CONSTRUCTION OF PROTEIN-BASED HYDROGELS VIA PROTEIN FRAGMENT RECONSTITUTION OF GB1

by

Haoquan Lin

B.Sc. (Hons), University of Waterloo, 2018
B.Eng., Soochow University, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Construction of Protein-based Hydrogels via Protein Fragment Reconstitution of GB1

submitted by Haoquan Lin in partial fulfillment of the requirements for

the degree of Master of Science

in Chemistry

Examining Committee:

Dr. Hongbin Li, Department of Chemistry, UBC
Supervisor

Dr. Russ Algar, Department of Chemistry, UBC
Supervisory Committee Member

Dr. Suzana Straus, Department of Chemistry, UBC
Supervisory Committee Member

Abstract

Tremendous progress has been made in the development of biomaterials as biomimetic extracellular matrices. Protein-based hydrogels are appealing candidates as artificial extracellular matrices due to their high absorption of water, reactive side chains, and tunable physical/chemical properties. Herein, we develop three proteins (G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N) containing FN3 that is critical for cell adhesion. These proteins can self-assemble into protein polymers with high molecular weight via protein fragment reconstitution of a small protein GB1 which can be spontaneously reassembled from its two split fragments G_N and G_C. The resultant polymerized G_C-FN3-G_N and G_C-GB1-FN3-G_N have been successfully used to construct hydrogels through a well-developed photochemical crosslinking approach. The G_C-GB1-FN3-G_N polyprotein hydrogels can be used as extracellular matrices for the cell culture of human lung fibroblasts. These hydrogels exhibit thermo- and redox-responsive features and support cell adhesion with high cell viability in 2D cell culture, which hence demonstrate excellent potential for cell culture. Moreover, the use of protein fragment reconstitution of GB1 allows the rational design of functional biomaterials.

Lay Summary

Hydrogels formed by the crosslinking of polymers are soft matters that contain a high level of water. Protein-based hydrogels have many advantages such as high biodegradability. This thesis provides a general introduction to the preparation of protein hydrogels and their biomedical applications. Also, protein hydrogels constructed by a novel and appealing method are reported in this thesis, which are potential candidates mimicking an environment that allows cell culture.

Preface

This study is based on unpublished work. My supervisor, Dr. Hongbin Li, designed the project. I prepared the sample using molecular biology techniques and did all the experiments except the imaging done by Dr. Yongliang Wang. Also, I analyzed the data and wrote this thesis.

Table of Contents

List of Tables

List of Figures

Glossary

2D	two-dimensional
3′	3′-end in DNA sequence
3D	three-dimensional
5′	5′-end in DNA sequence
AA	SpyTag-ELP-RGD-ELP-SpyTag
AAA	SpyTag-ELP-SpyTag-ELP-SpyTag
Abs_{280nm}	absorbance at 280 nm
APS	ammonium persulfate
bRGD-CUBE	ELP-D_{88}-RGD-CL
BSA	bovine serum albumin
CD	circular dichroism
CL	Coil-LL peptide
CnaB2	the second immunoglobulin-like collagen adhesin domain
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECMs	extracellular matrices
ELPs	elastin-like polypeptides
FbaB	*Streptococcus pyogenes* fibronectin-binding protein
FN3	the third fibronectin type III domain of tenascin-C
FPLC	fast protein liquid chromatography

G′	storage modulus
G″	loss modulus
GB1	the B1 binding domain of protein G from *Streptococcus*
GB1-L5/GL5	a mutant of GB1 with -GGGLG- inserted into the second loop of GB1
G_C	C-terminal fragment of GB1
GFP	green fluorescent protein
GL5−CC	bi-cysteine mutant of GL5
G_N	N-terminal fragment of GB1
HLFs	human lung fibroblasts
I27	the 27th Ig domain of human titin
I27w34f/I27F	a mutant of I27 with a Trp34Phe mutation
IPTG	isopropyl-1-β-D-thiogalactoside
K_d	dissociation constant
kDa	kilodalton
kPa	kilopascal
LB	Luria-Bertani broth
LCST	lower critical solution temperature
LZ	Leucine zipper
m_0	fresh-made sample mass
MEP	mutually exclusive protein
Mn	number average molecular weight
m_s	swollen sample mass
Mw	weight average molecular weight

n	number of residues
N_i	number of moles of polymer that have a molecular weight M_i
O-CUBE	ELP-D_{88}-CL
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDI	polydispersity index
r	swelling ratio
RGD	Arg-Gly-Asp
RGDS	Arg-Gly-Asp-Ser
RNA	ribonucleic acid
rpm	revolutions per minute
$Ru(II)(bpy)_3^{2+}$	ruthenium (II) tris-bipyridyl dication
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
T_m	melting temperature
UCST	upper critical solution temperature
V_e	elution volume
w/v%	weight by volume percent

Acknowledgements

I would like to sincerely thank my supervisor, Dr. Hongbin Li, for offering me the opportunity to explore protein-based hydrogels and providing me guidance and inspiration during my graduate study. Without his help, I would not be able to finish this program.

I would like to thank Dr. Russ Algar and Dr. Suzana Straus for their thoughtful comments and reviews.

I thank all the members of Dr. Li's research group for their help and encouragement. Dr. Ruidi Wang taught me molecular biology experiments and helped me initiate the first research project. Dr. Linglan Fu and Dr. Yongliang Wang taught me the techniques for cell culture. Dr. Jiahao Xia, Dr. Han Wang, Tianyu Duan, Jiayu Li, Adam Xiao, Jiacheng Zuo, Qingyuan Bian, Jess Fung, and Guojun Chen provide help and suggestions.

I would also thank Dr. Elena Polishchuk and Jessie Chen for their technical support in Bioservice facilities. I thank Dr. Katherine Ryan and Dr. Russ Algar for their generous help in FPLC measurement and fluorescence microscopic imaging.

I want to thank my family for their selfless love and support. Also, I would like to thank my girlfriend Wing Sum Tam for always caring, supporting, and encouraging me. Last but not least, I would like to thank my friends Tian Zhang, Jiahong Hu, and Jun Dai for their support during the pandemic.

Dedication

This thesis is dedicated to my parents, my family, and my love.

Chapter 1: Introduction

1.1 Introduction of hydrogels

Hydrogels are three-dimensional networks formed by hydrophilic polymer chains with cross-linkable functional groups, which can absorb and retain high water content. Hydrogels possess a large number of characteristics including porous structure, broadly tunable mechanical properties, and high water absorption.[1-4] Remarkably, some hydrogels can undergo a sol-gel phase transition when responding to external stimuli such as a change in temperature, magnetic fields, light intensity, pH, and ionic strength.[5-12] Sol-gel transition is a change from a liquid state to a gel state. Such transitions are mostly reversible, meaning that after the removal of stimuli the hydrogel can return to its initial state. The first hydrogel reported by Wichterle and Lim in 1960 was formed by the crosslinking of 2-hydroxyalkyl methacrylate, which showed promising results in developing contact lenses.[13] Since this revolutionary work, more and more studies about the application of hydrogels have been published including masks[14, 15], sensors[16-18], extracellular matrix materials[19-25], and drug delivery[26-32].

Hydrogels can be classified based on a myriad of properties. One example is to categorize them as synthetic and natural hydrogels based on their source of origins.[33] Synthetic hydrogels such as poly(vinyl alcohol) and poly(acrylamide) have exceptional mechanical strength; however, some synthetic hydrogels with poor biodegradability might become a threat to the environment.[34-36] In comparison to synthetic hydrogels, natural hydrogels possess inherently environmentally friendly and biodegradable characteristics but low mechanical strength.[34, 37] In addition to the classification

based on the origins, hydrogels can be either physically crosslinked or chemically crosslinked, depending on the nature of crosslinking. Hydrogels can be divided into three categories: neutral, ionic, and zwitterionic.[38] Also, hydrogels can be classified by the composition of their building blocks, such as protein-based hydrogels, DNA/RNA-based hydrogels, polysaccharides-based hydrogels, etc.

1.2 Hydrogels based on recombinant proteins

Proteins have many intrinsic advantages over other materials for hydrogel construction. With many functional groups that can serve as the reactive sites, including -NH2, -OH, -COOH, and -SH, proteins can be crosslinked or post-translationally modified. Proteins are environmentally friendly and biodegradable. Some proteins and peptides, such as proteins containing Arg-Gly-Asp (RGD) and Arg-Gly-Asp-Ser (RGDS) motifs, are frequently integrated into hydrogels to create biomaterials for tissue engineering as they are natural parts of the extracellular matrix.[39] The use of recombinant DNA technologies facilitates protein production to achieve distinct features, i.e., controllable amino acid sequence and chain length, as well as designed physicochemical properties.[40] Recently, a temperature-responsive hydrogel for 3D angiogenesis has been developed from genetically engineered proteins, such as coiled-coil unit bound elastin-like polypeptides. It was designed to exhibit the capacity of undergoing controllable sol-gel transitions, high transparency, adjustable mechanical and bio-functional properties, and growth factor-delivering activity.[41] Due to their potential biomedical applications, engineered protein-based hydrogels with desirable properties have attracted intense interest for years.[42-47] In the following subsections, I

will discuss the crosslinking methods that are needed to construct protein-based hydrogels and the applications of protein-based hydrogels in the field of biomedicine.

1.2.1 Crosslinking methods of protein-based hydrogels

Protein hydrogels are typically constructed via either physical or chemical crosslinking of proteins dispersed in aqueous solution. Physically crosslinked protein hydrogels are transiently held by either polymer chain entanglements or non-covalent physical interactions, including ionic interactions, hydrophobic interactions, van der Waals forces, and hydrogen bonding. In contrast, chemically crosslinked protein networks are stabilized by covalent bonding that permanently strengthens mechanical integrity. By using physical crosslinking approaches, hydrogelation can be easily achieved and even reversed if needed. Nevertheless, the physical methods limit the possibility of the hydrogel being fine-tuned, as they are mostly dependent on the intrinsic characteristics of the protein. Conversely, chemical crosslinking methods are more controllable and precise but require altering the protein sidechains.[1, 48] Generally, chemically crosslinked hydrogels are stronger than physically crosslinked ones.

1.2.2 Physical crosslinking methods of protein-based hydrogels

1.2.2.1 Thermally induced entanglement of proteins

Proteins can form thermally driven hydrogels, during which physical entanglement arises in response to a change in temperature. This temperature change alters the solubility of the proteins and results in packed and rigid protein backbones.[49, 50] Increasing or decreasing temperature may

lead to hydrogelation, and the transition temperatures are denoted as upper critical solution temperature (UCST) and lower critical solution temperature (LCST), respectively.[51, 52] The mechanism of thermally induced hydrogelation differs depending on the type of protein. Proteins exhibiting UCST behavior form hydrogels when the temperature drops below their respective UCSTs, whereas those exhibiting LCST transition gel above their UCSTs.

Gelatin, a mixture of proteins derived from hydrolysis of collagen, has an UCST of about 30 – 35 °C.[53-58] With sufficient concentration, it gels through physical entanglement at a temperature below the UCST, during which its conformation changes from a random coil to a triple helix. This gelation process is easily reversed through the dissociation of intermolecular hydrogen bonds at a temperature higher than 30 – 35 °C. Physical gelatin hydrogels possess low stability and poor mechanical properties. As they are unstable at the physiological temperature of 37 °C, their biomedical applications are limited.[59] Covalently crosslinking a gelatin hydrogel with small molecules such as formaldehyde[60] and glutaraldehyde[61] can strengthen its stability and mechanical properties which results from the formation of stable amide bonds between amino groups and carboxyl groups. Also, the introduction of other polymers such as oxidized cellulose nanowhiskers[62] and polyethylene glycol[63] can further broaden the applicability of gelatin. In a recent study, Laronda et al. developed a microporous hydrogel scaffold by 3D printing thermally tunable and partially crosslinked gelatin (Figure 1.1a), which demonstrated high potential as a bioprosthetic ovary. It underwent a sol-gel transition at about 33 °C, the temperature where the storage modulus (G′) equaled to loss modulus (G″) (Figure 1.1b). At a temperature lower than 33 °C, the formation of triple helices in gelatin led to the crosslinking of the polypeptides, resulting in gelatin hydrogel (Figure 1.1c). The scaffold was found to support the adhesion and growth of

ovarian murine follicles (Figure 1.1d & e). Furthermore, it promoted hormone production, oocyte maturation, and ovulation *in vitro*.[64]

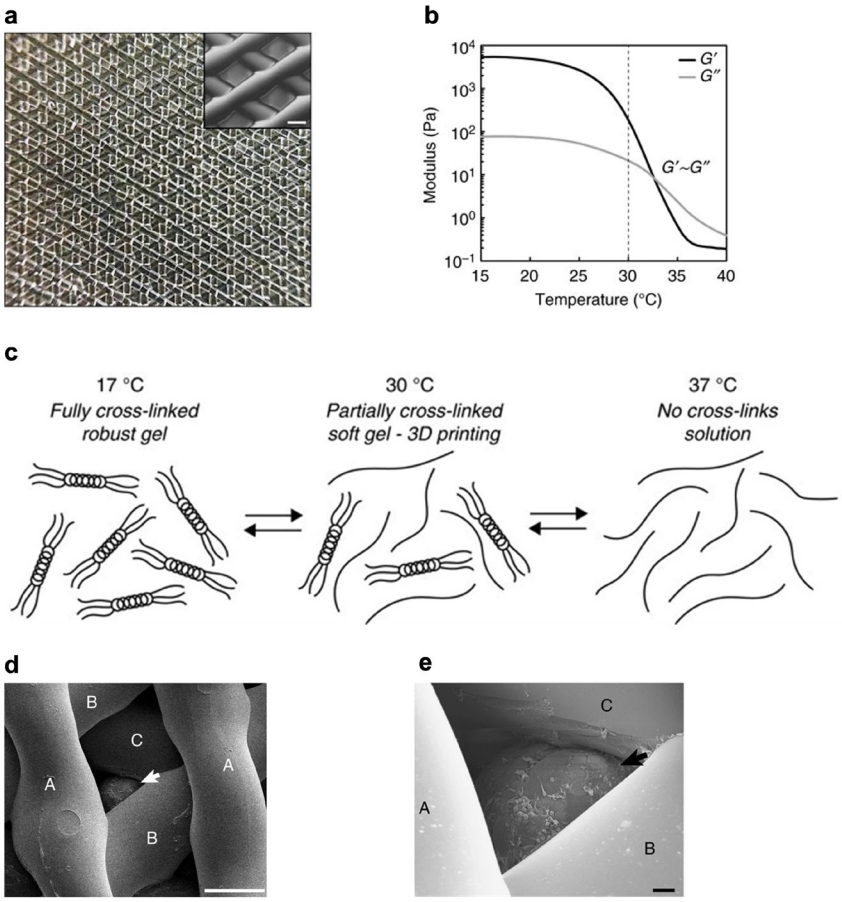

Figure 1.1. A 3D scaffold printed from partially crosslinked gelatin and designed as a bioprosthetic ovary. (a) Photography of a 3D gelatin scaffold with five layers and 100 μm nozzle. Insert: magnified scaffolds with scale bar 250 μm. (b) Rheology of 10% gelatin solution at different temperatures. The transition temperature is at about 33 °C, where $G' = G''$. (c) Thermo-reversible gelation of gelatin. Above 33 °C, gelatin is soluble, and polypeptide chains are separate. Physical gelation occurs spontaneously at a temperature below 33 °C, as triple helices start to form. At 17 °C, the gel is fully crosslinked. (d, e) SEM images of the follicle (arrows) underneath three layers (A, B & C) of 60° scaffold struts after two-day culture. Scale bars: (d) 100 μm; (e) 10 μm.[64] Copyright 2017, Springer Nature.

A notable example of peptides bearing LCST behavior is elastin-like polypeptides (ELPs) consisting of a repeating VPGXG sequence, where the guest residue X can be any amino acid except proline. Below its LCST, soluble ELPs remain as disordered random coils. Above its LCST, ELPs chains assemble into a β-spiral structure, which further aggregate through hydrophobic

interactions (Figure 1.2a). The LCST of ELPs varies depending on the molecular weight of the polymer, the concentration of ELPs in the solution, and the ionic strength of the solution.[65] Due to the cytocompatibility and precisely tunable LCST, ELPs have been widely used to construct hydrogels for tissue engineering. Figure 1.2b shows 4-armed coiled-coil unit bound ELPs, designed by Mizuguchi and his colleagues, which can assemble to form a hydrogel at a temperature above the LCST of the protein polymer. The hydrogel containing RGD peptide promoted the adhesion of human umbilical vein endothelial cells (Figure 1.2c). Moreover, this hydrogel achieved precise control of cellular functions by the incorporation of growth factors.[41]

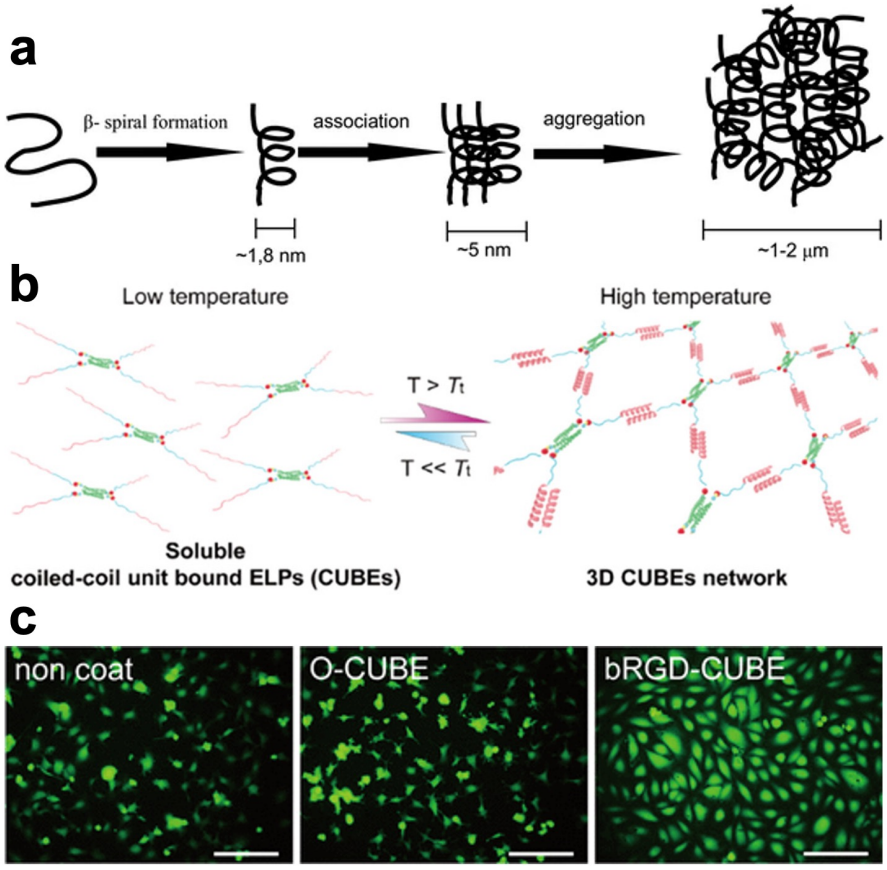

Figure 1.2. ELPs aggregation and their applications in tissue engineering (a) Mechanism of ELPs aggregation.[65] Copyright 2014, Springer Nature. (b) Schematic of temperature-responsive 3D hydrogel formed by 4-armed coiled-coil unit bound ELPs. T_t is the inverse phase-transition temperature. The green coils represent the coiled coil, and the pink coils are ELPs. (c) Images of cell-hydrogel matrices by fluorescent microscopy. Cells were seeded onto a surface coated with ELP-D_{88}-CL (O-CUBE) and ELP-D_{88}-RGD-CL (bRGD-CUBE) and stained with Calcein-AM. Scale bar: 200 μm.[41] Copyright 2020, American Chemical Society.

1.2.2.2 Non-covalent self-assembly of proteins

Self-assembly is a ubiquitous process through which individual constituents spontaneously erect a highly ordered entity, caused by internal specific interactions among the components themselves. Non-covalent molecular self-assembly is a powerful approach for the fabrication of protein-based hydrogels, directed through weak non-covalent bonding that favors the folding of peptide chains into well-organized structures with functionality.[1, 66] Biorecognition-driven self-assembly has inspired scientists to develop self-assembling protein hydrogels[40], and among them, Tirrell's research group are considered as the pioneers, who have substantially contributed towards successful engineering of protein-based hydrogels.

The association of coiled-coils motifs in biorecognition has been thoroughly studied for the development of protein-based hydrogels. The coiled-coil, a protein folding pattern, contains two or more α-helices that self-assemble into a superhelix by twisting around each other. The most common type of coiled coil is parallel, dimeric and left-handed. A regular α-helix has approximately 3.6 amino acid residues per turn, whereas this value is lowered to 3.5 for each helix within a left-handed coiled coil because of the imposed distortion.[67-69] The sequence of coiled coil, therefore, presents a periodicity of 7, which can be modeled as $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$ in one helix and $(a'\text{-}b'\text{-}c'\text{-}d'\text{-}e'\text{-}f'\text{-}g')_n$ in the other (Figure 1.3). Within this heptad repeat sequence, b, c, and f are typically hydrophilic so that they are exposed to the outside environment, while a and d are nonpolar core residues at the hydrophobic interface between two α-helices. The remaining two residues e and g are ionic and exposed to solvent, which can participate in interhelical electrostatic

interactions.[69] Self-assembly of two or more α-helices has been reported[70], including a seven-helix coiled coil[71].

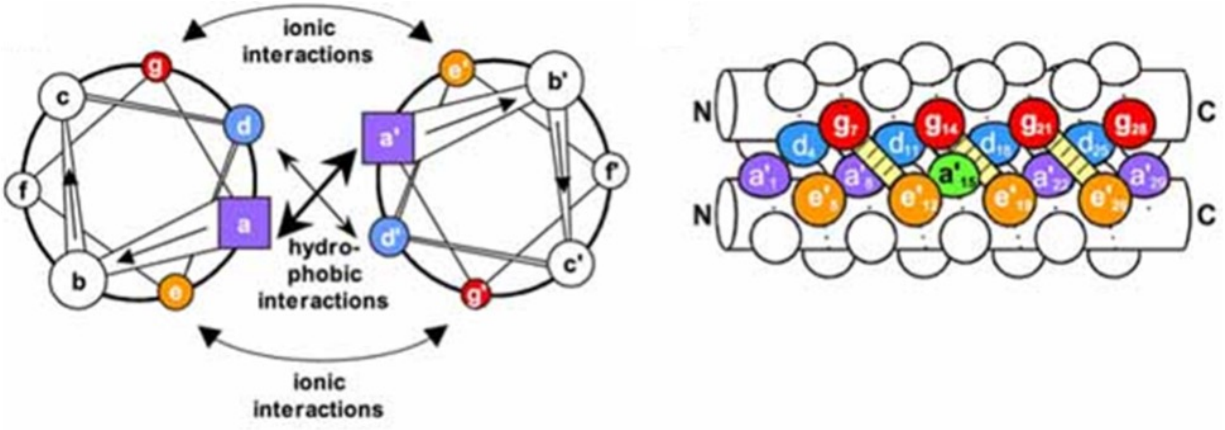

Figure 1.3. Top (left) and side (right) diagrams of a parallel two-stranded coiled coil. The coiled coil is formed from the heptad repeat sequence a-b-c-d-e-f-g. Position a and a' are analogous positions of the helices.[69] Copyright 2007, Humana Press.

Many researchers have utilized this fascinating self-assembling function of the coiled coil motif to construct protein hydrogels. Huang and his co-workers have created a durable leucine zipper based hydrogel with tunable properties for use in tissue engineering. Originally discovered in the amino acid sequences of several DNA binding proteins, leucine zipper domains are dimetric coiled coils, in which a and d of the heptad repeat are often occupied by leucine or other nonpolar residues. Figure 1.4a demonstrates the designed peptide sequences, which can form a robust hydrogel network through self-association of the leucine zipper and the formation of disulfide bonds. This hydrogel could keep human marrow stem cells viable and support their adhesion. Although the existence of the RGD motif was beneficial for initial cell attachment, the cell viability was not

significantly affected over 7 days (Figure 1.4b). These results revealed that the leucine zipper based hydrogels could serve as potent artificial extracellular matrices for tissue engineering.[72]

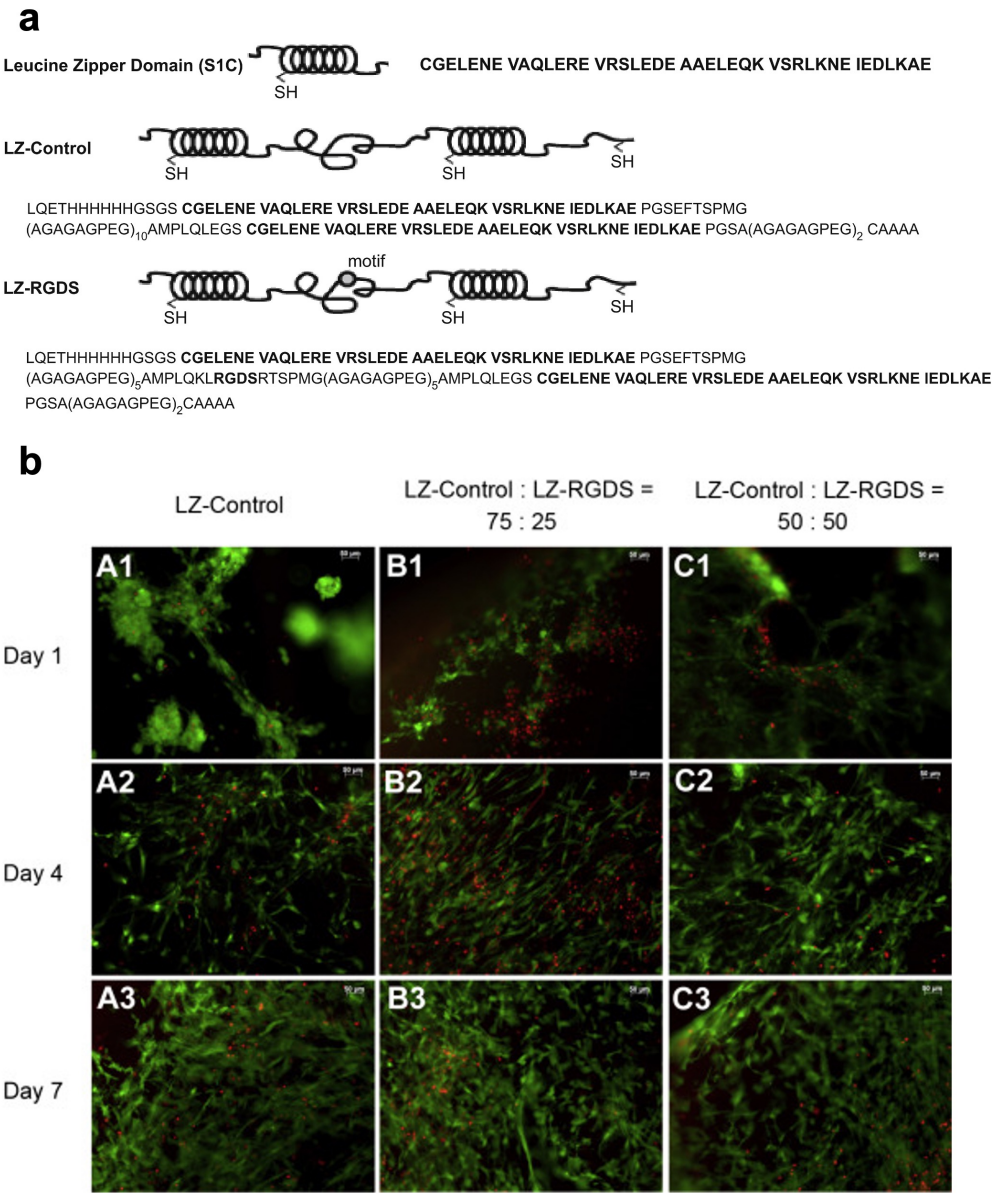

Figure 1.4. Leucine zipper (LZ) based hydrogel for tissue engineering. (a) Schematic of the peptide sequences. (b) Live-Dead assay of human marrow stem cells seeded on 7% w/v hydrogels with different LZ/LZ-RGD ratios. Live cells are stained in green, whereas dead cells are in red.[72] Copyright 2014, Elsevier Ltd.

Another prevalent self-assembling motif is β-sheet secondary structure formed by adjacent β-strands and stabilized by hydrogen bonds of which there are two kinds: parallel (Figure 1.5a) or

anti-parallel (Figure 1.5b). The conformational shift of the 42-amino-acid β-amyloid from an α-helix or random coil to a β-sheet structure leads to the self-assembly of insoluble amyloids fibrils (i.e., abnormal protein aggregation).[73]

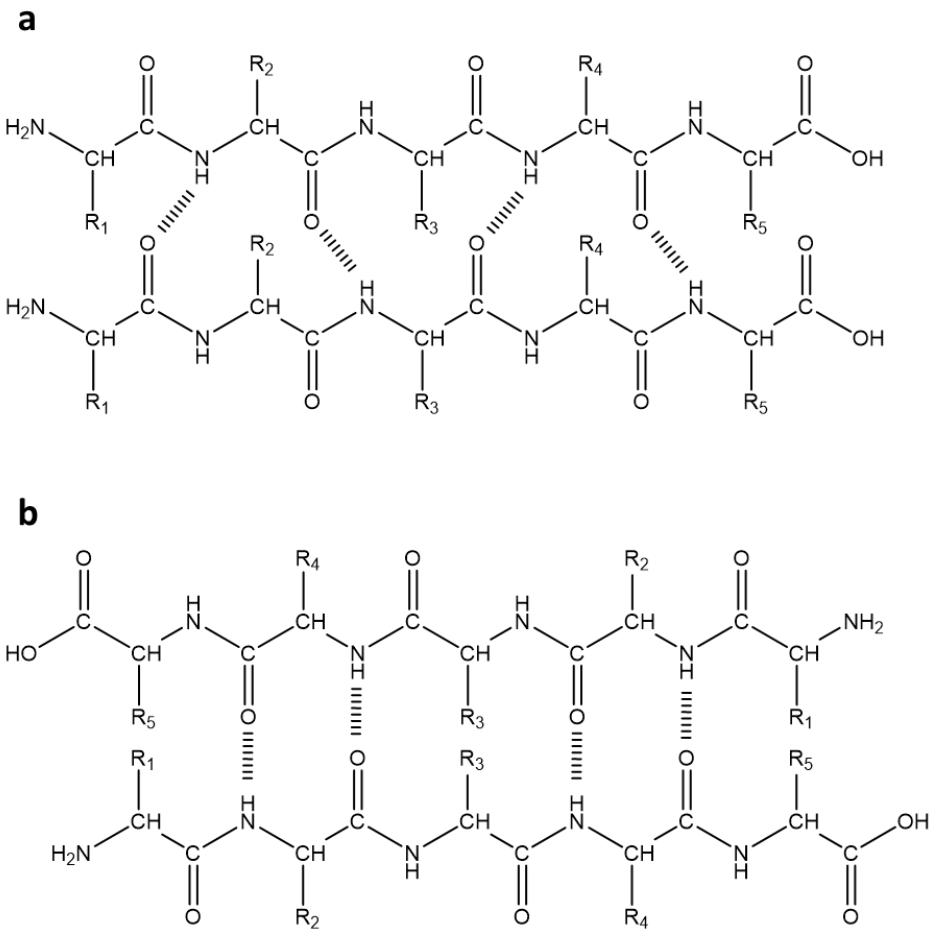

Figure 1.5. β-sheet structure of two adjacent peptide chains linked by hydrogel bonds. (a) Parallel β-sheet structure. (b) Anti-parallel β-sheet structure.

A robust and pH-sensitive hydrogel was formed by the self-assembly of three pentapeptides, as shown by Clarke, Parmenter, and Scherman (Figure 1.6). Also, its stiffness could be easily tuned. Remarkably, this hydrogel could heal itself via the re-assembly of the β-sheet structure. These

promising properties might allow the use of this hydrogel for tissue engineering and injectable delivery in the near future.[74]

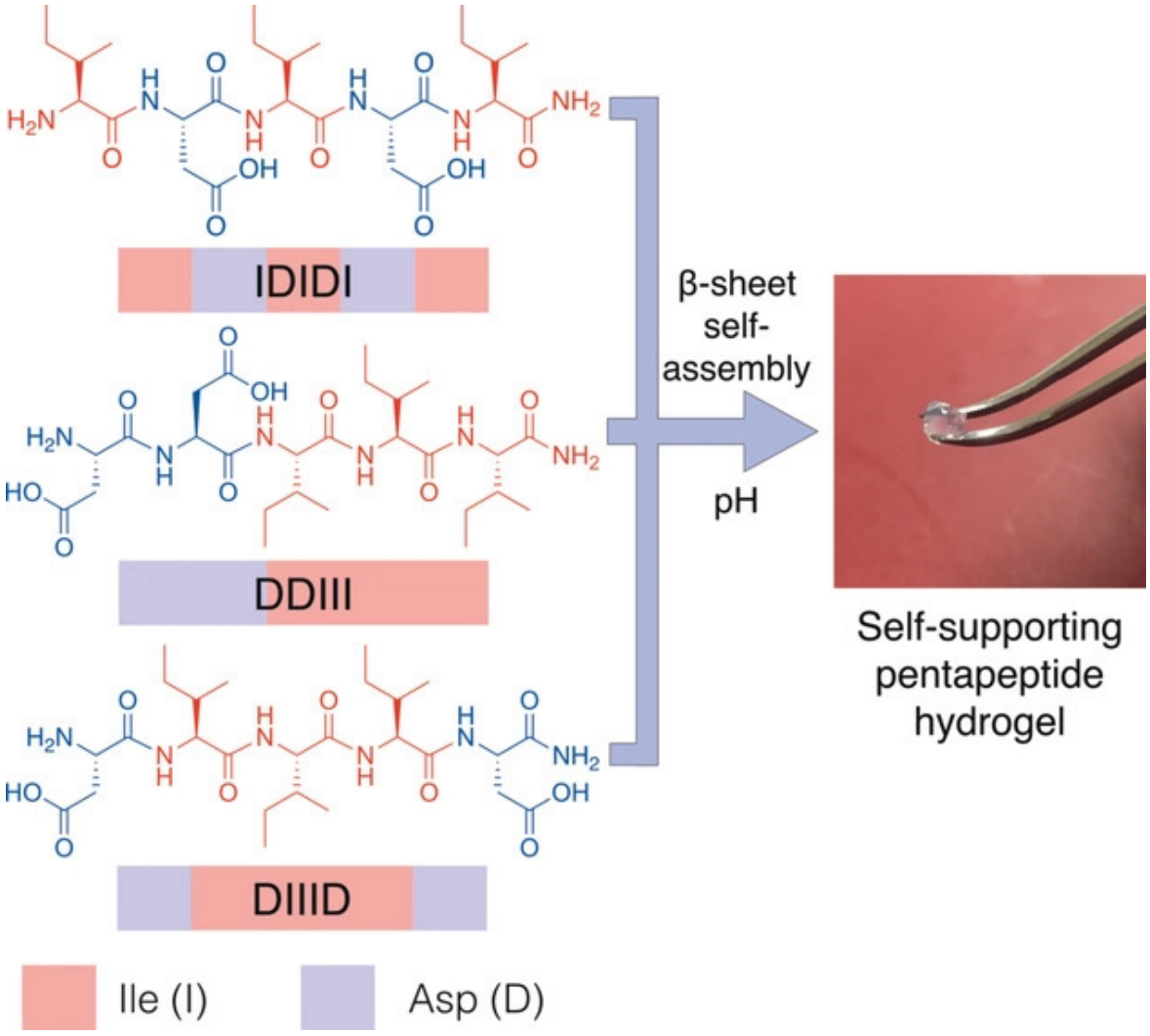

Figure 1.6. Hydrogel formed by three pentapeptides. The hydrogel is formed through the self-assembly of β-sheet structure.[74] Copyright 2018, John Wiley & Sons, Inc.

Many proteins can be split into fragments, and these fragments can spontaneously reassemble into a functional protein without forming covalent bonds. An example is split green fluorescent protein (GFP), the most frequently used genetically engineered fluorescent reporter. GFP can be split into

two polypeptides, GFP_{1-10} (residues 1-214; the detector) and GFP_{11} (residues 215-230; the tag). These two fragments are non-fluorescent on their own; however, they form a fluorescent GFP after spontaneous association (Figure 1.7a).[75-77] The split GFP technique can also be used to construct protein hydrogels. Sun and his co-workers developed a 4-arm star-like protein, $(SpyCatcher)_4GFP$, based on *in situ* split GFP reconstitution (Figure 1.7b). This 4-arm protein could form hydrogel networks with either the bifunctional protein, SpyTag-ELP-RGD-ELP-SpyTag (AA), or the trifunctional protein, SpyTag-ELP-SpyTag-ELP-SpyTag (AAA) through SpyTag-SpyCatcher chemistry (Figure 1.7c&d). This work has enabled the possibility of constructing protein hydrogels by directly assembling proteins with unusual frameworks.[77]

Another example of split proteins is GB1, which can be split into two fragments that can spontaneously reconstitute folded GB1. This example will be introduced in section 1.3.1.

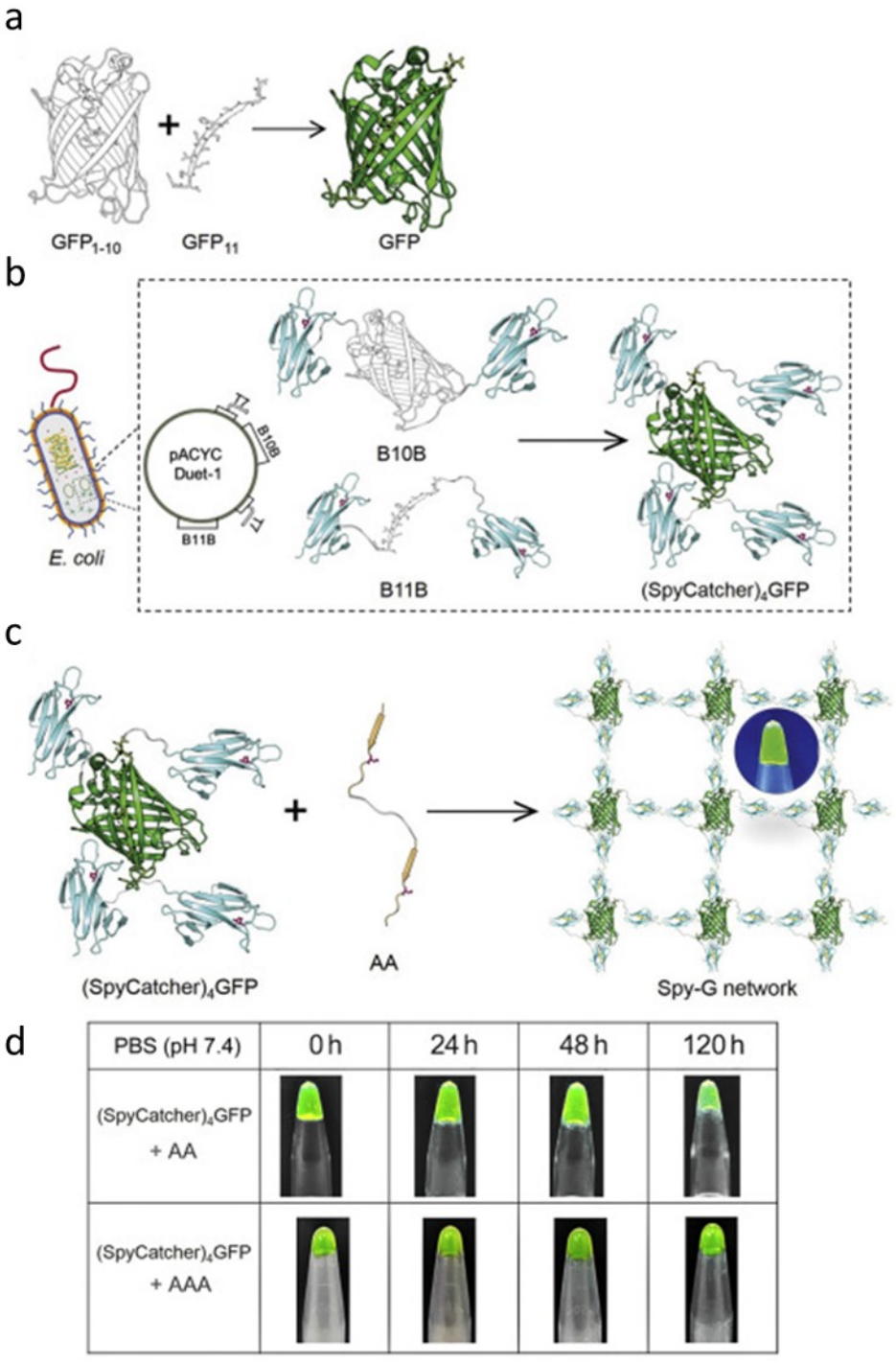

Figure 1.7. Synthesis of protein hydrogels by using the 4-arm star-like protein (SpyCatcher)$_4$GFP. (a) The formation of an intact GFP via the association of GFP$_{1-10}$ and GFP$_{11}$. (b) Schematic showing cellular synthesis of the 4-arm protein. The co-expression plasmid, pACYC-Duet-1, harbors two constructs, SpyCatcher-ELP-GFP$_{1-10}$-ELP-SpyCatcher (B10B) and SpyCatcher-ELP-GFP$_{11}$-ELP-SpyCatcher (B11B). (c) Schematic of the assembly of a protein hydrogel network via SpyTag/SpyCatcher chemistry. (d) Images of hydrogel networks consisting of (SpyCatcher)$_4$GFP + AA and (SpyCatcher)$_4$GFP + AAA in PBS (pH 7.4).[77] Copyright 2019, Elsevier Inc.

1.2.3 Chemical crosslinking methods of protein-based hydrogels

Chemically crosslinked protein hydrogels are achieved through the establishment of covalent bonds under appropriate circumstances, which mainly involve the sidechain of the protein residues. Compared to physical protein hydrogels, chemical ones are considered permanent, possessing more robust mechanical properties and stability. A variety of chemical crosslinking approaches that can effectively form protein-based hydrogels have been reported, some of which will be highlighted in the following subsections.

1.2.3.1 Chemical crosslinking via the formation of disulfides bonds

In biology, the formation of a disulfide bond between the thiol groups of two cysteine residues can stabilize protein structure. Inspired by this phenomenon, scientists have utilized the disulfide bond crosslinks for hydrogelation. For example, Zhang et al. constructed a soft bovine serum albumin (BSA) hydrogel crosslinked by disulfide bonds (Figure 1.8a), which can potentially be utilized in tissue engineering. BSA contains 583 amino acids, of which 35 are cysteine residues. Among these 35 cysteine residues, 34 are oxidized to form 17 disulfide bonds, leaving one free thiol group. As the disulfide bonds are reduced to free thiol groups, BSA unfolds, allowing the establishment of a hydrogel network through the recombination of disulfide bonds (Figure 1.8b). Surprisingly, this hydrogel can quickly self-repair a cut in the presence of H_2O_2, which significantly accelerates the process of recreating the disulfide bonds. Based on this characteristic, the hydrogel shows great injectability which could be made to different shapes. This study also illustrates the hydrogel

caused minimal death of human breast cancer MCF-7 cells that were seeded onto the hydrogel matrix.[78]

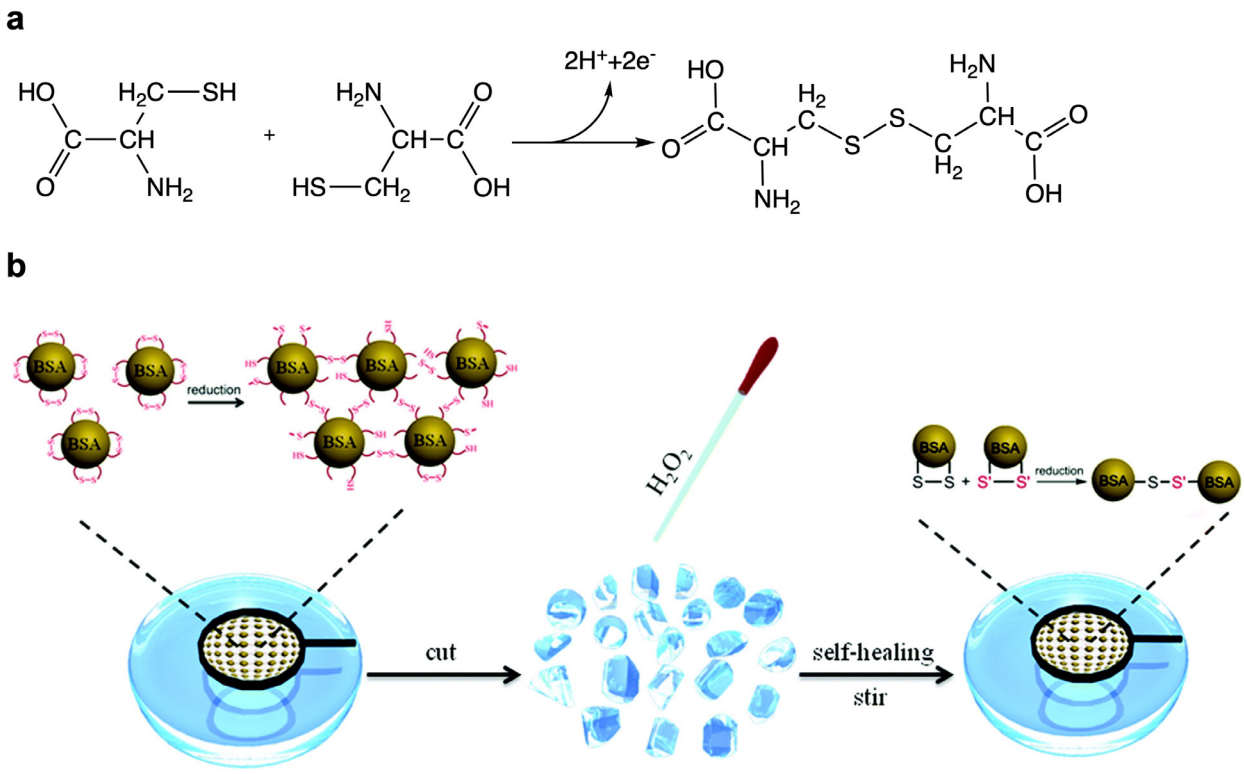

Figure 1.8. The formation of a disulfide bond and a BSA hydrogel. (a) A disulfide bond formed between two cysteine residues. (b) Schematic representing the formation of the hydrogel. Reduced BSA can form a hydrogel network via the formation of disulfides bonds. The hydrogel exhibits a self-healing property under H_2O_2 stimulation.[78] Copyright 2019, The Royal Society of Chemistry.

1.2.3.2 Ru(II)(bpy)$_3^{2+}$-mediated photochemical crosslinking

Under visible light illumination, two tyrosine residues in close proximity can be crosslinked to a dityrosine adduct in the presence of the catalyst tris(2, 2'-bipyridine)ruthenium(II) ion and the electron acceptor ammonium persulfate, which crosslinks the protein chains of interest (Figure

1.9). This reaction was designed by Fancy and Kodadek for the analysis of protein-protein interactions.[79] Subsequently, Elvin and his collaborators successfully adopted this crosslinking method to construct a rubber-like hydrogel based on a recombinant pro-resilin (or resilin-like protein) comprising 17 copies of a 15 amino acid sequence, GGRPSDSYGAPGGGN, which is recognized as an elastic repeat motif.[80] Resilin is an elastomeric protein found in cuticle regions of most insects, acting as soft rubber. This revolutionary work provided a *de novo* way to engineer protein-based hydrogels.

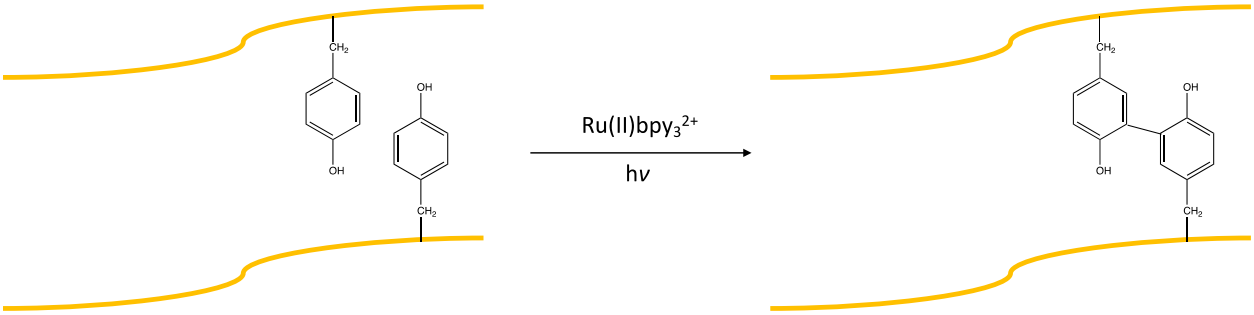

Figure 1.9. Dityrosine crosslinking catalyzed by ruthenium complex. With the presence of ruthenium complex and ammonium persulfate (APS), two adjacent tyrosine residues form a dityrosine adduct under light illumination.

By utilizing the adhesive property of mussel protein as well as the Ru(II)$(bpy)_3^{2+}$-mediated photochemical crosslinking strategy, Jeon et al. devised a rapidly light-activated surgical protein glue and applied it for *in vivo* rat skin incision wound closure (Figure 1.10). The results of their study elucidated that this bio-adhesive glue not only could promptly close a bleeding and open wound on the back of a rat via the strong adhesion to the wound but also could effectively facilitate tissue regeneration with minimal inflammation. Therefore, this hydrogel provided a medical application for structureless wound closures.[81]

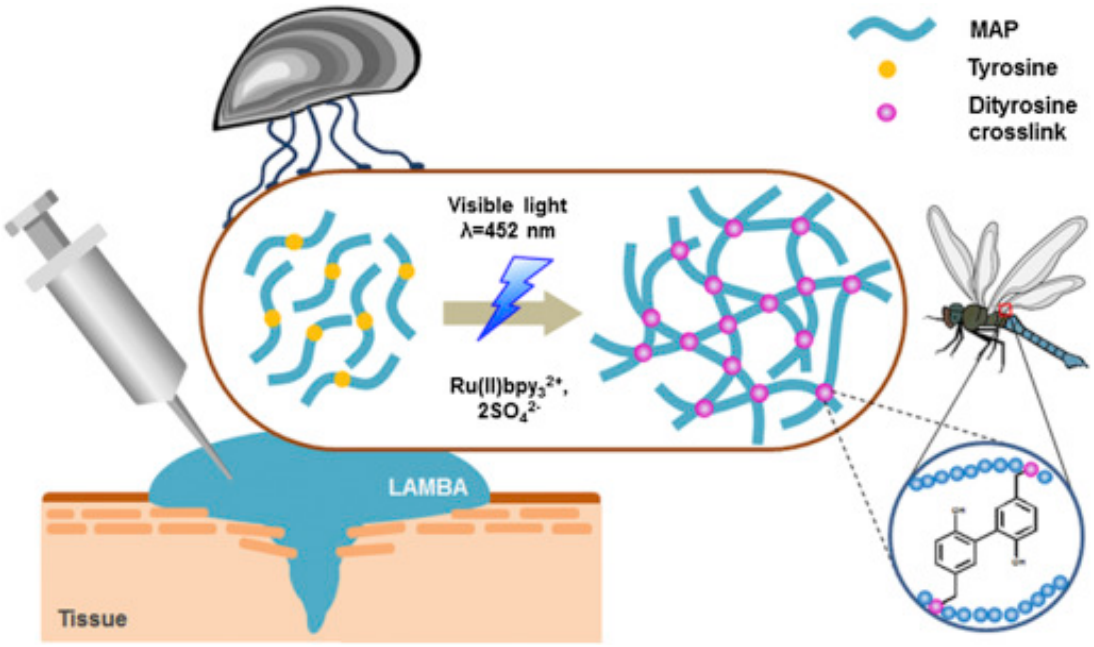

Figure 1.10. Schematic representation of a light-activated mussel protein-based bioadhesive hydrogel. The crosslinking via dityrosine bonds is light-inducible in the presence of $Ru(II)(bpy)_3^{2+}$ and SO_4^{2-}.[81] Copyright 2015, Elsevier Ltd.

Li and his colleagues reported a general and robust approach to functionalize protein-based hydrogels by utilizing $Ru(II)(bpy)_3^{2+}$-mediated photochemical crosslinking method and SpyTag-SpyCatcher conjugation, in which hydrogels were photochemically crosslinked and functionalized via the binding with protein ligand (Figure 1.11).[82]

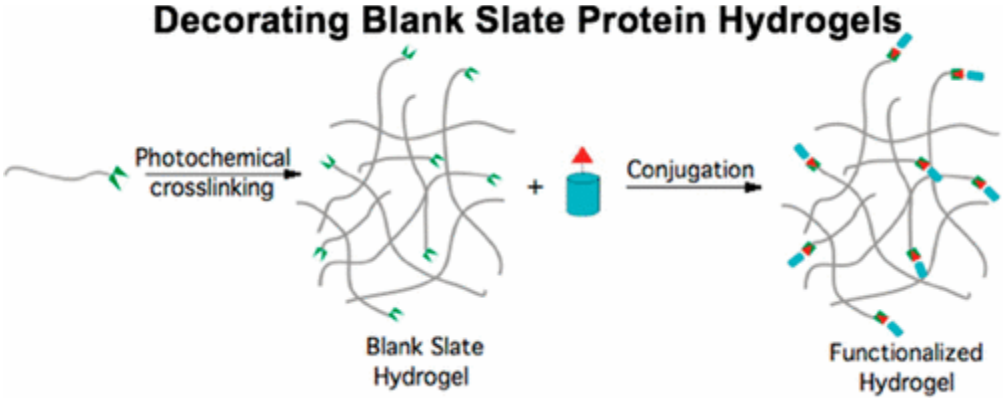

Figure 1.11. Schematics of hydrogel formation and functionalization. The blank state hydrogel was formed by $Ru(II)(bpy)_3^{2+}$-mediated photochemical crosslinking and functionalized via SpyTag-SpyCatcher chemistry.[82] Copyright 2017, American Chemical Society.

1.2.3.3 Chemical crosslinking via SpyTag-SpyCatcher conjugation

The SpyTag-SpyCatcher system was designed by Zakeri and his co-workers for protein ligation in 2012, in which SpyTag can spontaneously recognize its protein partner SpyCatcher reconstituting a complete second immunoglobulin-like collagen adhesin domain (CnaB2) with the formation of an irreversible covalent isopeptide bond between the side chains of aspartic acid in SpyTag and lysine in SpyCatcher in minutes (Figure 1.12). SpyTag (13 amino acids) and SpyCatcher (138 amino acids, 15 kDa) fragments were split from CnaB2 of *Streptococcus pyogenes* fibronectin-binding protein FbaB and rationally engineered.[83] The exceptionally robust and irreversible SpyTag-SpyCatcher linkage provides a practical and reliable module for constructing novel protein architectures.

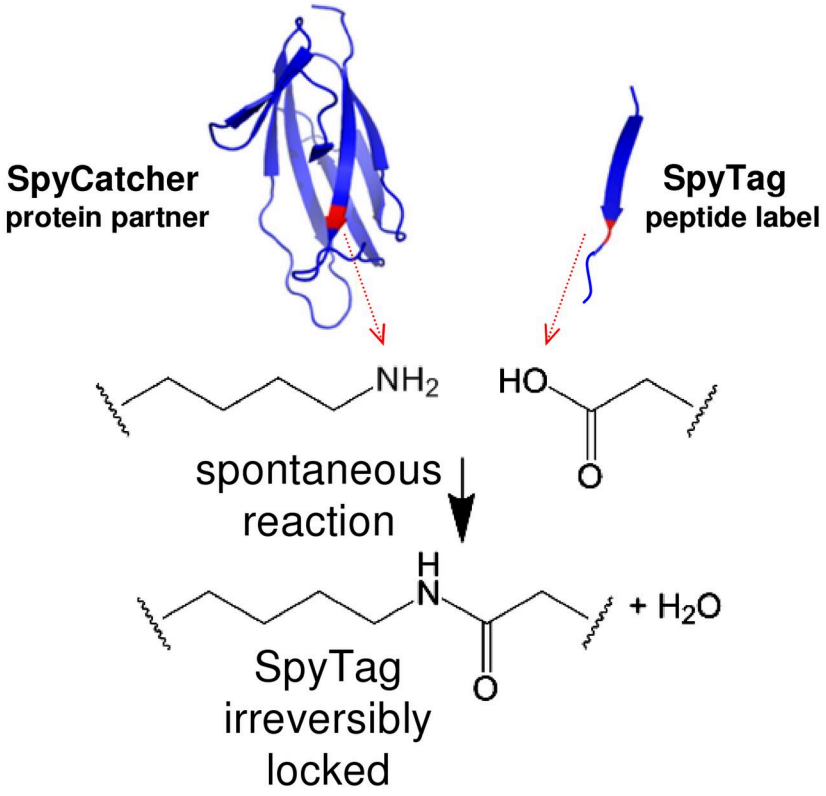

Figure 1.12. Spontaneous intermolecular amide bond formed between SpyTag and SpyCatcher. The reaction is irreversible. Lys and Asp are in positions 31 and 117 in the CnaB2, respectively.[84] Copyright 2012, National Academy of Sciences.

Recently, Tirrell and co-workers have employed the SpyCatcher-SpyTag approach to engineer elastin-like protein hydrogels that support 3D stem cell culture. This pioneering work has inspired many scientists to construct protein-based biomaterials in the same way.[85] Later, Gao et al. successfully constructed a soft protein hydrogel from engineered tandem modular elastomeric proteins by using SpyCatcher-SpyTag chemistry (Figure 1.13). The hydrogel showed extraordinary biocompatibility in the encapsulation and culture of human lung fibroblasts (HLFs). Also, it could be used as a controlled drug delivery vehicle.[86]

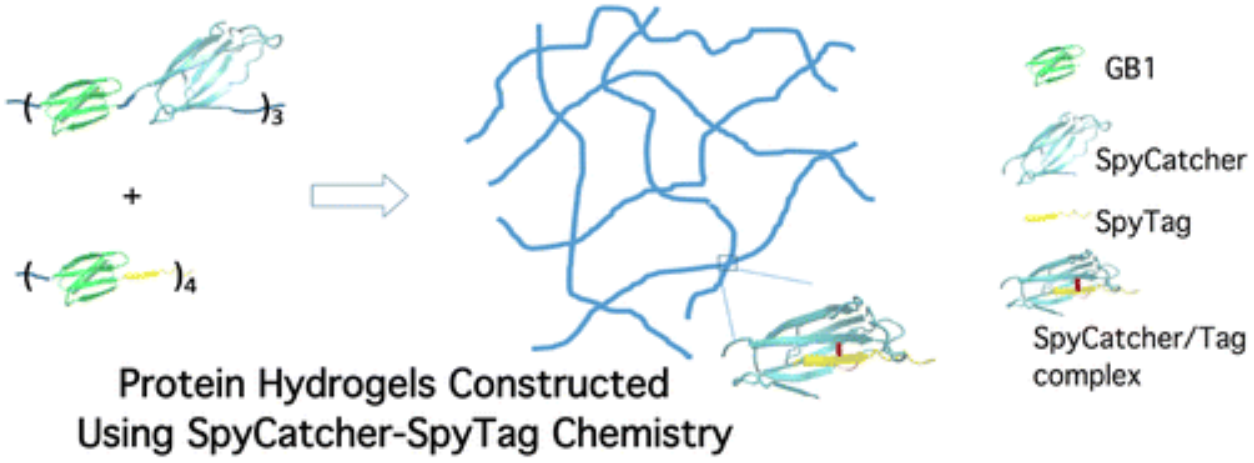

Figure 1.13. Cartoon showing the construction of a protein hydrogel based on SpyCatcher-SpyTag complex.[86] Copyright 2016, American Chemical Society.

1.3 Introduction of protein fragment reconstitution of GB1

1.3.1 From mutually exclusive protein to protein fragment reconstitution

Mutually exclusive proteins (MEP) are a particular type of proteins engineered with the insertion of a domain.[87] In MEPs, one guest domain with a longer N-C termini distance than the loop of a host protein is inserted into the loop. GB1 engineered with loop insertion is an example of mutually exclusive proteins. The B1 binding domain of protein G from Streptococcus (GB1) is a small protein that contains 56 amino acids with an α helix in-between a four-strand β sheet. Our group previously has investigated the effect of loop insertions on the mechanical stability of protein GB1. There are two loops linking the α helix to β strands 2 and 3, in which loop 2 with five amino acid residues 37-41 connects the α helix to β strand 3 (Figure 1.14). Four amino acid residues (2, 5, 24, and 46 amino acids long) acting as guest domains were inserted between residues 39 and 40 of loop 2 through protein engineering techniques, which elongated the flexible loop. GB1 was inserted with -GGGLG- sequence (GB1-L5), in which the codons of LG (CTC GGG) were a

nonpalindromic AvaI restriction site. In the end, far-UV circular dichroism (CD) spectroscopy was utilized to determine the structural stability of the engineered GB1, which confirmed that loop two could tolerate loop elongation without changing the native structure of GB1.[88]

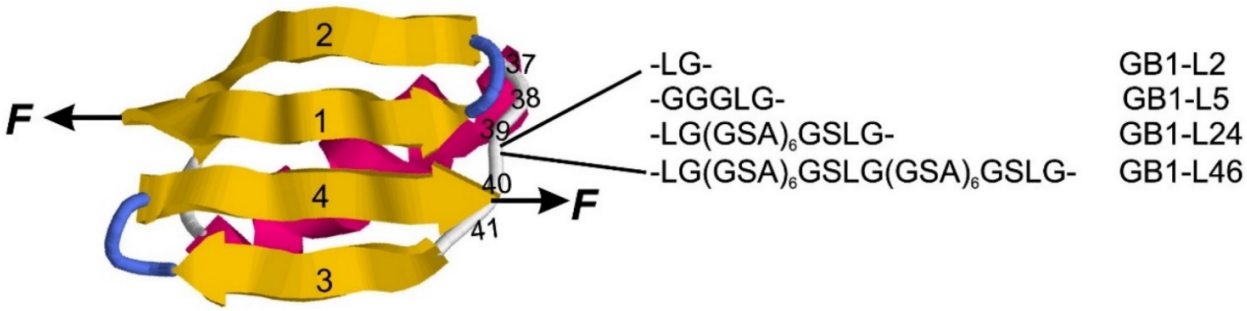

Figure 1.14. Insertion of amino acid residues into the loop connecting the α helix to β strand 3 in GB1. GB1 has two loops that connect α helix to β strands. Loop 2 (the white loop on the right of the figure) links α helix to β strand 3, which is lengthened by four various linkers (sequence shown on the right) between residues 39 and 40.[88] Copyright 2008, Elsevier Ltd.

Peng and Li discovered that the 27th Ig domain of human titin (I27) with a Trp34Phe mutation (I27w34f) could be constructed as a guest protein, which was inserted into GB1-L5 (or GL5) via PCR in order to create a mutually exclusive protein GL5/I27w34f (Figure 1.15). Though GL5 was separated by an 89-residue guest domain I27w34f, GL5 could refold to its native structure.[89]

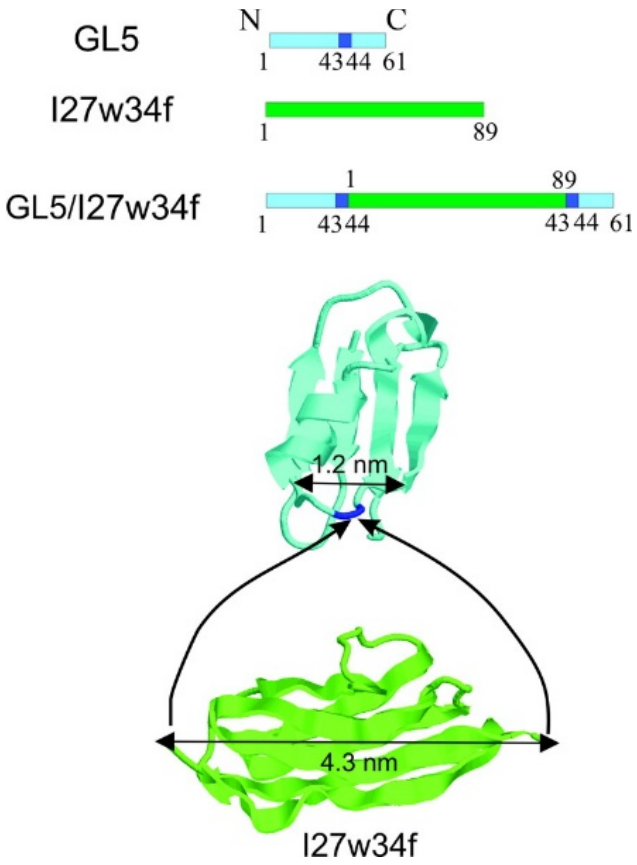

Figure 1.15. Mutually exclusive protein GL5/I27w34f. GL5/I27w34f is created by the insertion of I27w34f into loop 2 of GL5. Residues 43 and 44 are duplicated from the AvaI site in PCR.[89] Copyright 2009, American Chemical Society.

Protein fragment reconstitution, also known as fragment complementation, is a self-assembling phenomenon – proteins can be split into two half fragments (N- and C-terminal) and reconstitute to the folded conformation of the native protein either spontaneously or helped by assistant proteins.[90-93] Kobayashi et al. first discovered the protein fragment reconstitution of GB1. It was found that after being split, two segments of GB1 could associate at 1:1 ratio to regenerate a stable native structure of GB1 with a dissociation constant $K_d = 9 \times 10^{-6}$ M.[94] Furthermore, the melting temperature T_m of reconstituted GB1 was determined to be 42 °C.[95]

1.3.2 Hydrogelation based on protein fragment reconstitution of GB1

Inspired by the reconstituted GB1 and the GL5/I27w34f work, Kong and Li developed a self-assembling reversible hydrogelation approach driven by protein fragment reconstitution. GB1−GGGLG (GL5) was mutated to be GB1−GCGCG (GL5−CC) with two cysteines, which made it feasible that after separation, two fragments of G_N (1–42) and G_C (43–61), each carrying one cysteine, were able to spontaneously reconstitute GL5−CC and form a covalently linked disulfide bond between those two cysteines in the folded GL5-CC under oxidation (Figure 1.16A). Notably, G_N-I27 and I27-G_C (for simplicity, I27w34f is abbreviated as I27F) were capable of reconstituting I27F-GL5CC–I27F with a T_m of 23 ºC under a reduction state and be oxidized to form a disulfide bond. Two engineered proteins, (I27F-G_N-I27F)$_4$ and (I27F$_3$-G_C)$_3$, were utilized to form a hydrogel via two-component hydrogelation method (Figure 1.16B), which was temperature-dependent; when melting at a temperature above 23 ºC (T_m for the reconstituted G_N/G_C), the hydrogel became a viscous solution. However, after being chemically crosslinked to establish a disulfide bond upon oxidation, the hydrogel remained in a transparent gel state even though the temperature was increased to 85 °C. A tetra-functional protein GB1-Reslin-(GB1-G_N-I27F-G_C-Reslin)$_2$ can self-assemble into a physically crosslinked hydrogel (Figure 1.16C) with similar thermal responsiveness, which also could undergo chemical crosslinking upon oxidation.[90] These examples provide a novel way to synthesize protein-based hydrogels via protein fragment reconstitution. Other researchers also developed protein hydrogels based on the protein fragment reconstitution of GB1.[96, 97]

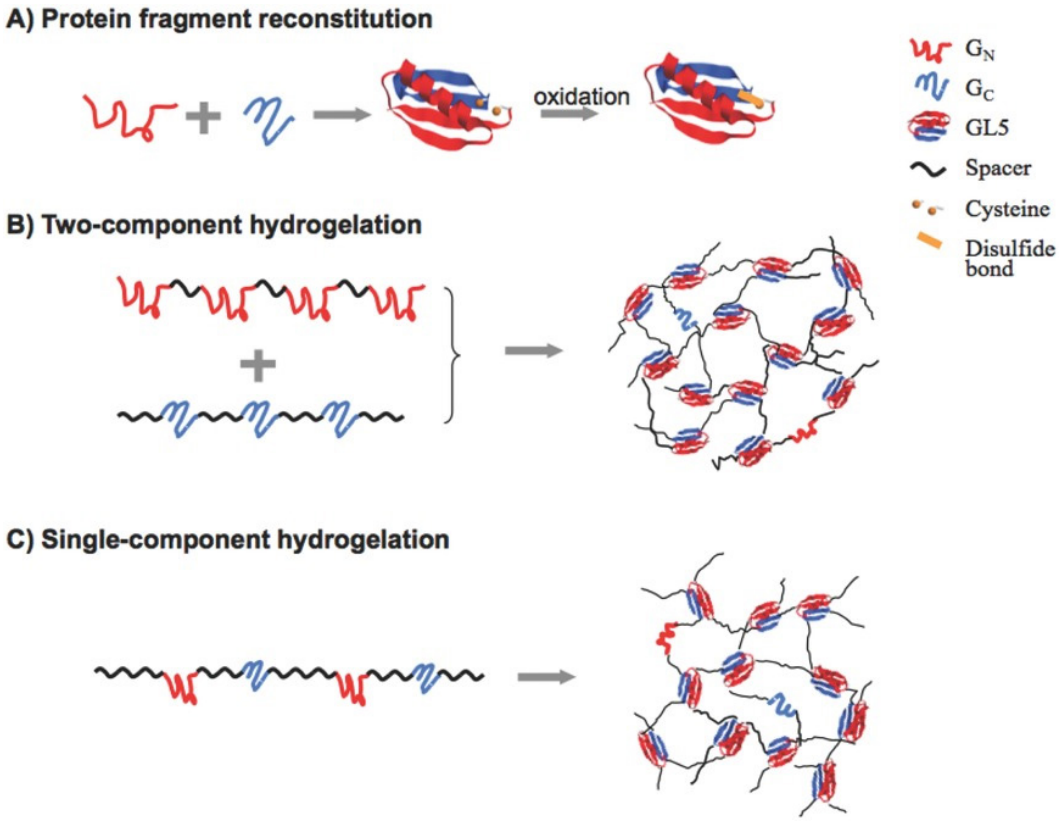

Figure 1.16. Hydrogelation caused by protein fragment reconstitution. (A) Reconstitution of GL5-CC from GN (1–42) and GC (43–61). After oxidation, two fragments are covalently linked by a disulfide bond. GL5 represents GL5-CC in this figure. (B) Mechanism of fragment reconstitution based two-component hydrogelation. Two polyproteins, each with either G_N or G_C, act as multiple functional precursors for hydrogelation. (C) Mechanism of fragment reconstitution based one-component hydrogelation. A single polyprotein with both G_N and G_C serves as a multi-functional precursor for hydrogelation.[90] Copyright 2015, WILEY-VCH Verlag GmbH & Co.

Recently, Wang et al. engineered a protein macromonomer G_C-I27F-G_N by inserting I27w34f into GL5-CC, which can undergo supramolecular polymerization via protein fragment reconstitution of GL5-CC. (Figure 1.17). This protein polymer was used to create a hydrogel via the photo-crosslinking method. This protein fragment reconstitution-based polymerization of G_C-macromonomer-G_N opened up a window for building ultra-high molecular weight protein polymers, which could lead to the development of protein hydrogels.[96]

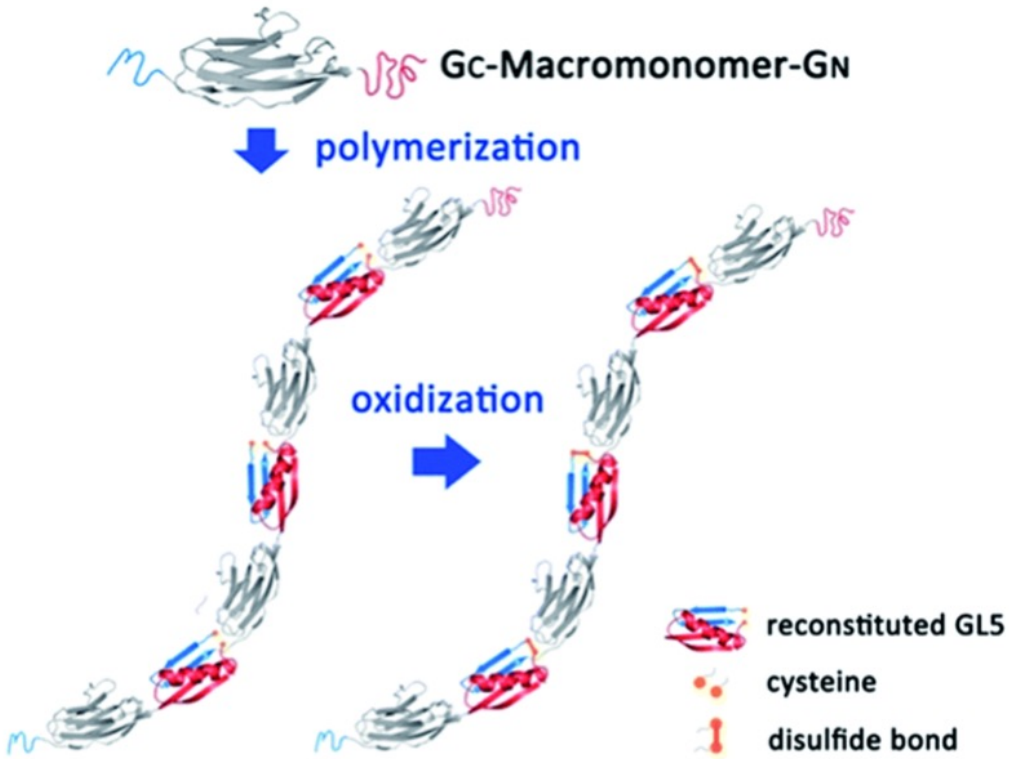

Figure 1.17. Supramolecular polymerization via protein fragment reconstitution. The polymerization is identical to condensation polymerization. After oxidation, the polyprotein chain is linked by disulfide bonds stabilizing the structure. GL5 represents GL5-CC in the figure.[96] Copyright 2019, Royal Society of Chemistry.

Cao et al. developed a hydrogel network based on the protein fragment constitution of GB1, which underwent polymerization (due to protein fragment reconstitution of G_C-G_N complex) and assembly (due to metal coordinate bonds) induced by Cu^{2+} simultaneously (Figure 1.18).[97]

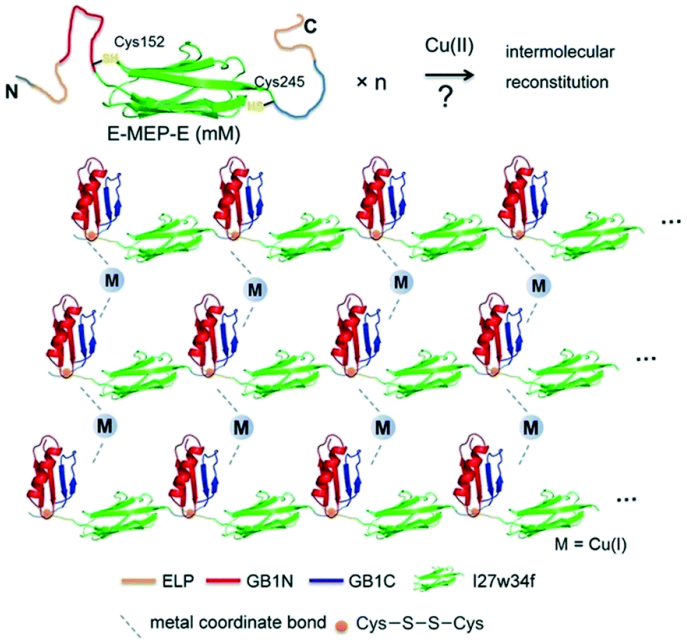

Figure 1.18. Schematic illustration of the formation of hydrogel network via protein fragment reconstitution of GB1 and metal coordinate bonds. ELP-MEP-ELP may highly polymerize due to the intermolecular disulfide bonds between Cys152 and Cys245 formed after oxidation. Also, interchain interactions could be established by metal coordination with side chains of the residues, such as Cys and His.[97] Copyright 2020, Royal Society of Chemistry.

1.4 Thesis aims

Due to high water content and tunable physical/chemical properties, protein-based hydrogels have been utilized as extracellular matrices (ECMs) for research in cellular behavior.[98-101] The mechanical stimuli from ECMs are of importance to cellular activity and function, which are transduced to cells through cell-ECM interaction to induce biochemical signaling cascades controlling several cellar behaviors, such as cell adhesion and growth.[102, 103] In order to achieve the desired biological signaling functions, protein hydrogels should be functionalized with designed signaling molecules via a variety of conjugation methods. Fibronectin type-III domain derived from the ECMs protein tenascin contains an arginine–glycine–aspartic acid (RGD) pattern, which

is responsible for the cell adhesion function of tenascin.[104] It has been applied to build protein hydrogels that help cell adhesion and spreading.[98, 99]

Protein fragment reconstitution of GB1 is a novel and promising method to construct protein polymers. By utilizing protein fragment reconstitution of G_N/G_C complex, $Ru(II)(bpy)_3^{2+}$-mediated photochemical crosslinking method, and the properties of FN3, it is possible to engineer a hydrogel that can act as an ECM but also has thermal and redox dependence based on the features of G_C-G_N association. G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N are constructed, which should have the capacity to polymerize in a way similar to condensation polymerization. From previous research, the GB1 tag is known to increase the expression level and solubility of the protein with the tag.[105, 106] Therefore, G_C-GB1-FN3-G_N is expected to have a higher expression level and solubility. With the inclusion of I27F, protein G_C-I27F-FN3-G_N should be more elastic than G_C-FN3-G_N.

Chapter 2: Methods and materials

2.1 Protein engineering

Polyproteins G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N were constructed using standard molecular biology techniques. GL5-CC with double-point mutation mutants 41C and 43C was created by site-directed mutagenesis, as previously reported.[90] Subsequently, G_N (1–42) flanked with 5′ BamHI and 3′ KpnI restriction sites, and G_C (43–61) flanked with 5′ BamHI and 3′ BglII-KpnI restriction sites were amplified via PCR of GL5-CC. The restriction sites of BamHI, BglII, and KpnI are given in Table 2.4. G_C was then digested by restriction endonuclease BamHI and KpnI to create overhanging "sticky ends" whose sequence corresponded to that of the pQE80L vector digested with BamHI and KpnI. Therefore, the G_C insert was ligated with the pQE80L vector to create pQE80L-G_C.[90] Then, the FN3 insert digested with the enzymes BamHI and KpnI was cloned into pQE80L-G_C vector digested with BglII and KpnI to form pQE80L-G_C-FN3, which was later digested with BglII and KpnI and ligated with G_N insert (digested with BamHI and KpnI) to form pQE80L-G_C-FN3-G_N (Figure 2.9). Similarly, pQE80L-G_C-GB1-FN3-G_N and pQE80L-G_C-I27F-FN3-G_N were constructed by cloning respective inserts into a pQE80L-G_C vector.

Table 2.1. Enzymes used in the protein engineering and their cut sites.

Enzyme	BamHI	BglII	KpnI
Restriction site	5′ GGATCC 3′ 3′ CCTAGG 5′	5′ AGATCT 3′ 3′ TCTAGA 5′	5′ GGTACC 3′ 3′ CCATGG 5′
Cut site	5′ G|GATCC 3′ 3′ CCTAG|G 5′	5′ A|GATCT 3′ 3′ TCTAG|A 5′	5′ GGTAC|C 3′ 3′ C|CATGG 5′

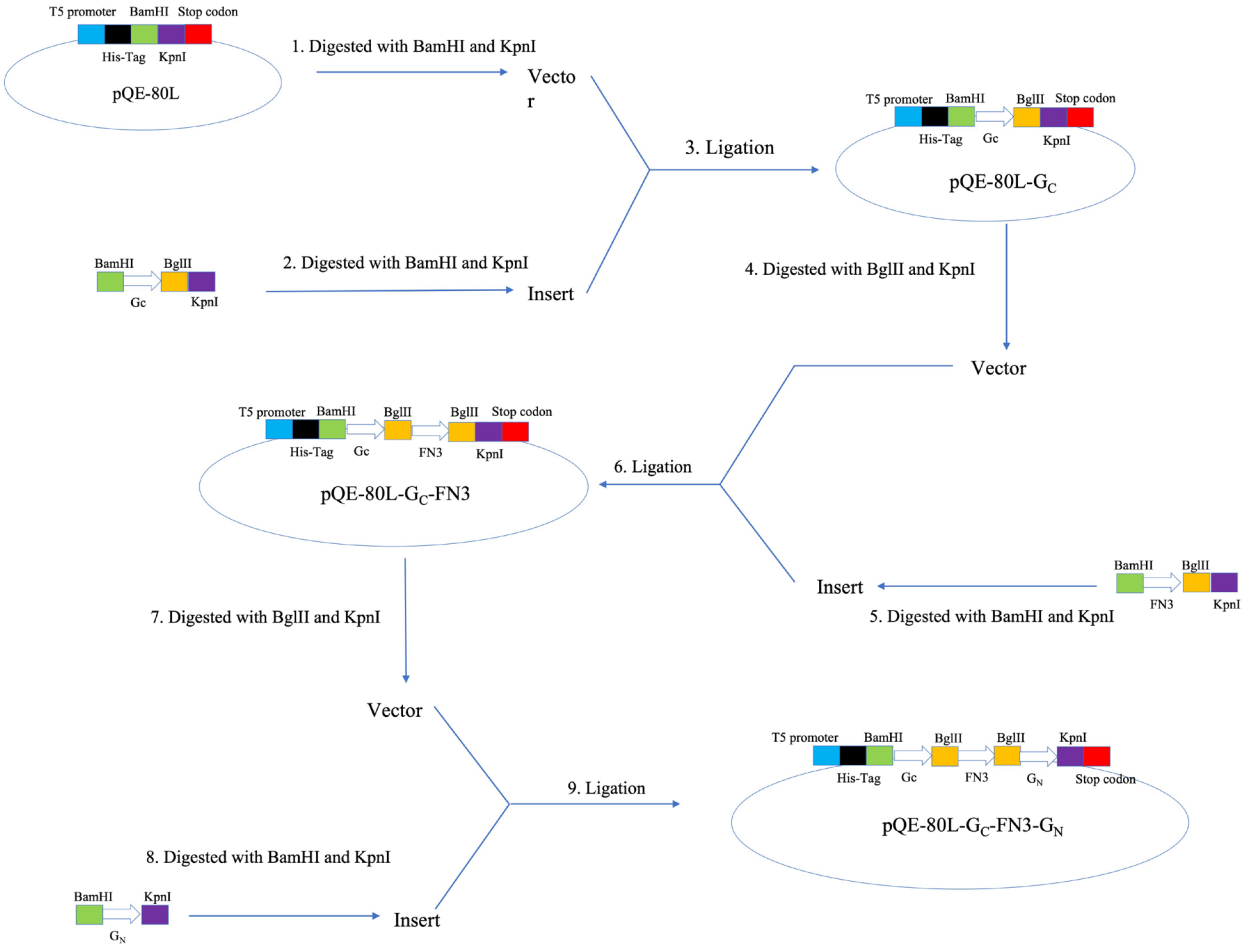

Figure 2.1. Schematic of the construction of pQE80L-GC-FN3-GN.

The plasmids pQE80L-G_C-FN3-G_N, pQE80L-G_C-GB1-FN3-G_N, and pQE80L-G_C-I27F-FN3-G_N were transformed into *Escherichia coli* strain DH5α competent cells for protein overexpression. Each starter was cultured overnight in a 20 mL mixture of 2.5 % Luria-Bertani broth (LB) medium and 100μg/mL ampicillin at 225 rpm and 37 ºC, which was later transferred to 800 mL LB liquid medium containing 100μg/mL ampicillin for incubation at 225 rpm and 37 ºC. After about three hours of incubation to reach OD600 of 0.6-0.8, 1 mM isopropyl-1-β-D-thiogalactoside (IPTG) was introduced to induce the protein overexpression, which lasted for approximately four hours. When the protein expression was over, the cells were harvested by centrifugation at 5000 rpm for 10 minutes at 4 ℃ and then stored in the -80 ℃ freezer. After thawed cells were treated with 100 mg/mL lysozyme for 30 mins, DNA and RNA of the cells were removed by adding 1 mg/mL DNase and RNase and subsequent centrifugation at 12000 rpm for 60 minutes. The supernatant containing soluble monomers after centrifugation was collected and purified by the Co^{2+} affinity column. The purified samples were dialyzed against deionized water for 24 hours, and later the dialyzed samples were lyophilized. Amino acid sequences of all constructs are demonstrated in Table S1 in the Appendix.

2.2 Supramolecular polymerization

Lyophilized G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N were dissolved in 1× phosphate-buffered saline (PBS; pH = 7.4; 1× PBS contains 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4) to 30 μM, which later underwent supramolecular polymerization via protein fragment reconstitution overnight at 4 °C to assemble high molecular weight polymers. The solutions were allowed to be oxidized by the oxygen in the air so that the protein polymers

were stabilized by covalent crosslinking. In order to evaluate the time course of polymerization, the protein solutions were reduced by 3 mM DTT at room temperature for 2 hours, and then DTT was removed after the protein solutions went through desalting columns. The protein solutions were then oxidized at a 4 ºC fridge for 10 minutes, 70 minutes, 3 hours, 6 hours, 1 day, and 2 days. 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to determine the degree of polymerization.

2.3 Fast protein liquid chromatography (FPLC)

FPLC experiments of G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N were carried out on an Akta FPLC system equipped with a HiLoad Superdex200 pg preparative size exclusion chromatography column (GE Healthcare) in Dr. Katherine Ryan's lab. The samples were dissolved in a 20 mM phosphate buffer (pH = 7.5) containing 100 mM NaCl to reach a concentration of 5 mg/mL, and after 5 mL samples were ejected, they were eluted with buffer at a constant flow rate of 1 mL/min. The absorbance of the elution was measured at 280 nm by a UV detector. The size-exclusion chromatography calibration was established by using the following protein standards: (1) amylase: 200 kDa; (2) alcohol dehydrogenase: 150 kDa; (3) bovine serum albumin: 66 kDa; (4) carbonic anhydrase: 29 kDa; (5) cytochrome c: 12.4 kDa.

2.4 Hydrogel preparation

Protein hydrogels were prepared by using the well-developed Ru(II)$(bpy)_3^{2+}$-mediated photochemical crosslinking approach.[107] The proteins were dissolved in 1× PBS to obtain the

desired concentration of protein solutions, and the protein solutions were allowed to be oxidized by oxygen in the air. Subsequently, protein solutions with 50 mM ammonium persulfate (APS) and 0.2 mM $Ru(II)(bpy)_3^{2+}$ were quickly transferred to a custom-made plexiglass square-shaped mold (4mm length × 4mm width × 1mm height). Later, the aqueous mixtures were irradiated by a 150W Fiber-Lite MI-150 light source (Dolan-Jenner) for 10 minutes at the height of 10 cm so that the proteins were crosslinked to form hydrogels. To re-dissolve the hydrogels, the hydrogels were immersed in PBS with 100 mM DTT at room temperature overnight and then at 60 °C for 1 hour.

2.5 Swelling ratio measurement

Swelling ratio measurements were carried out with the square-shaped hydrogels, which were carefully blotted and then weighed to obtain the fresh-made sample mass m_0 after being taken out from the mold. After being stored in PBS at 4 ºC overnight, the hydrogels were blotted and weighed again to measure swollen sample mass m_s. The swelling ratio was defined as $r = 100\% \times (m_s - m_0)/m_0$. The hydrogels were then transferred to PBS with 100 mM DTT overnight at 4 ºC to calculate the swelling ratio of reduced hydrogels.

2.6 Rheology measurement

Rheology measurements were performed by using a Discovery HR-2 Rheometer (TA Instruments) equipped with an 8 mm flat plate and an optical fiber illumination system in time-sweep mode at room temperature and a frequency of 10 rad/s, to yield measured viscoelastic moduli, storage modulus G′ and loss modulus G″. The oxidized protein solutions with 50 mM APS and 0.2 mM

$Ru(II)(bpy)_3^{2+}$ were mixed and placed on the center of the plate where the light can pass. The light was turned on after rheology had been started for one minute, and then the mixture was allowed for photo-crosslinking for 10 minutes. After the rheology of oxidized hydrogels was measured, the hydrogels were reduced by 100 mM DTT for at least 30 minutes at 4 ºC so that rheology measurements could be obtained under these conditions as well.

G′ and G″ can provide information about the amount of structure in a material. G′ accounts for the energy stored in the elastic structure, whereas G″ represents the energy dissipated in the viscous part. If G′ is larger than G″, the material can be considered as mainly elastic.

2.7 *In vitro* 2D cell culture using G_C-GB1-FN3-G_N hydrogels

Cell culture experiments were conducted under sterile conditions. To prepare hydrogels for cell culture, lyophilized G_C-GB1-FN3-G_N protein was first dissolved in 1× PBS at the desired concentration and oxidized at 4 ºC overnight. Then, the protein solution was mixed with 50 mM APS and 0.2 mM $Ru(II)(bpy)_3^{2+}$, which was placed between a clean hydrophobic surface and a glass coverslip. The hydrogel was formed by the photo-crosslinking method described previously. It had a thickness of 2 mm and a surface area of 50 mm^2.

Human lung fibroblasts were purchased from the American Type Culture Collection (ATCC; Manassas, VA) and cultured following the recommended ATCC protocol. HFL were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM; Gibco) with 10% fetal bovine serum (Hyclone) and 1× penicillin-streptomycin. Cells were kept at 37 °C with 5% CO_2 for growth and

passaged. HFL cells were suspended in DMEM, and each suspension that contained around 10,000 cells was pipetted to an Eppendorf tube. After the suspension was centrifugated, the supernatant was removed. The cells were resuspended with medium and loaded onto the hydrogel. The cell-hydrogel matrix was kept at 37 °C with 5% CO_2 overnight to allow the growth of HLFs.

2.9 *In vitro* 3D cell culture using G_C-GB1-FN3-G_N hydrogels

Lyophilized G_C-GB1-FN3-G_N protein was dissolved in DMEM at the 3 w/v% concentration and oxidized at 4 ºC overnight. HFLs were suspended in DMEM, and each suspension that contained around 10,000 cells was pipetted to an Eppendorf tube. After the suspension was centrifugated, the supernatant was removed, and the cells were resuspended with G_C-GB1-FN3-G_N solution, to which 5 mM APS and 0.2 mM $Ru(II)(bpy)_3^{2+}$ were then added. Subsequently, the mixture was pipetted into a 96-well plate, and the gelation was initiated by the illumination, which lasted for 10 minutes. After gelation, the cell-hydrogel matrix was immersed in DMEM, which was changed every two hours until most APS and $Ru(II)(bpy)_3^{2+}$ were washed off. The matrix was kept at 37 °C with 5% CO_2 overnight to allow the growth of HLFs.

2.10 LIVE/DEAD assay

A LIVE/DEAD Viability Kit for Mammalian Cells was purchased from Thermo Fisher Scientific, which was used to assess the viability of the cell culture. LIVE/DEAD assay solution was prepared following the manufacturer's protocols. After the cell medium was removed, the LIVE/DEAD assay solution was used to stain the cells inside the matrix for 30 minutes at room temperature,

which later was washed using PBS three times. Then, the matrix was imaged at 10× magnification using an IX83 Inverted fluorescence microscope (Olympus) in Dr. Russ Algar's lab. The live HLFs were captured with a green fluorescent protein filter (Excitation/Emission wavelength: 494/517 nm), whereas dead HLFs were captured with a red fluorescent protein filter (Excitation/Emission wavelength: 528/617 nm).

Chapter 3: Results and discussion

3.1 Protein fragment reconstitution of G_N-G_C complex leads to the polymerization of proteins

Hypothetically, proteins G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N interact with themselves through the protein fragment reconstitution of G_N and G_C, leading to the self-assembly of protein polymers with high molecular weight, which is similar to condensation polymerization. Moreover, the oxidation between the two cysteines of the reconstituted G_N-G_C complex is expected to generate a stable disulfide bond, which can convert the physically crosslinked polymers to chemically crosslinked polymers. To examine these hypothesizes, we conducted SDS-PAGE of G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N. The SDS-PAGE of G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N reduced by 2-mercaptoethanol indicated a molecular weight of approximately 20 kDa, 27 kDa, and 30 kDa, respectively (Figure 3.1). It is evident from the SDS-PAGE in Figure 3.2 that all monomeric G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N can self-assemble into polyproteins with a high degree of polymerization within one day. As the reaction proceeded, the number of protein monomers decreased while that of protein polymers increased as a function of time, and this characteristic confirms that the polymerization through protein fragment reconstitution of G_N and G_C. The impure bands may arise from the self-association of N-terminal G_C and C-terminal G_N to GB1 complexes; for example, G_C-FN3-G_N might form GB1-FN3. This conformational change could affect protein migration rate in SDS-PAGE.[108]

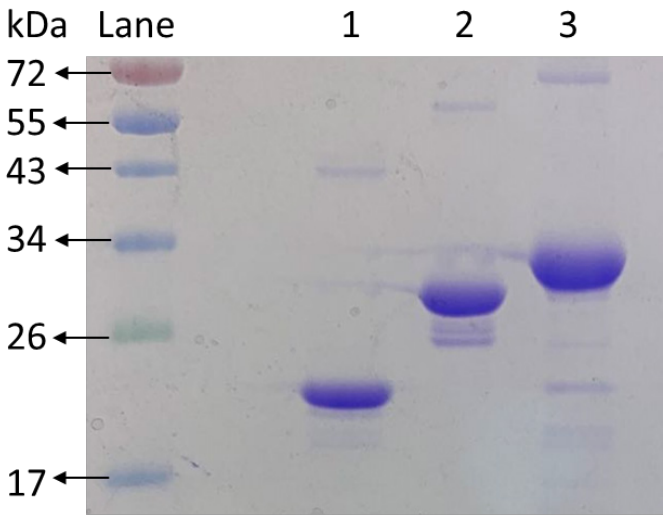

Figure 3.1. 12% SDS-PAGE of reduced G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N. Lanes 1 to 3 are reduced G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N, which have a MW of 20 kDa, 27 kDa, and 30 kDa, respectively.

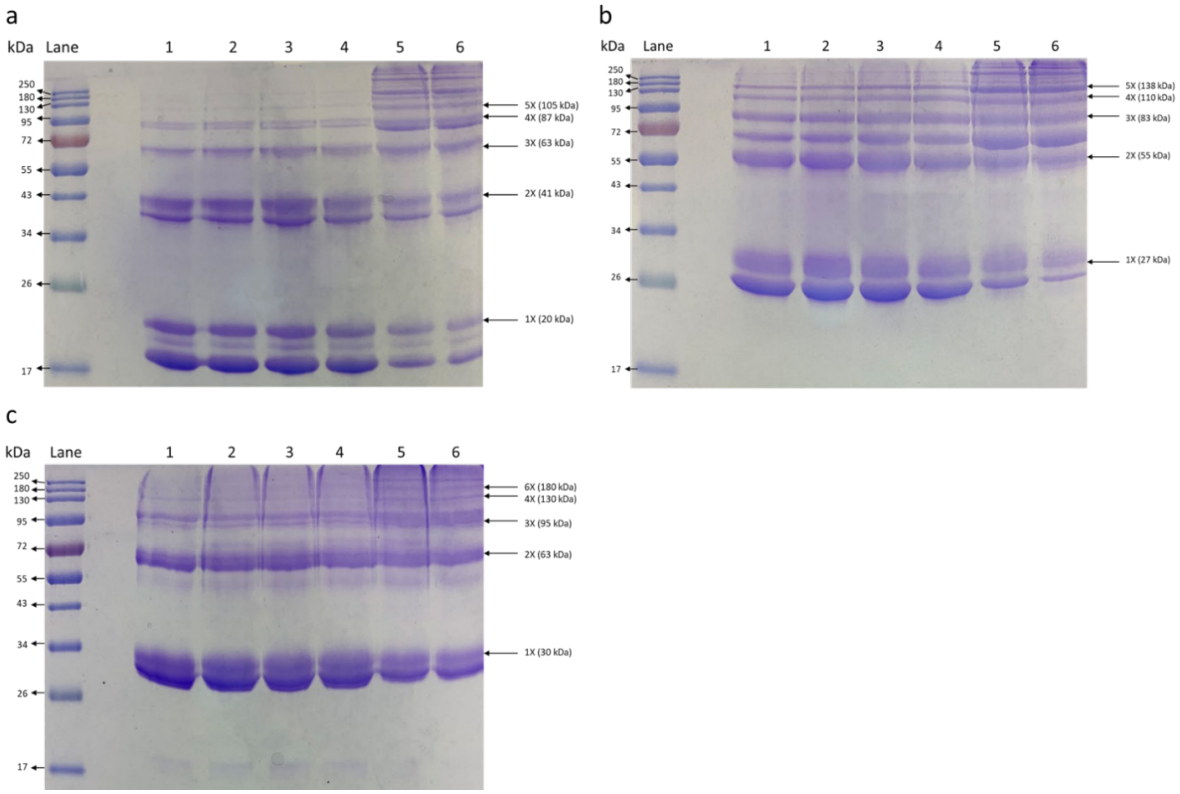

Figure 3.2. Time course of polymerization of non-reducing G_C-FN3-G_N (a), G_C-GB1-FN3-G_N (b), and G_C-I27F-FN3-G_N (c) by 12% SDS-PAGE. Lanes 1-6 are samples after air oxidation for 10 min, 70 min, 3 h, 6 h, 1 d, and 2 d.

Since SDS-PAGE has low resolution for protein polymers with a high degree of polymerization, FPLC was conducted to evaluate the molecular weight of the protein polymers. G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N protein polymers were mostly eluted before 70 minutes but had a small peak at ~80 minutes (Figure 3.3a) that corresponded to the molecular weight of monomer residues (Figure 3.3b). Absorbance at 280 nm offers an accurate determination of protein concentration, as it arises from tryptophan, tyrosine, and phenylalanine residues.[109-111] Therefore, the weight average molecular weight (M_w) and the number average molecular weight (M_n) can be determined (Appendix A.18). Mw of G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N were approximately 175.5 kDa, 180.1 kDa, and 221.6 kDa, respectively, while Mn of G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N were around 97.2 kDa, 106.0 kDa, and 113.7 kDa, respectively. The polydispersity index (PDI) was calculated by dividing Mw by Mn, i.e., 1.81, 1.70, and 1.95 for G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-FN3-G_N, respectively. PDI for these three GC/GN proteins was around two, suggesting that the polymerization of the GC/GN complexes was linear.

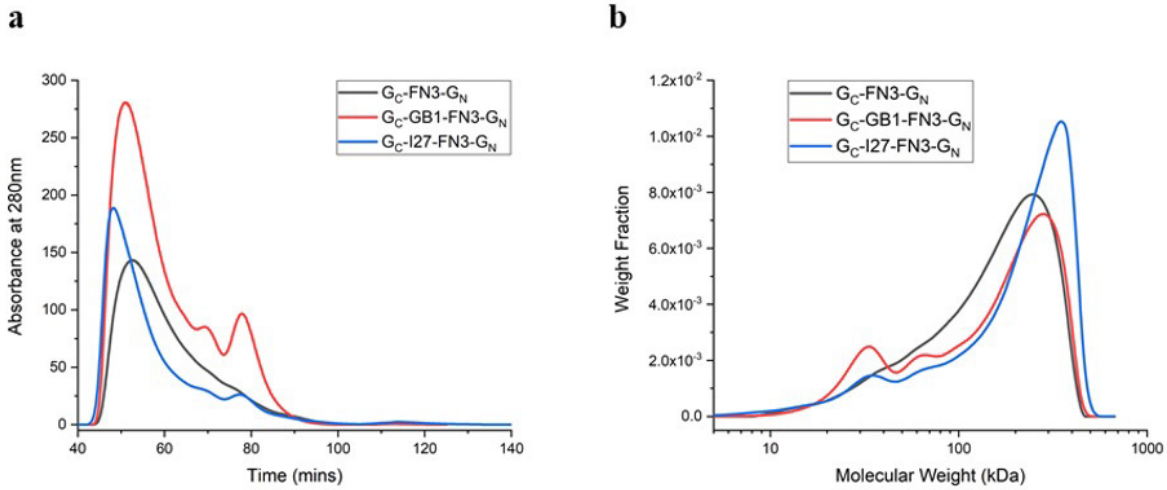

Figure 3.3. FPLC profile (a) and molecular weight distribution (b) of G_C-FN3-G_N (black), G_C-GB1-FN3-G_N (red), and G_C-I27F-FN3-G_N (blue). The samples were dissolved and eluted by 20 mM phosphate buffer (pH = 7.5) containing 100 mM NaCl. All samples were polymerized overnight, and their concentration were 5 mg/mL.

3.2 G_C-FN3-G_N and G_C-GB1-FN3-G_N can form hydrogels at low concentrations via photo-crosslinking

After Ru(II)$(bpy)_3^{2+}$-mediated photochemical crosslinking, 5 w/v% oxidized G_C-FN3-G_N and G_C-GB1-FN3-G_N formed hydrogels (Figure 3.4), which had a swelling ratio of ~8.5% and ~13.2%, respectively. Moreover, both G_C-FN3-G_N and G_C-GB1-FN3-G_N could create hydrogels at a lower concentration, even at 1 w/v%. Unexpectedly, 1 w/v% G_C-I27F-FN3-G_N after 24-hour oxidation was too viscous to be pipetted to the mold for further photo-crosslinking.

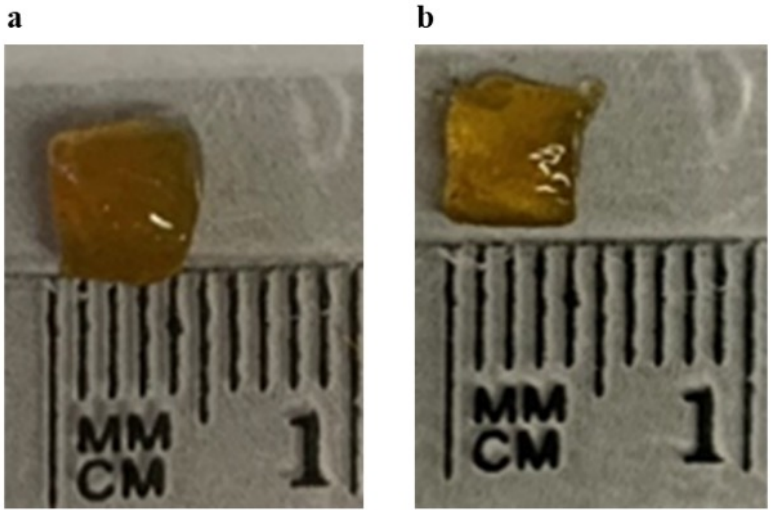

Figure 3.4. Formation of protein hydrogels by using G_C-FN3-G_N (a) and G_C-GB1-FN3-G_N (b). The hydrogels had an area of ~1 cm^2.

The storage moduli and loss moduli of the hydrogels were monitored by rheology. G' dramatically increased at the beginning of the photo-crosslinking and then reached a plateau in about 3 minutes (Figure 3.5). The plateau of G' indicated the crosslinking was reaching completion. Also, there was an apparent trend for G_C-FN3-G_N hydrogels: the higher concentration of the protein solution, the greater the storage modulus of the protein hydrogel. The storage modulus of 1 w/v% G_C-FN3-

G_N hydrogel was ~0.2 kPa, whereas those of 2 w/v% and 5 w/v% G_C-FN3-G_N hydrogels elevated more to about 10 kPa and 21 kPa, respectively (Figure 3.5a). The same trend applied to GC-GB1-FN3-GN hydrogels, and the storage moduli were approximately 0.8 kPa, 7 kPa, and 20 kPa for 1 w/v%, 3 w/v%, and 5 w/v% hydrogels, respectively (Figure 3.5b). Furthermore, G′ were larger than G″ for both G_C-FN3-G_N and G_C-GB1-FN3-G_N with various concentrations, suggesting these hydrogels are elastic.

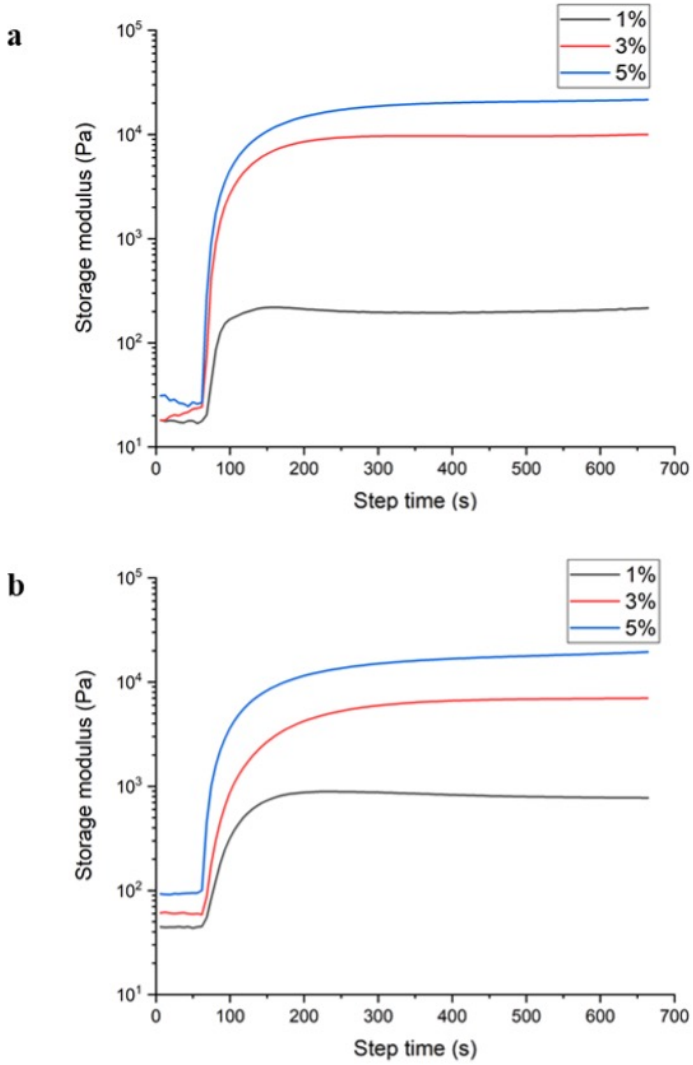

Figure 3.5. Rheology of G_C-FN3-G_N (a) and G_C-GB1-FN3-G_N (b) protein hydrogels. Protein concentrations are 1 w/v% (black), 3w/v% (red) and 5 w/v% (blue)**.** After the test started for one minute, the light was turned on, and the sample was allowed to undergo photo-crosslinking to form a hydrogel.

3.3 G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels are sensitive to redox potential and temperature

A unique property of G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels is their responsiveness to redox potential. After 100 mM DTT reduction overnight, the swelling ratio of 5 w/v% G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels escalated to ~54.4% and 56.9%, respectively, suggesting that the hydrogels absorbed more water due to the decreased crosslinking degree of the hydrogels caused by the reduction. In the reduced state, the protein polymer is associated by the non-covalent protein fragment reconstitution of G_C-G_N at a temperature below 23 ºC, T_m for the reconstituted G_N/G_C, whose structure should be more loosen than the oxidized polyprotein. The responsiveness to redox potential is also reflected in the subsequent change in the mechanical property of the hydrogels. The storage moduli for 1 w/v%, 3 w/v%, and 5 w/v% G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels dramatically dwindled by at least 75% after the hydrogels were reduced by 100 mM DTT for at least 30 minutes (Figure 3.6). This situation can be explained by the non-covalent G_C-G_N association being weaker than the covalent one.

Another feature of the protein hydrogels is that the reduced hydrogels are temperature-responsive. When temperature is higher than 23 ºC, the protein polymers will begin to depolymerize due to the dissociation of G_C-G_N complex. Both 5 w/v% G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels were still gel-like after overnight 100 mM reduction at 4 ºC. When the reduced hydrogels were heated at 60 ºC, they started to dissolve and eventually completely dissolved after one hour.

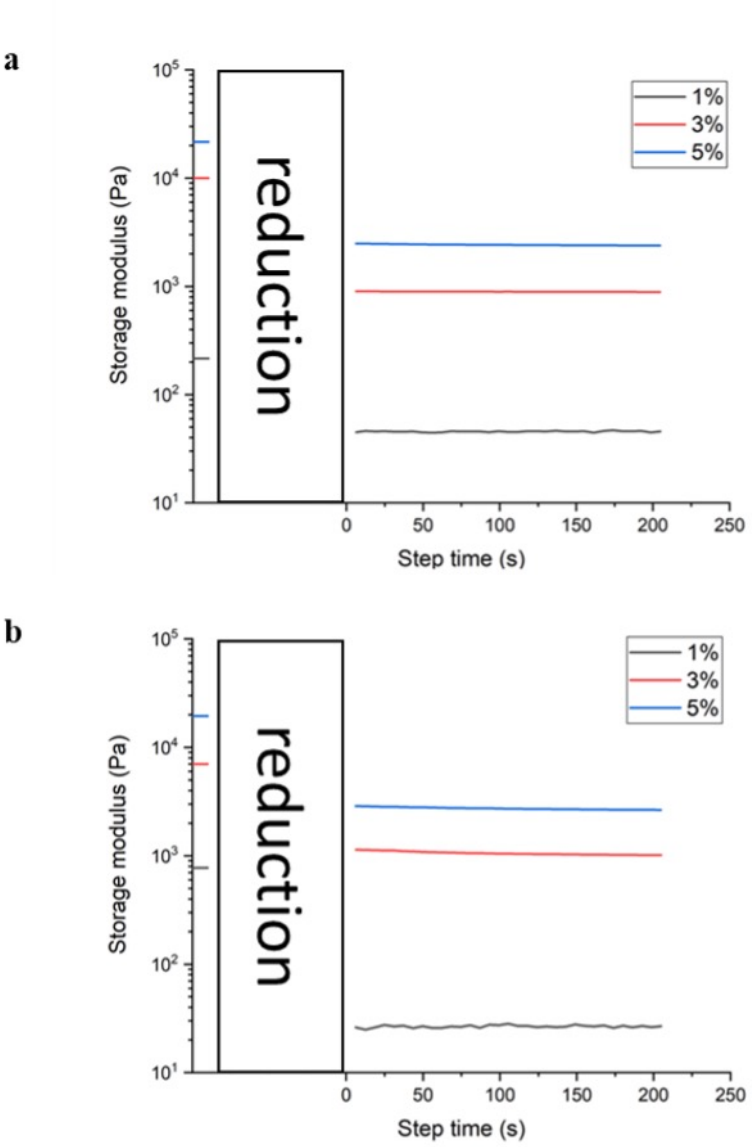

Figure 3.6. Rheology of G_C-FN3-G_N (a) and G_C-GB1-FN3-G_N (b) protein hydrogels after reduction. Protein concentrations are 1 w/v% (black), 3w/v% (red) and 5 w/v% (blue)**.** The hydrogels were reduced by 100 mM DTT for more than 30 minutes after the photo-crosslinking was complete. Therefore, the storage moduli before reduction are the same as those at the end of photo-crosslinking in Figure 3.5.

3.4 G_C-GB1-FN3-G_N hydrogels support cell adhesion in 2D cell culture

Cytocompatibility is defined as the property of allowing cell adhesion and proliferation without causing cell death.[112] To test the cytocompatibility of G_C-GB1-FN3-G_N hydrogels, HLFs were

seeded to the surface of oxidized G_C-GB1-FN3-G_N hydrogels for viability analysis by Live/Dead assay (Figure 3.7). The living cells would be green, whereas those dead should be in red. Both 3 w/v% and 5 w/v% G_C-GB1-FN3-G_N hydrogels supported cell adhesion with extremely high cytocompatibility. At least 99% HLFs were viable after overnight incubation (Figure 3.7b & d). Intriguingly, the cells in the 3 w/v% hydrogels were 30% more than that in 5 w/v% hydrogels, suggesting that cell adhesion was more favorable in 3 w/v% hydrogels. These results suggest that the hydrogels could be potentially used as ECMs.

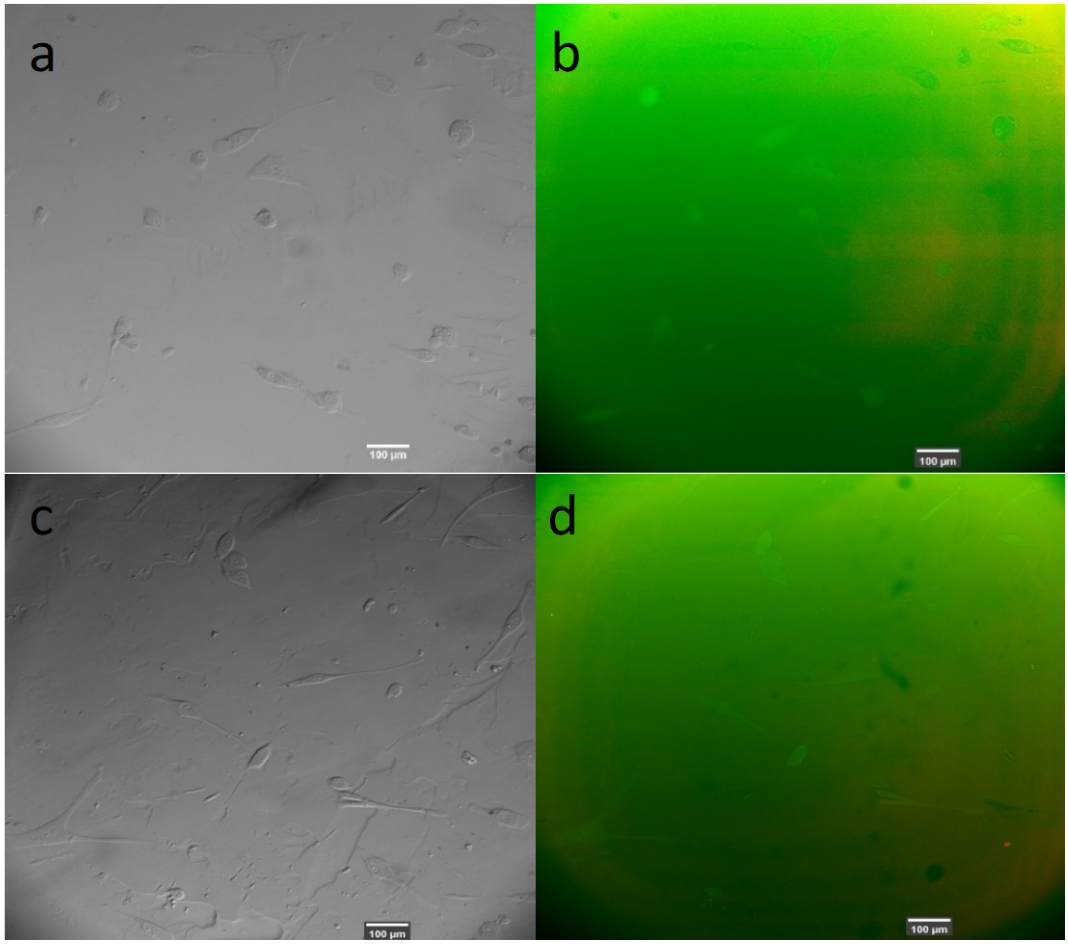

Figure 3.7. LIVE/DEAD staining of 3 w/v% and 5 w/v% G_C-GB1-FN3-G_N protein hydrogels. (a) Light microscopic image of a 3 w/v% hydrogel with HLFs seeded overnight. (b) Fluorescent imaging of s 3 w/v% hydrogel with HLFs seeded overnight. (c) Light microscopic image of a 5 w/v% hydrogel with HLFs seeded overnight. The right bottom corner showed an air bubble. (d) Fluorescent imaging of a 5 w/v% hydrogel with HLFs seeded overnight. All scale bars are 100 µm. The live HLFs are green, whereas dead HLF cells are red.

3.4 3D cell culture using G_C-GB1-FN3-G_N hydrogels is not as good as 2D cell culture

Next, the performance of 3 w/v% G_C-GB1-FN3-G_N hydrogels as 3D ECMs was explored. In contrast to 2D cell culture that only permits cells to grow on the surface, 3D cell culture provides an environment that allows cells to grow and interact with their surroundings in all directions. Hence, it is widely considered more representative of the *in vivo* environment than simple 2D cell culture.[113, 114] Figure 3.8 demonstrated that the amount of live HLFs was almost the same as that of dead HLFs with an average live/dead ratio of 1.25. Lv et al. found that the Ru complex- was not cytotoxic even with concentrations up to 500 μM and that 50 mM APS caused a huge toxic effects on HLFs.[115] Although the APS concentration was lowered to 5mM to reduce APS toxicity, excessive free radicals that attacked the HLFs were generated during photo-crosslinking. It inevitably became a dilemma: the hydrogel would not form if the APS concentration was too low, or the cells would die if the APS concentration was too high. Even though the hydrogel matrix causes cell death, protein fragment reconstitution of G_C-G_N complex opens up a window to construct hydrogel with potential applications in cell culture.

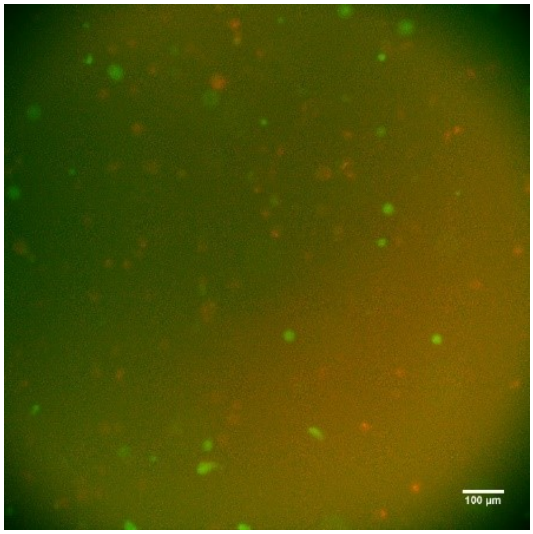

Figure 3.8. LIVE/DEAD cell viability analysis of HLFs after overnight 3D culture. 3 w/v% G_C-GB1-FN3-G_N hydrogel-HLF matrix was created by photo-crosslinking and then incubated overnight before LIVE/DEAD analysis.

Chapter 4: Conclusions and future works

Driven by protein fragment reconstitution of GB1, we have successfully engineered G_C-FN3-G_N, G_C-GB1-FN3-G_N, and G_C-I27F-GB1-FN3-G_N protein polymers with high molecular weight. It is of great success that G_C-FN3-G_N and G_C-GB1-FN3-G_N polyproteins have been used to construct protein hydrogels through $Ru(II)(bpy)_3^{2+}$-mediated photochemical crosslinking method. The resultant protein hydrogels can form at a low protein concentration, even at 1 w/v%, and exhibit thermo- and redox-responsive characteristics. Under oxidizing conditions, the polyprotein chains are covalently linked by the disulfide bonds. Conversely, these disulfide bonds vanish in the reduced state, and the G_C-G_N association relies on physical protein fragment reconstitution of GB1. More importantly, the G_C-G_N complex dissociates at a temperature above 23 ºC. These characteristics of G_C-G_N interaction entail the thermo- and redox-responsiveness of the G_C-FN3-G_N and G_C-GB1-FN3-G_N hydrogels. By adjusting temperature and/or redox potential, it would be feasible to tune the mechanical properties of the hydrogels.

Moreover, G_C-GB1-FN3-G_N hydrogels demonstrated extremely high cytocompatibility in 2D cell culture, in which 99% remained viable. Although the performance of G_C-GB1-FN3-G_N hydrogels in 3D cell culture is not as good as in 2D cell culture, the hydrogels have manifested a great potential for cell culture. Overall, this study points to an appealing approach for engineering polyproteins with a high degree of polymerization via protein fragment reconstitution of G_C-G_N complex, leading to the development of novel protein-based biomaterials with tunable physical and mechanical properties.

A possible future direction is to integrate protein fragment reconstitution of GB1 with the SpyTag-SpyCatcher chemistry (section 1.2.3.3) to create a biocompatible protein-based hydrogel as ECMs for 3D cell culture. As in this study, the biocompatibility of G_C-GB1-FN3-G_N hydrogels in 3D cell culture is limited by the attack of free radicals generated during the Ru(II)$(bpy)_3^{2+}$-mediated photochemical crosslinking method. In contrast to the crosslinking approach used in this project, the hydrogelation using protein fragment reconstitution of GB1 with the SpyTag-SpyCatcher chemistry would not create any toxic free radicals that cause cell death.

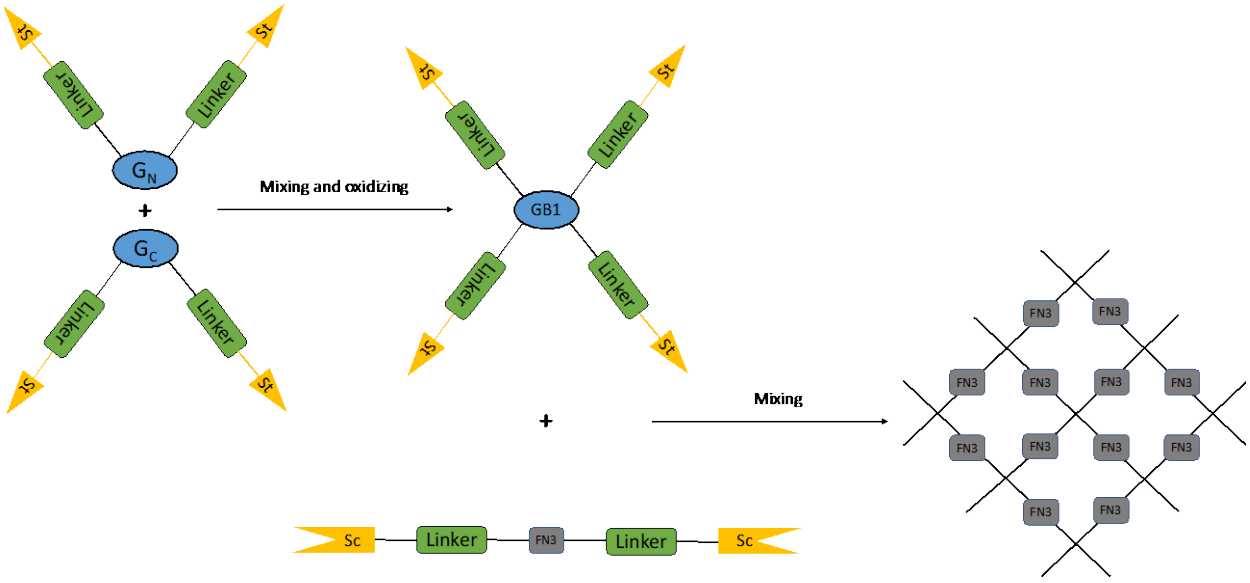

Figure 4.1. Schematics of hydrogel formation through SpyTag-SpyCatcher chemistry and protein fragment reconstitution of GB1. A 4-arm protein is formed through protein fragment reconstitution of GB1, which contains SpyTag domains at the end of each arm. SpyTag can recognize SpyCatcher, forming a complete CnaB2. A linker protein is designed with a FN3 domain in the middle and SpyCatcher at N- and C-terminals. By mixing, a hydrogel network with multiple FN3 domains should be formed.

We constructed the plasmid and expressed SpyTag-I27F-G_N-I27F-SpyTag, SpyTag-I27F-G_C-I27F-SpyTag, and SpyCatcher-I27F-FN3-I27F-SpyCatcher proteins. Their DNA and protein sequences are shown in Appendix. Theoretically, it would be feasible for SpyTag-I27F-G_N-I27F-

SpyTag and SpyTag-I27F-G_C-I27F-SpyTag to fuse a 4-armed protein dimer through the association of the G_C-G_N complex. Furthermore, SpyCatcher-I27F-FN3-I27F-SpyCatcher proteins are expected to link the 4-armed protein dimers via the irreversible SpyTag-SpyCatcher interaction, thereby forming a hydrogel network in a timely manner. FN3 is vital for cell adhesion in the 3D cell culture using protein-based hydrogel. To achieve 3D cell culture by using this hydrogel, the method in section 2.9 could be followed. This 3D cell culture model should have great potential for biomedical applications, such as artificial organ culture. We also constructed plasmids pQE80L-SpyCatcher-I27F-G_N-I27F-SpyCatcher, pQE80L-SpyCatcher-I27F-G_C-I27F-SpyCatcher, and pQE80L-SpyTag-I27F-FN3-I27F-SpyTag, which can be used to express their respective proteins for experiments in the future. These proteins are used for comparison. Their DNA and protein sequences are also shown in Appendix.

Inspired by the 4-armed protein structure, proteins that constitute a 3-armed structure were designed. SpyCatcher-I27F-G_C and SpyCatcher-I27F-G_N-I27F-SpyCatcher could associate a 3-armed protein, which also mixes with SpyTag-I27F-FN3-I27F-SpyTag to form a hydrogel for 3D cell culture. The DNA and protein sequences of SpyCatcher-I27F-G_C are also in Appendix.

Bibliography

1. Zhang, Y. S.; Khademhosseini, A. *Science* **2017**, *356* (6337), eaaf3627.
2. Seliktar, D. *Science* **2012**, *336* (6085), 1124-1128.
3. Burdick, J. A.; Murphy, W. L. *Nat. Commun.* **2012**, *3* (1), 1269.
4. Chen, Y. Chapter 1 - Properties and development of hydrogels**.** In *Hydrogels Based on Natural Polymers*; Chen, Y, Ed.; Elsevier: Amsterdam, 2020; pp 3–16.
5. Sidorenko, A.; Krupenkin, T.; Taylor, A.; Fratzl, P.; Aizenberg, J. *Science* **2007,** *315* (5811), 487–490.
6. Osada, Y.; Matsuda, A. *Nature* **1995**, *376* (6537), 219.
7. Beebe, D. J.; Moore, J. S.; Bauer, J. M.; Yu, Q.; Liu, R. H.; Devadoss, C.; Jo, B.-H. *Nature* **2000**, *404* (6778), 588–590.
8. Kopeček, J. *Biomaterials* **2007**, *28* (34), 5185–5192.
9. Liu, S.; Wang, P.; Huang, G.; Wang, L.; Zhou, J.; Lu, T. J.; Xu, F.; Lin, M. *Soft Matter* **2015**, *11* (3), 449–455.
10. Sidorenko, A.; Krupenkin, T.; Taylor, A.; Fratzl, P.; Aizenberg, J. *Science* **2007**, *315* (5811), 487–490.
11. Palleau, E.; Morales, D.; Dickey, M. D.; Velev, O. D. *Nat. Commun.* **2013**, *4* (1), 2257.
12. Shi, Q.; Liu, H.; Tang, D.; Li, Y.; Li, X.; Xu, F. *NPG Asia Mater.* **2019**, *11* (1), 64.
13. Wichterle, O.; Lím, D. *Nature* **1960**, *185* (4706), 117–118.
14. Okan, G.; Rendon, M. I. *J. Cosmet. Laser Ther.* **2011**, *13* (4), 162–165.
15. Parente, M. E.; Ochoa Andrade, A.; Ares, G.; Russo, F.; Jiménez-Kairuz, Á. *International Journal of Cosmetic Science* **2015**, *37* (5), 511-518.

16. Shigemitsu, H.; Fujisaku, T.; Tanaka, W.; Kubota, R.; Minami, S.; Urayama, K.; Hamachi, I. *Nat. Nanotechnol.* **2018**, *13* (2), 165–172.

17. Holtz, J. H.; Asher, S. A. *Nature* **1997**, *389* (6653), 829–832.

18. Nagamine, K.; Mano, T.; Nomura, A.; Ichimura, Y.; Izawa, R.; Furusawa, H.; Matsui, H.; Kumaki, D.; Tokito, S. *Sci. Rep.* **2019**, *9* (1), 10102.

19. Rowley, J. A.; Madlambayan, G.; Mooney, D. J. *Biomaterials* **1999**, *20* (1), 45–53.

20. Cui, Z.-K.; Kim, S.; Baljon, J. J.; Wu, B. M.; Aghaloo, T.; Lee, M. *Nat. Commun.* **2019**, *10* (1), 3523.

21. Formica, F. A.; Öztürk, E.; Hess, S. C.; Stark, W. J.; Maniura-Weber, K.; Rottmar, M.; Zenobi-Wong, M. *Adv. Healthc. Mater.* **2016**, *5* (24), 3129–3138.

22. Zhou, T.; Li, X.; Li, G.; Tian, T.; Lin, S.; Shi, S.; Liao, J.; Cai, X.; Lin, Y. *Sci. Rep.* **2017**, *7* (1), 10553.

23. Wolf, M. T.; Daly, K. A.; Brennan-Pierce, E. P.; Johnson, S. A.; Carruthers, C. A.; D'Amore, A.; Nagarkar, S. P.; Velankar, S. S.; Badylak, S. F. *Biomaterials* **2012**, *33* (29), 7028–7038.

24. Zhao, X.; Wu, H.; Guo, B.; Dong, R.; Qiu, Y.; Ma, P. X. *Biomaterials* **2017**, *122*, 34–47.

25. Qi, C.; Liu, J.; Jin, Y.; Xu, L.; Wang, G.; Wang, Z.; Wang, L. *Biomaterials* **2018**, *163*, 89–104.

26. Luo, Y.; Kirker, K. R.; Prestwich, G. D. *J. Control. Release* **2000**, *69* (1), 169–184.

27. Jiang, T.; Wang, T.; Li, T.; Ma, Y.; Shen, S.; He, B.; Mo, R. *ACS Nano* **2018**, *12* (10), 9693–9701.

28. Nguyen, L. H.; Gao, M.; Lin, J.; Wu, W.; Wang, J.; Chew, S. Y. *Sci. Rep.* **2017**, *7* (1), 42212.

29. Zhang, Y.; Liu, J.; Huang, L.; Wang, Z.; Wang, L. *Sci. Rep.* **2015**, *5* (1), 12374.

30. Nguyen, L. H.; Gao, M.; Lin, J.; Wu, W.; Wang, J.; Chew, S. Y. *Sci. Rep.* **2017**, *7* (1), 42212.

31. Pertici, V.; Pin-Barre, C.; Rivera, C.; Pellegrino, C.; Laurin, J.; Gigmes, D.; Trimaille, T. *Biomacromolecules* **2019**, *20* (1), 149–163.

32. Wu, J.; Wei, W.; Wang, L.-Y.; Su, Z.-G.; Ma, G.-H. *Biomaterials* **2007**, *28* (13), 2220–2232.

33. Zhao, W.; Jin, X.; Cong, Y.; Liu, Y.; Fu, J. *J. Chem. Technol. Biotechnol.* **2013**, *88* (3), 327–339.

34. Panahi, R.; Baghban-Salehi, M. Protein-Based Hydrogels. In *Cellulose-Based Superabsorbent Hydrogels*; Mondal, M., Ed.; Springer: Cham, 2019; pp 1561–1600.

35. Shi, W.; Dumont, M.-J.; Ly, E. B. *Eur. Polym. J.* **2014**, *54*, 172–180.

36. Hwang, D.-C.; Damodaran, S. *J. Agric. Food Chem.* **1996**, *44* (3), 751–758.

37. Cui, X.; Lee, J. J. L.; Chen, W. N. *Scientific Reports* **2019**, *9* (1), 18166.

38. Ahmed, E. M. *J. Adv. Res.* **2015**, *6* (2), 105–121.

39. Hosoyama, K.; Lazurko, C.; Muñoz, M.; McTiernan, C. D.; Alarcon, E. I. *Frontiers in Bioengineering and Biotechnology* **2019**, *7* (205).

40. Li, H.; Kong, N.; Laver, B.; Liu, J. *Small* **2016**, *12* (8), 973–987.

41. Mizuguchi, Y.; Mashimo, Y.; Mie, M.; Kobatake, E. *Biomacromolecules* **2020**, *21* (3), 1126−1135.

42. Langer, R.; Tirrell, D. A. *Nature* **2004**, *428* (6982), 487–492.

43. Lutolf, M. P.; Gilbert, P. M.; Blau, H. M. *Nature* **2009**, *462* (7272), 433–441.

44. Silva, R.; Fabry, B.; Boccaccini, A. R. *Biomaterials* **2014**, *35* (25), 6727–6738.

45. Maskarinec, S. A.; Tirrell, D. A. *Curr. Opin. Biotechnol.* **2005**, *16* (4), 422–426.

46. Liu, W.; Sun, J.; Sun, Y.; Xiang, Y.; Yan, Y.; Han, Z.; Bi, W.; Yang, F.; Zhou, Q.; Wang, L.; Yu, Y. *Chem. Eng. J.* **2020**, *394*, 124875.

47. Jiang, B.; Liu, X.; Yang, C.; Yang, Z.; Luo, J.; Kou, S.; Liu, K.; Sun, F. *Sci. Adv.* **2020**, *6* (41), eabc4824.

48. Gomez-Florit, M.; Pardo, A.; Domingues, R. M. A.; Graça, A. L.; Babo, P. S.; Reis, R. L.; Gomes, M. E. *Molecules* **2020**, *25*(24), 5858.

49. Djabourov, M.; Leblond, J.; Papon, P. *J. Phys. Fr.* **1988**, *49* (2), 319–332.

50. Djabourov, M.; Leblond, J.; Papon, P. *J. Phys. Fr.* **1988**, *49* (2), 333–343.

51. Gasperini, L.; Mano, J. F.; Reis, R. L. *J. R. Soc. Interface* **2014**, *11* (100), 20140817.

52. Ward, M. A.; Georgiou, T. K. *Polymers* **2011**, *3*(3), 1215-1242.

53. Mohanty, B.; Bohidar, H. B. *Biomacromolecules* **2003**, *4* (4), 1080–1086.

54. Mohanty, B.; Bohidar, H. B. *Int. J. Biol. Macromol.* **2005**, *36* (1), 39–46.

55. Bode, F.; da Silva, M. A.; Drake, A. F.; Ross-Murphy, S. B.; Dreiss, C. A. *Biomacromolecules* **2011**, *12* (10), 3741–3752.

56. Peña, C.; de la Caba, K.; Eceiza, A.; Ruseckaite, R.; Mondragon, I. *Bioresour. Technol.* **2010**, *101* (17), 6836–6842.

57. Van Den Bulcke, A. I.; Bogdanov, B.; De Rooze, N.; Schacht, E. H.; Cornelissen, M.; Berghmans, H. *Biomacromolecules* **2000**, *1* (1), 31–38.

58. Billiet, T.; Gevaert, E.; De Schryver, T.; Cornelissen, M.; Dubruel, P. *Biomaterials* **2014**, *35* (1), 49–62.

59. Xing, Q.; Yates, K.; Vogt, C.; Qian, Z.; Frost, M. C.; Zhao, F. *Sci. Rep.* **2014**, *4* (1), 4706.

60. Liu, W. G.; Yao, K. De; Wang, G. C.; Li, H. X. *Polymer* **2000**, *41* (20), 7589–7592.

61. Gold, T. B.; Buice, R. G.; Lodder, R. A.; Digenis, G. A. *Pharm. Res.* **1997**, *14* (8), 1046–1050.

62. Dash, R.; Foston, M.; Ragauskas, A. J. *Carbohydr. Polym.* **2013**, *91* (2), 638–645.

63. Carthew, J.; Frith, J. E.; Forsythe, J. S.; Truong, V. X. *J. Mater. Chem. B* **2018**, *6* (9), 1394–1401.

64. Laronda, M. M.; Rutz, A. L.; Xiao, S.; Whelan, K. A.; Duncan, F. E.; Roth, E. W.; Woodruff, T. K.; Shah, R. N. *Nat. Commun.* **2017**, *8* (1), 15261.

65. Kowalczyk, T.; Hnatuszko-Konka, K.; Gerszberg, A.; Kononowicz, A. K. *World J. Microbiol. Biotechnol.* **2014**, *30* (8), 2141–2152.

66. Zhang, S. *Nat. Biotechnol.* **2003**, *21* (10), 1171–1178.

67. Landschulz, W. H.; Johnson, P. F.; McKnight, S. L. *Science* **1988**, *240* (4860), 1759–1764.

68. Lupas, A. *Trends Biochem. Sci.* **1996**, *21* (10), 375–382.

69. Mason, J. M.; Müller, K. M.; Arndt, K. M. Considerations in the Design and Optimization of Coiled Coil Structures. In *Protein Engineering Protocols*; Arndt, K. M., Müller, K. M., Eds.; Humana Press: Totowa, NJ, 2007; pp 35–70.

70. Kohn, W. D.; Mant, C. T.; Hodges, R. S. *J. Biol. Chem.* **1997**, *272* (5), 2583–2586.

71. Liu, J.; Zheng, Q.; Deng, Y.; Cheng, C.-S.; Kallenbach, N. R.; Lu, M. *Proc. Natl. Acad. Sci.* **2006**, *103* (42), 15457–15462.

72. Huang, C.-C.; Ravindran, S.; Yin, Z.; George, A. *Biomaterials* **2014**, *35* (20), 5316–5326.

73. Hamley, I. W. *Chem. Rev.* **2012**, *112* (10), 5147–5192.

74. Clarke, D. E.; Parmenter, C. D. J.; Scherman, O. A. *Angew. Chemie.* **2018**, *57* (26), 7709–7713.

75. Romei, M. G.; Boxer, S. G. *Annu. Rev. Biophys.* **2019**, *48* (1), 19–44.

76. Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. *Proc. Natl. Acad. Sci.* **2017**, *114* (11), E2146 LP-E2155.

77. Yang, Z.; Yang, Y.; Wang, M.; Wang, T.; Fok, H. K. F.; Jiang, B.; Xiao, W.; Kou, S.; Guo, Y.; Yan, Y.; Deng, X.; Zhang, W.-B.; Sun, F. *Matter* **2020**, *2* (1), 233–249.

78. Zhang, X.; Jiang, S.; Yan, T.; Fan, X.; Li, F.; Yang, X.; Ren, B.; Xu, J.; Liu, J. *Soft Matter* **2019**, *15* (38), 7583–7589.

79. Fancy, D. A.; Kodadek, T. *Proc. Natl. Acad. Sci.* **1999**, *96* (11), 6020–6024.

80. 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. *Nature* **2005**, *437* (7061), 999–1002.

81. Jeon, E. Y.; Hwang, B. H.; Yang, Y. J.; Kim, B. J.; Choi, B.-H.; Jung, G. Y.; Cha, H. J. *Biomaterials* **2015**, *67*, 11–19.

82. Gao, X.; Lyu, S.; Li, H. *Biomacromolecules* **2017**, *18* (11), 3726–3732.

83. Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M. *Proc. Natl. Acad. Sci.* **2012**, *109* (12), E690-E697.

84. Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M. *Proc. Natl. Acad. Sci.* **2012**, *109* (12), 4347-4348.

85. Sun, F.; Zhang, W.-B.; Mahdavi, A.; Arnold, F. H.; Tirrell, D. A. *Proc. Natl. Acad. Sci.* **2014**, *111* (31), 11269 LP – 11274.

86. Gao, X.; Fang, J.; Xue, B.; Fu, L.; Li, H. *Biomacromolecules* **2016**, *17* (9), 2812–2819.

87. Radley, T. L.; Markowska, A. I.; Bettinger, B. T.; Ha, J. H.; Loh, S. N. *J. Mol. Biol.* **2003**, *332* (3), 529–536.

88. Li, H.; Wang, H. C.; Cao, Y.; Sharma, D.; Wang, M. *J. Mol. Biol.* **2008**, *379* (4), 871–880.

89. Peng, Q.; Li, H. *J. Am. Chem. Soc.* **2009**, *131* (37), 13347–13354.

90. Kong, N.; Li, H. *Adv. Funct. Mater.* **2015**, *25* (35), 5593–5601.

91. Galarneau, A.; Primeau, M.; Trudeau, L. E.; Michnick, S. W. *Nat. Biotechnol.* **2002**, *20* (6), 619–622.

92. Kerppola, T. K. *Chem. Soc. Rev.* **2009**, *38* (10), 2876–2886.

93. Rossi F.; Charlton, C. A.; Blau, H. M. *Proc. Natl. Acad. Sci.* **1997**, *94* (16), 8405-8410.

94. Kobayashi, N.; Hondo, S.; Yoshii, H.; Uedaira, H.; Munekata, E. *FEBS Lett.* **1995**, *366* (2-3), 99-103.

95. Honda, S.; Kobayashi, N.; Munekata, E.; Uedaira, H. *Biochemistry*. **1999**, *38* (4), 1203–1213.

96. Wang, R.; Li, J.; Li, X.; Guo, J.; Liu, J.; Li, H. *Chem. Sci.* **2019**, *10* (40), 9277–9284.

97. Cao, Y.; Wei, X.; Lin, Y.; Sun, F. *Mol. Syst. Des. Eng*. **2020**, 5 (1), 117–124.

98. Fu, L.; Haage, A.; Kong, N.; Tanentzapf, G.; Li, H. *Chem. Commun.* **2019**, *55* (36), 5235–5238.

99. Wang, R.; Fu, L.; Liu, J. *Chem. Commun. 2019*, *55* (84),12703–12706.

100. Seliktar, D. *Science* **2012**, *336* (6085), 1124–1128.

101. Benoit, D. S. W.; Schwartz M. P.; Durney A. R.; Anseth K. S. *Nat. Mater.* **2008**, *7* (10), 816–823.

102. Discher D. E.; Janmey P.; Wang Y. L. *Science.* **2005**, *310* (5751), 1139–1143.

103. Discher D. E.; Mooney D. J.; Zandstra P. W. *Science*. **2009**, *324* (5935), 1673–1677.

104. Leahy D. J.; Hendrickson W. A.; Aukhil I.; Erickson H. P. *Science*. **1992**, *258* (5084), 987–991.

105. Liu, Y.; Cherry, J. J.; Dineen, J. V; Androphy, E. J.; Baleja, J. D. *J. Mol. Biol.* **2009**, *386* (4), 1123–1137.

106. Zhou, P.; Wagner, G. *J. Biomol. NMR*. **2010**, *46* (1), 23–31.

107. Lv, S.; Dudek, D. M.; Cao, Y.; Balamurali, M. M.; Gosline, J.; Li, H. *Nature* **2010**, *465* (7294), 69–73.

108. Gallagher, S. R. *Curr. Protoc. Protein Sci.* **2012**, *68* (1), 10.1.1-10.1.44.

109. Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182* (2), 319–326.

110. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Sci.* **1995**, *4* (11), 2411–2423.

111. Edelhoch, H. *Biochemistry* **1967**, *6* (7), 1948–1954.

112. Kretsinger, J. K.; Haines, L. A.; Ozbas, B.; Pochan, D. J.; Schneider, J. P. *Biomaterials* **2005,** *26* (25), 5177-5186.

113. Rimann, M.; Graf-Hausner, U. *Curr. Opin. Biotechnol.* **2012**, *23* (5), 803–809.

114. Pampaloni, F.; Reynaud, E. G.; Stelzer, E. H. K. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (10), 839–845.

115. Lv, S.; Bu, T.; Kayser, J.; Bausch, A.; Li, H. *Acta Biomater.* **2013**, *9* (5), 6481–6491.

Appendix

A.1 G_C

Protein:

CGDGEWTYDDATKTFTVTE

DNA:

TGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGA

A

A.2 G_N

Protein:

MDTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVGCGLG

DNA:

ATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAAACCACCAC

CGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGCTAACGACA

ACGGTGTCGGTTGCGGACTCGGG

A.3 FN3

Protein:

TRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDENQYSIG

NLKPDTEYEVSLISRRGDMSSNPAKETFTTG

DNA:

ACACGCTTGGATGCCCCCAGCCAGATCGAGGTGAAAGATGTCACAGACACCACTGC

CTTGATCACCTGGTTCAAGCCCCTGGCTGAGATCGATGGCATTGAGCTGACCTACGG

CATCAAAGACGTGCCAGGAGACCGTACCACCATCGATCTCACAGAGGACGAGAACC

AGTACTCCATCGGGAACCTGAAGCCTGACACTGAGTACGAGGTGTCCCTCATCTCCC

GCAGAGGTGACATGTCAAGCAACCCAGCCAAAGAGACCTTCACAACAGGC

A.4 wild type GB1

Protein:

MDTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE

DNA:

ATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAAACCACCAC

CGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGCTAACGACA

ACGGTGTCGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACC

GAA

A.5 I27F

Protein:

LIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILHN

CQLGMTGEVSFQAANTKSAANLKVKEL

DNA:

CTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGC

CCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGG

ACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATAT

TCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGC

TAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTG

A.6 SpyTag

Protein: AHIVMVDAYKPTK

DNA: GCTCATATTGTCATGGTTGATGCTTACAAGCCAACTAAG

A.7 SpyCatcher

Protein:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST

WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHI

DNA:

GGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGA

CATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAG

ACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACC

ATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAG

TATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCAC

CTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCG

ACGCACACATC

A.8 G_C-FN3-G_N

Protein:

CGDGEWTYDDATKTFTVTERSTRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGI

KDVPGDRTTIDLTEDENQYSIGNLKPDTEYEVSLISRRGDMSSNPAKETFTTGRSMDTYK

LILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVGCGLG

DNA:

TGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGA

AAGATCCACACGCTTGGATGCCCCCAGCCAGATCGAGGTGAAAGATGTCACAGACA

CCACTGCCTTGATCACCTGGTTCAAGCCCCTGGCTGAGATCGATGGCATTGAGCTGA

CCTACGGCATCAAAGACGTGCCAGGAGACCGTACCACCATCGATCTCACAGAGGAC

GAGAACCAGTACTCCATCGGGAACCTGAAGCCTGACACTGAGTACGAGGTGTCCCT

CATCTCCCGCAGAGGTGACATGTCAAGCAACCCAGCCAAAGAGACCTTCACAACAG

GCAGATCCATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAA

ACCACCACCGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGC

TAACGACAACGGTGTCGGTTGCGGACTCGGG

A.9 G_C-GB1-FN3-G_N

Protein:

CGDGEWTYDDATKTFTVTERSMDTYKLILNGKTLKGETTTEAVDAATAEKVFKQYAN

DNGVDGEWTYDDATKTFTVTERSTRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTY

GIKDVPGDRTTIDLTEDENQYSIGNLKPDTEYEVSLISRRGDMSSNPAKETFTTGRSMDT

YKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVGCGLG

DNA:

TGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGA

AAGATCCATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAAA

CCACCACCGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGCT

AACGACAACGGTGTCGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCAC

GGTTACCGAAAGATCCACACGCTTGGATGCCCCCAGCCAGATCGAGGTGAAAGATG

TCACAGACACCACTGCCTTGATCACCTGGTTCAAGCCCCTGGCTGAGATCGATGGCA

TTGAGCTGACCTACGGCATCAAAGACGTGCCAGGAGACCGTACCACCATCGATCTC

ACAGAGGACGAGAACCAGTACTCCATCGGGAACCTGAAGCCTGACACTGAGTACGA

GGTGTCCCTCATCTCCCGCAGAGGTGACATGTCAAGCAACCCAGCCAAAGAGACCT

TCACAACAGGCAGATCCATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTG

AAAGGTGAAACCACCACCGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAA

ACAGTACGCTAACGACAACGGTGTCGGTTGCGGACTCGGG

A.10 G_C-I27F-FN3-G_N

Protein:

CGDGEWTYDDATKTFTVTERSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKG

QPLAASPDCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSTRLDAPS

QIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDENQYSIGNLKPDTE

YEVSLISRRGDMSSNPAKETFTTGRSMDTYKLILNGKTLKGETTTEAVDAATAEKVFKQ

YANDNGVGCGLG

DNA:

TGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGA

AAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTG

AAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGC

TGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAG

AAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTC

CAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTGAGATCCAC

ACGCTTGGATGCCCCCAGCCAGATCGAGGTGAAAGATGTCACAGACACCACTGCCT

TGATCACCTGGTTCAAGCCCCTGGCTGAGATCGATGGCATTGAGCTGACCTACGGCA

TCAAAGACGTGCCAGGAGACCGTACCACCATCGATCTCACAGAGGACGAGAACCAG

TACTCCATCGGGAACCTGAAGCCTGACACTGAGTACGAGGTGTCCCTCATCTCCCGC

AGAGGTGACATGTCAAGCAACCCAGCCAAAGAGACCTTCACAACAGGCAGATCCAT

GGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAAACCACCACCG

AAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGCTAACGACAAC

GGTGTCGGTTGCGGACTCGGG

A.11 SpyCatcher-I27F-G_N-I27F-SpyCatcher

Protein:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST

WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIR

SLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILH

NCQLGMTGEVSFQAANTKSAANLKVKELRSMDTYKLILNGKTLKGETTTEAVDAATAE

KVFKQYANDNGVGCGLGLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLA

ASPDCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSGAMVDTLSGL

SSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFY

LYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHI

DNA:

GGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGA

CATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAG

ACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACC

ATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAG

TATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCAC

CTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCG

ACGCACACATCAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTG

TTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGC

CAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAG

GATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGA

GGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATT

GAGATCCATGGACACCTACAAACTGATCCTGAACGGTAAAACCCTGAAAGGTGAAA

CCACCACCGAAGCTGTAGACGCTGCTACTGCAGAAAAAGTTTTCAAACAGTACGCT

AACGACAACGGTGTCGGTTGCGGACTCGGGCTAATAGAAGTGGAAAAGCCTCTGTA

CGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACC

TGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTG

TGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGG

TATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAA

AGTGAAAGAATTGAGATCCGGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCG

AACAAGGCCAAAGCGGCGACATGACGATTGAAGAAGACTCGGCTACCCACATTAAA

TTTAGCAAGCGTGATGAAGACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCG

CGATAGCTCTGGCAAGACCATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTT

TTATCTGTACCCGGGCAAGTATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGA

AGTTGCGACGGCAATCACCTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATG

GCAAGGCTACGAAAGGCGACGCACACATC

A.12 SpyCatcher-I27F-G$_C$

Protein:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST

WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIR

SLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILH

NCQLGMTGEVSFQAANTKSAANLKVKELRSCGDGEWTYDDATKTFTVTE

DNA:

GGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGA

CATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAG

ACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACC

ATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAG

TATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCAC

CTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCG

ACGCACACATCAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTG

TTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGC

CAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAG

GATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGA

GGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATT

GAGATCCTGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGG

TTACCGAA

A.13 SpyCatcher-I27F-G$_C$-I27F-SpyCatcher

Protein:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST

WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIR

SLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILH

NCQLGMTGEVSFQAANTKSAANLKVKELRSCGDGEWTYDDATKTFTVTERSLIEVEKP

LYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILHNCQLGM

TGEVSFQAANTKSAANLKVKELRSGAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKR

DEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAIT

FTVNEQGQVTVNGKATKGDAHI

DNA:

GGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGA

CATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAG

ACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACC

ATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAG

TATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCAC

CTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCG

ACGCACACATCAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTG

TTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGC

CAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAG

GATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGA

GGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATT

GAGATCCTGCGGGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGG

TTACCGAAAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTT

GTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAG

TTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAGGAT

GGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGAGGTT

TCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTGAGA

TCCGGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGG

CGACATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATG

AAGACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAG

ACCATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGC

AAGTATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAAT

CACCTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAG
GCGACGCACACATC

A.14 SpyCatcher-I27F-FN3-I27F-SpyCatcher

Protein:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST
WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIR
SLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILILH
NCQLGMTGEVSFQAANTKSAANLKVKELRSTRLDAPSQIEVKDVTDTTALITWFKPLAE
IDGIELTYGIKDVPGDRTTIDLTEDENQYSIGNLKPDTEYEVSLISRRGDMSSNPAKETFTT
GRSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDGKKHILI
LHNCQLGMTGEVSFQAANTKSAANLKVKELRSGAMVDTLSGLSSEQGQSGDMTIEEDS
ATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPD
GYEVATAITFTVNEQGQVTVNGKATKGDAHI

DNA:

GGTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGA
CATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAG
ACGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACC
ATTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAG
TATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCAC
CTTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCG
ACGCACACATCAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTG

TTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGC

CAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAG

GATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGA

GGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATT

GAGATCCACACGCTTGGATGCCCCCAGCCAGATCGAGGTGAAAGATGTCACAGACA

CCACTGCCTTGATCACCTGGTTCAAGCCCCTGGCTGAGATCGATGGCATTGAGCTGA

CCTACGGCATCAAAGACGTGCCAGGAGACCGTACCACCATCGATCTCACAGAGGAC

GAGAACCAGTACTCCATCGGGAACCTGAAGCCTGACACTGAGTACGAGGTGTCCCT

CATCTCCCGCAGAGGTGACATGTCAAGCAACCCAGCCAAAGAGACCTTCACAACAG

GCAGATCCCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGT

GAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAG

CTGAAAGGACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAGGATGGAAA

GAAGCATATTCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTT

CCAGGCTGCTAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTGAGATCCG

GTGCGATGGTTGATACCCTGAGCGGTCTGAGCAGCGAACAAGGCCAAAGCGGCGAC

ATGACGATTGAAGAAGACTCGGCTACCCACATTAAATTTAGCAAGCGTGATGAAGA

CGGCAAAGAACTGGCAGGTGCTACCATGGAACTGCGCGATAGCTCTGGCAAGACCA

TTAGTACGTGGATCTCCGATGGTCAGGTCAAAGACTTTTATCTGTACCCGGGCAAGT

ATACCTTCGTGGAAACGGCGGCCCCGGACGGTTACGAAGTTGCGACGGCAATCACC

TTCACGGTGAATGAACAGGGTCAGGTTACGGTGAATGGCAAGGCTACGAAAGGCGA

CGCACACATC

A.15 SpyTag-I27F-G_N-I27F-SpyTag

Protein:

AHIVMVDAYKPTKRSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASP

DCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSMDTYKLILNGKTL

KGETTTEAVDAATAEKVFKQYANDNGVGCGLGLIEVEKPLYGVEVFVGETAHFEIELSE

PDVHGQFKLKGQPLAASPDCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKV

KELRSAHIVMVDAYKPTK

DNA:

GCTCATATTGTCATGGTTGATGCTTACAAGCCAACTAAGAGATCCCTAATAGAAGTG

GAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATT

GAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCA

GCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCAT

AACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCT

GCAGCCAATCTGAAAGTGAAAGAATTGAGATCCATGGACACCTACAAACTGATCCT

GAACGGTAAAACCCTGAAAGGTGAAACCACCACCGAAGCTGTAGACGCTGCTACTG

CAGAAAAAGTTTTCAAACAGTACGCTAACGACAACGGTGTCGGTTGCGGACTCGGG

CTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGC

CCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGG

ACAGCCTTTGGCAGCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATAT

TCTGATCCTTCATAACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGC

TAATACCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTGAGATCCGCTCATATTGT

CATGGTTGATGCTTACAAGCCAACTAAG

A.16 SpyTag-I27F-G$_C$-I27F-SpyTag

Protein:

AHIVMVDAYKPTKRSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASP

DCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSCGDGEWTYDDATK

TFTVTERSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASPDCEIIEDG

KKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSAHIVMVDAYKPTK

DNA:

GCTCATATTGTCATGGTTGATGCTTACAAGCCAACTAAGAGATCCCTAATAGAAGTG

GAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATT

GAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCA

GCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCAT

AACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCT

GCAGCCAATCTGAAAGTGAAAGAATTGAGATCCTGCGGGGACGGTGAATGGACCTA

CGACGACGCTACCAAAACCTTCACGGTTACCGAAAGATCCCTAATAGAAGTGGAAA

AGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAAC

TTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTT

CCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACT

GTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAG

CCAATCTGAAAGTGAAAGAATTGAGATCCGCTCATATTGTCATGGTTGATGCTTACA

AGCCAACTAAG

A.17 SpyTag-I27F-FN3-I27F-SpyTag

Protein:

AHIVMVDAYKPTKRSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLAASP

DCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSTRLDAPSQIEVKDV

TDTTALITWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDENQYSIGNLKPDTEYEVSLISR

RGDMSSNPAKETFTTGRSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQFKLKGQPLA

ASPDCEIIEDGKKHILILHNCQLGMTGEVSFQAANTKSAANLKVKELRSAHIVMVDAYK

PTK

DNA:

GCTCATATTGTCATGGTTGATGCTTACAAGCCAACTAAGAGATCCCTAATAGAAGTG

GAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATT

GAACTTTCTGAACCTGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCA

GCTTCCCCTGACTGTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCAT

AACTGTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCT

GCAGCCAATCTGAAAGTGAAAGAATTGAGATCCACACGCTTGGATGCCCCCAGCCA

GATCGAGGTGAAAGATGTCACAGACACCACTGCCTTGATCACCTGGTTCAAGCCCCT

GGCTGAGATCGATGGCATTGAGCTGACCTACGGCATCAAAGACGTGCCAGGAGACC

GTACCACCATCGATCTCACAGAGGACGAGAACCAGTACTCCATCGGGAACCTGAAG

CCTGACACTGAGTACGAGGTGTCCCTCATCTCCCGCAGAGGTGACATGTCAAGCAAC

CCAGCCAAAGAGACCTTCACAACAGGCAGATCCCTAATAGAAGTGGAAAAGCCTCT

GTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGA

ACCTGATGTTCACGGCCAGTTTAAGCTGAAAGGACAGCCTTTGGCAGCTTCCCCTGA

CTGTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTGTCAGCT

GGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATACCAAATCTGCAGCCAATCT

GAAAGTGAAAGAATTGAGATCCGCTCATATTGTCATGGTTGATGCTTACAAGCCAAC

TAAG

A.18 Determination of Mw and Mn

$$\mathrm{Mn} = \sum \mathrm{N_i M_i} \Big/ \sum \mathrm{N_i};$$

$$\mathrm{Mw} = \sum \mathrm{N_i M_i^2} \Big/ \sum \mathrm{N_i M_i};$$

$$\mathrm{Abs}_{280nm}(\mathrm{i}) \propto \mathrm{N}_i$$

where N_i is the number of moles of polymer that have a molecular weight M_i and Abs_{280nm} is the absorbance at 280 nm.

M_i is calculated from the calibration curve formed by using standard proteins.

$$\log \mathrm{M_i} = -1.3766 \frac{\mathrm{V_e}}{\mathrm{V_o}} + 4.2051$$

where V_e is the elution volume and V_o is set to be 40 mL.