tissue engineering applications.

Many naturally occurring protein polymers contain structurally or functionally distinct domains. Likewise, native resilin, 60 kDa, is encoded by three exons of the *Drosophila melanogaster* CG15920 gene which are composed of N-terminal pentadecapeptide repeat domain, putative chitin binding domain and C-terminal tridecapeptide repeat domain (81). The repeat sequence, 18 repeats of pentadecapeptide consensus sequence (GGRPSDSYGAPGGGN), is dominantly identified at the N-terminal region. Elvin *et al.* successfully cloned and expressed the gene encoding the first exon region (rec1-

resilin) as a soluble form in *E. coli* (79). The recombinant rec1-resilin could be chemically crosslinked via the formation of di-tyrosine crosslinks (21%) using Ru(II)-mediated photo-crosslinking ($[Ru(bpy)_3]^{2+}$). Native resilin from locust wing hinge contains 25% di-tyrosine. This crosslinked rec1-resilin cast to rod and strip displayed similar properties as native resilin including a pH-dependent fluorescence and high resilience (90-92%), being able to extend to over 300% of its original length without breaking. Synthetic rubber-like polymers, such as chlorobutyl rubber and polybutadiene rubber, showed much lower resilience, 56 and 80%, respectively.

Additionally, Lyon's *et al.* generated recombinant resilin-mimetic polypeptides consisting of repeat sequences, Dros16 ($[GGRPSDSYGAPGGGN]_{16}$) from the *D. melanogaster* gene and An16 ($[AQTPSSQYGAP]_{16}$) from the *Anopheles gambiae* gene (82). Both resilin-mimetic biomaterials were prepared by Ru-mediated photo-crosslinking with di-tyrosine level of 46 and 14%. These resilin-mimetic polypeptides exhibited similar mechanical properties such as low stiffness (5.7kPa for An16) and high resilience (91% for Dros16 and 98% for An16) which are comparable to those of rec1-resilin. Kiick and coworkers have designed and synthesized the resilin-like polypeptide containing the twelve repeats of resilin pentadecapeptide (RLP12, $[GGRPSDSFGAPGGGN]_{12}$) where tyrosine is replaced with phenylalanine and incorporating bioactive motif (GRGDSP) after six resilin repeats (83). The repeat sequences are flanked by lysine residues which could be crosslinked via THPP ([tris(hydroxymethyl)-phosphino]propionic acid). The resulting resilin-like films exhibit Young's moduli of approximately 30-60kPa with stretching to average 180% before breaking. In addition, incorporation of RGD motifs facilitated RLP12 films to have the ability to cell adhesion and proliferation of NIH-3T3. The pentadecapeptide sequence was used to mimic the unstructured and elastic features of a muscle protein titin which serves as a molecular spring composing of folded domains and unordered sequences. Shanshan *et al.* generated a titin-mimetic protein through fusion of elastomeric GB1 for folded domain and resilin as a random coil structure (84). The photo-crosslinked GB1-resilin materials exhibited high resilience (>99%) at low strain and could be stretched to 135% without breaking with Young's moduli of 50-70 kPa.

Recombinant resilin proteins have recently emerged as promising biomaterials and its outstanding mechanical properties provide significant opportunities for biomedical and tissue engineering applications such as the medical device including spinal disc and artificial artery and injectable hydrogels for drug delivery.

4.3 Recombinant protein polymers from marine species

4.3.1 Mussel adhesive protein

In nature, a variety of biological systems often produce adhesive materials with the strength and adhesion to surface that synthetic materials are hard to emulate under the same condition. Of those, marine mussels have special ability to stick to a solid surface through mussel byssus thread under water (2). Similar to spider silk, byssus contains a variety of protein-based materials. Byssal thread is a collection of natural protein-based block copolymers known as PreCol-P in proximal thread and PreCol-D in distal thread, both similarly consisting of a central collagen block flanked by elastin or silk blocks and histidine-rich domains at both ends (85). The difference between two copolymers is that PreCol-P has an elastin block and PreCol-D has a silk block at both ends. Elastin and silk are the structural proteins and widely studied as versatile protein-based biomaterials. Repetitive elastin and silk sequences provide distinctive properties to the fibril thread such as elasticity and stiffness, respectively. Two histidine-rich domains at both end blocks have a role in acting as metal binding domains. Although these structural proteins would be useful as new biomaterials, there have not been attempts to synthesize analogous materials due to the synthetic difficulty of collagen mimics.

At the interface between the byssus thread and the substratum where attachment takes place, byssal plaque contains several mussel adhesive proteins (MAP) that play an important role in underwater adhesion. With pioneering studies by Waite, mussel foot protein-1 (fp-1) is found in the byssal thread and major protein component coating the thread (86,87). In contrast, mussel adhesive foot protein-3 and -5 (fp-3 and -5) are frequently found in the adhesion plaque and relatively small proteins (88,89). Significantly, amino acid composition analysis shows that these proteins contain high levels of post-translationally modified amino acids such as 3,4-dihydroxyphenylalanine, (4R)-hydroxyproline and (3R,4S)-3,4-dihydroxyproline. The mussel foot proteins commonly have high 3,4-dihydroxyphenylalanine (DOPA) content ranging from 10 to 30 mol%. Of the foot proteins, fp-3 and -5 are known to have the highest levels of DOPA. DOPA residues are expected to play a critical role in surface adhesion and intermolecular crosslinking via oxidation of catechol group of DOPA residues to reactive o-quinone (90,91). It has been reported that mussel adhesive protein with the regular tyrosine residues displayed reduced adhesion property (91). Moreover, the adhesion ability of foot proteins is proportionally correlated with DOPA content.

The mussel adhesive protein has been successfully commercialized as Cell-Tak™ (BD Bioscience Clontech) and MAP™ (Swedish BioScience Lab) for biomedical applications. The adhesive proteins were initially obtained from the natural source, mussel, through protein extraction techniques. However, direct extraction from mussels has proven to be inefficient because of a low yield of extraction (about 10,000 mussels to obtain 1 g of adhesive protein), technical difficulty to obtain a pure protein and high cost required for extraction (92). To improve these limitations, the recombinant adhesive proteins have been produced in bacterial expression system. There have been several attempts to synthesize recombinant foot protein-1 (fp-1) derived from cDNA coding sequence (93,94). Although the recombinant fp-1 was successfully expressed, it was found that fp-1 is insoluble and expression yield is low due to codon bias between Mussel and *E. coli*. For relatively small adhesive proteins but higher DOPA content, Hwang *et al.* have

recently cloned and expressed cDNA coding for foot protein-3 and 5 (fp-3 and -5) of *Mytilus galloprovincialis* in a soluble form in *E. coli* (95,96). Since unhydroxylated recombinant proteins showed weak adhesive property, the purified proteins were treated with an enzyme, tyrosinase, to generate hydroxylated fp-3 and 5 with conversion of tyrosine residues to DOPA. The recombinant foot proteins showed to have adhesive abilities comparable to those of the commercially available adhesive product, Cell-Tak^{TM}.

Salerno and Goldberg have successfully synthesized and expressed repetitive synthetic genes encoding the twenty repeats of the decapeptide unit (AKPSYPPTYK) derived from fp-1 of *Mytilus edulis* (93). Later, Kitamura *et al.* have explored the synthesis of short decapeptide repeats to investigate its secondary structure and adhesive property (97). The CD study indicated that the mussel-mimetic polypeptide dominantly exist in a random coil structure. Although the production of mussel-mimetic polypeptide was successful, there still remain issues such as insufficient yield and insolubility limiting its use in biomedical and industrial application. Hwang *et al.* have designed and expressed recombinant mussel adhesive proteins, fp151 and fp-353, consisting of a central fp-5 flanked by identical decapeptide repeats derived from fp-1 and more recently fp-5 flanked by identical fp-3, respectively (98,99). The fusion approach improved the yield and solubility of recombinant mussel adhesive proteins. The fp-151 was further modified by fusion of a cell-binding motif, RGD, to the end of the recombinant protein polymer to enhance cell adhesion on the RGD-containing protein-coated surface (100).

Recombinant mussel adhesive proteins have several possible applications. As medical and dental adhesives, adhesive protein polymers have shown good biocompatibility and therefore if appropriately processed as material platforms, they may play a significant role in the repair of damaged organs or as dental sealants. With notion of remarkably powerful adhesive property to various solid surfaces, another potential application can be the development of protein-based coating materials as new anti-biofouling materials in marine environments as well as coatings for medical device.

4.3.2. Reflectin

Another marvelous protein polymer occurring in nature is the structural protein of reflective tissues that functions in camouflage by modulating an incoming light in the skin and eye. For example, cephalopods can rapidly alternate the color and reflectance of their skin in response to the environmental stimuli.

This protein has been first identified from the dermis of the squid *Lolliguncula brevis* (101). Recently, McFall-Ngai and coworkers have shown that the protein from the reflective platelets of the Hawaiian bobtail squid *Euprymna scolopes*, named as reflectin, contains conserved oligopeptide sequence, $[M/FD(X)_5MD(X)_5MD(X)_{3/4}]$, found in five separate domains (102). The recombinant reflectin was successfully cloned and expressed in *E. coli*. Kramer *et al.* has revealed that the recombinant reflectin (reflectin a1) could undergo self-assembly into diverse morphologies under a variety of conditions including pH, additives, buffers, osmolytes and redox environment and be processed into films with very high refractive index (1.591±0.002) and fibers. Tao *et al.* has generated another recombinant reflectin protein (Reflectin A1 (RA1)) found in the iridophore cells of the squid *Loligo pealeii* (104). The RA1 contains the similar repeat sequence, $[PER(Y/W)MDMSGYQMDMQGRWMD(X)_3R(X)_3P]$, to reflectin a1, but less variable sequences. Mechanism for color change was proposed as the self-assembled protein film forms alternating layer with extracellular space around cells and by changing the spacing and thickness through phosphorylation this multilayer interacts with light differently such that the color of the skin can be modulated.

So far little has been known about the reflectin-mimetic biomaterials. However, based on previous pioneering studies on the reflectin proteins, synthesis of reflectin-mimetic polypeptides based on repetitive sequences and mutilayered thin films is of significant interest in bottom-up fabrication of nanostructured devices potentially for spectroscopic and bio-optic application.

5. SUMMARY AND PERSPECTIVE

Naturally occurring polypeptide-based biomaterials show incredibly outstanding properties when compared to synthetic polymer materials, but for obtaining enough protein material growing natural sources including animals, insects and marine organisms may not be ideal. Researchers have been seeking alternative ways to generate this material in large scale. With advances in protein engineering, the design of protein polymer consisting of naturally occurring oligopeptide building block with defined structures and properties is possible and design of protein block copolymer can fine tune the material property. Recombinant DNA approaches were useful for introducing structurally and functionally unique elements such as cell-binding motif to the repetitive gene. Furthermore, metabolic engineering approaches would be of value for increasing the production yield of highly repetitive protein polymer. As illustrated by biological studies, these protein polymers show good biocompatibility and biodegradation when implanted. As new generation biomaterials, recombinant polypeptide-based materials are being extensively studied for applications including drug delivery and tissue engineering.

6. ACKNOWLEDGEMENTS

The author would like to thank Dr. Chaikof for the support in the preparation of manuscript.

7. REFERENCES

1. van Hest, J. C., Tirrell, D. A., Protein-based materials, toward a new level of structural control. *Chem Commun (Camb)* 19, 1897-904 (2001)

2. Waite, J. H., Adhesion in byssally attached bivalves. *Biol. Rev.* 58, 209-231 (1983)

3. Kaplan, D. L., Adams, W. W., Farmer, B., Viney, C., Silk polymers: Materials science and biotechnology, *ACS Symp. Ser.* vol. 544 (1994)

4. Indik, Z., Yeh, H., Ornstein-Goldstein, N., Sheppard, P., Anderson, N., Rosenbloom, J. C., Peltonen, L., Rosenbloom, J., Alternative splicing of human elastin mRNA indicated by sequence analysis of cloned genomic and complementary DNA. *Proc Natl Acad Sci U S A, 84* (16), 5680-5684 (1987)

5. Kim, W., Conticello, V. P., Protein engineering methods for investigation of structure-function relationships in protein-based elastomeric materials. *Polym. Rev. 47, 93-119 (2007)*

6. Haider, M., Megeed, Z., Ghandehari, H., Genetically engineered polymers: status and prospects for controlled release. *J. Control. Release* 95(1), 1-26 (2004)

7. McGrath, K. P., Kaplan, D. L., Eds, *Protein-Based Materials*. 1997. 429.

8. Cappello, J., Crissman, J., Dorman, M., Mikolajczak, M., Textor, G., Marquet, M., Ferrari, F., Genetic engineering of structural protein polymers. *Biotechnol. Prog.* 6(3), 198-202 (1990)

9. Prince, J. T., McGrath, K. P. DiGirolamo, C. M., Kaplan, D. L., Construction, cloning, and expression of synthetic genes encoding spider dragline silk. *Biochemistry* 34(34), 10879-10885 (1995)

10. Urry, D. W., Molecular machines: how motion and other functions of living organisms stem from reversible chemical changes. *Angew Chem Int Ed Engl* 32(6), 819-841(1993)

11. John, D. C., Watson, R., Kind, A. J., Scott, A. R., Kadler, K. E., Bulleid, N. J., Expression of an engineered form of recombinant procollagen in mouse milk. *Nat. biotechnol.* 17(4), 385-389 (1999)

12. Lyons, R. E., Lesieur, E., Kim, M., Wong, D. C., Huson, M. G., Nairn, K. M., Brownlee, A. G., Pearson, R. D., Elvin, C. M., Design and facile production of recombinant resilin-like polypeptides: gene construction and a rapid protein purification method. *Protein Eng Des Sel 20* (1), 25-32 (2007)

13. Wright, E. R., Conticello, V. P., Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv Drug Deliv Rev* 54, 1057–1073 (2002)

14. Meyer, D. E., Chilkoti, A., Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules 3* (2), 357-367 (2002)

15. Kim, W., Chaikof, E. L., Recombinant elastin-mimetic biomaterials: Emerging applications in medicine. *Adv Drug Deliv Rev 62* (15), 1468-1478 (2010)

16. Rabotyagova, O. S., Cebe, P., Kaplan, D. L., Protein-based block copolymers. *Biomacromolecules 12* (2), 269-289 (2011)

17. Holmgren, S. K., Bretscher, L. E., Taylor, K. M., Raines, R. T., A hyperstable collagen mimic. Chem.Biol. 6(2), 63-70 (1999)

18. Prockop, D. J., Kivirikko, K. I., Tuderman, L., Guzman, N. A., The biosynthesis of collagen and its disorders (first of two parts) *N Engl J Med 301* (1), 13-23 (1979)

19. Kivirikko, K. I., Myllyharju, J., Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol 16* (7), 357-368 (1998)

20. Berg, R. A., Prockop, D. J., The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen. *Biochem Biophys Res Commun 52* (1), 115-120 (1973)

21. Fertala, A., Sieron, A. L., Ganguly, A., Li, S. W., Ala-Kokko, L., Anumula, K. R., Prockop, D. J., Synthesis of recombinant human procollagen II in a stably transfected tumour cell line (HT1080) *Biochem J 298 (Pt 1)*, 31-37 (1994)

22. Nokelainen, M.; Helaakoski, T.; Myllyharju, J.; Notbohm, H.; Pihlajaniemi, T.; Fietzek, P. P.; Kivirikko, K. I., Expression and characterization of recombinant human type II collagens with low and high contents of hydroxylysine and its glycosylated forms. *Matrix Biol 16* (6), 329-338 (1998)

23. Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T., Kivirikko, K. I., Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *EMBO J 16* (22), 6702-6712 (1997)

24. Goldberg, I., Salerno, A. J., Patterson, T., Williams, J. I., Cloning and expression of a collagen-analog-encoding synthetic gene in Escherichia coli. *Gene 80* (2), 305-314 (1989)

25. Yao, J., Yanagisawa, S., Asakura, T., Design, expression and characterization of collagen-like proteins based on the cell adhesive and crosslinking sequences derived from native collagens. *J Biochem 136* (5), 643-649 (2004)

26. Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H. P., Engel, J., Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol 308* (5), 1081-1089 (2001)

27. Buechter, D. D., Paolella, D. N., Leslie, B. S., Brown, M. S., Mehos, K. A., Gruskin, E. A., Co-translational

incorporation of trans-4-hydroxyproline into recombinant proteins in bacteria. *J Biol Chem 278* (1), 645-650 (2003)

28. John, D. C., Bulleid, N. J., Prolyl 4-hydroxylase: defective assembly of alpha-subunit mutants indicates that assembled alpha-subunits are intramolecularly disulfide bonded. *Biochemistry 33* (47), 14018-14025 (1994)

29. Kersteen, E. A., Higgin, J. J., Raines, R. T., Production of human prolyl 4-hydroxylase in Escherichia coli. *Protein Expr Purif 38* (2), 279-291 (2004)

30. Neubauer, A., Neubauer, P., Myllyharju, J., High-level production of human collagen prolyl 4-hydroxylase in Escherichia coli. *Matrix Biol 24* (1), 59-68 (2005)

31. Fields, G. B., Lauer, J. L., Dori, Y., Forns, P., Yu, Y. C., Tirrell, M., Protein-like molecular architecture: biomaterial applications for inducing cellular receptor binding and signal transduction. *Biopolymers 47* (2), 143-151 (1998)

32. Martin, R., Waldmann, L., Kaplan, D. L., Supramolecular assembly of collagen triblock peptides. *Biopolymers 70* (4), 435-444 (2003)

33. Kotch, F. W., Raines, R. T., Self-assembly of synthetic collagen triple helices. *Proc Natl Acad Sci U S A 103* (9), 3028-3033 (2006)

34. Yin, J., Lin, J., Li, W., Wang, D. I., Evaluation of different promoters and host strains for the high-level expression of collagen-like polymer in Escherichia coli. *J Biotechnol 100* (3), 181-191 (2003)

35. Kurihara, H., Morita, T., Shinkai, M., Nagamune, T., Recombinant extracellular matrix-like proteins with repetitive elastin or collagen-like functional motifs. *Biotechnol Lett 27* (9), 665-670 (2005)

36. Scheibel, T., Parthasarathy, R., Sawicki, G., Lin, X. M., Jaeger, H., Lindquist, S. L., Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proc Natl Acad Sci U S A 100* (8), 4527-4532 (2003)

37. Wilson, D. L.; Martin, R.; Hong, S.; Cronin-Golomb, M.; Mirkin, C. A.; Kaplan, D. L., Surface organization and nanopatterning of collagen by dip-pen nanolithography. *Proc Natl Acad Sci U S A 98* (24), 13660-13664 (2001)

38. Urry, D. W., Shaw, R. G., Prasad, K. U., Polypentapeptide of elastin: temperature dependence of ellipticity and correlation with elastomeric force. *Biochem Biophys Res Commun 130* (1), 50-57 (1985)

39. Reiersen, H., Clarke, A. R., Rees, A. R., Short elastin-like peptides exhibit the same temperature-induced structural transitions as elastin polymers: implications for protein engineering. *J Mol Biol 283* (1), 255-264 (1998)

40. Urry, D. W., Long, M. M., Harris, R. D., Prasad, K. U., Temperature-correlated force and structure development in elastomeric polypeptides: the Ile1 analog of the polypentapeptide of elastin. *Biopolymers 25* (10), 1939-1953 (1986)

41. Kim, W., George, A., Evans, M., Conticello, V. P., Cotranslational incorporation of a structurally diverse series of proline analogues in an Escherichia coli expression system. *Chembiochem 5* (7), 928-936 (2004)

42. Kim, W., McMillan, R. A., Snyder, J. P., Conticello, V. P., A stereoelectronic effect on turn formation due to proline substitution in elastin-mimetic polypeptides. *J Am Chem Soc 127* (51), 18121-18132. (2005)

43. Kim, W., Hardcastle, K. I., Conticello, V. P., Fluoroproline flip-flop: regiochemical reversal of a stereoelectronic effect on peptide and protein structures. *Angew Chem Int Ed Engl 45* (48), 8141-8145 (2006)

44. Luan, C. H., Parker, T. M., Gowda, D. C., Urry, D. W., Hydrophobicity of amino acid residues: differential scanning calorimetry and synthesis of the aromatic analogues of the polypentapeptide of elastin. *Biopolymers 32* (9), 1251-1261 (1992)

45. Reguera, J., Lagaron, J. M., Alonso, M., Reboto, V., Calvo, B., Rodriguez-Cabello, J. C., Thermal behavior and kinetic analysis of the chain unfolding and refolding and of the concomitant nonpolar solvation and desolvation of two elastin-like polymers. *Macromolecules 36* (22), 8470-8476 (2003)

46. Nagapudi, K., Brinkman, W. T., Thomas, B. S., Park, J. O., Srinivasarao, M., Wright, E., Conticello, V. P., Chaikof, E. L., Viscoelastic and mechanical behavior of recombinant protein elastomers. *Biomaterials 26* (23), 4695-4706 (2005)

47. Sallach, R. E., Wei, M., Biswas, N., Conticello, V. P., Lecommandoux, S., Dluhy, R. A., Chaikof, E. L., Micelle density regulated by a reversible switch of protein secondary structure. *J Am Chem Soc 128* (36), 12014-12019 (2006)

48. Kim, W., Thevenot, J., Ibarboure, E., Lecommandoux, S., Chaikof, E. L., Self-assembly of thermally responsive amphiphilic diblock copolypeptides into spherical micellar nanoparticles. *Angew Chem Int Ed Engl 49* (25), 4257-4260 (2010)

49. Sallach, R. E., Cui, W., Balderrama, F., Martinez, A. W., Wen, J., Haller, C. A., Taylor, J. V., Wright, E. R., Long, R. C., Jr., Chaikof, E. L., Long-term biostability of self-assembling protein polymers in the absence of covalent crosslinking. *Biomaterials 31* (4), 779-791 (2010)

50. Raucher, D., Chilkoti, A., Enhanced uptake of a thermally responsive polypeptide by tumor cells in response to its hyperthermia-mediated phase transition. *Cancer Res 61* (19), 7163-7170 (2001)

51. Chilkoti, A., Dreher, M. R., Meyer, D. E., Design of thermally responsive, recombinant polypeptide carriers for

targeted drug delivery. *Adv Drug Deliv Rev 54* (8), 1093-1111 (2002)

52. Dreher, M. R., Simnick, A. J., Fischer, K., Smith, R. J., Patel, A., Schmidt, M., Chilkoti, A., Temperature triggered self-assembly of polypeptides into multivalent spherical micelles. *J Am Chem Soc 130* (2), 687-694 (2008)

53. Herrero-Vanrell, R., Rincon, A. C., Alonso, M., Reboto, V., Molina-Martinez, I. T., Rodriguez-Cabello, J. C., Self-assembled particles of an elastin-like polymer as vehicles for controlled drug release. *J Control Release 102* (1), 113-122 (2005)

54. Bessa, P. C., Machado, R., Nurnberger, S., Dopler, D., Banerjee, A., Cunha, A. M., Rodriguez-Cabello, J. C., Redl, H., van Griensven, M., Reis, R. L., Casal, M., Thermoresponsive self-assembled elastin-based nanoparticles for delivery of BMPs. *J Control Release 142* (3), 312-318 (2010)

55. Chen, T. H., Bae, Y., Furgeson, D. Y., Intelligent biosynthetic nanobiomaterials (IBNs) for hyperthermic gene delivery. *Pharm Res 25* (3), 683-91 (2008)

56. Hatefi, A., Amsden, B., Biodegradable injectable *in situ* forming drug delivery systems. *J Control Release 80* (1-3), 9-28 (2002)

57. Adams, S. B., Jr., Shamji, M. F., Nettles, D. L., Hwang, P., Setton, L. A., Sustained release of antibiotics from injectable and thermally responsive polypeptide depots. *J Biomed Mater Res B Appl Biomater 90* (1), 67-74 (2009)

58. Shamji, M. F., Whitlatch, L., Friedman, A. H., Richardson, W. J., Chilkoti, A., Setton, L. A., An injectable and *in situ*-gelling biopolymer for sustained drug release following perineural administration. *Spine (Phila Pa 1976) 33* (7), 748-754 (2008)

59. Hart, D. S., Gehrke, S. H., Thermally associating polypeptides designed for drug delivery produced by genetically engineered cells. *J Pharm Sci 96* (3), 484-516 (2007)

60. Heilshorn, S. C., DiZio, K. A., Welsh, E. R., Tirrell, D. A., Endothelial cell adhesion to the fibronectin CS5 domain in artificial extracellular matrix proteins. *Biomaterials 24* (23), 4245-4252 (2003)

61. Heilshorn, S. C., Liu, J. C., Tirrell, D. A., Cell-binding domain context affects cell behavior on engineered proteins. *Biomacromolecules 6* (1), 318-323 (2005)

62. Martinez-Osorio, H., Juarez-Campo, M., Diebold, Y., Girotti, A., Alonso, M., Arias, F. J., Rodriguez-Cabello, J. C., Garcia-Vazquez, C., Calonge, M., Genetically engineered elastin-like polymer as a substratum to culture cells from the ocular surface. *Curr Eye Res 34* (1), 48-56 (2009)

63. Vollrath, F.; Knight, D. P., Liquid crystalline spinning of spider silk. *Nature 410* (6828), 541-548 (2001)

64. Bini, E., Knight, D. P., Kaplan, D. L., Mapping domain structures in silks from insects and spiders related to protein assembly. *J Mol Biol 335* (1), 27-40 (2004)

65. Gosline, J., Lillie, M., Carrington, E., Guerette, P., Ortlepp, C., Savage, K., Elastic proteins: biological roles and mechanical properties. *Philos Trans R Soc Lond B Biol Sci 357* (1418), 121-132 (2002)

66. Anderson, J. P., Cappello, J., Martin, D. C., Morphology and primary crystal structure of a silk-like protein polymer synthesized by genetically engineered Escherichia coli bacteria. *Biopolymers 34* (8), 1049-1058 (1994)

67. Krejchi, M. T., Atkins, E. D., Waddon, A. J., Fournier, M. J., Mason, T. L., Tirrell, D. A., Chemical sequence control of beta-sheet assembly in macromolecular crystals of periodic polypeptides. *Science 265* (5177), 1427-1432 (1994)

68. Winkler, S., Wilson, D., Kaplan, D. L., Controlling beta-sheet assembly in genetically engineered silk by enzymatic phosphorylation/dephosphorylation. *Biochemistry 39* (41), 12739-12746 (2000)

69. Prince, J. T., McGrath, K. P., DiGirolamo, C. M., Kaplan, D. L., Construction, cloning, and expression of synthetic genes encoding spider dragline silk. *Biochemistry 34* (34), 10879-10885 (1995)

70. Zhou, Y., Wu, S., Conticello, V. P., Genetically directed synthesis and spectroscopic analysis of a protein polymer derived from a flagelliform silk sequence. *Biomacromolecules 2* (1), 111-125 (2001)

71. Bini, E., Foo, C. W., Huang, J., Karageorgiou, V., Kitchel, B., Kaplan, D. L., RGD-functionalized bioengineered spider dragline silk biomaterial. *Biomacromolecules 7* (11), 3139-3145 (2006)

72. Wong Po Foo, C., Patwardhan, S. V., Belton, D. J., Kitchel, B., Anastasiades, D., Huang, J., Naik, R. R., Perry, C. C., Kaplan, D. L., Novel nanocomposites from spider silk-silica fusion (chimeric) proteins. *Proc Natl Acad Sci U S A 103* (25), 9428-9433 (2006)

73. Huang, J., Wong, C., George, A., Kaplan, D. L., The effect of genetically engineered spider silk-dentin matrix protein 1 chimeric protein on hydroxyapatite nucleation. *Biomaterials 28* (14), 2358-2367 (2007)

74. Numata, K., Subramanian, B., Currie, H. A., Kaplan, D. L., Bioengineered silk protein-based gene delivery systems. *Biomaterials 30* (29), 5775-5784 (2009)

75. Numata, K., Hamasaki, J., Subramanian, B., Kaplan, D. L., Gene delivery mediated by recombinant silk proteins containing cationic and cell binding motifs. *J Control Release 146* (1), 136-143 (2010)

76. Weis-Fogh, T. A., A rubber-like protein in insect cuticle. *J Exp Biol* 37, 887-907 (1960)

77. Qin, G. K., Lapidot, S., Numata, K., Hu, X., Meirovitch, S., Dekel, M., Podoler, I., Shoseyov, O., Kaplan, D. L., Expression, cross-linking, and characterization of recombinant chitin binding resilin. *Biomacromolecules* 10 (12), 3227-3234 (2009)

78. Andersen, S. O. The crosslinks in resilin identified as dityrosine and trityrosine. *Biochim Biophys Acta* 93, 213-215 (1964)

79. Elvin, C. M., Carr, A. G., Huson, M. G., Maxwell, J. M., Pearson, R. D., Vuocolo, T., Liyou, N. E., Wong, D. C. C., Merritt, D. J., Dixon, N. E., Synthesis and properties of crosslinked recombinant pro-resilin. *Nature* 437, 999-1002 (2005)

80. Weis-Fogh, T. A., Molecular interpretation of the elasticity of resilin, a rubber-like protein. *J. Mol. Biol.* 3, 648-667 (1961)

81. Ardell, D. H., Andersen, S. O., Tentative identification of a resilin gene in Drosophila melanogaster. *Insect Biochem Mol Biol* 31, 965-970 (2001)

82. Lyons, R. E., Nairn, K. M., Huson, M. G., Kim, M., Dumsday, G., Elvin, C. M., Comparisons of recombinant resilin-like proteins: repetitive domains are sufficient to confer resilin-like properties. *Biomacromolecules* 10, 3009-3014 (2009)

83. Charati, M. B., Ifkovits, J. L., Burdick, J. A., Linhardt, J. G., Kiick, K. L., Hydrophilic elastomeric biomaterials based on resilin-like polypeptides. *Soft Matter* 5, 3412-3416 (2009)

84. Lv, S., Dudek, D. M., Cao, Y., Balamurali, M. M., Gosline, J., Li, H. B., Designed biomaterials to mimic the mechanical properties of muscles. *Nature* 465, 69–73 (2010)

85. Coyne, K. J., Qin, X. X., Waite, J. H., Extensible collagen in mussel byssus: a natural block copolymer. *Science 277* (5333), 1830-1832 (1997)

86. Waite, J. H., Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, Mytilus edulis L. *J Biol Chem 258* (5), 2911-2915 (1983)

87. Waite, J. H., Tanzer, M. L., Polyphenolic Substance of Mytilus edulis: Novel Adhesive Containing L-Dopa and Hydroxyproline. *Science 212* (4498), 1038-1040 (1981)

88. Papov, V. V., Diamond, T. V., Biemann, K., Waite, J. H., Hydroxyarginine-containing polyphenolic proteins in the adhesive plaques of the marine mussel Mytilus edulis. *J Biol Chem 270* (34), 20183-20192 (1995)

89. Waite, J. H., Qin, X., Polyphosphoprotein from the adhesive pads of Mytilus edulis. *Biochemistry 40* (9), 2887-2893 (2001)

90. Deming, T. J., Mussel byssus and biomolecular materials. *Curr Opin Chem Biol 3* (1), 100-105 (1999)

91. Yu, M., Deming, T. J., Synthetic Polypeptide Mimics of Marine Adhesives. *Macromolecules 31* (15), 4739-4745 (1998)

92. Strausberg, R. L., Link, R. P., Protein-based medical adhesives. *Trends Biotechnol 8* (2), 53-57 (1990)

93. Salerno, A. J., Goldberg, I., Cloning, expression, and characterization of a synthetic analog to the bioadhesive precursor protein of the sea mussel Mytilus edulis. *Appl Microbiol Biotechnol 39* (2), 221-226 (1993)

94. Wong, H. H., Kim, Y. C., Lee, S. Y., Chang, H. N., Effect of post-induction nutrient feeding strategies on the production of bioadhesive protein in Escherichia coli. *Biotechnol Bioeng 60* (3), 271-276 (1998)

95. Hwang, D. S., Yoo, H. J., Jun, J. H., Moon, W. K., Cha, H. J., Expression of functional recombinant mussel adhesive protein Mgfp-5 in Escherichia coli. *Appl Environ Microbiol 70* (6), 3352-3359 (2004)

96. Hwang, D. S., Gim, Y., Cha, H. J., Expression of functional recombinant mussel adhesive protein type 3A in Escherichia coli. *Biotechnol Prog 21* (3), 965-970 (2005)

97. Kitamura, M., Kawakami, K., Nakamura, N., Tsumoto, K., Uchiyama, H., Ueda, Y., Kumagai, I., Nakaya, T., Expression of a model peptide of a marine mussel adhesive protein in Escherichia coli and characterization of its structural and functional properties. *J. Polymer Sci*. 37, 729-736 (1999)

98. Hwang, D. S., Gim, Y., Yoo, H. J., Cha, H. J., Practical recombinant hybrid mussel bioadhesive fp-151. *Biomaterials 28* (24), 3560-3568 (2007)

99. Gim, Y., Hwang, D. S., Lim, S., Song, Y. H., Cha, H. J., Production of fusion mussel adhesive fp-353 in Escherichia coli. *Biotechnol Prog 24* (6), 1272-1277 (2008)

100. Hwang, D. S., Sim, S. B., Cha, H. J., Cell adhesion biomaterial based on mussel adhesive protein fused with RGD peptide. *Biomaterials 28* (28), 4039-4046 (2007)

101. Cooper, K. M., Hanlon, R. T., Budelmann, B. U., Physiological color change in squid iridophores. *Cell Tissue Res* 259 (1), 15–24 (1990)

102. Crookes, W. J., Ding, L-L, Huang, Q. L., Kimbell, J. R., Horwitz, J., McFall-Ngai, M. J., Reflectins: the unusual proteins of squid reflective tissues. *Science*, 303(5655), 235–238 (2004)

103. Kramer, R. M., Crookes-Goodson, W. J., Naik, R. R., The self-organizing properties of squid reflectin protein. *Nat Mater* 6 (7), 533–538 (2007)

104. Tao, A. R., DeMartini, D. G., Izumi, M., Sweeney, A. M., Holt, A. L., Morse, D. E., The role of protein assembly in dynamically tunable bio-optical tissues. *Biomaterials* 31, 793-801 (2010)

Abbreviations: RDL: recursive directional ligation, P4H: Prolyl 4-hydroxylase, EMP: elastin-mimetic polymer, HUVEC: human vein endothelial cells, DMP: Dexamethasone phosphate, BMP: bone morphogenetic protein, pLL: poly(L-lysine), MAP: mussel adhesive proteins, fp: foot protein, DOPA: 3,4-dihydroxyphenylalanine

Key Words: Protein Polymer, Genetic Engineering, Biomaterial, Self-Assembly, Biomedical, Review

Send correspondence to: Wookhyun Kim, Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, and the Wyss Institute of Biologically Inspired Engineering of Harvard University, Boston, MA, Tel: 617-735-4645, Fax: 617-735-4650, E-mail: wkim2@bidmc.harvard.edu