EXPLORING AND EXPLOITING COLLAGEN MIMETIC PEPTIDES

By

Aubrey J. Ellison

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

Date of final oral examination: 12/13/2017
The dissertation has been approved by the following members of the Final Oral Committee:
Ronald T. Raines, Professor, Biochemistry, Organic Chemistry, Chemical Biology
Helen E. Blackwell, Professor, Chemical Biology, Organic Chemistry
Sandro Mecozzi, Associate Professor, Organic Chemistry, Pharmaceutical Sciences
Doug B. Weibel, Professor, Chemical Biology, Biochemistry

EXPLORING AND EXPLOITING COLLAGEN MIMETIC PEPTIDES

Aubrey J. Ellison

Under the supervision of Professor Ronald T. Raines

At the University of Wisconsin–Madison

Collagen is the most abundant protein in humans and is a major structural component of the extracellular matrix. Damage to its triple helix structure occurs during wounding or through abnormal remodeling caused by pathological diseases. Collagen mimetic peptides (CMPs) are an emerging biomedical tool to that can be used to identify denatured collagen through hybridization with the host tissue by the creation of a triple helix with the damaged strands. This thesis aims to address deficiencies in the field of invasive CMPs by improving the synthesis, expanding the scope, and evaluating the mechanism of binding.

A facile synthetic method is crucial to enable rapid access to these complex molecules. Time and material intensive chromatography is commonly used in both the solution-phase and solid-phase synthesis of peptides. By utilizing coupling reagents that generate readily separable byproducts, chromatographic separations can be avoided. This methodology was used to access tripeptides related to collagen with purity suitable for solid-phase segment condensation to generate CMPs.

CMPs have been used for payload delivery to bind their cargo to host collagen. This payload delivery approach can be applied to the development of long-lasting sunscreens. This work provides evidence for a generalizable strategy using CMP conjugates for broad spectrum skin-adherent UV protection.

Nature might utilize an invasive strand mechanism for adherence to a collagen surface. Group A *Streptococcus* (GAS) displays cell-surface proteins that resemble human collagen. We

hypothesize that the collagen-like strands of GAS form triple helices with denatured host collagen in a wound bed. With CMPs as tools, we have provided the first evidence that bacterial proteins Scl1 and Scl2 can adhere to mammalian-like collagen.

An important corollary of our work with invasive CMPs is understanding whether their invasion occurs in wounded tissue through completion of a triple helix. A collagen duplex, termed the “nest,” was envisioned and synthesized as a wound-bed mimic. Known and novel CMPs were compared for their propensity to bind to the nest to form a stable hetero-triple helix.

The strategies and results reported in this dissertation expand the uses of invasive CMPs and explore their mechanism of binding.

Acknowledgements

First and foremost, I am grateful to Professor Ron Raines for being my mentor and research advisor. I appreciate the time that Professor Raines has put towards helping me develop as a scientist. He promotes scientific independence and has allowed the freedom to pursue my own areas of scientific interest while still receiving guidance when needed. Professor Raines also provided an outstanding environment to conduct science and set me up for success by giving me access to outstanding instrumentation, facilities, and lab full of terrific colleagues.

I would like to thank the many staff members that work in at the on campus facilities. They run the instruments that are essential to productive science. I would particularly like to thank Dr. Elle Grevstad of the Biochemistry Optical Core, Dr. Darrell R. McCaslin of the Biophysics instrumentation Facility, Dr. Rachael Sheridan of the UWCCC Flow Cytometry Laboratory, Dr. Martha M. Vestling of the Mass Spectrometry Facility, and Dr. Melissa Boersma and Nina Porcaro of the former UW Peptide Synthesis Facility. All of which have not only helped with instrumentation but also insightful discussions.

I am also grateful to my collaborators for my various projects. Professor Daniel Kearns and Dr. Felix Dempwolff at Indiana University created the knock-in *B. subtilis* model discussed in Chapter 4. Professor Angela Gibson and Edgar Ocotl provided human surgical samples and laboratory space to process samples for developing and *ex vivo* model for testing the sunscreen conjugates discussed in Chapter 3. Much of the work contained within would not be achievable without their participation and contribution.

I would like to thank my thesis committee, Helen Blackwell, Sandro Mecozzi, and Doug Weibel. All of which whom have provided support by serving on committees for degree requirements and providing helpful direction though out my time in graduate school. I am

particularly appreciative to Doug Weibel for providing me with laboratory space and resources for the final stretch of my graduate experience.

I am very grateful for my time as part of the Chemical biology training grant. Not only did this provide funding to allow me the freedom to focus on my research efforts, but the grant provided a community of graduate students that worked to support each other's research. The monthly meeting and literature discussion course helped develop my scientific communication skills. Through the training grant I was able to partake in an internship working at Saint Olaf college teaching organic chemistry.

My colleagues within the Raines group have also been of tremendous help in my development as a scientist. I would like to thank Dr. Brett VanVeller and Dr. I. Caglar Tanrikulu for their guidance and mentorship during the early years of my studies and for teaching me the fundamentals of peptide chemistry. I would like to give special thanks Kalie Mix who joined the Raines group in the same year and was my partner through all the requirements of graduate school. I also need to give a special thanks to Kristen Andersen, Emily Garnett, Trish Hoang, and Val Ressler for their advice and guidance on everything biochemistry. I had the opportunity to work with many talented scientists in the Raines group such as Matt Aronoff, Christine Bradford, Ho-Hsuan Chou, Wen Chyan, Kevin Desai, Chelcie Eller, Jesus Dones, Brian Gold, Brian Graham, Sean Johnston, Henry Kilgore, Joelle Lomax, John Lukesh, Nick McGrath, Robert Newberry, Lindsey Orgren, Rob Presler, Thom Smith, Kaylee Underkofler, Jim Vasta, and Ian Windsor.

My friends and family have helped to keep me sane over the course graduate school. They remind me to have fun. I am very lucky to have made some great friendships while in school that have made Madison a great place to call home.

I will forever be indebted to my parents, Paul and Jenny. I can never appreciate enough their constant love and support and being such terrific role models throughout my life. From a young age, they taught me the importance of education, a strong work ethic, and to persevere through the difficult and challenging times.

Most importantly, I am so grateful for my husband's love and support throughout graduate school. Brad has been with me through every step of the program and I look forward to the new adventures that we will create after our time as students.

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations

aa	amino acid
Ac	acetyl
AHX	6-aminohexanoic acid
amu	atomic mass unit
ApoB	apolipoprotein B
Arg, R	arginine
Asp, D	aspartic acid
AUC	analytical ultracentrifugation
BH	brain heart
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
Cbz	carboxybenzyl
CD	circular dichroism
cDNA	complementary deoxyribonucleic acid
CHP	collagen hybridizing peptide

CMP	collagen mimetic peptide
C-terminal	carboxy-terminal
Cys. C	cysteine
DCE	dichloroethane
DCM	dichloromethane
DI	deionized
DIC	*N*,*N*'-diisopropylcarbodiimide
DIEA	diisopropylethyl amine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sufoxide
DNA	deoxyribonucleic acid
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
ELISA	enzyme-linked immunosorbent assay
ESI	electrosplay ionization

flp	(2*S*,4*S*)-fluoroproline
Flp	(2*S*,4*R*)-fluoroproline
Fmoc	9*H*-fluoren-9-ylmethoxycarbonyl
FRET	Förster resonance energy transfer
GAS	group A *Streptococcus*
Glu, E	glutamic acid
Gly, G	glycine
HATU	1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HBTU	*0*-benzotriazole-*N*,*N*,*N ',N '*-tetrameth yl-uronium-hexafluorophosphate
HGF	human growth factor
HOBt	1-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
Hyp, O	(2*S*,4*R*)-hydroxyproline
Ile, I	isoleucine
ITPG	Isopropyl-β-D-1-thiogalactopyranoside
LB	Lysogeny broth

LDH	lactate dehydrogenase
LDL	low-density lipoprotein
Leu, L	leucine
Lys, K	lysine
MALDI–TOF	matrix-assisted laser desorption/ionization–time-of-flight
Met, M	methionine
MMP	matrix metalloproteinase
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MSCRAMM	microbial surface components recognizing adhesive matrix molecules
MTT	(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)
NHS, -Osu	*N*-hydroxysuccinimide
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
N-terminal	nitrogen-terminal
PBS	phosphate-buffered saline

PCR	polymerase chain reaction
Pfp	pentaflurophenyl ester
Pro, P	proline
PyBrOP	bromo-tris-pyrrolidino-phosphonium hexafluorophosphate
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
sal	salicylic acid
Scl	*Streptococcal* collagen-like protein
Ser, S	serine
SHG	second harmonic generation
SPPS	solid-phase peptide synthesis
Sub P	substance P
*t*Bu	*tertiary*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPSH	triisopropylsilane

T_m	melting temperature
Trp, W	tryptophan
Trt	trityl
Tyr, Y	tyrosine
UPLC	ultra-performance liquid chromatography
UV	ultraviolet
Val, V	valine
VEGF	vascular endothelial growth factor
vis	visible
WT	wild type

Chapter 1

Targeting and understanding atypical collagen

1.1 Abstract

Collagen is the most abundant protein in humans and is a major structural component of the extracellular matrix (ECM). Collagen can become damaged when wounded or through abnormal remodeling caused by pathological diseases. These damaged tissues can be targeted for imaging and drug delivery by recognition of the unique triple-helical structure of collagen. Specific recognition of collagen that has been denatured from a triple helix can shed further light on disease progression. Collagen mimetic peptides (CMPs) are an emerging biomedical tool to that can be used to identify denatured collagen by hybridization with the host tissue through the creation of a triple helix. Atypical collagen can also be found outside of pathological disease states. Non-metazoan collagen maintains the canonical triple helix, but stability is achieved through different means than with mammalian collagen. This chapter provides a review of recent developments in the field of collagen targeting for studying atypical collagen.

1.2 Introduction

Collagen, a structural protein found in all connective tissue, is the most abundant protein in animals and is thus the most studied biomolecule in the extracellular matrix (ECM). Collagen plays a crucial role in tissue development and regeneration. Collagen can become damaged when skin is wounded or through abnormal remodeling caused by pathological diseases such as cancer, osteoporosis, arthritis, and fibrosis.[1-7] These damaged tissues can be targeted for imaging and drug delivery by recognition of the unique triple helical structure of collagen. Specific recognition of collagen that has been denatured from a triple helix can shed further light on disease progression. Additionally, targeting denatured collagen allows for delivery of therapeutics in high local concentrations to the area most in need.

Developing targeting agents requires an understanding of the structure and stabilizing factors of the collagen triple helix. Collagen mimetic peptides (CMPs) are an emerging biomedical tool that can be used to identify denatured collagen by hybridization with the host tissue through the creation of a triple helix (Figure 1.1). Incorporation of non-natural amino acids in to CMPs can modulate triple-helix hybridization.

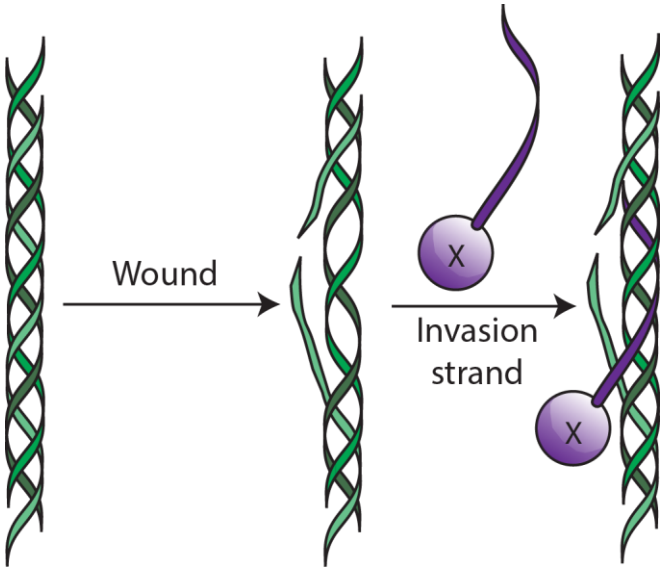


Figure 1.1. Scheme of invasive strand mechanism by CMPs to damaged host collagen.

Atypical collagen can also be found outside of pathological disease states. Non-metazoan collagen maintains the canonical triple helix, but stability is achieved through different means than mammalian collagen. These collagen sequences introduce new ways to stabilize triple helices and have inspired new ways to study collagen sequences and biomaterials. Still, the biological function of non-metazoan collagens is comparatively poorly understood.

1.3 Collagen structure and stability

Collagen is the most abundant protein in vertebrates and invertebrates. The role of collagen is primarily structural in connective tissues such as skin, bone, tendons, or cartilage. Mechanical stability and elasticity is derived from the cable-like protein arranging in fibrils. There are 28 types of collagen classified by their function, domain architecture, and supramolecular organization.[8] The main fibril forming collagens (I, II, III, V and XI) provide tensile strength in animal skin, bones, cartilage, and blood vessels; whereas other types of collagen are network forming and present in basement membranes (type IV) and vascular sub-endothelial matrices (type VIII).

All collagens are trimeric proteins formed by the association of three polypeptide chains in polyproline-II–like helices that come together in a right-handed triple helix.[9] Each strand of collagen is a repeat of Xaa-Yaa-Gly, where (2*S*)-proline (Pro) and (2*S*, 4*R*)-hydroxyproline (Hyp) are most common in the Xaa and Yaa positions, respectively.[10] Hydroxylation is achieved biologically by the post-translational modification of proline by prolyl hydroxylase.

Coiling of the triple helix occurs with a one residue offset of the polypeptide strands such that a cross-section contains Xaa, Yaa, and Gly from the three strands (Figure 1.2). As such, interstrand N-H···O=C hydrogen bonds create an intermolecular backbone holding adjacent strands together.[11,12] The repetitive sequence is also important for imparting triple-helix stability. Glycine at every third residue allows close packing around a central axis for trimerization.[11] Substitution with bulkier amino acids in this position are not tolerated and disrupt triple helix formation.[13] Additionally, an abundance of Pro residues predisposes the φ and ψ angles of the amino acid for triple-helix formation (Figure 1.3).[14]

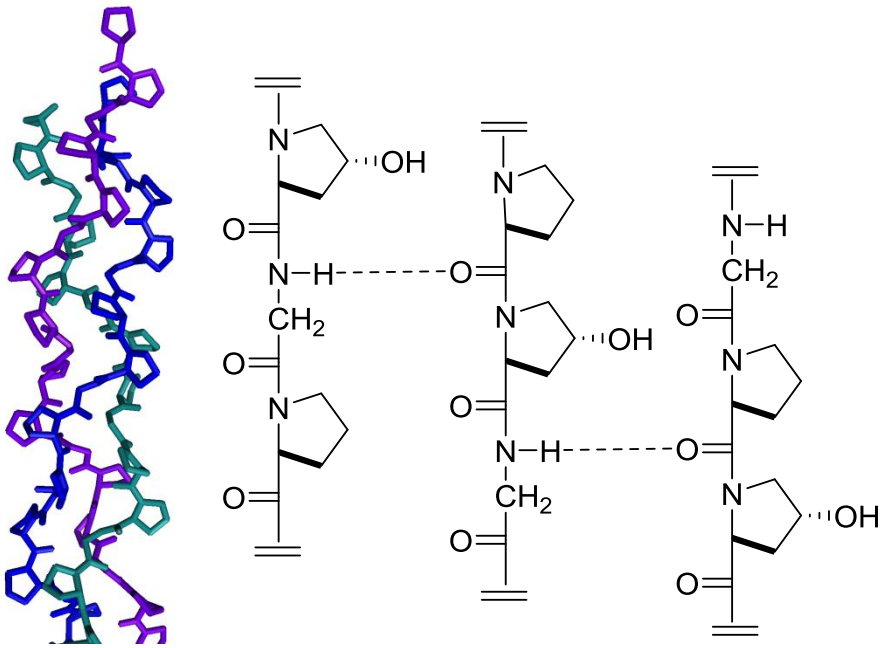


Figure 1.2. Collagen triple helix (PDB ID: 1K6F) and a representation of intrastrand hydrogen-bonding.

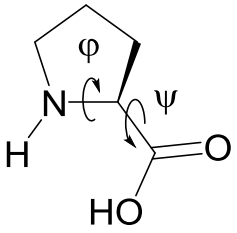


Figure 1.3. Restricted rotation of proline leads to preorganization of the φ and ψ bond angles of proline.

Hyp in the Yaa position also contributes to complex stability.[15] Like Pro, Hyp is conformationally restricted and thus preorganized to favor triple-helix formation. Nonetheless,

Hyp has an additional stabilizing effect, as demonstrated by the differences in thermal stability between hydroxylated (T_m = 39 °C) and non-hydroxylated (T_m = 24 °C) human type I collagen.[16]

Originally, the added stability of Hyp was solely attributed to hydrogen bonding with surrounding water to support triple-helix structure.[11,17-20] Ramachandran also suggested that pyrrolidine ring pucker caused by hydroxyproline helped stabilize collagen.[20] The crystal structure of collagen revealed a conformational preference of the pyrrolidine ring pucker in the Xaa and Yaa positions.[21] Pro can adopt two conformations C^{γ}-endo and C^{γ}-exo. Xaa adopts a C^{γ}-endo conformation, whereas Yaa prefers C^{γ}-exo (Figure 1.4).[22] Hyp induces a C^{γ}-exo ring pucker through the gauche effect.[23,24] As Hyp is found in the Yaa position, it provides stability to the folded structure through entropic effects. The induced C^{γ}-exo ring pucker is better preorganized to the torsion angles required at that position in the triple helix.

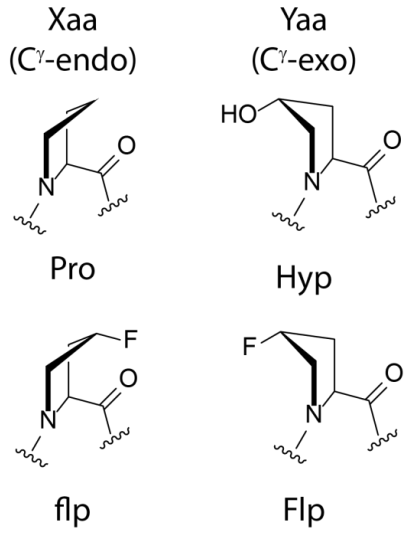


Figure 1.4. Pyrrolidine ring puckers preferred in the Xaa and Yaa position of natural and fluoro substituted proline. Image adapted from Chattopadhyay *et al.* 2014.

Although Pro and Hyp are highly prevalent in the collagen sequence, other amino acids are also incorporated into the Xaa and Yaa positions.[10] Incorporation of non-pyrrolidine side chains can play a major role in the structure and stability of collagen fibers. Met and Lys for example, can spawn crosslinks in the fibrils.[25] Integration of other amino acids can tune stability further in localized areas or provide areas for enzyme recognition.

The Brodsky lab has systematically studied the effect of incorporating different amino acids in the collagen sequence on the stability of the triple helix.[26-28] These studies were accomplished though incorporation of the guest Xaa-Yaa-Gly trimer of interest into a host collagen peptide. Most notably, Arg in the Yaa position showed stability similar to that of Hyp.[27] From these studies, an algorithm was derived to predict collagen triple-helix stability based on sequence (Collagen Stability Calculator, http://compbio.cs.princeton.edu/csc).[29]

1.4 Hyper stable collagen mimetic peptides

Many substituted prolines have been of interest to tease out an understanding of the stability of the collagen triple helix. In studies to deconvolute the potential contributions of Hyp to the collagen triple-helix stability, Hyp was changed to (2*S*,4*R*)-fluoroproline (Flp).[23,24] The fluoro moiety on the 4-position maintains electronegativity while no longer being capable of participating in hydrogen bonding. The peptide $(Pro\text{-}Flp\text{-}Gly)_7$ not only retained triple-helix stability, but also increased the melting temperature to 45 ºC compared to 36 °C for Hyp-containing collagen mimetic peptide.[30]

The work done by the Raines group showed that the extra stability is derived from the gauche effect preorganizing the pyrrolidine to C^{γ}-exo ring pucker as preferred in the Yaa position for triple-helix formation.[30,31] Fluorine, being more electronegative than oxygen, further

preorganizes the ring pucker. Substitution in the Xaa position with diastereomeric (2*S*,4*S*)-fluoroproline (flp) also increased stability, and this increase is attributable to C^{γ}-endo ring pucker.[32,33] The ring pucker can also be controlled by a steric effect.[34]

A hyperstable collagen triple helix can also be achieved through molecular engineering of an additional hydrogen-bonding network. Although Gly substitution is typically invariable, the Chenoweth group introduced aza-Gly into a guest trimer of a host CMP.[35,36] Replacing the -CH_2 group with an –NH group maintains the steric requirements for close packing while adding a interchain hydrogen bond that stabilizes the collagen triple helix.[37] Incorporating aza-Gly into every trimeric repeat, a peptide with 12 residues, (Pro-Hyp-aza-Gly)$_4$, represents the shortest peptide to date that is capable of triple helix self-assembly at a physiologically relevant temperature (T_m = 36 °C).[36]

Addition of fluoroproline and aza-Gly in the collagen triple helix present a unique way to modulate stability of collagen mimetic peptides. Using understanding of collagen stability and structure, collagen mimics containing triple-helix stabilizing factors could improve and lead to new biomaterials for restorative therapies, drug delivery, or as new chemical tools for exploring fundamental aspects of collagen biology and recognition.

1.5 Non-metazoan collagen

An emerging area in collagen study is collagen found in non-metazoan sources, such as prokaryotes and viruses. Collagens are present in the most primitive animals, such as sponges, and therefore are considered intrinsic to the evolution of metazoans.[38,39] 18,874 collagen-like proteins in bacteria, 695 in viruses, and 157 in archaea have been identified by genome mining.[40] Björck and coworkers proposed horizontal transfer of collagenous sequences from eukaryotes to

prokaryotes.[41] However more recently work by McConkey and coworkers. suggested convergent evolution by via simple Xaa-Yaa-Gly repeat amplification.[42]

Recombinate expression of these genes have allowed for confirmation and evaluation of collagen triple-helix formation. Prokaryotic collagens showed thermostabilities close to those of vertebrate collagens. Still, sequence composition differs greatly from the eukaryote counterparts. In bacteria, the Xaa position contains Pro in more than 30% of the trimeric repeats, similar to human type I collagen.[43] The Yaa position is contains Pro in only 5% of repeats, and there is absolutely no Hyp there, as bacteria lack the enzyme to hydroxylate. Pro and Hyp are important for triple-helix stability in mammalian collagen (*vida supra*). There is an abundance of charged residues: Glu/Asp in the Xaa position and Lys/Arg in the Yaa position.[44,45] In the triple helix formed by the non-metazoan collagens, the thermostability derives from favorable electrostatic interactions and hydration-mediated hydrogen bonding networks.[46,47]

These new unique sources of collagen triple-helix stability are of interest as new tools to study the collagen triple helix. Bacterial collagen-like proteins have been expressed recombinantly with genetic modification to incorporate human collagen interaction motifs. The bacterial collagen sequence acts as a triple-helical template that helps to fold embedded human collagen interaction sites that are too short to form a triple helix on their own. Integrin, fibronectin, and heparin binding as well as matrix metalloproteinase (MMP) cleavage sequences have been identified with this strategy.[48-52] This expression system acts to complement peptide studies.

Mammalian collagen has been used for developing biomaterials.[53] Still, bulk recombinant expression systems have proved challenging with the need to hydroxylate Pro in the Yaa

position. As stability of the triple helix from non-metazoan collagen arises from interactions independent of hydroxylated proline, these collagen sequences can be produced in large scale by recombinant expression.[54,55] Non-metazoan collagen provides a new and different way for engineering biomaterials.[56]

Collagens in higher eukaryotes are evolved as components of the ECM that support tissue structure and provide an essential network for cell function. Still, the role of collagen in non-metazoans is less well understood. *Streptococcal* collagen-like (Scl) protein from *S. pyogenes* was reported in 2000 by Lukomski *et al.*[45] In subsequent years, collagen-like proteins from pathogenic *E. coli*,[44] *Bacillus anthracis*,[57] *Legionella pneumophila*,[58] *Burkholderia* spp.,[59] and other *Streptococcal* bacteria[40] have been identified and studied. In the most well-studied Scl from *S. pyogenes*, the collagen-like sequence is sandwiched between a membrane-display sequence and a variable domain.[45,60] These proteins have been evaluated extensively as adhesins and play a role in virulence.[40] These Scl proteins have been implicated in binding to a wide variety of host surface factors including fibronectin, laminin, integrins α2β1 and α11β1, thrombin-activable fibrinolysis, LDL, ApoB, and Factor H.[61-69] In addition to binding host factors, Scl proteins may also play a role in biofilm formation, neutrophil evasion, and protection against host-defense peptides.[40,70,71]

The virulence activity has been attributed to the variable domain. The collagen sequence is hypothesized to again play a structural role and act as a ridged assembly on which the variable domain is projected away from the cell surface. Even the binding of fibronectin and laminin to Scl proteins has been shown to be mediated though the variable domain.[61,72] Still, in some serotypes the collagen sequence contains a binding motif for integrins.[62-65,73] There are other

adhesins that have similar host targets that do not contain a collagen sequence,[74,75] which could mean that the collagen sequence evolved for more than just a structural role.

Although initial work in the field has been focused on collagen-like proteins in bacterial pathogens, only a small proportion of predicted non-metazoan collagen proteins have been investigated thus far. There is still much to understand about the role of collagen sequences in bacteria as well as other non-metazoans. Techniques and tools used to evaluate mammalian collagen described below could be applied here for imaging or pathogen targeting.

1.6 Collagen binding and detection

1.6.1 Triple-helix targeting

The triple-helix structure is unique to collagen, barring a few exceptions of small sub-domains of proteins.[76-82] This distinction allows for the specific targeting of collagen through recognition of its triple helix. As collagen is found in most tissue types, it can play a role in different diseases associated with collagen degradation and remodeling, such as cancer, osteoporosis, arthritis, and fibrosis.[1-7] Understanding how collagen manifests in normal tissues and in a pathological state is of interest for understanding how a disease progresses and for disease detection.

First used in microscopy in 1974, second harmonic generation (SHG) became prevalent in the imaging of collagen in the early 2000s.[83] SHG is a nondestructive method in which 3D images of intact, non-fixed, and non-stained tissues can be obtained.[84,85] The ability to generate second harmonics is peculiar to molecules that are not centrosymmetric, such as the collagen triple helix. The crystalline arrangement of collagen enhances this signal, making collagen fibers

ideal for imaging with SHG. This imaging technology is capable of providing pictures of the structural arrangement of collagen in many different tissues.

SHG has been used to study the normal and disease states of fibrillar collagen. Using SHG, tumor-associated collagen signatures were identified in many cancers,[86] increases in collagen fibers in different tissues with fibrosis can be monitored,[87] and depletion of collagen due to photo-aging is detectable.[88] A more comprehensive discussion of resent uses of SHG in collagen imaging can be found in a review by Hackett and coworkers.[89] Although SHG has been utilized often for detection of abnormal collagen arrangement, the collagen imaged is still in the canonical triple helix. SHG lacks sensitivity for detecting collagen as it is reordering or when it is damaged and denatured.

Selective recognition by binding collagen also allows for tissue imaging, but more importantly potentiates itself toward drug delivery. There are many biological factors that interact and bind to the collagen triple helix. Well-known collagen binders in the human body are fibronectin, integrins, and laminin.[90-92] The natural binding of the triple helix can be exploited through the creation of biological hybrids for imaging and targeted delivery to collagen. One such example is through the fusion of the collagen-binding domain of fibronectin and growth factors EGF or HGF.[93-95] Both fusions showed prolonged retention and improved wound healing.[96,97] De Crescenzo and coworkers have written a comprehensive review of utilizing collagen binding domains for the delivery of growth factors.[98]

Pathogens have also developed means to bind to the ECM as a method for infection.[99] These surface adhesins that bind to the ECM are known as microbial surface components recognizing adhesive matrix molecules (MSCRAMMs).[100] The most studied collagen-binding

MSCRAMMs are CNA from *Staphylococcus aureus*[101] and ACE from *Enterococcus faecalis.*[102] Collagen-binding MSCRAMMs have also been found in the pathogens *Enterococcus faecium,*[103] *Streptococcus equi,*[104] *Streptococcus mutans,*[105] *Streptococcus pyogenes,*[106] *Erysipelothrix rhusiopathiae,*[107] and *Bacillus anthracis.*[108]

CNA binds the collagen triple helix through a collagen "hug" mechanism.[109] Recombinantly expressed CNAs conjugated to fluorophores have been used to image collagen in different tissues.[110,111] This method of detection was applied to the evaluation of different pathologies, such as fibrosis and atherosclerotic sites.[112,113] The Merkx group developed a multivalent interaction by displaying CNA35 on a micelle.[114] Not only did this approach showed improved affinity, but the micelle could be used to deliver drugs to the collagen surface.

CNA-binding is sensitive to collagen stretching.[115] The work done by Mofrad and coworkers suggests that CNA would bind stronger to collagen that is relaxed such as that found in areas of inflammation or a wound site. This differential binding means that that CNA conjugates for imaging or delivery could be used to target areas of collagen remodeling. Binding is still dependent on the presence of a fully formed triple helix and would not represent collagen that is denatured or actively undergoing remodeling.

Another difficulty with collagen-binding protein conjugates is that they are large structures that suffer from low diffusion into tissues. Peptides, being more compact, have improved diffusion. Using phage display and the incorporation of nonnatural amino acids, Caravan and coworkers created a type I collagen-targeting peptide.[116] Appended with an MRI contrast agent, myocardial scarring was visualized in mice.[116,117] Peptides that bind to collagen have also been derived from its binding-domain proteins.[98] Peptides based on decorin show

strong adherence to collagen matrices, can modulate fibrillogenesis, and reduce dermal scarring.[118-120] As a way to improve peptide binding affinity and to avoid incorporation of nonnatural amino acids, Meijer and coworkers attached collagen-targeting peptides by dendrimeric display to allow for multivalent interactions with a surface.[121]

Antibodies are highly versatile for recognition and selective detection of an antigen. Using soluble collagen as immunogens, antibodies have been developed to recognize the triple helix of different types of collagen.[122,123] Through detection of the triple helix, Phan and coworkers. monitored production of collagen in fibrosis from cultured cells.[124] ELISA and immunoblotting methods for detection and quantification of collagen are also accessible by using antibodies.[125]

The methods of detection described above are limited to fully formed collagen triple helices. As such, these methods cannot readily distinguish healthy collagen from diseased tissue. Additionally, binding of the triple helix is limited to the detection for collagen in many pathological conditions in which the triple helix has been damaged or is undergoing remodeling. Methods directed at recognizing denatured collagen could allow for early detection of different pathological states.

1.6.2 Recognition of damaged collagen

The recognition of damaged collagen proves to be more difficult than the recognition of intact collagens as the canonical triple helix is no longer present for selective recognition. Historically, efforts for detection have been through indirect means, such as the evaluation of biological markers that are representative of collagen degradation or remodeling.[7,126] Direct detection through binding of damaged collagen can provide a new, more detailed view of the

diseased landscape. Sensitivity to denatured collagen also allows for therapeutic delivery to the area most in need.

Remodeling of the ECM is carried out by MMPs. In many pathological diseases, detection of protease activity can be indicative of damaged collagen. Merkx and coworkers modified CNA35 to be cyclic, blocking the ability to bind collagen until proteolytically cleaved and activated by MMP1.[127] As an alternative method, a collagen triple-helix mimic was conjugated to CNA35 via a long flexible linker.[128] The tethered triple helix acts as an inhibitor that is cleaved and released by MMP1 activity, freeing CNA35 to bind native collagen. This semi-synthetic method could be applied to other collagen-binding proteins. Protease sensitive collagen-binding allows for selectivity to those tissue areas undergoing collagen remodeling.

The use of denatured collagens as immunogens is attractive, as antibodies could be used to target remodeling collagen directly.[3,122] As a means to detect and specifically bind denatured collagen, monoclonal antibodies were created with thermally denatured type I and IV collagens.[129] The resulting antibodies, HUIV26 and HUI77, bind to denatured and proteolyzed collagen with little activity toward collagen in a triple helix. Binding is selective for cryptic sites that are exposed only after disruption of the triple helix. HU177 and the humanized version, D93, showed affinity for damaged collagen of types I, II, III, IV, and V.[130] D93 has been used to target cryptic sites associated with angiogenesis and tumor growth inhibiting activity.[130,131]

Peptide TLTYTWS, selected by phage display, binds to proteolyzed type IV collagen.[132] Like with antibodies, the peptide recognizes a cryptic site that is revealed when the triple helix is denatured and that is associated with angiogenesis and tumor progression. Upon binding, TLTYTWS inhibits angiogenesis *in vivo*.

Targeting denatured collagen through immunization and phage display approaches have led to new knowledge on the role that denatured collagen plays in pathological progression. Cryptic sites behave as biological signals affecting the fate of tissue repair. Binding of damaged collagen not only allows for new ways to image disease states but might also prove useful as a therapeutic strategy.

1.6.3 Invasive CMPs

Natural collagen contains loops or interruptions in its triple helix.[133-136] Further disruption to the triple helix can come about through reordering in different disease states. Additionally, the collagen triple helix can be disrupted through physical damage, such as a wound, and then reordered in the healing process.[137] All of these domains in which the triple helix has been broken are accessible to hybridization with collagen mimetic peptides (CMPs).[138] CMPs are hypothesized to act by invading and intercalating into the host collagen to complete a triple helix, thereby binding the CMP to the collagen surface. These invasive CMPs, also called collagen hybridizing peptides (CHPs), provide a unique way to target collagen for detecting damage or drug delivery.

An early example of invasive CMPs was from the Yu group who worked with gold nanoparticles covered with CMPs of Pro-Hyp-Gly repeats.[139] As it is hypothesized that the CMPs act by completing a triple helix, it was expected that the coated gold nanoparticles would congregate in the gaps in collagen fibers. Nanoparticle labeling was clustered to only where gaps in the collagen triple helix were expected. Upon heating of the collagen fiber to denature the collagen, the nanoparticles indiscriminately labeled the fiber. The nanoparticles were able to demonstrate selectivity for gaps and denatured collagen fibers with no additional nonspecific binding.[140] Nevertheless, the CMPs did have a propensity for self-assembly when the

nanoparticles were coated with a high density of CMPs, limiting their utility as an imaging agent.[140,141]

Building off of this initial work with CMP coated nanoparticles, the Yu group moved to using fluorophore functionalized CMPs, again using Pro-Hyp-Gly repeats of 6–10, to label collagen liver sections.[142] These CMPs were able to detect collagen damage caused by detergents or mechanical stress in different tissues through binding to the damaged collagen.[143,144]

Furthermore, the invasive CMPs developed by the Yu group have been used to modify synthetic collagen surfaces. By conjugating the peptides to an anionic scaffold, they were able attract vascular endothelial growth factor (VEGF) to the surface by charge–charge interactions.[145] They were also able to immobilize an angiogenic signals directly on a collagen hydrogel and thereby promote cellular vascularization.[146-148]

As CMPs with Pro-Hyp-Gly repeats will form homo-triple helices spontaneously, the Yu group heat-denatured their CMPs prior to use to release monomeric CMPs that are then available to hybridize with a collagen surface. Heat denaturation to form monomeric CMPs needs to be done a short time before of application. The resulting hot solutions were then applied to samples, which could cause damage to the tissues.

The Yu group addressed this problem by creating a “caged” CMP.[149] These caged CMPs are conformationally locked but upon cleavage of the cage, the CMP is able to form a triple helix. To construct the cage, a glycine N is labeled with a nitrobenzyl group to form a tertiary amide. As substitutions to glycine that add steric bulk are not tolerated in the formation of a triple helix, the CMPs remain monomeric. The nitrobenzyl is cleaved with ultraviolet light, and the unhindered glycine is restored, allowing for triple-helix formation. Because hybridization

with a collagen surface only occurs upon cleavage of the cage, these CMPs were useful for highly tuned patterning on a synthetic collagen surface.[150]

The caged CMPs were able to label collagen rich tissues after cage-cleavage.[151] Labeling and detection of collagen showed improvement over other methods such as antibody recognition. Additionally, they were able to detect damaged collagen in osteoarthritis, myocardial infarctions, glomerulonephritis, pulmonary fibrosis, endochondral ossification, and skin aging.[152] With no necessity to heat CMP solutions prior to application and with proof of high serum stability, the Yu group was able to visualize collagen remodeling *in vivo*.[153] They were able to detect high MMP activity such as that found around tumors.[149] Any approach that employs Pro-Hyp-Gly repeats is, however, flawed because of the strong tendency of such CMPs to form inert trimers rapidly.

The Raines group took a different approach to creating invasive CMPs that do not homotrimerize but still have the capability to form a hetero-triple helix with other strands. To accomplish this goal, they utilized an understanding of stability provided by proline and substituted prolines in the collagen triple helix.

Lacking the stability from hydroxyproline, the peptide $(Pro\text{-}Pro\text{-}Gly)_7$ does not form a homo-triple helix. Fluoroproline is known to increases triple-helix stability (*vide supra*). Incorporation of fluoroproline in both the Xaa and the Yaa positions to make $(flp\text{-}Flp\text{-}Gly)_7$, has, however, a destabilizing effect.[154,155] Homo-triple helix formation does not occur because of steric interactions between the fluorine atoms on flp and Flp residues of component strands.[13,154] In contrast, a hetero-triple helix between $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ strands readily forms

in a 1:2 and 2:1 ratio, providing evidence that these peptides could interact and bind to host collagen surface.[154]

Both peptides are observed to bind to natural collagen surfaces *in vitro* and *ex vivo*.[156] $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ were the first examples of invasive CMPs that needed no pretreatment of heat or light to bind to host tissue. These peptides were immediately viable for *in vivo* application.[157] Skin acts as a protective barrier to the outside world, and any break must be repaired quickly and efficiently. As collagen is a major component in skin, invasive CMPs can be used to deliver healing factors to a wound. Substance P (SubP), which promotes wound healing, was conjugated CMP $(Pro\text{-}Pro\text{-}Gly)_7$. With a single application, excision wounds on mice showed faster wound healing than did applying SubP or CMP strands alone. As binding of the CMPs occurs where the host collagen is damaged, the CMPs will localize to where intervention is required. This strategy could be particularly useful for treating chronic wounds that need additional encouragement to properly heal.

Lee and coworkers also employed a similar method by engineering phages with Pro-Pro-Gly repeats that targeted abnormal collagen produced by lung carcinoma cells.[158] This work showed that genetically incorporated short collagen sequences can behave like invasive CMPs. Once again, the tethering of collagen repeats that can hybridize with host collagen lends itself to a useful method for imaging denatured and damaged collagen or for targeted drug delivery.

The Wennemers group has recently reported pH-responsive CMPs.[159,160] These CMPs are incapable of forming a triple helix until in a solution of proper pH. By the incorporation of (4*S*)-aminoproline or (4*R*)-aminoproline into a CMP, they are able to control the endo/exo ring pucker of proline. An ammonium group ($–NH_3^+$) directs the ring pucker by a gauche effect. In basic

conditions, however, the ammonium group becomes deprotonated to form an amino group (–NH_2), and the proline ring pucker is now governed by steric interactions. In addition to control of ring pucker, pH controls the creation of an intramolecular transannular hydrogen bond and hydration shell with the ammonium moiety. Although these peptides have yet to be tested by hybridization to a collagen surface, they show potential for use as invasive CMPs that would respond to the changes in pH such as that in a wound environment.

Invasive CMPs are a unique tool for imaging and drug delivery, and have the potential to act as biosensors. Because CMPs specifically bind damaged collagen, they provide new opportunities to detect, manage, and treat pathological conditions associated with collagen remodeling. Whereas many of the studies done so far support the hypothesis that CMPs bind collagen through the completion of a triple helix, the direct mechanism has yet to be elucidated. Study of the binding between damaged host collagen and CMP could allow for further optimization and control of the interaction. Efforts to expand the repertoire of CMPs and exploration of the mechanism of binding are discussed in the following chapters.

Chapter 2

Convenient synthesis of collagen-related tripeptides for segment condensation

This chapter has been published under the same title.

Ellison, A. J.; VanVeller, B.; Raines, R. T. *Peptide Science* **2015**, *106*, 674-681.

Contributions

Dr. Brett VanVeller proposed use of isobutyl chloroformate for tripeptide synthesis. Aubrey J. Ellison performed all experiments and drafted the original manuscript and figures. Aubrey J. Ellison, Dr. Brett VanVeller, and Ronald T. Raines planned experiments, analyzed data, and edited the manuscript and figures.

2.1 Abstract

Chromatography is a common step in the solution-phase synthesis of typical peptides, as well as peptide fragments for subsequent coupling on a solid support. Combining known reagents that form readily separable byproducts is shown to eliminate this step, which wastes time and other resources. Specifically, activating carboxyl groups with isobutyl chloroformate or as pentafluorophenyl esters and using *N*-methyl morpholine as a base enable chromatography-free synthetic routes in which peptide products are isolated from byproducts by facile evaporation, extraction, and trituration. This methodology was used to access tripeptides related to collagen, such as Fmoc-Pro-Pro-Gly-OH and Fmoc-Pro-Hyp(*t*Bu)-Gly-OH, in a purity suitable for solid-phase segment condensation to form collagen mimetic peptides.

2.2 Introduction

Synthetic peptides have been the basis for transformative advances in medicine, as well as the creation of new catalysts and materials.[161-168] Peptides are synthesized either in solution, as pioneered by Emil Fischer at the beginning of the 20th century,[169] or on a solid support, as developed by Bruce Merrifield in the 1960s.[170] Drawbacks occur with both methods. The synthesis of peptides in solution is plagued by challenging purification steps and other issues.[171] Solid-phase peptide synthesis (SPPS) facilitates purification[172,173] and continues to advance,[174-178] though overall yields necessarily diminish with each coupling step in the linear route and are intrinsically low for some sequences.[179,180]

A convergent strategy in which protected peptide fragments made in solution are coupled on a solid support offers an attractive alternative.[176,181,182] Such “segment condensation” requires fewer couplings, increasing efficiency and reducing waste. In part, however, the use of segment

condensation on a solid support merely displaces inefficiency and waste to the solution phase, and its requisite chromatographic purification.

Our research group and many others are interested in collagen mimetic peptides (CMPs).[183-185] CMPs have been used to reveal the structure of the collagen triple helix[10,11] along with the forces that underlie that structure,[13] and have been of particular relevance in biomedicine and materials science.[53,157,186-191] Like collagen itself, CMPs are composed of Xaa-Yaa-Gly triplets. This invariant repetition makes CMPs ideal for synthesis by segment condensation. These factors motivated us to seek synthetic routes to CMPs that avoid cumbersome purification steps.

Typical CMPs comprise 21–30 residues. The component Xaa-Yaa-Gly triplets for segment condensation are made by using solution-phase methods.[192-195] In our experience,[156,196-198] the synthesis of tripeptides with common reagents requires chromatography on a column of silica gel (or crystallization, which can be idiosyncratic) to obtain products of adequate purity from each coupling step and the final C-terminal deprotection step. Purification by chromatography is not only laborious, but also requires the use of chromatographic media and solvents that are typically discarded as waste. Finally, chromatography of longer peptides or on a large scale suffers from diminished resolution of products. Here, we report on methodology for the chromatography-free synthesis of collagen-related tripeptides that are poised for SPPS. Our methods are efficient and scalable, and could be applicable to other peptide fragments.

2.3 Materials and methods

2.3.1 General

Pentafluorophenyl (Pfp) esters[199] and other amino acid derivatives were from Chem-Impex International (Wood Dale, IL). Isobutyl chloroformate[200] and all other reagents were from

Sigma–Aldrich (St. Louis, MO) and were used without further purification. Tetrahydrofuran (THF) and *N*,*N*-dimethylformamide (DMF) were dried with a Glass Contour system from Pure Process Technology (Nashua, NH). In addition, DMF was passed through an associated isocyanate "scrubbing" column to remove any amines.

Volatile removal: The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials with a rotary evaporator at water aspirator pressure (<20 torr) while maintaining a water bath below 40 °C. The phrase "volatiles were removed under reduced pressure" refers to the removal of solvents and other volatile materials with a rotary evaporator at vacuum-pump pressure (~2 torr) while maintaining a water bath below 40 °C.

Trituration: Trituration was done by the "solvent swap" method. Briefly, a compound was dissolved in a mixture of two solvents, one being a more volatile solvent in which the compound has good solubility (*e.g.*, DCM), the other (the "anti-solvent") being a less volatile solvent in which the compound has poor solubility (*e.g.*, hexanes). The resulting solution was concentrated under reduced pressure to remove the more volatile solvent. The solids were isolated from the slurry by filtration through a Büchner funnel, and the filter cake was washed with the anti-solvent.

2.3.2 Solution-phase synthesis

Molecular mass was determined at high resolution by electrospray ionization (ESI) mass spectrometry with an LCT instrument from Waters (Milford, MA).

^{1}H and ^{13}C NMR spectra were acquired with an Avance III 400 spectrometer or Avance III 500I spectrometer equipped with a cryogenic probe from Bruker (Billerica, MA). Chemical

shifts are reported in units of δ (ppm) relative to tetramethylsilane as the internal standard. In the spectrum of each compound herein, multiple conformers are apparent due to cis−trans isomerism around amide and/or carbamide bonds.

Peptide purity was assessed by reverse-phase chromatography with a Discovery BIO Wide Pore C5-5 column from Supelco (Bellefonte, PA). The column was eluted with 5–95% v/v B over 10 min, 95% v/v B for 2 min, 95–5% over 2 min, and 5% v/v B for 5 min (A: H_2O containing 0.1% v/v formic acid; B: acetonitrile containing 0.1% v/v formic acid). Absorbance at 254 nm and molecular mass were monitored with an LCMS-2020 instrument from Shimadzu (Kyoto, Japan).

Cbz-Pro-Gly-O*t*Bu: Cbz-Pro-OH (3.01 g, 12 mmol) and NMM (4.0 mL, 36 mmol) were dissolved in 200 mL THF, and the resulting solution was cooled to –78 °C in a dry ice/acetone bath. Isobutyl chloroformate (1.5 mL, 12 mmol) was added dropwise, and a white precipitate (NMM·HCl) formed. Solid HCl·Gly-O*t*Bu (2.02 g, 12 mmol) was added in one portion, and the reaction mixture was allowed to warm to room temperature and then stirred for 12 h. The precipitate was removed by filtration, and the solution was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and volatiles were removed under reduced pressure to give Cbz-Pro-Gly-O*t*Bu (4.01 g, 92%) as a pale yellow solid, which was judged to be 99% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 363.1915; found, 363.1909. Integration of signals for the N–H proton revealed that the cis:trans ratio of the prolyl peptide bond was 1.0:1.5. ^{1}H NMR (500 MHz, $CDCl_3$, δ): 7.42–7.28 (m, 5H), 5.32–4.99 (m, 2H), 4.56–4.22 (m, 1H), 3.93 (dt, J = 16.0, 8.8 Hz, 2H), 3.66–3.39 (m, 2H), 2.43–2.11 (m, 2H), 2.03–1.80 (m, 3H), 1.46 (s, 9H). ^{13}C NMR (126

MHz, $CDCl_3$, δ): 171.75, 168.75, 136.38, 128.51, 128.09, 127.94, 82.08, 67.33, 60.57, 47.08, 42.09, 28.71, 28.05, 24.52.

Fmoc-Pro-Pro-Gly-O*t*Bu: A suspension of Cbz-Pro-Gly-O*t*Bu (4.01 g, 11 mmol) and Pd/C (1.17 g, 10% w/w) in 25 mL of MeOH was stirred under an atmosphere of H_2(g) for 16 h at room temperature. Any insoluble material was removed by filtration, and volatiles were removed under reduced pressure. (The weight of the crude intermediate was 2.48 g.) The residue was dissolved in 35 mL of THF, and NMM (3.47 mL, 32 mmol) and Fmoc-Pro-OPfp (4.00 g, 7.9 mmol) were added to the resulting solution. After stirring for 16 h, volatiles were removed under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity, material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Pro-Pro-Gly-O*t*Bu (4.44 g, 74%) as an off-white solid, which was judged to be 94% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 548.2756; found, 548.2742. ^{1}H NMR (500 MHz, $CDCl_3$, δ): 7.81–7.72 (m, 2H), 7.66–7.52 (m, 2H), 7.40 (t, J = 7.5 Hz, 2H), 7.31 (tt, J = 7.4, 1.3 Hz, 2H), 4.76–4.14 (m, 5H), 4.04–3.34 (m, 6H), 2.42–1.78 (m, 8H), 1.54–1.29 (m, 9H). ^{13}C NMR (126 MHz, $CDCl_3$, δ): 171.62, 155.21, 144.28, 143.97, 141.43, 127.81, 127.20, 127.15, 125.38, 125.24, 125.18, 120.12, 120.09, 82.10, 81.68, 77.16, 67.95, 67.69, 61.17, 59.97, 59.65, 58.48, 57.76, 47.33, 47.25, 47.17, 46.99, 42.33, 42.10, 29.57, 28.16, 28.14, 28.07, 27.24, 25.39, 24.88, 24.50, 23.18.

Fmoc-Pro-Pro-Gly-OH: Fmoc-Pro-Pro-Gly-O*t*Bu (4.44 g, 8.1 mmol) was dissolved in 12 mL 4 M HCl in dioxane, and the resulting solution was stirred for 14 h at room temperature.

The reaction mixture was concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity, material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Pro-Pro-Gly-OH (3.57 g, 90%) as an off-white solid, which was judged to be 85% pure by LCMS. ^{1}H NMR (500 MHz, $CDCl_3$, δ): 7.75 (ddd, J = 10.3, 7.2, 4.1 Hz, 2H), 7.65–7.49 (m, 2H), 7.43–7.36 (m, 2H), 7.30 (ttd, J = 7.4, 3.1, 1.4 Hz, 2H), 4.73–4.49 (m, 1H), 4.49–4.18 (m, 3H), 4.18–3.79 (m, 2H), 3.79–3.64 (m, 2H), 3.64–3.50 (m, 2H), 2.60–1.65 (m, 8H). ^{13}C NMR (126 MHz, $CDCl_3$, δ): 172.49, 172.24, 172.10, 171.91, 171.85, 171.80, 171.69, 155.21, 154.49, 144.28, 144.06, 143.98, 143.72, 143.57, 141.29, 141.25, 141.23, 127.77, 127.72, 127.69, 127.59, 127.08, 127.04, 125.21, 125.07, 124.78, 124.61, 119.96, 119.77, 67.93, 67.70, 66.28, 60.99, 60.06, 59.86, 59.70, 58.75, 58.44, 57.67, 52.27, 47.58, 47.25, 47.16, 47.01, 46.90, 41.33, 41.15, 31.84, 30.15, 29.34, 29.29, 27.68, 27.49, 27.16, 25.17, 25.02, 24.67, 24.39, 23.08, 22.02.

Cbz-D-Pro-Gly-O*t*Bu: Cbz-D-Pro-OH (3.02 g, 12 mmol) and NMM (4.0 mL, 36 mmol) were dissolved in 200 mL THF, and the resulting solution was cooled to –78 °C in a dry ice/acetone ice bath. Isobutyl chloroformate (1.5 mL, 12 mmol) was added dropwise and a white precipitate (NMM·HCl) formed. Solid HCl·Gly-O*t*Bu (2.04 g, 12 mmol) was added in one portion, and the reaction mixture was allowed to warm to room temperature and then stirred for 12 h. The precipitate was removed by filtration, and the solution was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and volatiles were removed under reduced pressure to give Cbz-D-Pro-Gly-O*t*Bu (4.27 g, 97%) as a pale yellow solid, which was judged to be 91% pure by

LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 363.1915; found, 363.1907. Integration of signals for the N–H proton revealed that the cis:trans ratio of the prolyl peptide bond was 1.0:1.7. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.45–7.30 (m, 5H), 5.29–5.03 (m, 2H), 4.37 (d, J = 22.7 Hz, 1H), 4.06–3.81 (m, 2H), 3.66–3.36 (m, 2H), 2.43–2.08 (m, 1H), 2.02–1.84 (m, 3H), 1.46 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$, δ): 172.30, 171.71, 168.75, 156.15, 136.36, 128.51, 128.10, 127.96, 82.12, 77.28, 77.03, 67.35, 60.90, 60.58, 47.53, 47.09, 46.93, 42.10, 41.68, 31.06, 29.72, 28.64, 28.05, 24.53, 23.63, 19.03.

Fmoc-D-Pro-D-Pro-Gly-O*t*Bu: A suspension of Cbz-D-Pro-Gly-O*t*Bu (4.267 g, 12 mmol) and Pd/C (1.25 g, 10% w/w) in 25 mL of MeOH was stirred under an atmosphere of H_2(g) for 18 h at room temperature. Any insoluble material was removed by filtration, and volatiles were removed under reduced pressure. The residue was dissolved in 80 mL of 5:3 THF/DMF. NMM (3.87 mL, 35 mmol), HBTU (4.45 g, 12 mmol), and Fmoc-D-Pro-OH (3.97 g 12 mmol) were added to the resulting solution. After stirring for 16 h, volatiles were removed under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity, material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-D-Pro-D-Pro-Gly-O*t*Bu (5.26 g, 82%) as an off-white solid, which was judged to be 95% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 548.2756; found, 548.2755. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.85–7.27 (m, 8H), 4.81–4.13 (m, 2H), 4.05–3.36 (m, 6H), 3.34–2.66 (m, 2H), 2.36–1.62 (m, 9H), 1.56–1.37 (m, 9H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 171.32, 168.87, 140.13, 138.01, 134.70, 129.09, 128.73, 127.05, 124.34, 121.00,

120.32, 119.74, 107.79, 82.10, 77.36, 77.24, 77.04, 76.72, 59.96, 59.24, 47.68, 46.79, 42.04, 41.98, 30.49, 28.04, 27.38, 26.28, 25.01.

Fmoc-D-Pro-D-Pro-Gly-OH: Fmoc-D-Pro-D-Pro-Gly-O*t*Bu (5.26 g, 9.6 mmol) was dissolved in 12 mL of 4 M HCl in dioxane, and the resulting solution was stirred for 14 h at room temperature. The reaction mixture was concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. Material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-D-Pro-D-Pro-Gly-OH (3.37 g, 71% yield) as an off-white solid, which was judged to be 90% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 492.2130; found, 492.2133. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.75 (t, J = 7.6 Hz, 2H), 7.57 (ddd, J = 25.7, 15.6, 7.3 Hz, 2H), 7.39 (td, J = 7.7, 3.1 Hz, 2H), 7.30 (tt, J = 7.5, 1.3 Hz, 2H), 4.73–4.48 (m, 1H), 4.47–4.18 (m, 3H), 4.18–3.80 (m, 2H), 3.80–3.63 (m, 2H), 3.63–3.48 (m, 2H), 2.61–1.65 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 172.34, 172.19, 172.15, 172.12, 171.91, 171.85, 171.82, 171.68, 155.43, 155.12, 154.45, 144.31, 144.11, 144.01, 143.78, 143.64, 141.28, 141.24, 141.22, 127.77, 127.72, 127.69, 127.59, 127.14, 127.10, 127.07, 127.01, 125.26, 125.12, 124.81, 124.65, 119.96, 119.77, 67.85, 67.63, 66.28, 60.94, 60.00, 59.67, 58.73, 58.40, 57.67, 47.58, 47.24, 47.16, 47.05, 46.88, 46.64, 41.42, 41.34, 41.27, 31.84, 30.14, 29.33, 29.28, 27.87, 27.66, 25.12, 24.99, 24.64, 24.40, 23.09, 22.03, 20.79.

Boc-Pro-Pro-OMe: Boc-Pro-OH (3.97 g, 18 mmol) and NMM (6.0 mL, 55 mmol) were dissolved in 200 mL of THF, and the resulting solution was cooled to –78 °C in a dry ice/acetone bath. Isobutyl chloroformate (2.4 mL, 18 mmol) was added dropwise, and a white precipitate (NMM·HCl) formed. Solid HCl·Pro-OMe (3.05 g, 18 mmol) was added in one portion, and the reaction mixture was allowed to warm to room temperature and then stirred for 12 h. The

precipitate was removed by filtration, and the solution was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and volatiles were removed under reduced pressure to give Boc-Pro-Pro-OMe (3.25 g, 54%) as a pale yellow solid. Also present is the byproduct of the carbamate formed by reaction of HCl·Pro-OMe and isobutyl chloroformate. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 327.1915; found, 327.1912. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 4.57 (ddd, J = 13.3, 8.5, 4.2 Hz, 0H), 4.50 (dd, J = 8.4, 3.1 Hz, 0H), 4.38 (td, J = 8.6, 3.9 Hz, 0H), 4.32 (dd, J = 8.7, 3.8 Hz, 0H), 3.95–3.85 (m, 1H), 3.82–3.75 (m, 0H), 3.75–3.69 (m, 3H), 3.66–3.32 (m, 3H), 2.32–1.74 (m, 7H), 1.45 (s, 3H), 1.39 (s, 2H), 0.94 (d, J = 6.7 Hz, 1H), 0.89 (d, J = 6.7 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 173.42, 173.26, 172.92, 172.63, 171.60, 171.12, 154.62, 154.57, 153.71, 79.44, 71.45, 59.01, 58.73, 58.67, 57.73, 57.68, 52.18, 52.08, 46.84, 46.77, 46.64, 46.46, 46.25, 30.94, 29.98, 29.90, 29.05, 28.81, 28.70, 28.49, 28.34, 28.02, 27.96, 25.03, 24.97, 24.31, 24.05, 23.55, 19.04, 18.92, 18.89.

Fmoc-Gly-Pro-Pro-OMe: Boc-Pro-Pro-OMe (3.25 g, 9.9 mmol) was dissolved in 9.0 mL of 4 M HCl in dioxane. The resulting solution was stirred for 14 h at room temperature. The reaction mixture was concentrated under reduced pressure. The residue was dissolved in 30 mL of THF, and NMM (3.28 mL, 29 mmol) and Fmoc-Gly-OPfp (4.56 g, 9.9 mmol) were added to the resulting solution. After stirring for 16 h, the reaction mixture was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×), saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity,

material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Gly-Pro-Pro-OMe (4.38 g, 87%) as an off-white solid, which was judged to be 90% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 506.2292; found, 506.2286. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.76 (dd, *J* = 7.6, 4.9 Hz, 2H), 7.59 (dd, *J* = 7.6, 3.6 Hz, 2H), 7.43–7.35 (m, 2H), 7.30 (tdd, J = 7.6, 6.0, 1.5 Hz, 2H), 5.73 (t, *J* = 4.4 Hz, 1H), 5.42 (d, *J* = 23.5 Hz, 1H), 4.70 (dd, *J* = 8.0, 3.7 Hz, 1H), 4.57 (dd, *J* = 8.6, 4.2 Hz, 1H), 4.41 (d, *J* = 7.2 Hz, 1H), 4.36 (dd, *J* = 7.3, 3.3 Hz, 1H), 4.22 (q, *J* = 7.7 Hz, 1H), 4.13 (dd, *J* = 17.2, 5.7 Hz, 1H), 4.05 (s, 1H), 4.03–3.91 (m, 1H), 3.83 (dt, *J* = 9.5, 7.1 Hz, 1H), 3.77 (s, 1H), 3.72 (s, 2H), 3.63 (dt, *J* = 9.7, 6.2 Hz, 2H), 3.57–3.43 (m, 1H), 2.22 (qd, *J* = 11.5, 10.3, 7.1 Hz, 2H), 2.13–1.93 (m, 4H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 172.57, 170.51, 167.21, 156.38, 143.89, 143.84, 141.29, 141.25, 127.74, 127.67, 127.09, 127.07, 127.05, 125.19, 125.08, 120.00, 119.97, 119.93, 67.24, 63.81, 58.90, 58.22, 53.58, 52.43, 52.28, 47.09, 46.79, 46.36, 43.34, 28.79, 28.20, 24.93, 24.61.

Fmoc-Gly-Pro-Pro-OH: Fmoc-Gly-Pro-Pro-OMe (4.38 g, 8.6 mmol) was dissolved in 50 mL of 80:20 isopropanol/THF. Powdered $CaCl_2$ (14.22 g, 128 mmol), $LiOH{\cdot}H_2O$ (1.42 g, 33 mmol), and H_2O (5 mL) were added to the resulting solution. The mixture was vigorously stirred for 48 h. The pH was adjusted to 2 with aqueous HCl (1% v/v), and the resulting solution was concentrated under reduced pressure. The aqueous layer was extracted with EtOAc, and the organic layer was washed with water and brine, dried over Na_2SO_4(s), and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity, material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Gly-Pro-Pro-OH (3.63 g, 86%) as an off-white solid, which was judged to be 87% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 492.2130; found, 492.2128. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.75 (d, *J* = 7.5 Hz, 2H),

7.64–7.54 (m, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.33–7.27 (m, 2H), 5.77 (s, 1H), 5.39 (s, 1H), 4.82–4.51 (m, 2H), 4.49–4.29 (m, 2H), 4.21 (t, *J* = 7.3 Hz, 1H), 4.14 (dd, *J* = 17.3, 5.8 Hz, 1H), 4.06–3.88 (m, 1H), 3.83 (dd, *J* = 9.4, 7.2 Hz, 1H), 3.74–3.45 (m, 3H), 2.43–1.79 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 174.38, 172.80, 171.53, 167.89, 156.72, 156.68, 143.88, 143.79, 143.74, 141.25, 127.74, 127.70, 127.09, 127.06, 125.20, 125.17, 125.10, 119.98, 119.95, 67.31, 59.31, 58.99, 58.26, 52.35, 47.13, 47.07, 47.03, 46.47, 43.27, 42.58, 28.77, 28.39, 28.19, 24.92, 24.88, 24.70.

Cbz-Hyp(*t*Bu)-Gly-OMe: Cbz-Hyp(*t*Bu)-OH (3.01 g, 9.4 mmol) and *N*-methyl morpholine (NMM) (3.1 mL, 28 mmol) were dissolved in 200 mL of THF, and the resulting solution was cooled to –78 °C in a dry ice/acetone bath. Isobutyl chloroformate (1.2 mL, 9.3 mmol) was added dropwise, and a white precipitate (NMM·HCl) formed. Solid HCl·GlyOMe (1.17 g, 9.3 mmol) was added in one portion, and the reaction mixture was allowed to warm to room temperature and then stirred for 12 h. The precipitate was removed by filtration, and the solvent was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and volatiles were removed under reduced pressure to give Cbz-Hyp(*t*Bu)-Gly-OMe (3.26 g, 88%) as a pale yellow solid, which was judged to be 99% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 393.2021; found, 393.2015. Integration of signals for the N–H proton revealed that the cis:trans ratio of the prolyl peptide bond was 1.0:2.1. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.39–7.28 (m, 3H), 7.20 (s, 1H), 6.32 (s, 1H), 5.34–4.93 (m, 2H), 4.58–4.20 (m, 2H), 4.03 (d, *J* = 5.4 Hz, 1H), 3.74 (s, 3H), 3.64 (dd, *J* = 10.5, 7.2 Hz, 1H), 3.26 (dd, *J* = 10.6, 6.6 Hz, 1H), 2.53–2.35 (m, 1H), 2.19 (d, *J* = 22.4 Hz, 1H), 1.94 (q, *J* = 10.5, 9.7 Hz, 1H), 1.72 (s, 1H), 1.18 (s, 9H). ^{13}C NMR (100 MHz,

$CDCl_3$, *δ*): 171.76, 170.06, 156.26, 136.30, 128.52, 128.12, 127.84, 69.48, 67.45, 58.95, 52.96, 52.34, 41.24, 35.99, 28.26.

Fmoc-Pro-Hyp(*t*Bu)-Gly-OMe: A suspension of Cbz-Hyp(*t*Bu)-Gly-OMe (3.26 g, 8.3 mmol) and Pd/C (877 mg, 10% w/w) in 25 mL of MeOH was stirred under an atmosphere of H_2(g) for 16 h at room temperature. Any insoluble material was removed by filtration, and volatiles were removed under reduced pressure. (The weight of the crude intermediate was 2.0 g.) The residue was dissolved in 35 mL of THF, and NMM (2.4 mL, 22 mmol) and Fmoc-Pro-OPfp (2.91 g, 5.8 mmol) were added to the resulting solution. After stirring for 16 h, the reaction mixture was concentrated under reduced pressure. The residue was dissolved in EtOAc, and the resulting solution was washed successively with 0.10 M $KHSO_4$ (2×) and saturated aqueous $NaHCO_3$ (2×). The organic layer was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. To enhance its solidity, material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Pro-Hyp(*t*Bu)-Gly-OMe (1.23 g, 25%) as an off-white solid, which was judged to be 96% pure by LCMS. HRMS-ESI (*m*/*z*): $[M + H]^+$ calcd, 578.2861; found, 578.2866. ^{1}H NMR (500 MHz, $CDCl_3$, *δ*): 7.77 (dd, *J* = 7.5, 3.9 Hz, 2H), 7.68–7.52 (m, 2H), 7.40 (t, *J* = 7.7 Hz, 2H), 7.36–7.29 (m, 2H), 4.64–4.19 (m, 6H), 4.10 (ddd, *J* = 23.8, 18.0, 6.6 Hz, 1H), 4.00–3.79 (m, 1H), 3.79–3.43 (m, 6H), 2.72–2.39 (m, 1H), 2.37–1.83 (m, 6H), 1.44–1.01 (m, 9H). ^{13}C NMR (126 MHz, $CDCl_3$, *δ*): 172.93, 171.86, 170.40, 141.65, 128.04, 127.44, 125.63, 125.49, 120.40, 120.34, 77.16, 74.48, 70.57, 67.92, 58.78, 52.63, 47.61, 47.21, 41.42, 34.82, 29.66, 28.67, 28.62, 25.15, 24.75.

Fmoc-Pro-Hyp(*t*Bu)-Gly-OH: Fmoc-Pro-Hyp(*t*Bu)-OMe (1.23 g, 2.1 mmol) was dissolved in 30 mL of 80:20 isopropanol/THF. Powdered $CaCl_2$ (3.53 g, 32 mmol), $LiOH·H_2O$

(355 mg, 8.5 mmol), and water (5 mL) were added to the resulting solution. The mixture was stirred vigorously for 12 h. The pH was adjusted to 2 with 1% w/v aqueous HCl, and the resulting solution was concentrated under reduced pressure. The aqueous layer was extracted with EtOAc, and the organic layer was washed with water and brine, dried over Na_2SO_4(s), and concentrated under reduced pressure. The residue was dissolved in the minimum volume of DCM and triturated with hexanes. Material insoluble in hexanes was dissolved in methanol, and volatiles were removed under reduced pressure to give Fmoc-Pro-Hyp(*t*Bu)-Gly-OH (978 mg, 83% yield) as an off-white solid, which was judged to be 94% pure by LCMS. ^{1}H NMR (500 MHz, $CDCl_3$, δ): 7.85–7.76 (m, 2H), 7.71–7.54 (m, 2H), 7.44 (td, J = 7.4, 3.0 Hz, 2H), 7.35 (ddt, J = 8.6, 7.4, 1.2 Hz, 2H), 4.91–3.45 (m, 12H), 2.78–1.59 (m, 7H), 1.50–0.92 (m, 9H). ^{13}C NMR (126 MHz, $CDCl_3$, δ): 172.27, 172.21, 172.09, 172.04, 172.00, 171.96, 171.65, 155.50, 155.19, 154.63, 144.18, 144.10, 144.00, 143.73, 143.57, 141.29, 141.26, 141.23, 141.14, 127.78, 127.71, 127.68, 127.13, 127.09, 127.05, 125.23, 125.08, 124.76, 124.72, 120.03, 119.99, 119.94, 74.50, 74.27, 74.18, 70.15, 69.91, 67.92, 67.69, 67.37, 67.10, 59.48, 58.86, 58.64, 58.44, 58.04, 57.78, 53.38, 52.74, 47.32, 47.15, 47.02, 46.87, 41.37, 41.25, 35.62, 35.35, 30.14, 29.12, 28.22, 28.20, 28.16, 24.62, 24.33, 23.02.

2.3.3 Solid-phase peptide synthesis

Solid-phase peptide synthesis was performed at the University of Wisconsin–Madison Biotechnology Center with a Prelude peptide synthesizer from Protein Technologies (Tucson, AZ). Synthetic peptide was purified by HPLC with an LC-20 instrument from Shimadzu. Molecular mass was determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid matrix with a

Voyager DE-Pro instrument at the Biophysics Instrumentation Facility at the University of Wisconsin–Madison.

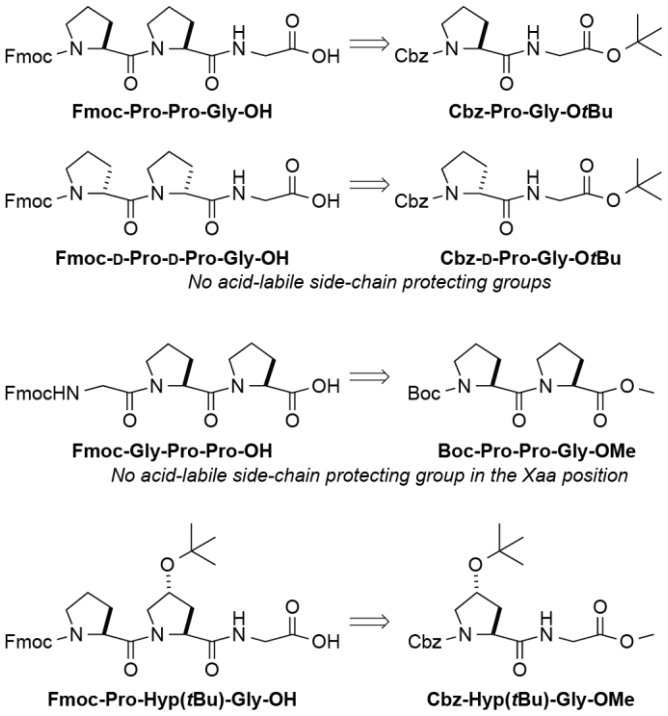
Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Using the Fmoc-D-Pro-D-Pro-Gly-OH tripeptide synthesized in solution without chromatography (*vide supra*) and an Fmoc-D-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g) from EMD Millipore (La Jolla, CA). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections.

Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at 0 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): $[M + H]^+$ calcd 2380.6, found 2380.0. A 50-μmol scale synthesis afforded 18.2 mg (15%) of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ after purification.

2.4 Results and discussion

The key to establishing a route for chromatography-free peptide synthesis is to mitigate byproduct formation. To do so, we coupled amino acids by using isobutyl chloroformate[200] to form transient anhydrides or by using stable pentafluorophenyl esters.[199] These known reagents[201] produce volatile and water-soluble byproducts, respectively. We applied variations of our method to four tripeptides that are used commonly in CMP strands (Scheme 2.1). These tripeptides are rich in glycine (which lacks a stereogenic center) and proline (which has virtually

no tendency to epimerize during coupling reactions).[202-204] Fmoc-Pro-Pro-Gly-OH and its permutation, Fmoc-Gly-Pro-Pro-OH, provide the simplest triplets in CMPs. Fmoc-D-Pro-D-Pro-Gly-OH is an enantiomer that can serve as a control for experiments in chiral environments. Finally, Fmoc-Pro-Hyp(*t*Bu)-Gly-OH provides the most common triplet found in natural collagen, and contains an acid-labile side-chain protecting group.[10] We were able to access each of these tripeptides without the use of chromatography (Table 2.1).



Scheme 2.1. Four Fmoc-protected tripeptides synthesized by chromatography-free methodology and their dipeptide precursors. Restrictions on utility are in italics typeface.

Table 2.1. Yield and purity of intermediates and tripeptide products made by chromatography-free synthesis.

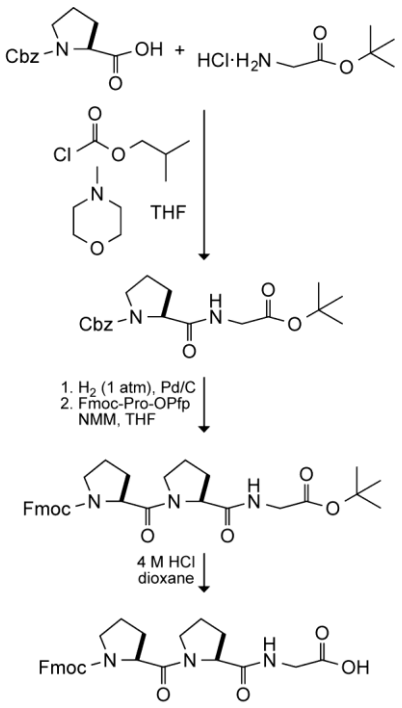
Protected Dipeptide Ester	Yield	Purity[a]	Fmoc Tripeptide Ester	Yield	Purity[a]	Fmoc Tripeptide	Yield	Purity[a]
Cbz-Pro-Gly-O*t*Bu	92%	99%	Fmoc-Pro-Pro-Gly-O*t*Bu	74%	94%	Fmoc-Pro-Pro-Gly-OH	90%	85%
Cbz-D-Pro-Gly-O*t*Bu	97%	91%	Fmoc-D-Pro-D-Pro-Gly-O*t*Bu	82%	95%	Fmoc-D-Pro-D-Pro-Gly-OH	71%	90%
Boc-Pro-Pro-OMe	54%	ND	Fmoc-Gly-Pro-Pro-OMe	87%	90%	Fmoc-Gly-Pro-Pro-OH	86%	87%
Cbz-Hyp(*t*Bu)-Gly-OMe	88%	99%	Fmoc-Pro-Hyp(*t*Bu)-Gly-OMe	25%	96%	Fmoc-Pro-Hyp(*t*Bu)-Gly-OH	83%	94%

[a]Purity was assessed with LCMS monitored at 254 nm. LCMS traces are shown in Figure 1 and the Supporting Information. The absorbance of Boc-Pro-Pro-OMe at 254 nm was not detectable (ND).

As an example, we elaborate on the synthesis of Fmoc-Pro-Pro-Gly-OH (Scheme 2.2). The first coupling step was accomplished by activating the carboxylic acid with isobutyl chloroformate and NMM. Isobutyl chloroformate is a relatively inexpensive reagent compared to many common coupling reagents (*e.g.*, HATU). Coupling with isobutyl choroformate enables ready access to a relatively pure product, as the two major byproducts—CO_2 and isobutanol—are volatile.[205] Insoluble NMM·HCl is removed by filtration, and any residual NMM (bp 115 °C) is removed by evaporation or by aqueous extraction along with unreacted starting material (*e.g.*, the free amine and free acid, which are non-volatile). This route provided Cbz-Pro-Gly-O*t*Bu in 92% yield.

Deprotection of the amino group of the dipeptide by hydrogenolysis was followed immediately by addition of the third amino acid, Fmoc-Pro-OPfp. Again, the byproduct, pentafluorophenol is soluble in the aqueous layer during an extraction. Pentafluorophenol (bp 143 °C) also has modest volatility. Initially, we attempted this coupling reaction with *N*,*N*-diisopropylethylamine as the base in DMF. These conditions provided the desired tripeptide,

but LCMS showed evidence of some proline oligomers, presumably from inadvertent Fmoc deprotection. Changing to a weaker base—NMM (which has a conjugate acid of pK_a 7.41[206])—in THF eliminated this problem and afforded Fmoc-Pro-Pro-Gly-O*t*Bu in 74% yield.



Scheme 2.2. Synthetic route to Fmoc-Pro-Pro-Gly-OH that requires no chromatography.

Finally, the *t*-butyl ester of Fmoc-Pro-Pro-Gly-O*t*Bu was removed with anhydrous HCl in dioxane, providing Fmoc-Pro-Pro-Gly-OH in 90% yield. In the analogous step en route to other tripeptides, selective base saponification of a methyl ester in a wet organic solvent containing LiOH (4 equiv) and $CaCl_2$ (15 equiv) afforded the desired carboxylic acid without removing the Fmoc group,[207] affording Fmoc-Gly-Pro-Pro-OH and Fmoc-Pro-Hyp(*t*Bu)-Gly-OH in 86% and 83% yield, respectively.

In lieu of chromatography, we used evaporation, extraction, and trituration to purify products. In the first coupling step, isobutyl chloroformate was the coupling agent, and evaporation and extraction were sufficient to provide a clean product. The second coupling and final deprotection steps also employed trituration with hexanes, which does not dissolve the tripeptide product. The analytical liquid chromatography traces for each step in the synthesis of Fmoc-Pro-Pro-Gly-OH exemplify the efficacy of these simple purification steps (Figure 2.1).

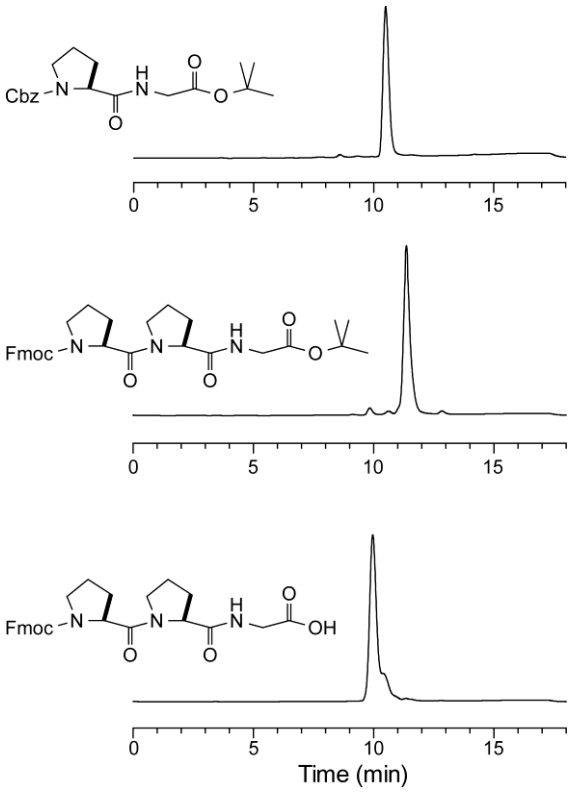


Figure 2.1. LCMS traces of products from each step in the chromatography-free synthesis of Fmoc-Pro-Pro-Gly-OH by the route in Scheme 2. Molecules were detected by asborance at 254 nm during LCMS with a reverse-phase column. Traces for the syntheses of other tripeptides are shown in Figure S1 of the Supporting Information.

Like Fmoc-Pro-Pro-Gly-OH, three other tripeptides were obtained through this chromatography-free methodology (Table 2.1). In general, coupling reactions proceeded in high yield and provided product of high purity. We did note with ^{1}H NMR spectroscopy that the reaction of HCl·Pro-OMe and isobutyl chloroformate generated a carbamate byproduct, which was not observed in any other coupling. The carbamate impurity was, however, not detrimental to subsequent steps and was readily removable by trituration, as shown by LCMS analysis and by NMR spectroscopy.

The reaction of Hyp(*t*Bu)-Gly-OMe with Fmoc-Pro-OPfp was atypical in providing a low (<50%) yield of product. We speculate that this deviation derives from particular steric and electronic attributes of Hyp(*t*Bu)-Gly-OMe (Figure 2.2). Its *t*-butyl protecting group is bulky and could interfere with amide bond formation. Moreover, the side-chain oxygen of Hyp is electron-withdrawing, thereby lowering the nucleophilicity of its amino group. Coupling to Hyp(*t*Bu) was likewise the lowest yielding step in a reported synthesis of Fmoc-Pro-Hyp-Gly-OH that was chromatography-free but used crystallization to purify most reactions products.[195] Yields might be improved by using an unprotected Hyp monomer.[195] Despite this one deviance, all other reactions with the secondary amine of proline were highly successful. Moreover, each tripeptide product was of suitable purity for SPPS.

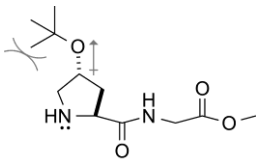


Figure 2.2. Steric and electronic effects that could diminish the reactivity of the amino group of Hyp(*t*Bu)-Gly-OMe.

In our work we have focused on the synthesis of amino acid trimers for CMPs. Nonetheless, we envision that the methodology herein could be applicable to other peptides as well. In this regard, the chromatography-free synthesis of Fmoc-Pro-Hyp(*t*Bu)-Gly-OH, which demonstrates tolerance for acid-labile protecting groups, is especially noteworthy. We note as well that the pentafluorophenyl ester of Fmoc-D-Pro is not available from commercial vendors. Accordingly, we used HBTU as the coupling agent in the chromatography-free synthesis of Fmoc-D-Pro-D-Pro-Gly-OH. These instances serve to showcase the generality of our procedures. Notably, our route avoids hydrogenation of a molecule containing an Fmoc protecting group, which is vulnerable to reduction.

Finally, we demonstrated the segment condensation of a tripeptide synthesized in solution without chromatography. Specifically, we coupled Fmoc-D-Pro-D-Pro-Gly-OH units to effect the solid-phase synthesis of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$-OH on a solid support. A mass spectrum of the purified 28-mer verified its integrity (Figure 8.2 of Appendix B).

2.5 Conclusion

We report on the synthesis of short peptides for use in SPPS by a novel combination of known reagents and reaction conditions. The methodology enables the isolation of products by using the facile techniques of evaporation, extraction, and trituration (Figure 2.3). Eliminating the need for purification by column chromatography reduces waste, reduces cost, and increases the ease of synthesis. This methodology is especially well suited for the synthesis of CMPs, which are in widespread use.[10,11,13,53,157,183-191]

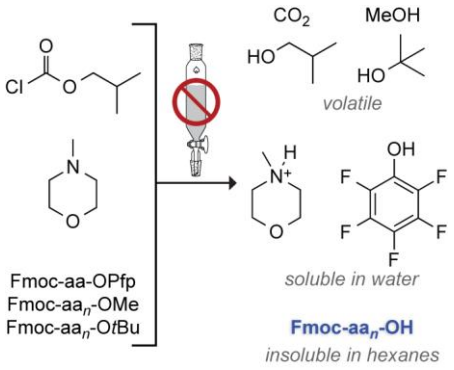


Figure 2.3. Summary of the physical properties of the products and byproducts from the reactions herein.

2.6 Acknowledgements

This work was supported by Grant R01 AR044276 (NIH) and P41 GM103399 (NIH). A. J. Ellison was supported by Chemistry-Biology Interface (CBI) Training Grant T32 GM008505 (NIH). Dr. Brett VanVeller was supported by Canadian Institute of Health Research (grant number: 289613).

Chapter 3

Collagen adherent sunscreen

Contributions

Professor Ronald T. Raines proposed creation of CMP–sunscreen conjugate. Aubrey J. Ellison synthesized and characterized CMP-sal. Professor Angela Gibson and Edgar Octol helped plan *ex vivo* experiments and obtained human tissue samples. Aubrey J. Ellison performed all other experiments and drafted the original manuscript and figures. Aubrey J. Ellison and Ronald T. Raines planned experiments, analyzed data, and edited the manuscript and figures.

3.1 Abstract

Collagen contains natural loops and breaks in the triple helix. As collagen prefers to form a triple helix, we can take advantage of the damaged collagen and introduce collagen mimetic peptides (CMPs) that will adhere to the collagen surface. These invasive strand CMPs can thereby deliver a payload to a collagen surface, such as skin. This method of delivery can be applied to the development of long-lasting sunscreens. The UV core of salicylate retained UVB protection when linked via an amide bond to a CMP. Using Vitro-skin® as an *in vitro* skin model, CMP-sal showed better retention to the skin-like surface after repeated washes compared to 2-ethylhexyl salicylate. This work provides evidence for a generalizable strategy using CMP conjugates for broad spectrum skin-adherent UV protection.

3.2 Introduction

Skin is our largest organ and acts as a barrier towards the environment. Ultraviolet (UV) light causes skin damage that manifests as skin burns, premature skin aging, immunosuppression, and skin cancer.[208-219] There are more cases of skin cancer than all other types of cancer combined.[208,220] As such, UV radiation is considered a public health threat. Much of the risk can, however, be reduced by using proper UV protection.[213,215-219,221]

UV radiation is broken into 3 types: A, B, and C based on wavelength. Type C is blocked by ozone, but UVA (320–400 nm) and UVB (290–320 nm) can penetrate human skin. UVA and UVB can cause damage to DNA[208-211] in addition to other types of skin damage. Thus, sunscreen contains ingredients, called filters, to absorb both UVA and UVB.[222,223] There exists both inorganic and organic filters. Inorganic filters include titanium dioxide and zinc oxide. Organic filters are small aromatic compounds that can be broken into types based on core UV active

structure: *para*-aminobenzoic acid derivatives, benzophenones, salicylates, cinnamates, and 'others'.[222,224]

Despite the abundance of available UV filters, public compliance remains a significant problem to proper use and application of sunscreen.[225,226] These issues can be addressed chemically by making sunscreens longer lasting and water resistant. Often UV-active organic filters are attached covalently to lipophilic tails. These tails instill hydrophobic interactions between the skin cells and the sunscreen, preventing water and sweat from easily washing it away. This hydrophobic modification often results in a greasy texture, which is one of the main reasons for low compliance in the first place.[227] This work examines whether a biomimetic tether in place of the hydrophobic tails could be used to create water-resistant UV filters.

Collagen is the most abundant protein in the human body and is the primary component of skin. Collagen forms triple helices, which are then arranged into fibrils. Natural collagen contains loops or interruptions in its triple helix[133-135] and these domains are accessible to collagen mimetic peptides (CMPs).[138,139,156] The Raines lab has been successful in attaching dyes and healing factors to CMPs and using these invasive CMPs to anchor the cargo to a collagen surface.[156,157] We hypothesize that attaching a UV active core to a CMP should behave similarly in anchoring to a collagen surface like skin and resist being washed away. Herein, we test an alternative method to creating water resistant and long lasting UV protective molecules through attachment of a UV active core to a CMP.

3.3 Results and discussion

3.3.1 Design of UV core tethered to a collagen mimetic peptide

To provide a proof of concept, a small-molecule UV core containing an easy-to-modify synthetic handle was selected. Salicylates are commonly used in sunscreens to protect against UVB light. Salicylic acid provides as an easily amendable UV active core with a carboxylic acid moiety that can be tethered to a CMP via an amide bond (Figure 3.1). Modified salicylate cores are also represented in sunscreens. 2-Ethylhexyl salicylate uses this same carboxylic acid handle to conjugate a lipophilic tail via an ester.

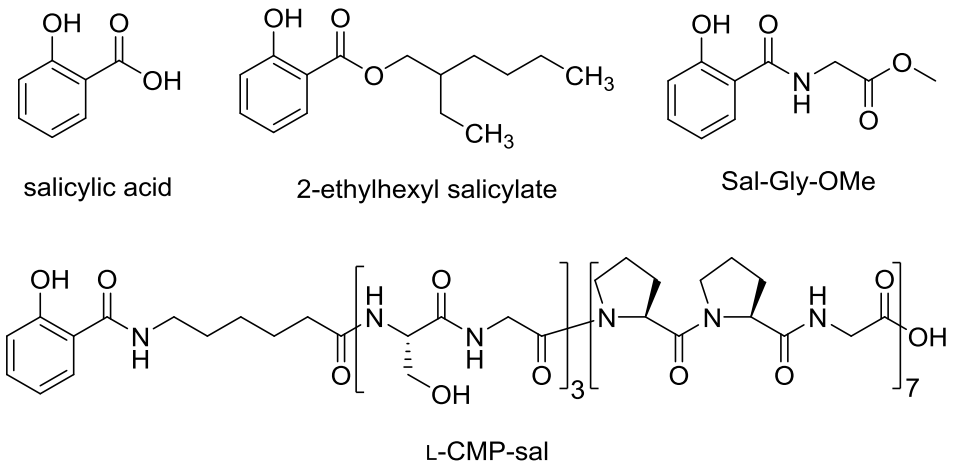


Figure 3.1. Structure of salicylic acid, 2-ethylhexyl salicylate, Sal-Gly-OMe, and CMP conjugated to salicylamide (L-CMP-sal)

Because the absorbances of UV-active molecules are sensitive to the electronics of substituents, changing from a carboxylic acid or ester to an amide could change the absorbance

profile of a salicylic acid. To make sure that an amide connection does not change the UV absorption, a test molecule was made by attaching glycine to salicylic acid (Sal-Gly-OMe, Figure 8). Because an amide is electronically similar to carboxylic acids and esters it would be expected that a Sal-Gly-OMe conjugate should maintain UVB absorbance. The solution absorbance in methanol was tested and compared to salicylic acid and 2-ethylhexyl salicylate (**Figure 3.2**). As expected, the absorption profile of the Sal-Gly-OMe conjugate still shows strong absorbance covering the UVB range.

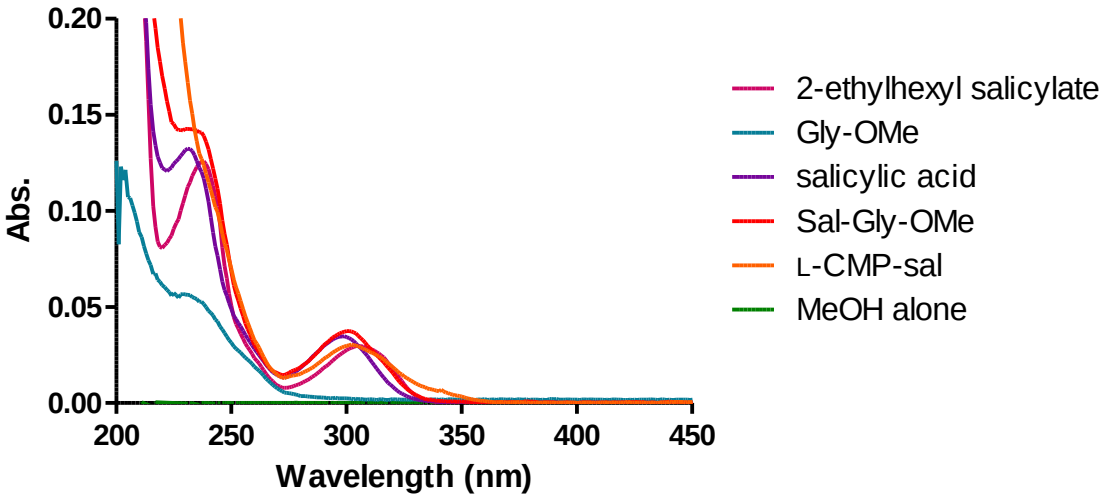


Figure 3.2. UV absorbance spectra of 10 μM solutions of analytes in methanol.

Confident that an amide bond will retain the UVB activity of the core, salicylic acid was attached to a CMP. The CMP salicylamide (CMP-sal) conjugate was made via solid-phase peptide synthesis. The collagen sequence $(Pro\text{-}Pro\text{-}Gly)_7$ was followed by a 6-aminohexanoic acid spacer. While still on resin, salicylic acid was appended to the N terminus. The UV

absorbance was tested again in a methanol solution. The collagen sequence does not impart any UV activity in the UVB or UVA range. This indicates that the salicylate core is responsible for all observed absorbance in the UVB range.

3.3.2 Detection of UV protection with the cyanotype process

With successful solution tests of UV absorbance, our effort shifted to evaluating UV activity of the CMP-sal conjugate on a collagen surface. Collagen-coated wells were chosen to act as a mimic to human skin for initial tests. To evaluate whether this collagen surface will bind the CMPs, a fluorescent CMP conjugate was first tested. A solution of CMP-Red was added to the wells and after incubation the solution was removed and the remaining CMP-Red adhered to the surface was measured by using a plate reader. A D-CMP fluorescent conjugate was used to ensure that any binding is due to triple helix formation with the collagen surface and not some other attractive force. Only the L-CMP is able to form a triple helix with natural collagen strands. By measuring the fluorescence of the material remaining on the collagen coated wells after a series of washes, we found that the L-CMP did indeed bind while the D-CMP was washed away (Figure 3.3).

Having demonstrated that CMPs will bind the collagen coated wells; the cyanotype process was used as a method to provide a visual readout of UV activity. The cyanotype process is an image-developing process invented in 1842 by Sir John Herschel. Today, cyanotype is most well-known for its use in construction blueprints. Cyanotype works by using paper coated with ferric ammonium citrate ($(NH_4)_5Fe(C_6H_4O_7)_2$) and potassium ferricyanide ($K_3Fe_3(CN)_6$). Both of these salts are soluble in water. Upon irradiation with UV light, Fe(III) from ferric ammonium citrate, is reduced to Fe(II). Fe(II) can then complex with ferricyanide to form insoluble ferric

ferrocyanide—iron(III) hexacyanoferrate(II), or Prussian blue.[228] The image is then developed in water where any of the starting soluble salts are washed away.

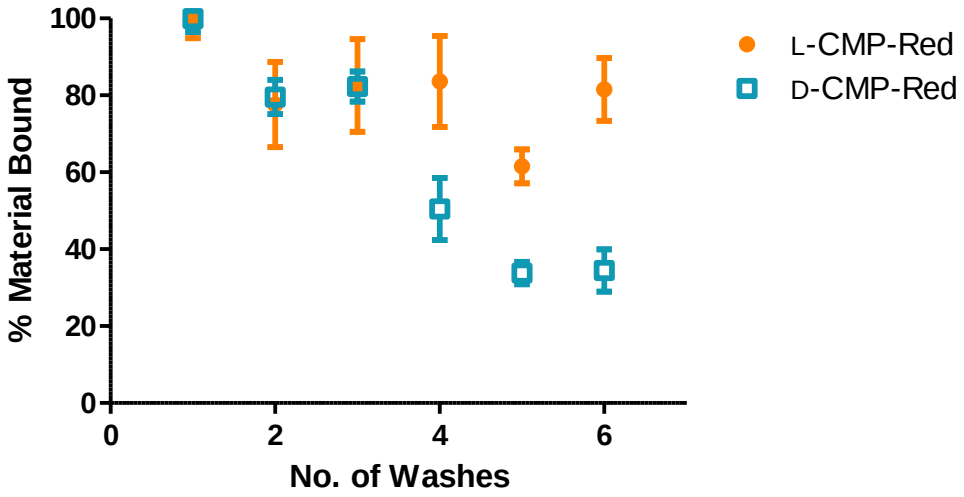


Figure 3.3. Adherence of L-CMP compared to D-CMP to collagen-coated wells over a series of washes.

In our experiment, the collagen-coated wells were UV-transparent. Accordingly, the wells were placed between the cyanotype paper and light source. CMP-sal or 2-ethylhexyl salicylate was added to the wells, which were then exposed to UV light. UV protection was represented by white areas, as in these regions, Fe(III) was not converted to Fe(II) and was thus washed away. Any area exposed to UV was blue from the production of Prussian blue.

Both CMP-sal and 2-ethylhexyl salicylate show protection from UV irradiation, as the paper remained white under wells containing those compounds (Figure 3.4, top). Because CMP-

sal should be able to bind to the collagen coated wells via a collagen triple helix, we expected that, after rinsing, the well treated with CMP-sal should maintain UV protection whereas 2-ethylhexyl salicylate will be washed away. Repeating UV exposure after washing shows that both CMP-sal and 2-ethylhexyl salicylate have diminished UV protection (Figure 3.4, bottom). The well once containing 2-ethylhexyl salicylate looks identical to an untreated well. In contrast, the well containing CMP-sal appears to be whiter, indicating that UV protection remained after washing.

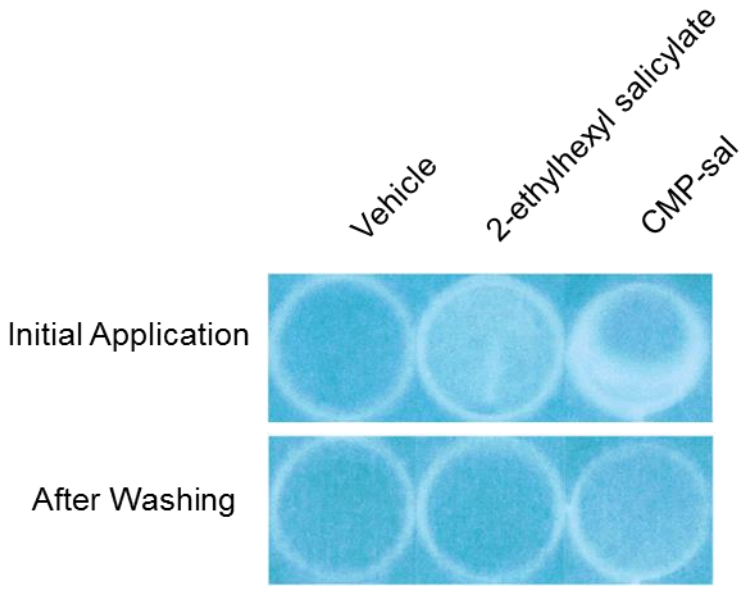


Figure 3.4. Cyanotype paper cover by collagen coated well treated with water/methanol vehicle, 2-ethylhexyl salicylate, or CMP-sal.

3.3.3 Detection of UV protection on Vitro-skin®

Using collagen-coated wells and cyanotype paper lacks sensitivity and quantification for clear evaluation. For a more quantitative measure of UV protection and water resistance, we moved to an alternative *in vitro* skin model that can be read on a solid-state UV–vis holder. Vitro-skin® is a substrate used to evaluate UV protection and the water resistance of sunscreen.[229,230] Vitro-skin® is designed to mimic human skin and contains collagen.[230] As with the collagen wells, L-CMP and D-CMP fluorescent conjugates were applied to the Vitro-skin® surface to test for binding to the collagen surface (Figure 3.5). After multiple washes, the D-CMP was washed away to a greater extent while the L-CMP remained adhered even at lower concentration. Once again, we are observed evidence of strand interacting with the collagen surface through the formation of a triple helix.

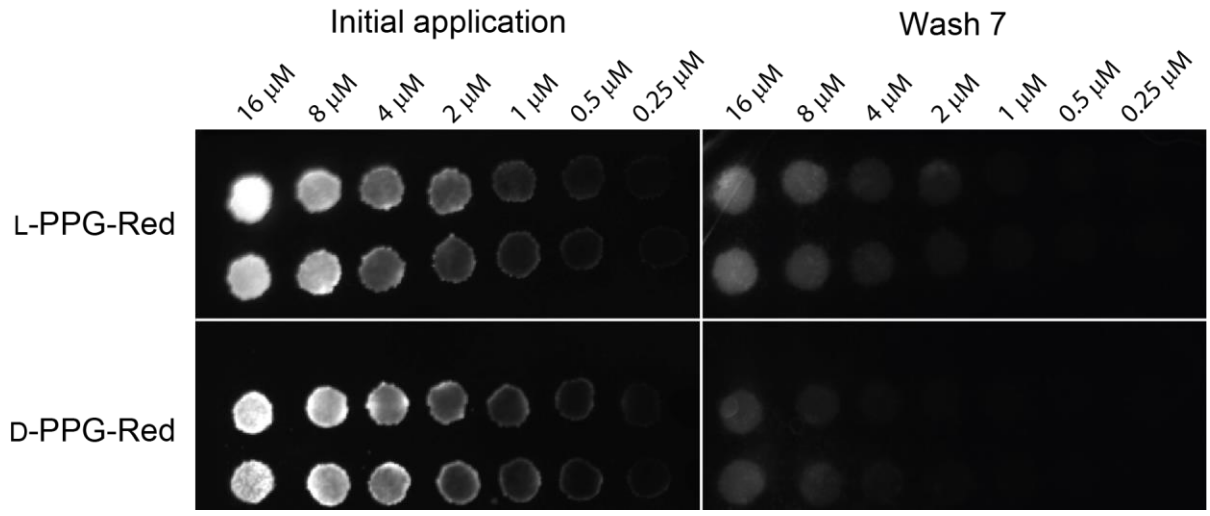


Figure 3.5. Adherence of a dilution series of L-CMP-Red compared to a dilution series of D-CMP-Red on Vitro-skin® after 7 washes as visualized by fluorescence.

CMP-sal and 2-ethylhexyl salicylate were spread onto the surface of Vitro-skin®. Using solid state UV–vis spectrometer, the absorption spectra of each UV core was measured. It was expected that, as in solution, both sunscreens would show absorbance activity for UVB light. Indeed we again observed comparable UV-absorption profiles, showing greatest activity over the UVB range (Figure 3.6).

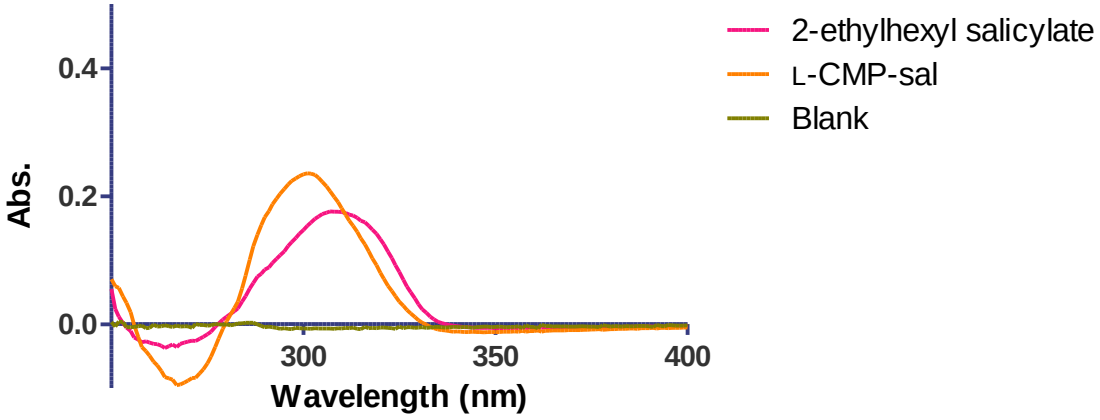


Figure 3.6. UV spectra of Vitro-skin® treated with a 133.5 nmol/cm^2 surface density of 2-ethylhexyl salicylate and CMP-sal.

With the Vitro-skin® model and the sensitivity of UV–vis we decided to test the longevity of the two sunscreens through a series of washes. The monitoring at 300 nm shows that CMP-sal and 2-ethylhexyl salicylate diminishes by 30% after the first wash (Figure 3.7). With further washes, CMP-sal maintains UV absorption whereas that of 2-ethylhexyl salicylate

continues to diminish. This result also indicates that the CMP-sal conjugate not only has similar protection against UVB, but improves skin adherence and longevity of the UV active core.

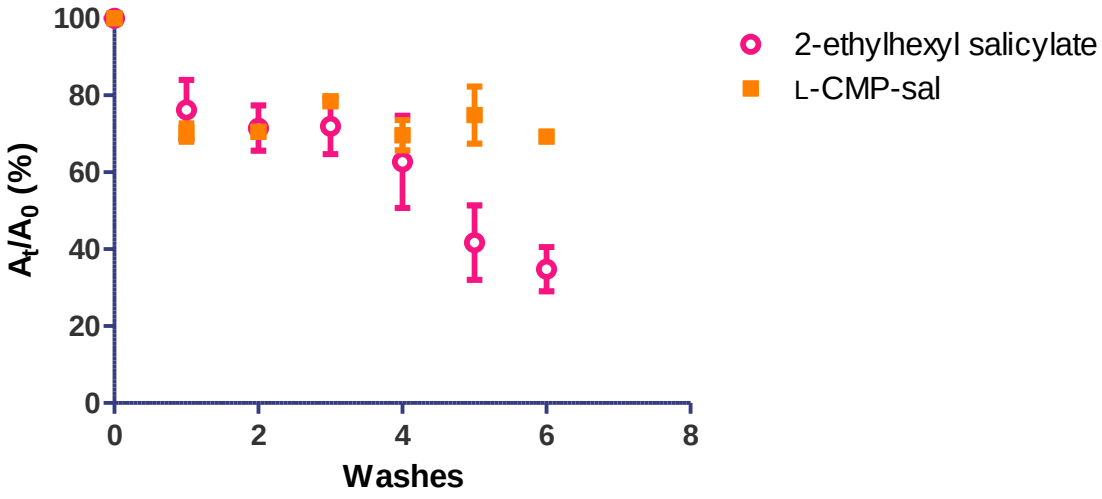


Figure 3.7. Adherence of CMP-sal compared to 2-ethylhexyl salicylate on Vitro-skin® over a series of washes determined by absorbance at 300 nm.

3.4 Conclusions

Here, we demonstrated the use of collagen mimetic peptides as invasive strands for the delivery of sunscreen to an *in vitro* skin model. The UV core of salicylic acid retained UVB protection when chemically linked by an amide bond to a CMP. The CMP-sal conjugate showed a similar absorbance profile to 2-ethylhexyl salicylate, a commercial sunscreen active ingredient, both in solution and on a solid surface.

Water-resistant sunscreens act as a way to improve retention of the UV protection on the skin surface. As a large barrier to sunscreen use is application and reapplication, new methods of retaining UV protection without the downside of greasy hydrophobic tails are important for public compliance. Using Vitro-skin®, CMP-sal showed better retention to the skin-like surface after repeated washes compared to 2-ethylhexyl salicylate. The use of CMP tethered sunscreens show potential as an alternative method to providing water resistance. With consideration to the chemical connectivity to CMPs, other UV cores can be accessed thus providing a broad spectrum of skin-adherent UV protection.

3.5 Future directions

While we were able to show *in vitro* proof of concept for use of CMP invasive strands conjugated to a UV core, the next step would be to evaluate an *ex vivo* model. With the assistance of Dr. Angela Gibson from the University of Wisconsin–Madison Department of Surgery, we were able to obtain and test human skin punch biopsies. The initial goal was to preform cytotoxicity assays on the biopsies after UVB irradiation to determine UV protection by either commercial or synthesized sunscreen.

Both lactate dehydrogenase and MTT assays did not show a difference in cytotoxicity for untreated tissue exposed to UVB compared to those kept in the dark. This equivalence remained true for UVB irradiation from 0.72 J/cm^2 to 7.2 J/cm^2. With some adjustment to light exposure or cytotoxicity assay, an *ex vivo* model with skin punch biopsies may yet be achieved.

Once an *ex vivo* model is optimized, UV protection can also be demonstrated by testing for common inflammation pathways and evaluating DNA damage. These objectives can be achieved using qPCR and comet assay, respectively.[212,231] Samples could also be evaluated with

histology to gather an understanding of the depth of damage into the skin dermis to compare the protection of CMP-bound sunscreens to commercial products.

3.6 Materials and methods

3.6.1 Instrumentation

Solid-phase peptide synthesis was performed at the University of Wisconsin–Madison Biotechnology Center with a Prelude peptide synthesizer from Protein Technologies (Tucson, AZ). Solid-phase peptide synthesis was also performed on a Liberty Blue Peptide Synthesizer from CEM (Matthews, NC). Synthetic peptide was purified by HPLC with Shimadzu Prominence instrument from Shimadzu (Kyoto, Japan) equipped with a VarioPrep 250/21 C18 column from Macherey–Nagel (Düren, Germany). Molecular mass was determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid with a Bruker microflex LRF™ (Billerica, MA) or Voyager DE-Pro instrument from Thermo Fischer Scientific (Waltham, MA) at the Biophysics Instrumentation Facility at the University of Wisconsin–Madison. Purity analyses were performed with an Acquity UPLC® H-Class system from Waters that was equipped with an Acquity photodiode array detector, Acquity quaternary solvent manager, Acquity sample manager with a flow-through needle, Acquity UPLC® BEH C18 column (2.1´ 50 mm, 1.7-µm particle size) and Empower 3 software. Colorimetric reads were done with a Tecan Infinite M1000 plate reader (Männedorf, Switzerland). GE Healthcare LAS 4010 Imaging System (Marlborough, MA) was used for image capture. Solution and solid sample spectra were obtained with a Cary 60 UV–vis spectrometer from Agilent (Santa Clara, CA). Cyanotype technique was accomplished with a UV lamp from a Fotodyne Gel Doc imager (Hartland, WI). For tissue exposure a Phillips Narrow Band 311 nm UVB Lamp (uvb-lamps.com) was used.

Irradiation dose was determined by use of Lutron Ultra-Violet Light Meter from Inspect USA (Marshall, NC).

3.6.2 General

Amino acid derivatives, Wang resin, and HOBt were from Chem-Impex International (Wood Dale, IL). Fmoc-Gly-2-chlorotrityl resin was from EMD Millipore (La Jolla, CA). Rhodamine Red™-X, succinimidyl ester, 5-isomer was obtained from Thermo Fischer Scientific (Waltham, MA). DIC and 4-methyl-piperidine were from Oakwood Chemical (Estill, SC). Cyanotype technique was done on NaturePrint™ paper (natureprintpaper.com). Solid state spectra were done on Vitro-skin® from IMS inc. (Portland, ME). Pierce LDH Cytotoxicity Assay kit and MTT were from ThermoFisher Scientific (Itasca, IL). All other reagents were from Sigma–Aldrich (St. Louis, MO) and were used without further purification. *N*,*N*-dimethylformamide (DMF) was dried with a Glass Contour system from Pure Process Technology (Nashua, NH). In addition, DMF was passed through an associated isocyanate "scrubbing" column to remove any amines. Water was purified by an Arium Pro from Sartorius (Goettingen, Germany).

3.6.3 Synthesis

Sal-Gly-OMe: A solution of salicylic acid (300.0 mg, 2.17 mmol), HATU (760.0 mg, 2.00 mmol) and DIEA (1.75 mL, 10.0 mmol) in DMF was allowed to stir for 15 min. HCl·Gly-OMe (271.3 mg, 2.17 mmol) was added and allowed to react for 12 h. The reaction mixture was concentrated under reduced presure. Crude product was purified via silica gel chromatography, eluting with 10% v/v EtOAc in hexanes to yield Sal-Gly-OMe (0.224 g, 54%). HRMS–ESI (*m*/*z*): $[M - 1]^-$ calcd, 208.07; found, 208.07. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.46 (dd, J = 8.0, 1.6 Hz, 1H), 7.38 (ddd, J = 8.6, 7.2, 1.6 Hz, 1H), 6.95 (dd, J = 8.4, 1.1 Hz, 1H), 6.84 (ddd, J =

8.2, 7.3, 1.2 Hz, 1H), 4.21 (d, J = 5.1 Hz, 2H), 3.80 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 170.34, 170.09, 161.39, 134.59, 125.84, 118.87, 118.48, 113.68, 52.68, 41.25.

Fmoc-6-Aminohexanoic acid: 6-Aminohexanoic acid (1.00 g, 7.62 mmol) was dissolved in a saturated aqueous solution of $NaHCO_3$ (50 mL). In a separate flask, Fmoc-OSu (2.82 g, 8.38 mmol) was dissolved in dioxane (50 mL). The two solutions were combined, and the solution became cloudy. The reaction mixture was stirred for 16 h, and then concentrated under reduced pressure. The residue was dissolved in EtOAc and washed with 1 N HCl and brine. The organic layer was dried over Na_2SO_4(s), decanted, and concentrated under reduced pressure. Crude product was purified via silica gel chromatography, eluting with 1% v/v acetic acid and 40% v/v EtOAc in hexanes to yield Fmoc-6-aminohexanoic acid (2.556 g, 95%) a white solid. HRMS–ESI (m/z): $[M + 1]^+$ calcd, 354.17; found, 354.17. ^{1}H NMR (400 MHz, MeOD, δ): 7.78 (d, J = 7.5 Hz, 2H), 7.63 (d, J = 7.5 Hz, 2H), 7.37 (t, J = 7.4 Hz, 2H), 7.33–7.26 (m, 2H), 4.33 (d, J = 6.8 Hz, 2H), 4.18 (t, J = 6.9 Hz, 1H), 3.08 (t, J = 7.0 Hz, 2H), 2.27 (t, J = 7.4 Hz, 2H), 1.60 (p, J = 7.5 Hz, 2H), 1.49 (p, J = 7.1 Hz, 2H), 1.33 (p, J = 10.1, 6.0 Hz, 2H). ^{13}C NMR (101 MHz, MeOD, δ): 157.49, 143.95, 141.20, 127.34, 126.71, 124.75, 119.50, 66.11, 47.13, 40.17, 33.71, 29.18, 25.97, 24.45.

Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Using the Fmoc-D-Pro-D-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-D-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections. Peptide was cleaved

from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): $[M + H]^+$ calcd, 2380.6; found, 2380.0. A 0.05-mmol scale synthesis afforded 18.2 mg (15%) of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ after purification.

Ac-Lys(Red)-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ (3.4 mg, 1.43 μmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (199.9 μL, 1.30 μmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (200 μL, 1.14 mmol) was added dropwise. The reaction mixture was allowed to stir for 8 h, and then diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 0.5 mg of peptide (11%). MALDI $[M + Na]^+$: calcd, 3056.42; found, 3056.36. Purity was assessed as >95% by UPLC.

Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$: Using the Fmoc-Pro-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–

50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): MALDI $[M + Na]^+$: calcd, 2402.61; found, 2403.15. A 0.05-mmol scale synthesis afforded 12.0 mg (10%) of Ac-Lys-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ after purification.

Ac-Lys(Red)-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$: Ac-Lys-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ (6.4 mg, 2.68 µmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (400 µL, 2.60 µmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (2.2 µL, 13.0 µmol) was added dropwise. The reaction mixture was allowed to stir for 8 h, and then was diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 1.0 mg of peptide (12%). MALDI $[M + Na]^+$: calcd, 3056.42; found, 3056.16. Purity was assessed as >95% by UPLC.

Sal-AHX-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$: Sal-AHX-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ was synthesized by single amino-acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g) from Chem-Impex International (Wood Dale, IL). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino-acid monomer or small molecule (5 equiv) was converted to an active ester using DIC and HOBt. The peptide was then cleaved from resin. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): MALDI $[M + Na]^+$: calcd, 2463.15; found, 2463.48. A 0.25-mmol scale synthesis afforded 93.7 mg (15.2%) of Sal-AHX-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ after purification.

3.6.4 UV Scan in Solution

Solution spectra were obtained with a UV–vis spectrometer. Spectra were acquired on solutions of analyte (10 μM) in methanol in a quartz cuvette with a path length of 1 cm.

3.6.5 CMP binding to collagen wells

Collagen in the wells of a 96-well plate were treated with 16 μM solutions of either L-CMP-Red or D-CMP-Red. Wells were agitated at 37 °C for 1 h. Solutions were then removed, and the fluorescence of the empty wells was measured with excitation at 560 nm and emission at 580 nm. For each wash, 50 μL of water was added to the wells and the wells were once again agitated at 37 °C for 1 h. Wells were emptied again and assessed with the plate reader. This procedure was repeated for each subsequent wash.

3.6.6 Cyanotype process

Collagen-coated wells of a 24-well plate were treated with 25 μL of 25 μM solutions of 2-ethylhexyl salicylate or L-CMP-sal in methanol or water, respectively. An additional well was treated with 25 μL of methanol and water. Wells were allowed to dry completely overnight at room temperature.

Using the UV lamp, the treated wells were placed between the light source and NaturePrint™. Samples and paper were irradiated for 2 min. Paper was washed immediately with water to develop the image. Wells were then washed 3 times with 200 μL of water. The wells were once again place between the UV light source and a new piece of NaturePrint™ paper and irradiated for 2 min. Paper was immediately washed with water to develop the image.

3.6.7 CMP binding to Vitro-skin®

Vitro-skin® was hydrated following the manufacturer's directions. Briefly, a mixture 298 grams of water and 52 grams of glycerin was added to a 2.5 gallon sealable container to make a hydration chamber. Vitro-skin® was placed on racks in the hydration chamber for 24 h prior to use.

After the hydration procedure, Vitro-skin® was spotted with 5 μL of a solution of L-CMP-Red and D-CMP-Red in water. The spots were allowed to dry. Spots were imaged using a GE Healthcare LAS 4010 Imaging System. The Vitro-skin® was washed under a stream of DI water. After washing, the Vitro-skin® was once again imaged. The same procedure was repeated for each subsequent wash.

3.6.8 UV Scan on Vitro-skin®

Vitro-skin® was cut (36.0 cm^2 squares) and hydrated following the manufacturer's directions. After hydration, the surface was coated with 2-ethylhexyl salicylate or CMP-sal to a surface density of 133.5 $nmol/cm^2$. A vehicle control was treated with water and isopropanol. Samples are then dried for 15 min followed by 2 h in the hydration chamber. Each sample was read in four different areas using a UV–vis spectrometer with an insert to read solid samples. The Vitro-skin® was then washed under a stream of water and dried for 15 min followed by 1 h in the hydration chamber. After rehydration, samples were read using UV–vis spectrometer. The same procedure was repeated for each subsequent wash. The resulting data are comprised of three experimental trials.

3.6.9 Developing a human skin model

Tissue was collected from plastic surgery cases and processed within 2 h of collection. Punch biopsies (6 mm) of dermis with the subcutaneous fat removed were washed 2x with sterile PBS and placed epidermal side up on a transwell inserts in a 24-well plate. 500 μL of media (Dulbecco's modified Eagle's culture medium containing 10% v/v fetal bovine serum supplemented with 0.625 μg/mL amphotericin B and 100 μg/mL streptomycin) was added to the wells such that the inserts were in contact with the medium without air bubbles. Tissue biopsies were incubated in culture dishes at 37 °C under 5% CO_2 for 20 h prior to experimental application.

Narrow Band 311 nm UVB Lamp was used for irradiation of samples. Irradiation dose was determined by use of Light Meter through the top of a 24-well culture plate. Samples that were kept dark were removed from the incubator for the same amount of time as irradiated samples. After irradiation samples were returned to incubator for 1 h or 24 h before LDH and MTT cytotoxicity assays.

LDH assays were performed on 50 μL of medium from the transwells following procedure as outlined in the Pierce LDH Cytotoxicity Assay kit. The absorbance was recorded at 490 nm with a plate reader.

Prior to MTT assay, all tissue samples were weighed. Tissues were then placed in a 24-well plate containing 300 μL of 1 mg/mL MTT in medium. Samples were incubated at 37 °C under 5% CO_2 for 3 h. Tissue samples were then transferred from MTT solution to a 24-well plate containing 2 mL of isopropanol. The plate was wrapped in parafilm and agitated in the dark

for 16 h. A 200-μL aliquot was removed and its absorbance was recorded at 550 nm with a plate reader.

3.7 Acknowledgements

We are grateful to Dr. Martha M. Vestling for facility support. MALDI–TOF mass spectrometer was obtained by the chemistry department with funding by the Bender gift to the University Wisconsin–Madison. This work was supported by Grant R01 AR044276 (NIH) and Instrumentation Grant GM044783 (NIH). A. J. Ellison was supported by Chemistry-Biology Interface (CBI) Training Grant T32 GM008505 (NIH).

Chapter 4

Collagen-like proteins on *S. pyogenes* show advantageous binding to mammalian collagen

Contributions

Professor Ronald T. Raines conceived of idea of studying collagen-like proteins with CMPs. Professor Daniel B. Kearns and Dr. Felix Dempwolff created the *B. subtilis* knock-ins. Aubrey J. Ellison performed all other experiments and drafted the original manuscript and figures. Aubrey J. Ellison and Professor Ronald T. Raines planned experiments, analyzed data, and edited the manuscript and figures.

4.1 Abstract

Group A *Streptococcus* (GAS) displays cell-surface proteins that resemble human collagen. We hypothesize that the collagen-like strands of GAS form triple helices with denatured host collagen in a wound bed. The ensuing adherence could be a basis for infection by GAS. With collagen mimetic peptides (CMPs) as tools, we hope to understand and ultimately control infections by GAS. To test our hypothesis, we use fluorophore-labeled CMPs to detect collagen strands on the surface of live GAS cells. With confocal microscopy, we observed selective labeling of GAS with a CMP-dye conjugate over that of *E. coli* and *B. subtilis*. Moreover, binding was not apparent when probing with the D entantiomer of the CMP. We also applied these methods to a strain of *Bacillus subtilis* in which we used reverse genetics to "knock-in" expression of collagen-like strands Scl1 and Scl2. Knock-in expression endows *B. subtilis* with the phenotypic response of GAS. Knock-in *B. subtilis* also showed improved adherence to a collagen surface. We have provided the first evidence that Scl1 and Scl2 can bind and adhere to mammalian-like collagen adding to the repertoire of GAS infection.

4.2 Introduction

Group A *Streptococcus pyogenes* (GAS) is a gram-positive bacterial pathogen that is responsible for a wide range of ailments, ranging from mild infections such as pharyngitis and impetigo to serious infections such as necrotizing fasciitis and streptococcal toxic shock syndrome.[233,234] Because GAS can manifest as many diseases, GAS requires targeted interactions to a variety of host environments. Much research has been done to tease out the molecular interactions between host and bacterium, as interrupting these interactions could provide new targets for disease treatment.[99,235-237] Of particular interest are cell-surface proteins known as adhesins, which are involved in the adherence and colonization during infection.[237]

Pathogenic bacteria display variety and redundancy in adhesin mechanism and tissue binding. Adhesins can also exhibit different actions in different environments, allowing the bacteria to adapt to the host.[233,236,238] Collagen-like proteins Scl1 and Scl2, discovered by the groups of Björck[239] and Musser[45,60], are no exception. The Scl1 and Scl2 proteins are displayed on the bacteria surface via a membrane-binding region. Immediately adjacent to the membrane-binding region is the collagen-like sequence that presents as the canonical repeat of Xaa-Yaa-Gly. Finally, on the N-terminus is a variable domain. The collagen-like sequence forms a triple helix that holds the variable domain in a trimeric complex, presenting as lollypop shape (Figure 4.1).[73,240,241] In a mouse infection model in which Scl1 is knocked out, GAS shows diminished virulence.[45] These Scl proteins have been implicated in binding to a wide variety of host-surface factors including fibronectin, laminin, integrins α2β1 and α11β1, thrombin-activable fibrinolysis, LDL, ApoB, and Factor H.[61-69] In addition to binding host factors, Scl proteins can also play a role in biofilm formation, neutrophil evasion, and protection against host defense peptides.[40,70,71]

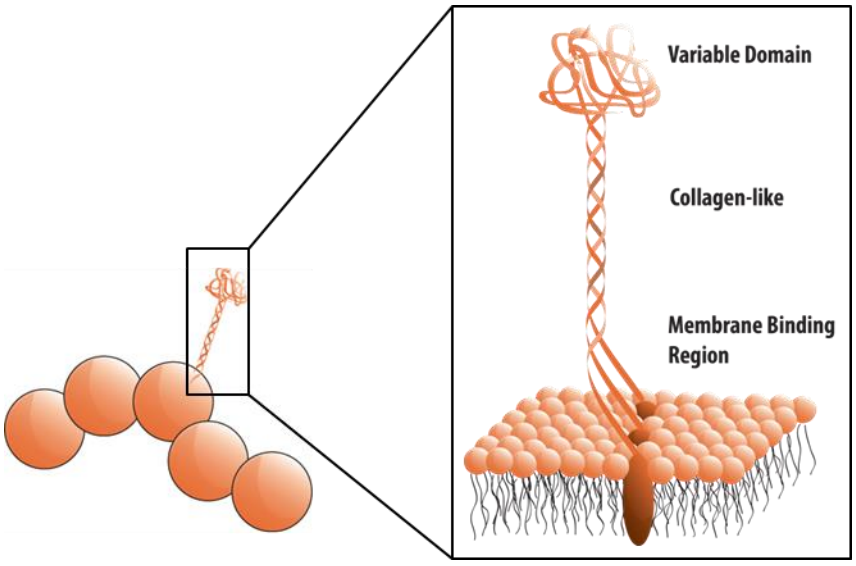


Figure 4.1. Anatomy of Scl proteins on GAS.

Most of the known interactions between host and Scl proteins occur through binding to the variable domain.[40] The collagen-like domain of Scl proteins is thought to primarily act as rigid support that orients the variable domain outward from the cell surface.[40,66,73] In some serotypes, however, the collagen sequence contains a binding motif for integrins.[62-65,73] As another way to adapt to the host environment, we hypothesize that the collagen-like sequence of Scl1 and Scl2 is not only for structure and binding to host-cell displayed integrins, but also mediate additional interactions with the extracellular matrix (ECM).

Collagen, a major component of the ECM, is the most abundant protein in the human body, comprising 1/3 of all protein by dry weight and 3/4 of protein found in the skin. Skin is the first barrier of protection against bacteria, and when the skin becomes damaged, the host is more vulnerable to infection. In wounds, the collagen that was once structured is now denatured from its native triple helix. As collagen spontaneously forms a triple helix, the Raines group has taken advantage of damaged collagen by introducing collagen mimetic peptides (CMPs) that adhere to the wound surface.[156,157] A payload of a small molecule can be delivered to a collagen surface through the use of an invasive CMP–small molecule conjugate.

Nature might also use this invasive strand mechanism as a form of host infection. We speculate that GAS takes advantage of the damaged collagen to intercalate and bind to the host surface using its collagen-like surface proteins, Scl1 and Scl2. We hypothesize that the collagen-like strands of GAS not only form triple helices with themselves, but can also form a hetero-triple helix with the denatured collagen in wound beds. The ensuing adherence could be a basis for infection by GAS.

Unlike mammalian collagen, the Xaa-Yaa-Gly sequence of the Scl collagen domain is proline poor and contains no hydroxylated proline. Proline and hydroxylated proline are important for the thermostability of mammalian collagen. In the triple helix formed by the bacterial proteins, the thermostability derives from electrostatic interactions and hydration-mediated hydrogen-bonding networks.[46,47] The collagen domain has a T_m value near 37 °C.[47,73] Thus at human body temperature, half of the Scl protein triple helix will be denatured to single strands. These strands are primed for adherence to host collagen. To test this hypothesis, we used fluorophore-labeled CMPs to detect collagen strands on the surface of live GAS cells. We also evaluated knock-in expression of collagen-like strands, Scl1 and Scl2, on the surface of *B. subtilis* to transfer the ability to anneal to our CMPs. With these tools we worked to understand and ultimately antagonize infections by GAS.

With this hypothesis and understanding of the stability of the collagen like proteins we have two primary questions to answer. First, can collagen proteins Scl1 and Scl2 bind to mammalian like collagen? Additionally, is the binding of Scl1 and Scl2 to mammalian collagen temperature dependent? And second, does the presence of Scl1 and Scl2 increase bacterial adherence to a collagen surface?

4.3 Results and discussion

4.3.1 Scl 1 and Scl 2 binding collagen mimetic peptides

In previous work, the Raines lab utilized CMPs that anneal to a collagen surface to deliver a growth factor.[157] A CMP–fluorophore conjugate was used to evaluate adherence to a collagen surface.[156] In the present work, we can picture the bacteria as a surface that is coated in collagen-like proteins. Using a rhodamine fluorophore conjugated to a CMP (L-CMP-Red), we

probed this collagen-like bacterial surface to discern if it can interact with and bind to CMPs that are mammalian-like.

Bacterial cells were grown for 24 h and incubated with CMP-Red for 1 h. Cells were then washed and visualized using confocal microscopy. Cells labeled red indicate a binding interaction between CMP and the cell surface. To show that this interaction is unique to GAS, we also evaluated binding to *Escherichia coli* and *Bacillus subtilis*. Selected strains of these bacteria were used as they are known not to express collagen-like proteins. GAS was the only bacterium that fluoresced red indicating that the interaction and adherence of the L-CMP-Red is specific to the GAS bacterium (Figure 4.2). Because *E. coli* and *B. subtilis* do not display a collagen-like protein, their lack of affinity for a CMP could be due to their lack of Scl adhesins. Alternatively, the specific labeling of GAS could be caused by another interaction of the CMP with the GAS surface.

In addition to the bacterial controls, we also used a CMP control. If the adherence to the bacterial surface is due to triple-helix formation between peptide and bacterial protein, then only the L-peptide should be able to form the appropriate left-handed triple helix. The enantiomeric D-peptide does not form the hetero-triple helix with the collagen-like proteins and thus should not label the cell. If, however, the labeling is due to electrostatic or hydrophobic interactions, then D-CMP should behave the same as the L-CMP.

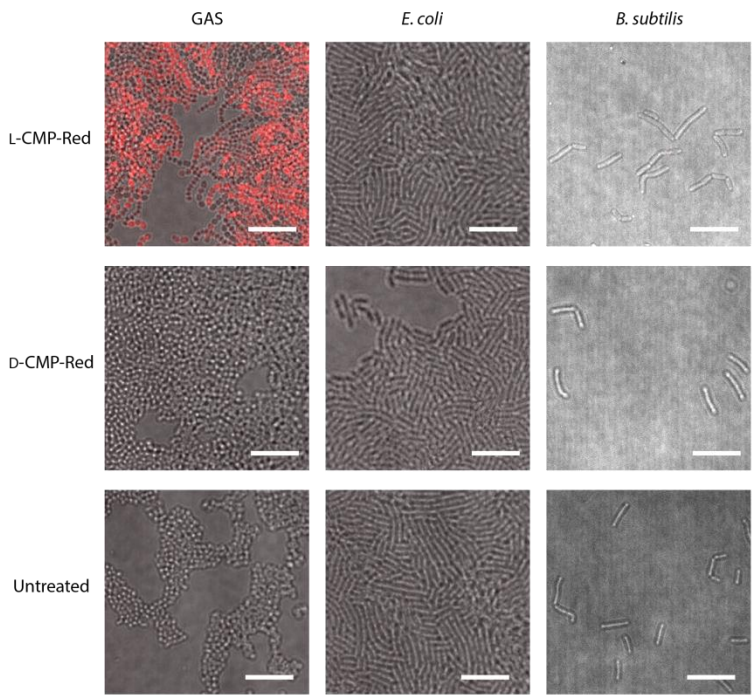


Figure 4.2. Confocal microscopy images of GAS, *E. coli*, and *B. subtilis* exposed to L-CMP-Red, D-CMP-Red, and Untreated. Scale bar: 10 μm.

When treated with the D-CMP-Red, GAS remains unlabeled. Because only L-CMP-Red labeled the GAS cells, we are assured that the CMP surface interaction is not simply due to attractive forces. The selectivity for L over D supports that binding requires proper collagen triple-helix formation. This result again suggests that Scl1 and Scl2 are interacting with the CMP as per our hetero-triple helix hypothesis. Next, we wanted to insure that the selective interaction is due to Scl1 and Scl2 and not another protein on the GAS surface.

4.3.2 Effects of temperature on Scl1 and Scl2 binding collagen mimetic peptides

In our initial prediction of the infection mechanism by GAS, we believed that the T_m of the collagen-like triple helix was important. Denaturing at 37 °C, the triple helix can be unwound

at body temperature, meaning that the proteins are primed for infection in the presence of a host. To evaluate whether this is the case, we repeated the CMP-Red labeling at different temperatures (Figure 4.3). If the interaction is temperature-dependent, then we would expect that the Scl proteins will be engaged in a triple helix with themselves at low temperatures and therefore not able to bind to L-CMP-Red, thus leaving the cell unlabeled. Above the T_m of 37 °C, we should see labeling increasing with temperature. We observed labeling at all temperatures.

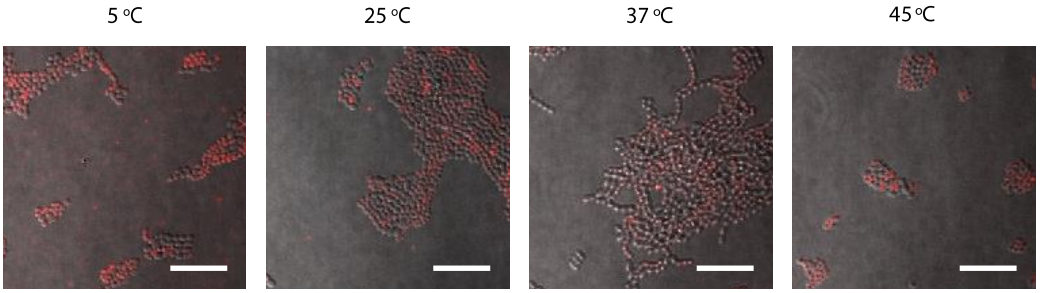


Figure 4.3. Confocal images of GAS labeling with L-CMP-Red at 5, 25, 37, and 45 °C. Scale bar: 10 μm.

The confocal images were corroborated by looking at a larger population of cells with flow cytometry (Figure 4.4). Again, GAS cells are labeled by L-CMP-Red at all temperatures. If the labeling of GAS is due to peptide and Scl proteins interacting, than the collagen-like proteins are still able to form a hetero-triple helix at low temperatures, suggesting that the T_m of the triple-helical domains of the Scl adhesins is < 37 °C *in vivo*.

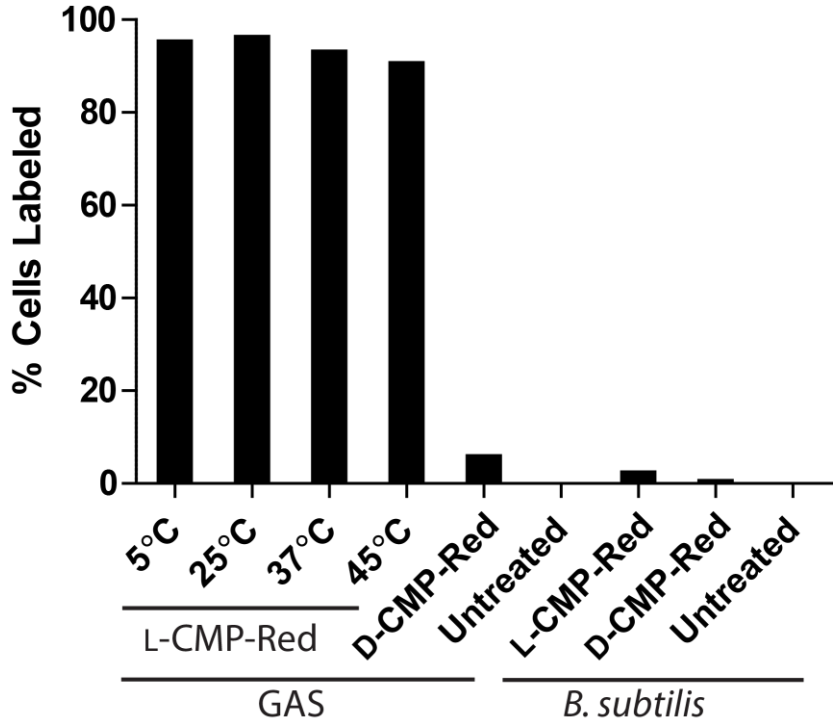


Figure 4.4. Flow cytometry of GAS and *B. subtilis* exposed to L-CMP-Red, D-CMP-Red, and untreated. Labeling represented as percentage of cell population within the gate for SYTO 9 and rhodamine red fluorescence.

4.3.3 Scl1 and Scl2 *B. subtilis* knock-ins

GAS has variety of adhesins that enable to binding to the host. Some of these adhesins can bind to collagen.[72,106,242-244] Although fluorescence shows that GAS binds to L-CMP-Red, we sought to demonstrate that the binding is due to the collagen-like proteins and not another factor on the surface of the GAS cell. Having tested *B. subtilis* and seen no labeling, with the help of Dan Kearns lab, we created an inducible *B. subtilis* knock-in of Scl1 and Scl2. The goal was to isolate these two adhesin factors and see if we could reproduce the labeling with CMPs that was observed in GAS.

Expression of Scl1 and Scl2 in *B. subtilis* was evaluated through mRNA production using qPCR. There is a background production of both Scl1 and Scl2 mRNA but upon induction we observed a 3.7-fold increase in the expression of Scl1 and a 4-fold increase in the expression of Scl2 (Figure 4.5). Because there is increase in expression upon induction, we expect that if Scl1 and Scl2 are able to interact with mammalian-like collagen, this change will also manifest in enhanced labeling with CMPs.

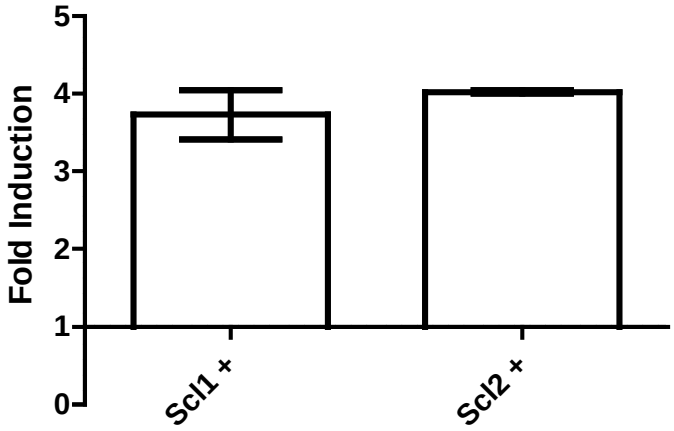


Figure 4.5. Change in expression of Scl1 and Scl2 upon induction as evaluated by qPCR.

Upon repeating the CMP-labeling experiment with the knock-in *B. subtilis* bacteria, we expected to see cells fluoresce red if the interaction is indeed due to the newly expressed collagen-like proteins. By detecting labeling with confocal microscopy (Figure 4.6) and flow cytometry (Figure 4.7), we indeed did observe labeling by L-CMP-Red. The Scl1 and Scl2 proteins confer the phenotype that was seen in GAS. This concurrence indicates that these collagen-like proteins are responsible for the interaction between GAS and the CMP-Red.

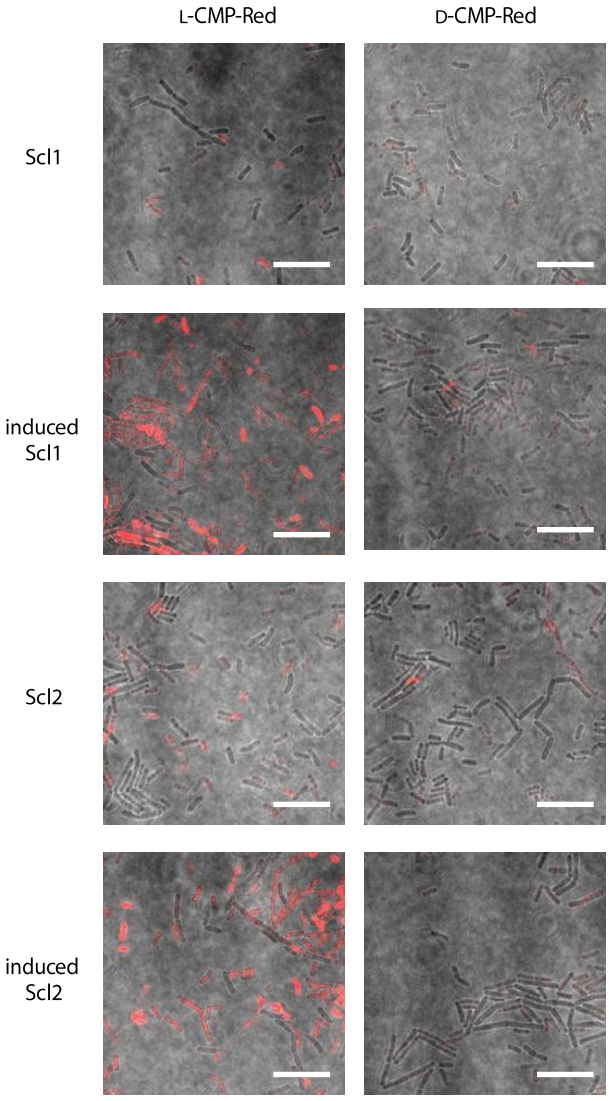


Figure 4.6. Confocal microscopy images of *B. subtilis* Scl1 and Scl2 knock-ins exposed to L-CMP-Red and D-CMP-Red. Comparison images with and without the induction of Scl expression. Scale bar: 10 μm.

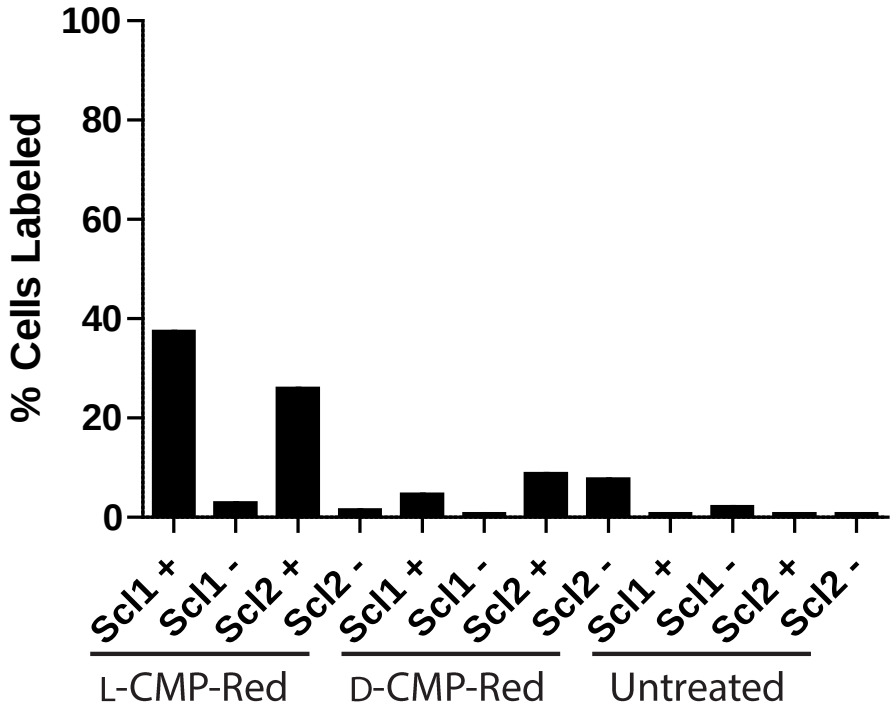


Figure 4.7. Flow cytometry of *B. subtilis* Scl1 and Sc12 knock-ins exposed to L-CMP-Red, D-CMP-Red and untreated. Populations with (+) and without (–) the induction of Scl expression are compared. Labeling is represented as a percentage of cell population within the gate for SYTO 9 and rhodamine red fluorescence.

Binding between the CMPs and the knock-ins show that there is a specific interaction that occurs between the protein and peptide. When induced, there is greater expression and thus greater CMP binding. The increase in binding resembles the increase in expression as found by qPCR. The interaction between CMPs and *B. subtilis* knock-ins is the first evidence that the Scl proteins can interact with and bind to mammalian-like collagen.

4.3.4 Scl1 and Scl2 binding to a collagen surface

Although we have shown that the collagen-like proteins on the surface of GAS can interact with a single-stranded of CMP, our data do not report on how the proteins would behave in the presence of a collagen surface. For example, GAS could adhere to disorganized or damaged collagen, like that would be found in a wound.

To evaluate adhesion to a collagen surface we employed an assay using crystal violet stain.[245] This assay is performed by growing cells on the surface of interest. The surface is then washed, leaving only adhered bacteria. Then, the surface is submerged in a 0.1% w/v solution of crystal violet. Excess crystal violet is washed away, leaving behind crystal violet that has been internalized by bacteria. Ethanol extracts the dye from the cells and allows for a colorimetric readout on the relative amount of cells that have remained on the surface. In our assay, we used collagen-coated wells that had *B. subtilis* with knock-ins of Scl1 and Scl2 grown in them. Unmodified *B. subtilis* acted as a negative control. As *B. subtilis* is free of any other collagen-binding adhesins, any improved surface adherence would be attributable to the presence of Scl1 and Scl2.

Using crystal violet to compare bacteria that remained on the collagen wells, we found that the Scl1 and Scl2 knock-ins showed greater adherence by having more dye extracted by the ethanol wash than the control of wild-type *B. subtilis* (Figure 4.8). Additionally, the binding only occurs in wells that are coated with collagen. Thus, the improved bacterial adherence is due to an interaction between the displayed collagen-like proteins and the collagen surface. As seen in the CMP labeling, there is an increase in bacterial adherence to the collagen surface upon induction. The increase in binding is once again similar to the increase in expression levels as determined

by qPCR. This correlation shows that Scl proteins are responsible for binding to a collagen surface.

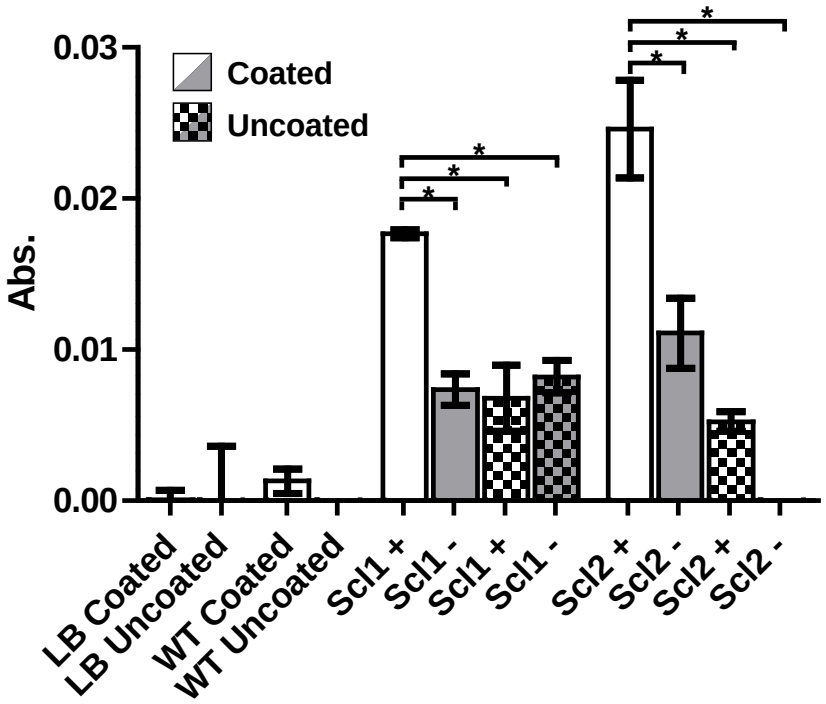


Figure 4.8. Adherence of wild-type *B. subtilis* and *B. subtilis* Scl1 and Scl2 knock-ins to uncoated and collagen-coated wells represented as absorbance of extracted crystal violet. Adherence with (+) and without (–) the induction of Scl expression are compared. * indicates $P < 0.05$.

4.4 Conclusion

Collagen-like Scl1 and Scl2 proteins on GAS have for many years now been recognized as an adhesin and virulence factor.[40] The variable domain has been shown to interact with many

host factors, allowing the bacteria to adhere to the host surface.[40] A known role of the triple helix formed by the collagen domain is to bind integrins and provide structure to project the variable domain away from the cell surface, to better interact with the host cells.[40,62-66,73] We hypothesized that the collagen-like domain could have further adherence properties in the environment of the ECM.

We have demonstrated the use of collagen mimetic peptides as probes to elucidate an adhesin mechanism of GAS. Our probes showed selectivity for bacteria expressing collagen-like proteins Scl1 and Scl2 over other bacterial cell surfaces. Additionally, the knock-in of Scl1 and Scl2 in *B. subtilis* improves binding to a collagen surface, giving credence to our hypothesis that these surface collagen-like proteins anneal to host collagen.

Herein, we have provided the first evidence that Scl1 and Scl2 can bind and adhere to mammalian-like collagen. This additional mechanism adds to the bacteria's repertoire of binding and the ability to adapt to the host environment. Although the exact nature of the interaction between Scl proteins and host collagen is not understood, we hypothesize that the interaction between collagen-like proteins and peptide is the formation of a hetero-triple helix that tethers GAS to the host surface.

4.5 Material and methods

4.5.1 Instrumentation

Solid-phase peptide synthesis was performed at the University of Wisconsin–Madison Biotechnology Center with a Prelude peptide synthesizer from Protein Technologies (Tucson, AZ). Synthetic peptide was purified by HPLC with Shimadzu Prominence instrument from Shimadzu (Kyoto, Japan) equipped with a VarioPrep 250/21 C18 column from Macherey–Nagel

(Düren, Germany). Molecular mass was determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid matrix with a Voyager DE-Pro instrument from Thermo Fischer Scientific (Waltham, MA) at the Biophysics Instrumentation Facility at the University of Wisconsin–Madison. Purity analyses were performed with an Acquity UPLC® H-Class system from Waters that was equipped with an Acquity photodiode array detector, Acquity quaternary solvent manager, Acquity sample manager with a flow-through needle, Acquity UPLC® BEH C18 column (2.1´ 50 mm, 1.7-µm particle size) and Empower 3 software. Cultures were imaged with a Nikon Eclipse Ti inverted confocal microscope (Melville, NY) at the Biochemistry Optical Core of the University of Wisconsin−Madison. Flow cytometry was performed with a BD Biosciences LSR Fortessa (San Jose, CA) at the Carbone Cancer Center at the University Wisconsin–Madison. Amplification for qPCR was accomplished with the Applied Biosystems QuantStudio 7 from Thermo Fischer Scientific (Carlsbad, CA). Colorimetric reads were done with a Tecan Infinite M1000 plate reader (Männedorf, Switzerland).

4.5.2 General

Amino acid derivatives and HOBt were from Chem-Impex International (Wood Dale, IL). Fmoc-Gly-2-chlorotrityl resin was from EMD Millipore (La Jolla, CA). Rhodamine Red™-X, succinimidyl ester, 5-isomer was obtained from Thermo Fischer Scientific (Waltham, MA). Nuclei lysis solution and protein precipitation solution were from Promega (Fitchburg, WI). miRCURY RNA isolation kit was obtained from Qiagen (Hilden, Germany). All other reagents were from Sigma–Aldrich (St. Louis, MO) and were used without further purification. *N*,*N*-dimethylformamide (DMF) was dried with a Glass Contour system from Pure Process Technology (Nashua, NH). In addition, DMF was passed through an associated isocyanate

"scrubbing" column to remove any amines. Water was purified by an Arium Pro from Sartorius (Goettingen, Germany).

4.5.3 Synthesis of peptides

Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Using the Fmoc-D-Pro-D-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-D-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): $[M+H]^+$ calcd 2380.6, found 2380.0. A 0.05-mmol scale synthesis afforded 18.2 mg (15%) of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ after purification.

Ac-Lys(Red)-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ (3.4 mg, 1.43 µmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (199.9 µL, 1.30 µmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (200 µL, 1.14 mmol) was added dropwise. The reaction mixture was allowed to stir for 8 h. The solution was then diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1%

v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 0.5 mg (11%). MALDI $[M + Na]^+$: calcd, 3056.42; found, 3056.36. Purity was assessed as >95% by UPLC.

Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$: Using the Fmoc-Pro-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): MALDI $[M + Na]^+$: calcd, 2402.61; found 2403.15. A 0.05-mmol scale synthesis afforded 12.0 mg (10%) of Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ after purification.

Ac-Lys(Red)-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$: Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ (6.4 mg, 2.68 µmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (400 µL, 2.60 µmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (2.2 µL, 13.0 µmol) was added dropwise. The reaction mixture was allowed to stir for 8 h. The solution was diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 1.0 mg (12%). MALDI $[M+Na]^+$: calcd, 3056.42; found, 3056.16. Purity was assessed as >95% by UPLC.

4.5.4 Growth of cells

Cultures of Group A *S. pyogenes* (SF 370) were grown for 24 hours in brain heart (BH) broth at 37 °C. Cultures of *E. coli* (RP437) were grown for 16 h in LB at 37 °C on a platform shaker. Cultures of *B. subtilis* (Ol 1085) were grown overnight in LB at 37 °C on a platform shaker. *B. subtilis* knock-in cultures were grown in LB. Knock-in cultures were induced after 4 h by addition ITPG to a final concentration of 1 mM. All experiments on knock-in cultures were done after 8 h of growth.

4.5.5 Cell labeling for confocal microscopy and flow cytometry

Aliquots (1.0 mL) of bacterial culture (*vide supra*) were taken and were treated with either Ac-Lys(Red)-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ or AcLys(Red)-$(Ser\text{-}Gly)_3$-$(D\text{-}Pro\text{-}D\text{-}Pro\text{-}Gly)_7$ to a final concentrations of 3.25 μM. Mixtures were allowed to incubate at 37 °C for an hour. Cultures were pelleted and the supernatant was removed and resuspended in water. This procedure was repeated three times to wash the cultures. Finally, cells were resuspended in 1 mL of water.

For confocal images, cultures (5 μL) were spotted on a thin layer of LB agarose on microscope slides. Cultures were imaged with a microscope.

For flow cytometry, after washing the cells, 1 μL of 5 mM Syto-9 in DMSO was added to the resuspended cells. Cells were counted by using 488 nm and 561 nm lasers with 530/30 nm and 586/15 nm filters.

4.5.6 Creation and genetic verification of *B. subtilis* knock-ins of Scl1 and Scl2

To achieve heterologous expression, we integrated the coding sequence for Scl1 (SPy1983) and Scl2 (SPy1054) from the M1 serotype of GAS at the *amyE* locus in the *B. subtilis*

chromosome (with translation of the mRNA being ensured by the addition of the *tufA* ribosome binding site) and controlled expression with the IPTG-inducible $P_{hyspanc}$ promoter. The presence of the genes was verified by PCR and chromosomal integration was confirmed by plating on starch containing agar. Integration of the construct leads to the disruption of the gene coding for amylase and hence to the loss of the ability to degrade starch.

For genomic DNA isolation; 1 ml of a turbid culture was harvested and pelleted (2 min at maximum rpm). Cells were resuspend in 480 µL 50 mM EDTA (pH 8.0). Lysozyme (60 µL of a 10 mg/mL solution) was added and the solution was mixed gently. The solution was then incubated for 30 min at 37 °C, and 600 µL of nuclei Lysis Solution (Promega A7943) was added. The resulting mixture was incubated for 2–3 min at room temperature, 200 µL of protein precipitation Solution (Promega A7951) was added, and the mixture was vortexed for 20 s. The mixture was incubated on ice for 5 min and then clarified by centrifugation for 3 min at maximum rpm. 900 µL of supernatant was transferred to a new tube containing 600 µL of isopropyl alcohol, and the tube was inverted gently. The resulting precipitate was collected by centrifugation for 2 min at maximum rpm. The pellet was washed with 500 µL of 70% v/v EtOH and collected by centrifugation again for 1 min at maximum rpm. The dry pellet was resuspended in 100 µL of H_2O, and the resulting solution was incubated for 30 min at 37 °C. The isolated genomic DNA was stored at –20° C.

PCR was run with a 1:30 min extension time (annealing temperature 58 °C, 30 cycles) with the following primer pairs:

Scl1:

TCAAAGCTTTAAGGAGGATTTTAGAATGTTGACATCAAAGCACCATAATC

TCAGCATGCTTAGTTGTTTTCTTTGCGTTTTG

Scl2:

TCAAAGCTTTAAGGAGGATTTTAGATTGCTGACCTTTGGAGGTGC

TCAGCATGCTTAGTTGTTTTCTTGACGTTTTGC

4.5.7 qPCR on B. subtilis knock-ins

Prior to qPCR, RNA was isolated from cultures using miRCURY RNA isolation kit and complementary DNA was prepared from the RNA by reverse transcription using Applied Biosystems High-Capacity cDNA Reverse Transcription kit. Samples were prepared in duplicate and added to Quanta PerfeCTa SYBR Green Fastmix Low Rox master mix. These samples were amplified by using an Applied Biosystems QuantStudio 7 kit. Relative gene expression was calculated with the ddCT method[246] normalized against *RspJ*, *RspE*, and *yoxA*.

Table 4.1. Primer sequences used in qPCR for Scl1, Scl2, and house-keeper genes used for normalization.

Genes	Primer sequence
Scl1	**F:** CTAACCAAACGCCAGAACGC **R:** GTTGGCTTGCTCACCAGTTG
Scl2	**F:** CAAAAGACGTGACACCTGCTC **R:** CTGCCGGTGTTTTTGCGAG
RspJ	**F:** TCGTGCGGTGCACAAATACA **R:** CAGTTTGTGGTGTTGGGTTCA
RspE	**F:** TTCGCAGCTCTTGTCGTTGT **R:** CAGCCTTACGGATCGCTTCT
yoxA	**F:** TTCCGGCAGAAAGCTCGTTA **R:** ACGGGACATCCATCAGCTTC

Primers were designed using a known algorithm[247] and are listed in Table 4.1. cDNA and primers were tested against each other and analyzed by gel electrophoresis on a gel to ensure purity and integrity.

4.5.8 Bacterial adherence

B. Subtilis knock-ins and wild-type Ol 1085 2-mL cultures were grown in collagen-coated wells of a 24-well plate. Gene expression was induced after 4 h of growth. All cultures were removed from the wells, which were washed three times with water. 1 mL of 0.1% w/v crystal violet solution was added to the wells and for 3 min. The crystal violet solution was then removed, and the wells washed three times with water. 500 μL of absolute ethanol was added to the wells. 200 μL from each well was transferred to a 96-well plate to be read on a plate reader at 620 nm.

4.6 Acknowledgements

We are grateful to Dr. Rachael Sheridan, Dr. Elle Grevstad, and Dr. Darrell McCaslin for facility support. We are also grateful to Professor Doug Weibel for contributive discussions. This work was supported by Grant R01 AR044276 (NIH) and Instrumentation Grant GM044783 (NIH). Aubrey J. Ellison was supported by Chemistry-Biology Interface (CBI) Training Grant T32 GM008505 (NIH). Work done at UWCCC Flow Cytometry Laboratory was supported by Grant P30 CA014520.

Chapter 5

Tethered collagen duplex for evaluating invasive CMP hetero-triple helix formation

Contributions

Professor Ronald T. Raines proposed creation of collagen duplex. Aubrey J. Ellison synthesized and characterized collagen duplexes. Dr. I. Caglar Tanrikulu performed analytical centrifugation and analyzed the resulting data. Jesus M Monroig Dones synthesized and characterized (flp-Hyp-Gly)$_7$. Aubrey J. Ellison performed all other experiments and drafted the original manuscript and figures. Aubrey J. Ellison, Dr. I. Caglar Tanrikulu, and Professor Ronald T. Raines planned experiments and analyzed data. Aubrey J. Ellison and Professor Ronald T. Raines edited the manuscript and figures.

5.1 Abstract

An important corollary of our work with invasive collagen mimetic peptides (CMPs) is understanding whether their invasion by occurs in wounded tissue through completion of a triple helix. Because these invasive CMPs are successful in adhering to collagen surfaces, knowing if the mechanism is hetero-triple helix formation will facilitate the creation of better invasive CMPs. A collagen duplex, here called the "nest" was envisioned and synthesized as a wound bed mimic. CMP binding to a nest was evaluated by the fluorescent overlap between a nest labeled bead and fluorescent CMP. The triple helicity of the resulting complex was confirmed by CD spectroscopy. Known and novel CMPs were compared for their propensity to bind to the nest and form a stable hetero-triple helix. $(flp\text{-}Hyp\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ peptides were the best at forming such a triple helix while thermal stability is determined by the nest.

5.2 Introduction

Using collagen mimetic peptides (CMPs) to interact with a collagenous surface, such as skin, centers on the concept of "invasive strands". In this modality, a CMP is able to intercalate into the collagen on a surface to recapitulate a triple helix. This idea has been exploited previously to adhere a growth factor[157] and sunscreen as well as probes for understanding bacterial infection, discussed in earlier chapters, to a collagen surface.

The design of an invasive strand requires a detailed understanding of the collagen triple helix. Each strand of collagen is a repeat of Xaa-Yaa-Gly units, where (2*S*)-proline (Pro) and (2*S*, 4*R*)-hydroxyproline (Hyp) are most common in the Xaa and Yaa positions, respectively.[10] The structure and stability of the triple helix comes from three sources. First, glycine at every third residue allows close packing around a central axis for trimerization.[11] Second, interstrand hydrogen bonds hold adjacent strands together.[11,12] Lastly, an abundance of Pro residues

predisposes the φ and ψ angles of the amino acid for triple-helix formation.[13] The Hyp in the Yaa position also contributes to stability.[15] Originally, the added stability of Hyp was solely attributed to hydrogen bonding with surrounding water to support the triple helix structure.[11,17-20] Though, Ramachandran first suggested that pyrrolidine ring pucker caused by hydroxyproline helped stabilize collagen.[20] The Raines group has demonstrated that changing the (2*S*,4*R*)-hydroxyproline (Hyp) residue to (2*S*,4*R*)-fluoroproline (Flp) creates hyperstable collagen strands.[23,24] Through this work, they demonstrated that the extra stability is derived from the *gauche* effect preorganizing the pyrrolidine to C^{γ}-exo ring pucker as preferred in the Yaa position for triple-helix formation.[30,31] Substitution of the Xaa position with diastereomeric (2*S*,4*S*)-fluoroproline (flp) also increased stability, and this increase is attributable to C^{γ}-endo ring pucker.[32,33]

In the design of invasive CMP strands, the strands must be incapable of forming homotrimeric triple helices but be able to heterotrimerize with host tissue. The Yu group accomplished this goal by using a photocleavable "cage" that, upon cleavage, a CMP would recoup the ability of to form a triple helix.[150,151] Additionally, two CMP strands developed by the Raines group were made and used as invasive strands: $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$.[154,156,157] The peptide $(Pro\text{-}Pro\text{-}Gly)_7$ lacks the electronegative groups that confer stability to triple helix. Substitution of one modified amino acid results in a gain in stability, but incorporation of fluoroproline to both Xaa and Yaa positions to make $(flp\text{-}Flp\text{-}Gly)_7$, has a destabilizing effect.[154,155] A homotrimeric triple helix does not form because of deleterious steric interactions between the fluorine atoms on flp and Flp residues on component strands.[13,154] A hetero-triple helix between $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ strands does, however, readily form in a 1:2 and 2:1 ratio.[154]

An important corollary of invasive strand CMP work is understanding whether strand invasion by collagen mimetic peptides occurs in wounded tissue through completion of a triple helix. Because these invasive strand CMPs are successful in adhering to collagen surfaces, knowing if the mechanism is hetero-triple helix formation will facilitate the creation of better invasive CMPs. Testing the hypothesis in an *ex-vivo* or *in-vivo* model is difficult. We could, however, approach this goal *in vitro* by a host–guest study of a hetero-triple helix.

Initial work in understanding triple-helix formation came from host–guest studies.[248] Substituting one of the Xaa-Yaa-Gly triplets in a 30-mer CMP, allowed for the evaluation of how tolerant the triple helix is to the incorporation of different amino acids.[26-29,249-251] The 30-mer CMPs provided a homo-triple helix for study, demonstrating the compounding interaction that a single substitution manifests in a triple helix. Previously, hetero-triple helix host–guest studies were carried out by tethering three CMP strands together and measuring the thermostability of the resulting triple helix.[195,252-262] Systematic substitution in one of the strands in a host–guest study illuminated the tolerance of changes to the Xaa-Yaa-Gly collagen sequence. Each new study required a new synthesis to incorporate the guest strand. To mitigate the synthetic burden, induced self-assembly mechanisms were utilized for host–guest studies.[263-265] Still, tethering all three strands together covalently or through induced self-assembly does not create the same interaction that would be expected with invasive CMPs and a collagen surface.

Mimicking the conditions of a wound bed and an exogenous CMP requires two independent components interacting (Figure 5.1). One component would consist of a collagen 'duplex' that could act as a wound-bed mimic, here called the “nest”. The second component would be the same invasive CMPs used to adhere and deliver a small molecule to a collagen

surface. Collagen mimetic peptides would be incubated with the double-stranded template, and triple-helix formation would be monitored *in vitro* with circular dichroism (CD) spectroscopy.

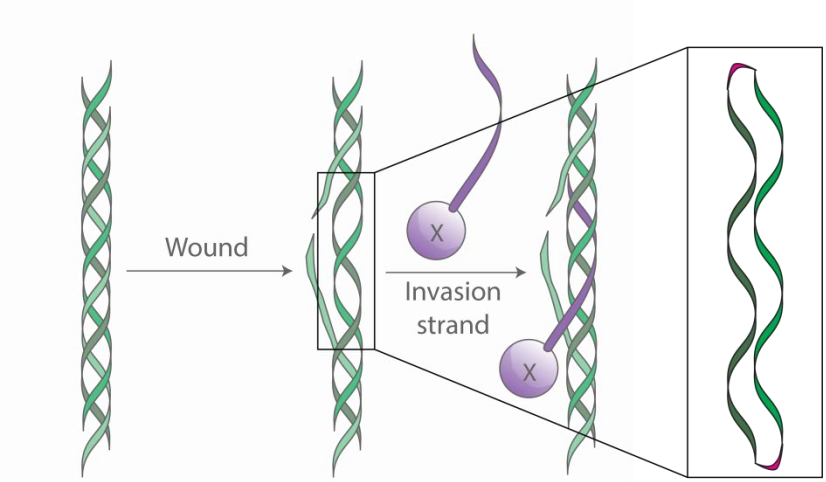


Figure 5.1. Scheme of collagen duplex that mimics damaged collagen in a wound.

Similar double-stranded duplexes have been synthesized recently to create hereto-triple helices.[198,266] In these examples, one or both ends of the peptides are free, allowing for many possible complexes to form. To simplify the environment in which the CMP is monitored, the two parallel strands of the nest will be tethered together at both N and C termini (Figure 5.2). This design will not only minimize the conformational entropy of the duplex, but also mimic host collagen strands that have a small break in the triple helix, as might be found in a wound.

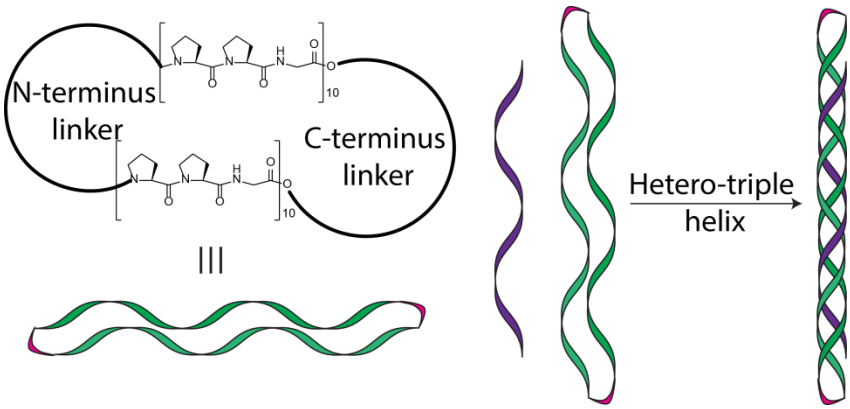


Figure 5.2. Design of duplex or "nest" for host–guest interactions tethered on both ends and able to bind to a CMP to form a hetero-triple helix.

5.3 Results and discussion

5.3.1 Nest design and synthesis

Synthesizing a collagen duplex that is tethered together and lacks a third strand will allow access to new host–guest studies. This nest will be used to evaluate the formation of a hetero-triple helix with a third invading strand from invasive CMPs. The nest design consists of two parallel CMP strands that are linked covalently on each end. Employing a C-terminus design similar to that of host–guest studies where three strands were tethered together at their C-termini, lysine was used as a template for the simultaneous synthesis of two strands on a resin bead (Figure 5.3).

Because the collagen triple helix forms with a one residue offset, the linkers on both ends of the nest need to be flexible to allow for proper folding. To provide flexibility, a 6-aminohexanoic acid spacer was added to both of the amino groups of lysine. These long carbon

chains should allow for conformational freedom. Then, two $(Pro\text{-}Pro\text{-}Gly)_{10}$ collagen peptides were synthesized by solid-phase methods from these amino groups.

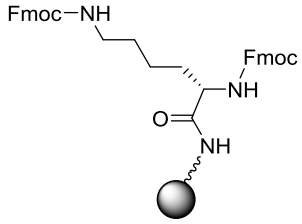


Figure 5.3. Resin-bound lysine residue for parallel peptide synthesis and use as a C-terminal linker for a collagen nest.

The same attention to flexibility was considered for designing the N-terminal linker and macrocycle closing. 6-Aminohexanoic acid spacers were also attached at the C-termini of the $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptides while still on the solid support. Because synthesis of both collagen peptides is done in tandem for synthetic ease, the chemical mechanism for linking the N-termini, is also symmetric. A few different linkers were considered for closing the macro-cycle. Initial work was focused on using a small molecule to react with both N-termini creating a bridge between collagen strands (Figure 5.4). First attempts were with glutaric acid to form two amide linkages with the N-termini of the two peptide strands while on resin. This approach resulted in a mixture of products with one or two glutaric acids added to each nest without the desired tethering of the two ends together. Maintaining the idea of a third component to bridge the ends together, the design was shifted to the use of N-terminal cysteine residues and dibromobimane. The thioether formation is known to occur quickly between a thiol group and dibromobimane[267]

and we hypothesized that a high reaction rate constant would be advantageous for ring closure. Still, like with glutaric acid, dibromobimane provided a mixture of products with no desired linked peptide observable by mass spectrometry.

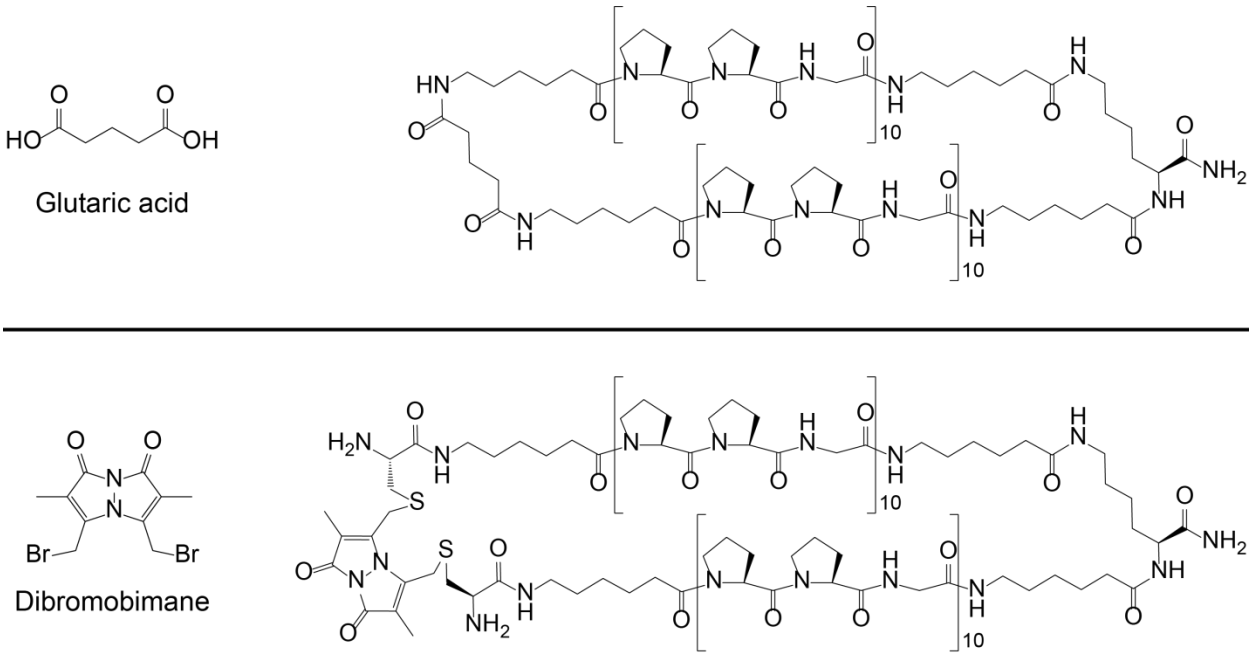


Figure 5.4. Possible chemical tethers for the N-terminus of a collagen nest using glutaric acid for amide bond formation or dibromobimane for nucleophilic substitution.

Through this experimentation, it was determined that using a small molecule to bridge together the ends of the nest provides a complicated mixture. To simplify the reactive components, a symmetric way to chemically cross link the ends to themselves was evaluated. Using the precedent of peptide stapling on resin,[268] 3-butenoic acid was added to the N terminus of the $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptides (Figure 5.5). Now with a terminal alkene, the macrocycle could

be formed by olefin metathesis. The open nest (nest-o) before metathesis was compared to the closed nest (nest-c) after metathesis, and a loss of 28 amu was observed by mass spectrometry, indicating the loss of ethylene and successful tethering of the two ends (Figure 5.6).

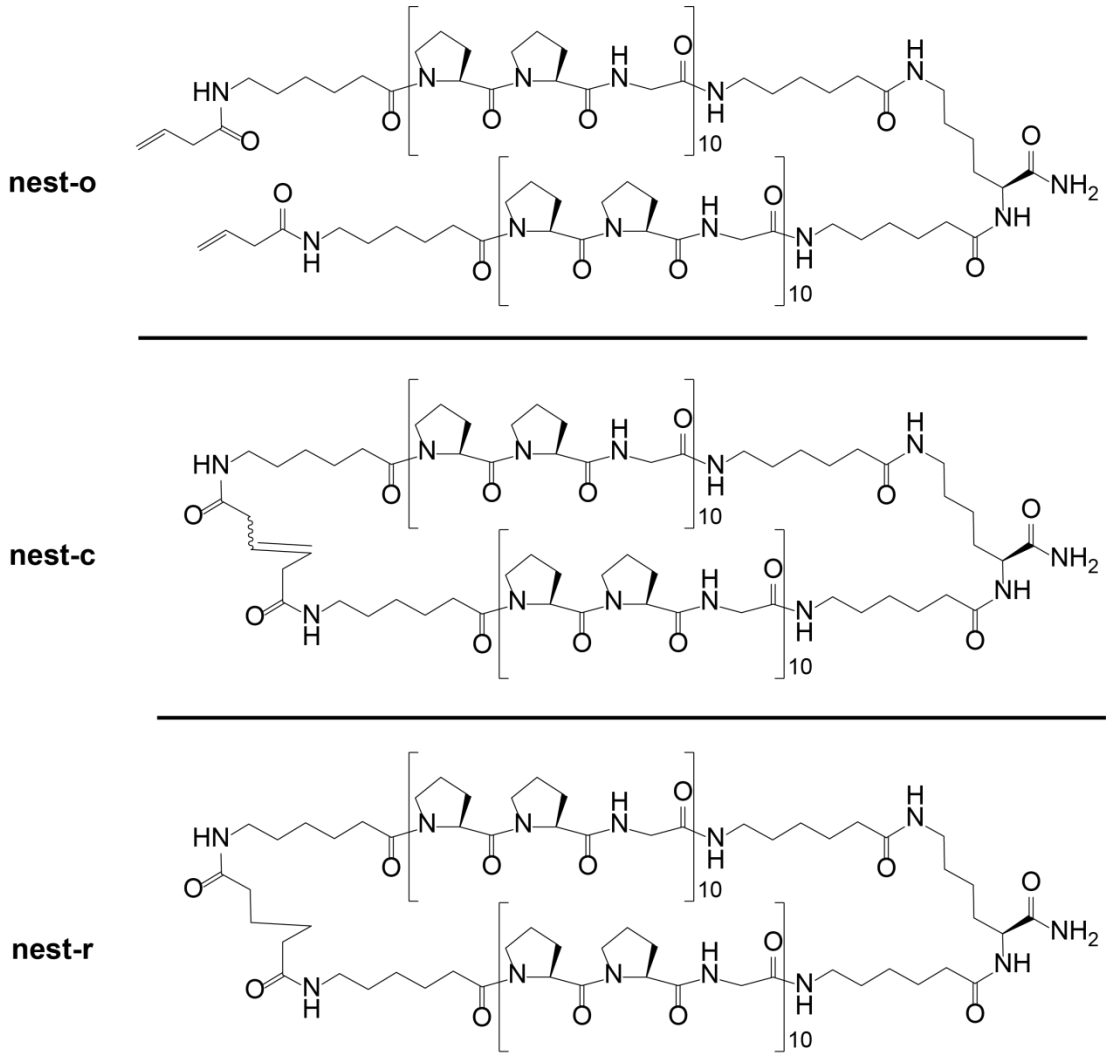


Figure 5.5. Final design of the N-terminal linkage in which olefin metathesis of nest-o is used to form nest-c. nest-r is achieved by alkene reduction of nest-c.

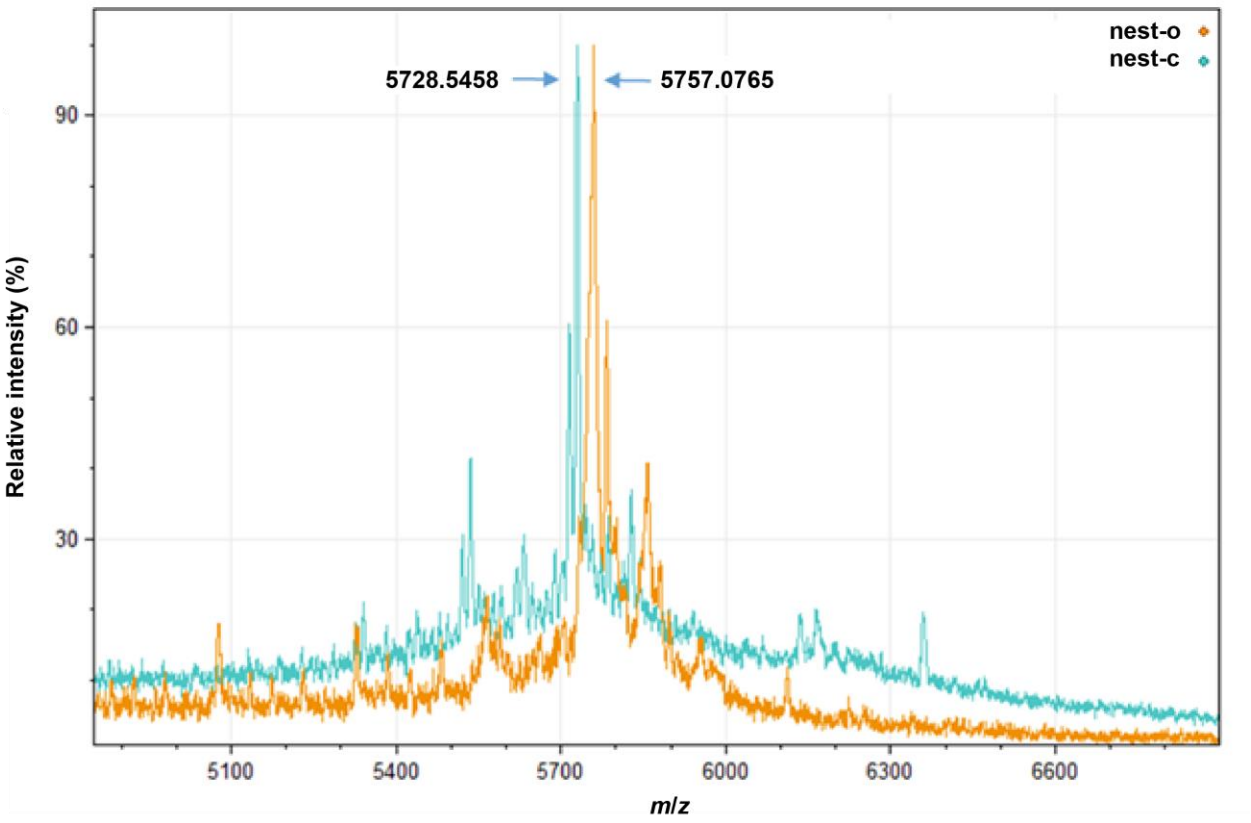


Figure 5.6. MALDI–TOF mass spectra of nest-o (5757.07 amu) and nest-c (5728.55 amu), which is consistent with the loss of ethylene after metathesis to form nest-c.

In an effort to further improve our wound-bed mimic, the newly formed alkene bond was reduced to the alkane, resulting in the reduced nest (nest-r) (Figure 5.5). When the alkene tether is formed by metathesis, the peptide nest is under harsh organic conditions that do not promote proper protein or peptide folding. Thus, the ridged alkene could result in a contorted structure that is unable to anneal in to a proper triple helix. To restore flexibility after metathesis, the resulting alkene was reduced while still on resin. The nascent alkane should now have the

rotational freedom to allow the collagen strands to anneal and form a triple helix with an additional CMP.

5.3.2 Analysis of nest–CMP complex formation with circular dichroism spectroscopy

CMPs of interest were invasive strands previously used for delivery, $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$.[156,157] These collagen sequences do not form homo-triple helix, we expected that, in the presence of the nest, a hetero-triple helix would form. Also of interest were two other CMPs: $(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$ (Figure 5.7). These peptides do form a homo-triple helix but in combination with the nest we expected a hetero-triple helix to form. The thermal denaturation curve provided by CD spectroscopy of such a hetero-triple helix should be distinct from that of the homo-triple helix. The $(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$, CMPs would thus act as a control to denote complex formation rather than homo-triple helix in the presence of the nest.

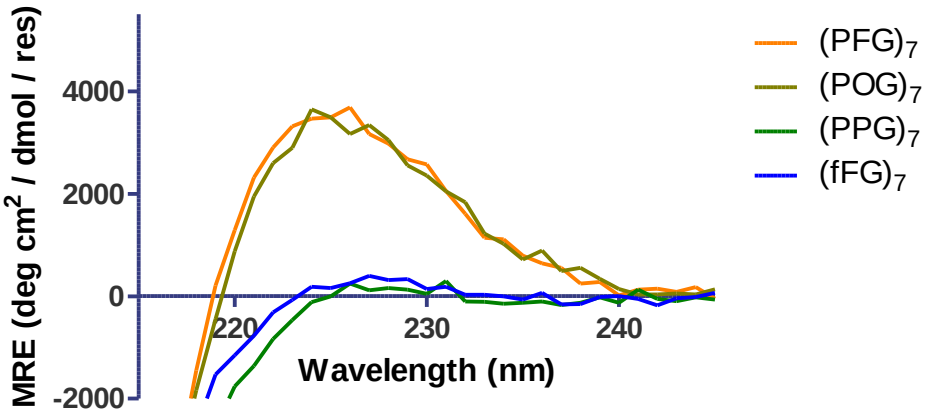


Figure 5.7. CD spectra of CMPs when annealed alone. Data were obtained in 50 mM HOAc at 4 °C.

Initially, the CMPs of interest were tested against nest-o and nest-c; before metathesis (open) and after metathesis (closed) respectively. For both nests, all CMPs of interest were added under annealing conditions. The nest-o exhibited triple helicity, indicated by a signal at 226 nm when in combination with all peptides (Figure 5.8). Notably $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ showed a drastic change in their CD spectra upon introduction of nest-o, suggestive hetero-triple helix formation. Yet, the open nest is not tethered at both ends and has the potential to splay open and create a variety of different complexes other than just the desired "wound-bed" mimic.

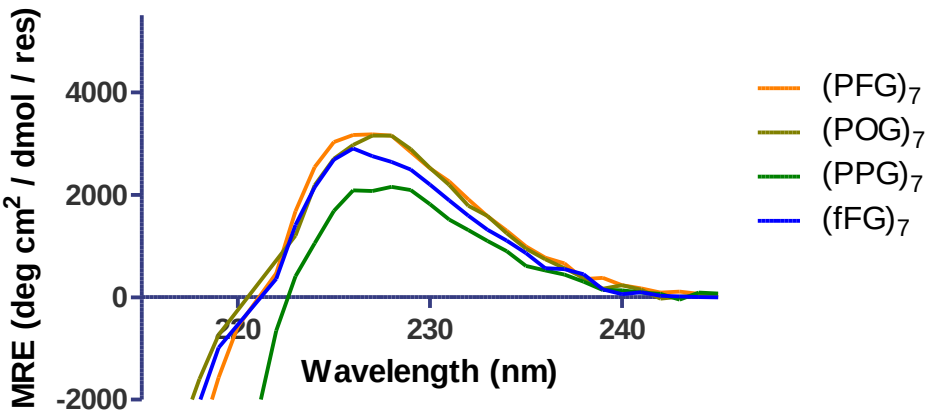


Figure 5.8. CD spectra of CMPs when annealed with an equimolar amount of nest-o**.** Data were obtained in 50 mM HOAc at 4 °C.

The nest-c has its two collagen strands in parallel. Yet, in the presence when of CMPs, we observed little triple-helical structure by CD spectroscopy (Figure 5.9). There is, however, a modest gain in triple-helix signal for $(Pro\text{-}Pro\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$ when added to the nest-c

compared to the individual peptides. The signal for these CMPs in the presence of nest-c, could be indicative of hetero-triple helix formation.

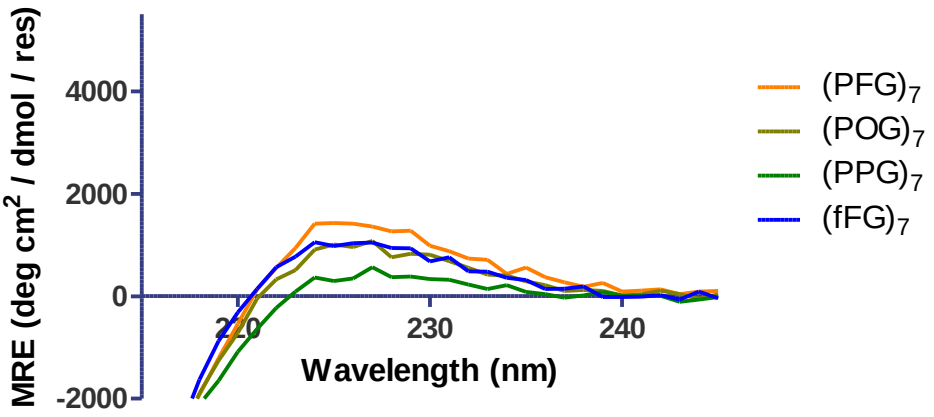


Figure 5.9. CD spectra of CMPs when annealed with an equimolar amount of nest-c**.** Data were obtained in 50 mM HOAc at 4 °C.

Repeating the annealing experiments with nest-r, which is the reduced version of nest-c, the CD shows improvement in triple helicity over the previous nest-c. Reducing the alkene functionality was hypothesized to improve flexibility within the N-terminal linker. As seen with nest-o, nest-r provides a distinct change in signal for the peptides (Pro-Pro-Gly)$_7$ and (flp-Flp-Gly)$_7$ compared to the peptides under the same conditions alone (Figure 5.10). There is restoration of CD signal for all peptides upon reduction of the nest-c to make nest-r, which is consistent with the enhanced flexibility of the linker.

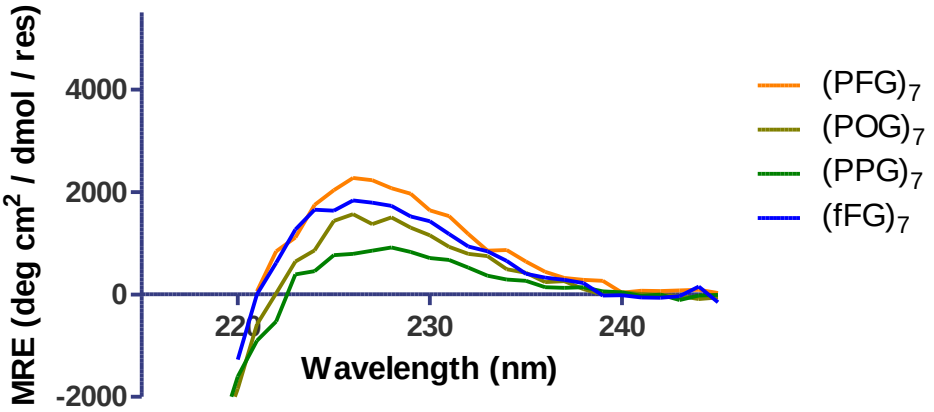


Figure 5.10. CD spectra of CMPs when annealed with an equimolar amount of nest-r**.** Data were obtained in 50 mM HOAc at 4 °C.

Before concluding that a hetero-triple helix is formed, it should be noted that both nest-o and nest-r display a diagnostic signal in their CD spectra when they are in isolation (Figure 5.11). The nest-o can splay open (*vide supra*), which means that the nest-o could adopt a conformation that would allow for a homo-triple helix. The nest-r however, is conformationally "locked" and with one turn of a triple helix being comprised of 3.3 residues,[269] it would be impossible for nest-r to make a complex with another nest-r in which a homo-triple helix is formed. The CD signal could then be from an incomplete triple helix in which two strands are wound around each other as if in a triple helix with a gap for a third strand. This winding could mimic enough of the triple helix structure to provide a reminiscent triple-helical signal in the CD spectra.

Regardless of the origin of the diagnostic triple-helical spectrum, its existence convolutes the interpretation of CD spectra in the presence of CMPs. One way that discern whether a hetero-triple helix is formed is to examine the thermal denaturation curves from $(Pro\text{-}Hyp\text{-}Gly)_7$ and

(Pro-Flp-Gly)$_7$ annealed to the nest-r (Figure 5.12). Both CMPs form a homo-triple helix, so if nest-r and CMP are creating two independent homo-triple helixes rather than forming a hetero-triple helix we would see two inflection points in the melting curves, as the two species would each likely have a distinct T_m. In combination, however, there is still one inflection point. Additionally, the curves of the mixtures are shallow like nest-r alone. If two independent complexes were forming, some of the curve steepness of the individual CMPs should manifest in the curves of the mixtures. These data indicate that there is one complex formed that is likely a hetero-triple helix—the one between the CMP and nest. Based on this interaction for (Pro-Hyp-Gly)$_7$ and (Pro-Flp-Gly)$_7$, we infer that (Pro-Pro-Gly)$_7$ and (flp-Flp-Gly)$_7$ are also forming a hetero-triple helix with the nest-r.

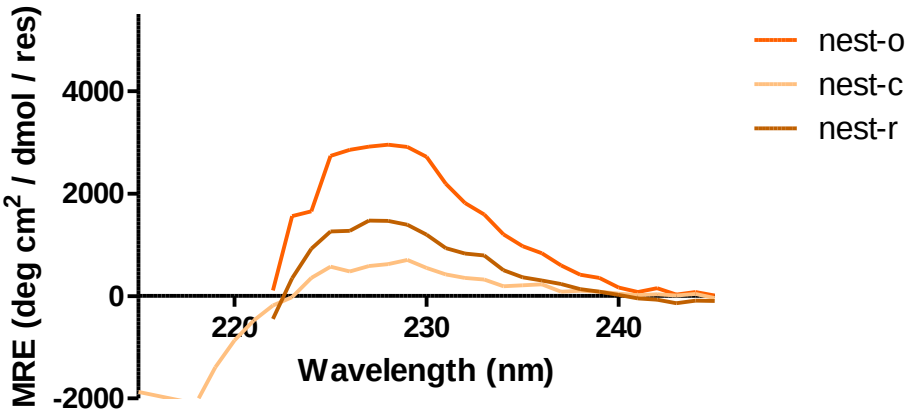


Figure 5.11. CD spectra of the nests when annealed alone**.** Data were obtained in 50 mM HOAc at 4 °C.

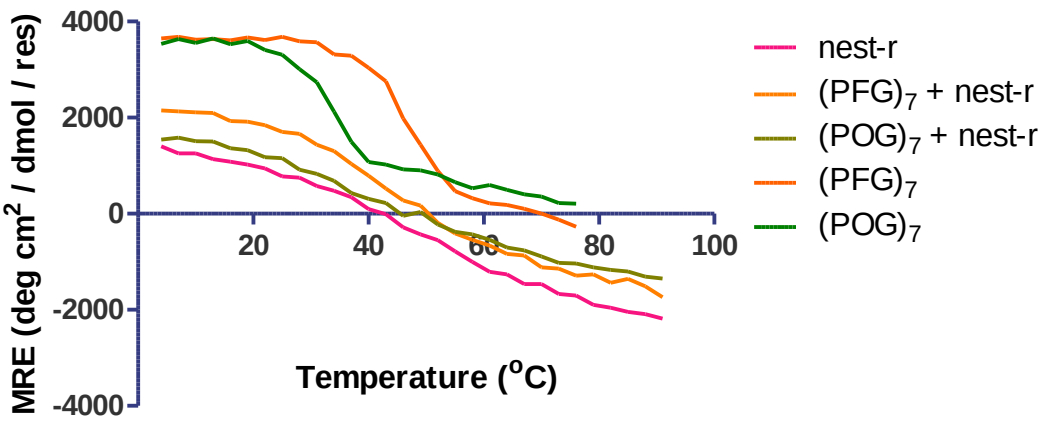


Figure 5.12. Thermal denaturation curves of nest-r, CMPs (Pro-Hyp-Gly)$_7$ and (Pro-Flp-Gly)$_7$, alone and in equimolar mixtures. Data were obtained in 50 mM HOAc.

5.3.3 Analysis of nest complex formation with ultracentrifugation

As a way to validate the assumptions about how nest-o and nest-r behave in solution independent of the addition of CMPs, solutions of the nests were subjected to analytical ultracentrifugation (AUC). Sedimentation will allow for evaluation of possible complex formation. Because both nests show the characteristic CD spectrum of a triple helix but only nest-o has the conformational flexibility to make complexes that contain a triple helix, sedimentation with AUC should show that complexes are formed by nest-o, but nest-r would remain as a monomer. Data collected at 34,000 rpm for nest-o and nest-r were fitted to monomer + pentamer and monomer + trimer, respectively. For nest-o there was a mixture of monomer and pentamer as is evident by the data strongly deviating from linearity (Figure 5.13). For nest-r, the data remained essentially linear with minor deviation from the slope of monomer sedimentation

(Figure 5.14). As predicted, AUC showed complex formation by nest-o, whereas nest-r remained in an uncomplexed state.

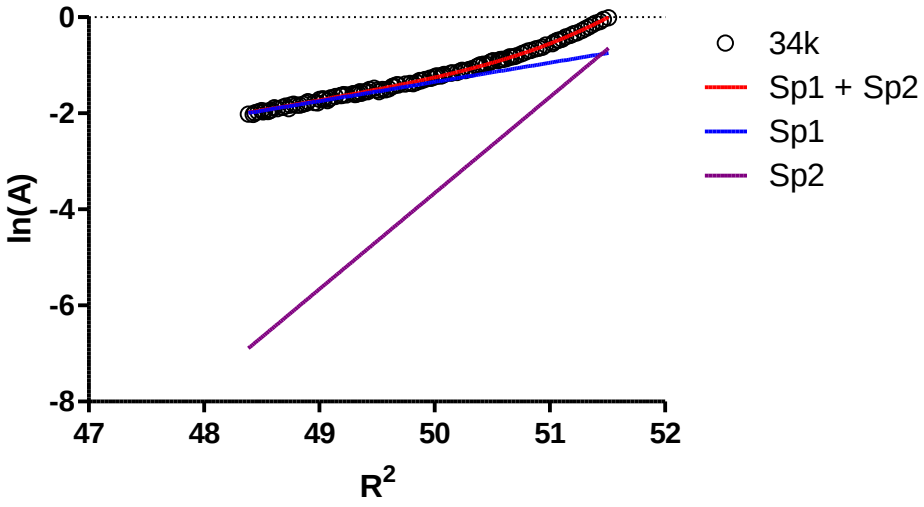


Figure 5.13. AUC data at 34k rpm of nest-o. Data were fitted to a monomer (Sp1) + pentamer (Sp2) mixture. Graphical representations of sedimentation of Sp1 or Sp2 are present as a comparison of the deviation of the AUC data from linearity.

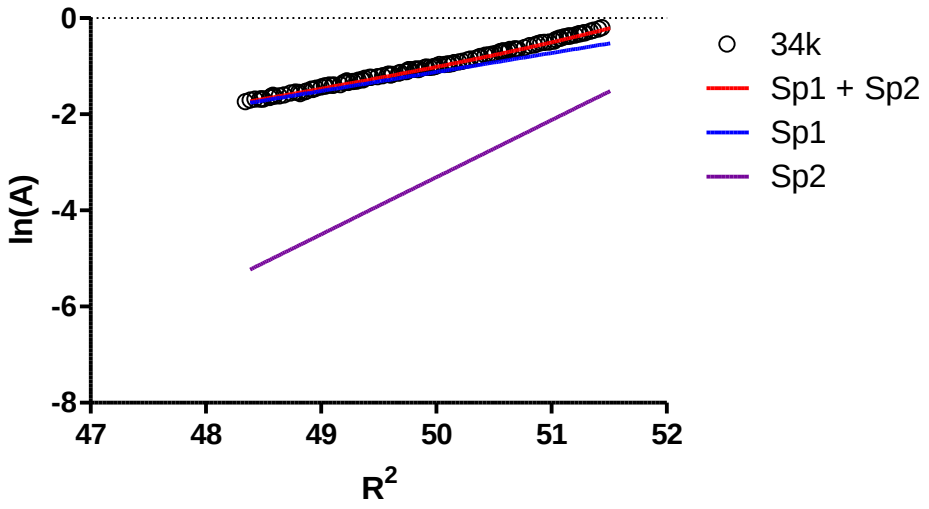


Figure 5.14. AUC data at 34k rpm of nest-r. Data were fitted to a monomer (Sp1) + trimer (Sp2) mixture. Graphical representations of sedimentation of Sp1 or Sp2 are present as a comparison of the deviation of the AUC data from linearity.

5.3.4 Analysis of nest–CMP complex formation with fluorescence spectroscopy

To confirm that there is a hetero-triple helix between the nests and the invasive CMPs, we employed a second method for detecting association. A fluorescein probe was tethered to the weaker binder of the two invasive strands, $(Pro\text{-}Pro\text{-}Gly)_7$. Nest-o and nest-r were then conjugated to a fluorescent bead. Upon mixing the nest–bead and CMP–fluorescein conjugates, coinciding fluorescence would indicate association and hetero-triple helix formation.

Attachment of the nests to the bead was accomplished through a biotin/streptavidin conjugation (Figure 5.15). Fluorescent beads were obtained pre-coated with streptavidin. To add biotin to the nests, an additional side-chain–protected lysine was added prior to the branching of

the two collagen strands on the C-terminus. Then, after cleavage of the nest from the resin, the newly revealed amino group on lysine would react with a biotin-NHS ester. A solution of purified biotin-nest-o or biotin-nest-r was added to the streptavidin-coated beads. Beads incubated with the nests, were then put under annealing conditions with the CMP–fluorescein conjugate. Any ensuing labeling of the beads was assessed with confocal microscopy and evaluated with flow cytometry.

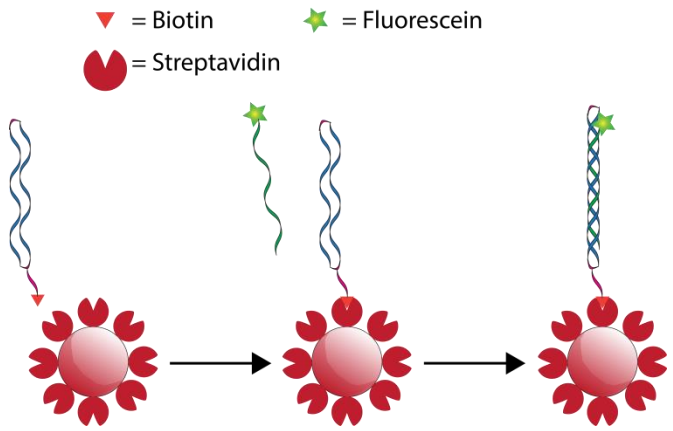


Figure 5.15. Cartoon representation of streptavidin-coated beads binding to biotin–nest-r followed by hetero-triple helix formation of nest-r with an invasive CMP labeled with fluorescein.

To show that labeling by the CMP–fluorescein conjugate is unique to when the nests are displayed on the bead, nest-free beads were also evaluated. In addition to the negative control of the nest free beads, a CMP control was employed. If adherence to the nest-coated beads were due to triple-helix formation with the CMP, then only the L-peptide would be able to form a proper

left-handed triple helix. The enantiomeric D-peptide should not be able to form a hetero-triple helix with the nests and therefore not label the nest-coated beads. If, however, the labeling was due to electrostatic or hydrophobic interactions, then D-CMP should behave the same as L-CMP.

A short $(Ser\text{-}Gly)_3$ peptide sequence was attached to both fluorescein and biotin, as a positive control, and then incubated with the beads. This peptide served as a model of an overlay of fluorescein and a fluorescent bead. Through the use of confocal microscopy, a green halo was observed superimposed upon the red fluorescence of the bead (Figure 5.16). The same pattern was is visible for nest-o and nest-r coated beads when mixed with L-CMP–fluorescein. As for beads uncoated with nests, treatment with L-CMP–fluorescein leaves the beads unlabeled with green florescence, meaning that the labeling occurred only when a nest is tethered to the surface of the bead (Figure 5.17). D-CMP–fluorescein mixed with the nest-coated beads also remained mostly unlabeled with a slight green halo. Because the L-CMP–fluorescein conjugate imbues the nest coated-beads with a distinct halo compared to the D-CMP-fluorescein conjugate, we conclude that L-CMP forms a hetero-triple helix with the nest-coated beads.

The confocal images were corroborated by examining a larger population of beads with flow cytometry (Figure 5.18). Again, the beads showed labeling with fluorescein in the positive control and in the nest-coated beads incubated with the L-CMP–fluorescein conjugate. There is also a small amount of labeling from D-CMP-fluorescein with the nest coated beads. Because the labeling from the D-CMP-fluorescein cannot be from hetero-triple helix formation, residual association must occur through other means. Still, for both nest-o and nest-r labeled beads, labeling with L-CMP-fluorescein is greater than with D-CMP-fluorescein.

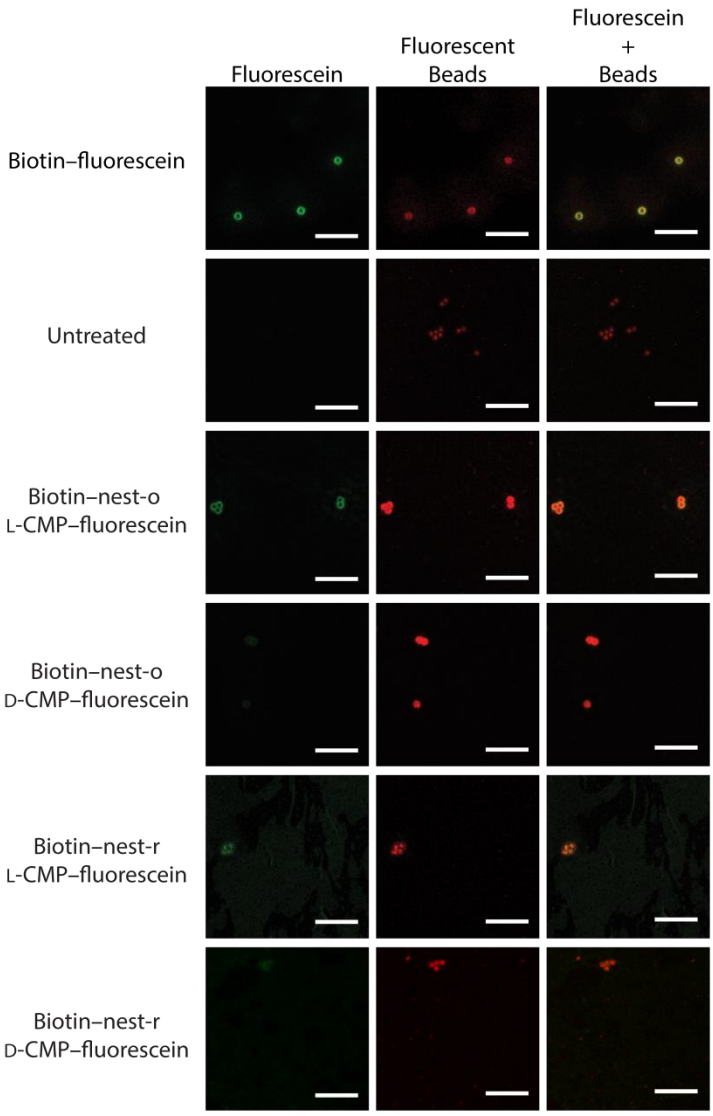


Figure 5.16. Confocal microscopy of streptavidin-coated beads treated with nest-o or nest-r and then annealed to L-CMP–fluorescein or D-CMP–fluorescein. Also included are a positive control of biotin–fluorescein and negative control of untreated beads. Scale bar: 10 μm.

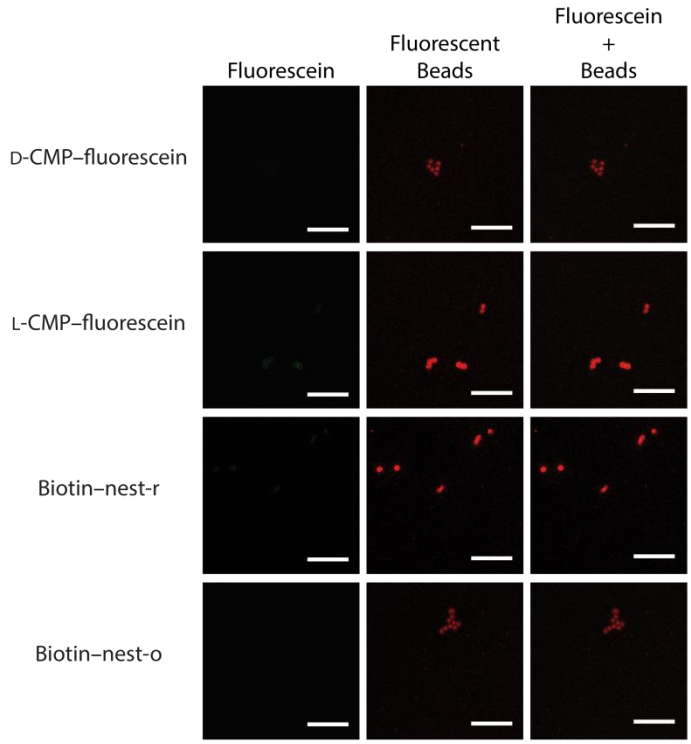


Figure 5.17. Confocal microscopy images of streptavidin-coated beads treated with nest-o, nest-r, L-CMP–fluorescein or D-CMP–fluorescein. Scale bar: 10 μm.

Because nest-o is not closed on its N-terminus, was expected its hetero-triple helix with a CMP to be more dynamic than that of the closed nest-r. Thus, if a hetero-triple helix is forming, then nest-r should bind tighter to the CMP over time; but if some other interaction is occurring, both nests should bind the CMP with similar kinetics. After annealing L-CMP–fluorescein to the two nest-coated beads, the beads were agitated at room temperature for 9 h. Again using flow cytometry, nest-r coated beads maintain labeling whereas nest-o–coated beads have diminished fluorescence. These data again support the hypothesis that the interaction is through the formation of a hetero-triple helix.

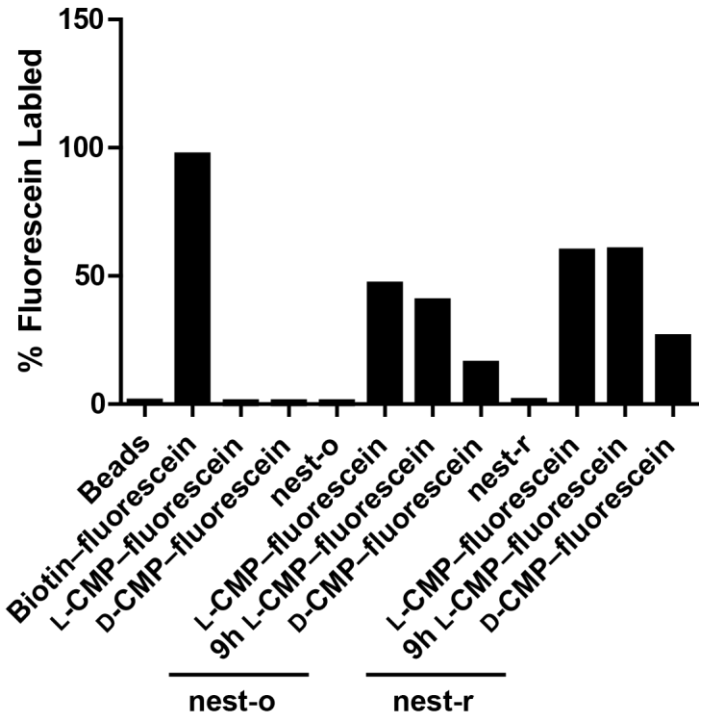


Figure 5.18. Flow cytometry of streptavidin-coated beads treated with nest-o or nest-r and then annealed with L-CMP–fluorescein or D-CMP–fluorescein. Beads were also treated with nest-o, nest-r, L-CMP–fluorescein or D-CMP–fluorescein in isolation. Also included are positive control biotin-fluorescein and negative control of untreated beads.

5.3.5 Comparison between CMPs

With support from fluorescent binding studies that nests form a hetero-triple helix with an invasive CMP, trends can be extracted from the signal for triple helicity and the thermal denaturation curves provided by CD spectroscopy. These comparisons can elucidate which invasive CMP, if any, is best for annealing to a surface of damaged collagen. In addition to (Pro-Pro-Gly)$_7$ and (flp-Flp-Gly)$_7$, potential invasive strands (Pro-Ile-Gly)$_7$ and (flp-Hyp-Gly)$_7$ were also annealed to nest-r (Figure 5.19).

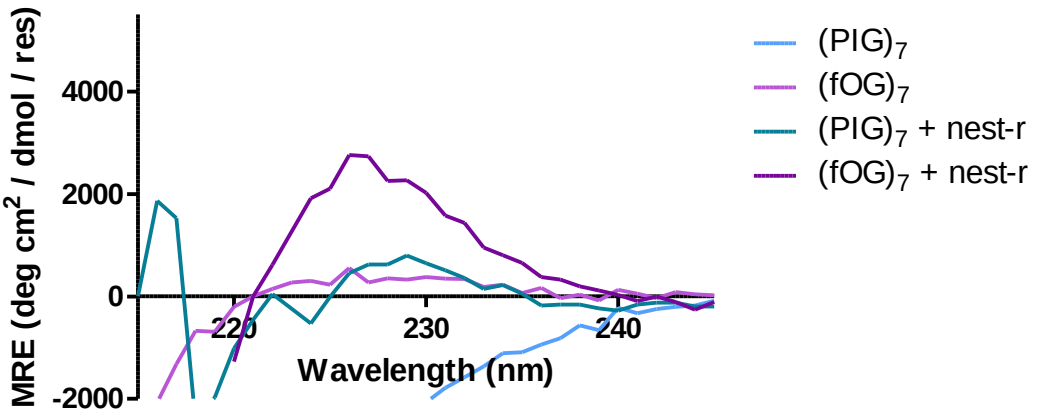


Figure 5.19. CD spectra of CMPs alone and annealed with nest-r**.** Data were obtained in 50 mM HOAc at 4 °C.

Molar ellipticity at 226 is diagnostic for a collagen triple helix.[184] Looking at signal intensity for the CMPs bound to the wound-bed mimic, nest-r, relative triple helicity can be gaged. Relative triple helicity provides a comparison of how well the CMP is able to bind to a wound-bed mimic. CMP (flp-Flp-Gly)$_7$, with the two fluoroprolines desirably influencing the equilibrium of the ring pucker should show the strongest signal. Closely following would be (flp-Hyp-Gly)$_7$, hydroxyproline not being as electronegative and therefore less influential in the ring pucker equilibrium. Next in predicted signal intensity would be (Pro-Pro-Gly)$_7$. Isoleucine is known to be destabilizing to a triple helix,[26,270] and thus (Pro-Ile-Gly)$_7$ should show the least hetero-triple helix formation.

For the four invasive strands CMPs bound to nest-r, (flp-Hyp-Gly)$_7$ showed the strongest signal followed by (flp-Flp-Gly)$_7$, (Pro-Pro-Gly)$_7$, and (Pro-Ile-Gly)$_7$ in order (Table 5.1). This order is close to the predicted order based on the constituent amino acid residues. The one

variant is the order of $(flp\text{-}Hyp\text{-}Gly)_7$ and $(flp\text{-}Flp\text{-}Gly)_7$. This variation could be explained by increased interstrand steric clashes caused by the stronger induced ring pucker by fluoroproline over hydroxylated proline. The correlation with CMP signal in CD with predicted triple-helix stabilizing effects further suggests hetero-triple helix formation.

Table 5.1. Molar ellipticity of CMPs annealed to nest-r at 226 nm. Data are derived from the CD spectra in Figure 5.10 and Figure 5.19.

Peptide	***MRE (deg cm^2/dmol/res)***
$(PPG)_7$ + nest-r	788.76
$(fFG)_7$ + nest-r	1836.01
$(fOG)_7$ + nest-r	2758.50
$(PIG)_7$ + nest-r	451.00

The thermal denaturation curve for the optimized wound-bed mimic, nest-r, is shallow with an inflection point at 49 °C (Figure 5.20). A shallow melting curve is not unexpected, as the two collagen strands are physically held together making disassociation through unwinding more difficult and, hence, slower. When incubated with the four invasive CMPs, $(flp\text{-}Hyp\text{-}Gly)_7$, $(flp\text{-}Flp\text{-}Gly)_7$, $(Pro\text{-}Pro\text{-}Gly)_7$, and $(Pro\text{-}Ile\text{-}Gly)_7$, the denaturation curves remain shallow and have T_m values within a few degrees of nest-r alone (Table 5.2). With the melting curves remaining shallow and the T_m values for all complexes being close, the data suggest that binding dynamics are largely controlled by the wound bed mimic, nest-r. Similar trends can be seen for both nest-o (Figure 5.21) and nest-c (Figure 5.22), corroborating that the unwinding rate of the whole hetero-triple helix structure is determined by that of its slowest component, the nests.

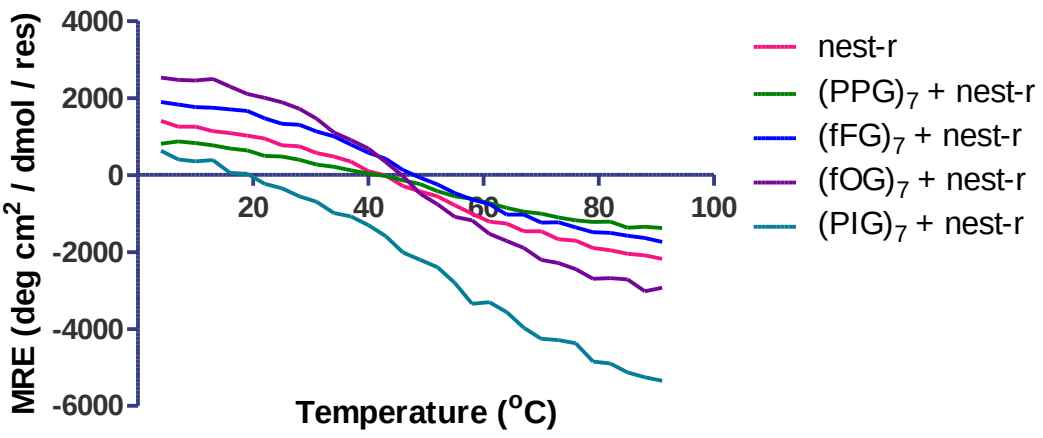


Figure 5.20. Thermal denaturation curves of nest-r alone and annealed to CMPs. Data were obtained in 50 mM HOAc.

Table 5.2. Values of T_m for nest-r alone and annealed to CMPs. Data are derived from the thermal denaturation curves in Figure 5.12 and Figure 5.20.

Peptide	**T_m (°C)**
nest-r	49.2
$(PFG)_7$ + nest-r	47.3
$(POG)_7$ + nest-r	42.3
$(PPG)_7$ + nest-r	48.5
$(fFG)_7$ + nest-r	46.7
$(fOG)_7$ + nest-r	47.3
$(PIG)_7$ + nest-r	52.6

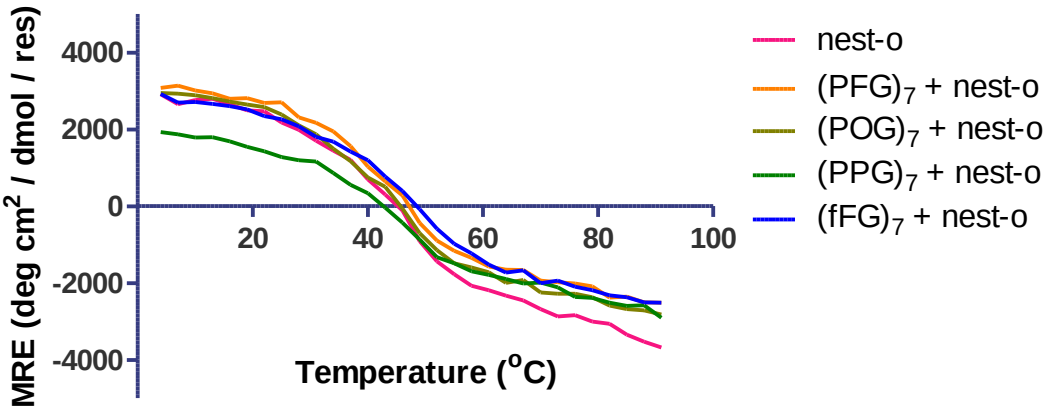


Figure 5.21. Thermal denaturation curves of nest-o alone and annealed to CMPs. Data were obtained in 50 mM HOAc.

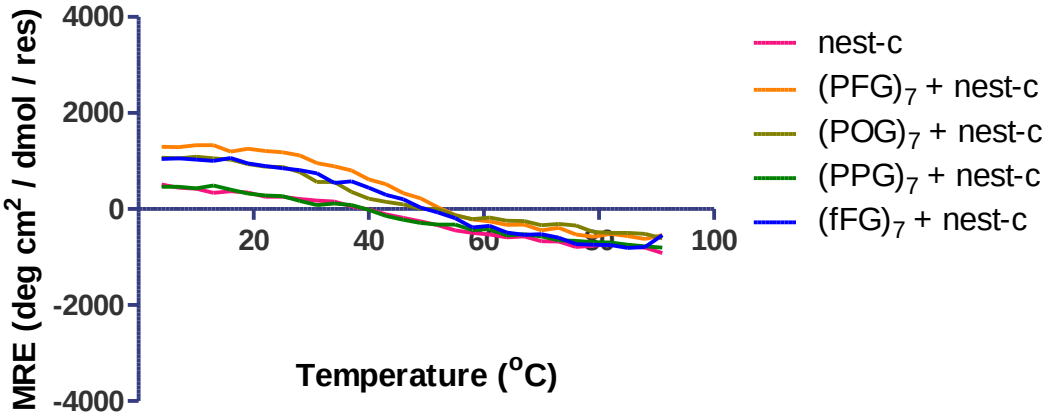


Figure 5.22. Thermal denaturation curves of nest-c alone and annealed to CMPs. Data were obtained in 50 mM HOAc.

5.4 Conclusions

Collagen duplexes have been of interest for studying potential biomaterials[198] and for possible use as a tag for fusion-protein recognition.[266] In a wound, collagen is damaged and denatured from its idealized triple helix. We reasoned that a collagen duplex could be developed as a wound-bed mimic. To optimize the duplex to behave as two long parallel collagen strands lacking a third strand, both N and C termini were tethered together. Tethering was accomplished on the C-terminus by building collagen mimetic peptides off of both amino groups of lysine. The N-terminal closing of the macrocycle was ultimately successful with olefin metathesis, followed by the reduction of the alkene. This synthetic route provided a novel collagen duplex, nest-r, which could participate in host–guest studies acting as a nest for a third strand to form a hetero-triple helix.

Interestingly, nest-r shows triple helix character by CD spectroscopy despite its topological inability to form triple helix. AUC confirms that nest-r exists primarily as a monomer in solution. Its "triple helix" signal could arise from a double helix that maintains the pitch of a triple helix lacking the third strand. The origin of this apparent "triple helicity" could provide insight into kinetics of triple-helix formation.

Using nest-r as a wound-bed mimic, fluorescent studies show preferential binding of the L-peptide, $(Pro\text{-}Pro\text{-}Gly)_7$, over the D-peptide. As predicted assuming hetero-triple helix formation, L-CMP-fluorescein showed slower dissociation from nest-r compared to nest-o. These data, in combination with the analysis of the melting curves when annealing $(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$, indicate that nest-r can act as a host, mimicking the environment of a wound for invasive binding through hetero-triple helix formation.

(flp-Hyp-Gly)$_7$ was the best strand at forming a hetero-triple helix with nest-r. The ability of hetero-triple helix formation can be attributed to the preorganization endowed by fluoroproline and hydroxyproline. In the context of a wound, (flp-Hyp-Gly)$_7$ will be best at initiating binding, though any of the tested strands will provide adherence.

5.5 Materials and methods

5.5.1 Instrumentation

Solid-phase peptide synthesis was performed on a Liberty Blue Peptide Synthesizer from CEM (Matthews, NC). Synthetic peptides were purified by HPLC with a Prominence instrument from Shimadzu (Kyoto, Japan) equipped with a VarioPrep 250/21 C18 column from Macherey–Nagel (Düren, Germany). Molecular mass was determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid or sinapic acid matrix with a Bruker microflex LRF™ (Billerica, MA) Purity analyses were performed with an Acquity UPLC® H-Class system from Waters that was equipped with an Acquity photodiode array detector, Acquity quaternary solvent manager, Acquity sample manager with a flow-through needle, Acquity UPLC® BEH C18 column (2.1´ 50 mm, 1.7-μm particle size) and Empower 3 software. ^{1}H and ^{13}C NMR spectra were acquired with an Avance III 400 spectrometer from Bruker (Billerica, MA). All CD data were acquired with a 420 CD spectrophotometer from Aviv Biomedical at the UW Biophysics Instrumentation Facility (BIF). Beads were imaged using a Eclipse Ti inverted confocal microscope from Nikon (Melville, NY) at the Biochemistry Optical Core of the University of Wisconsin−Madison. Flow cytometry was done with an Accuri Flow Cytometer with C-Sampler from BD (San Jose, CA) at the UW BIF. Sedimentation equilibrium experiments were performed at the UW BIF with an

XL-A analytical ultracentrifuge from Beckman Coulter (Brea, CA) equipped with an An-60 Ti rotor.

5.5.2 General

Amino acid derivatives, resins, Fmoc-OSu, and HOBt were from Chem-Impex International (Wood Dale, IL). Boc-Flp-OH, Boc-flp-OH, Fmoc-Flp-OH, and Fmoc-flp-OH were obtained from OmegaChem (Lévis, Québec). DIC and 4-methyl-piperidine were from Oakwood Chemical (Estill, SC). Streptavidin-coated fluorescent blue particles were from Spherotech (Lake Forest, IL). 6-Aminohexanoic acid and all other reagents were from Sigma–Aldrich (St. Louis, MO) and were used without further purification. *N*,*N*-dimethylformamide (DMF) was dried with a Glass Contour system from Pure Process Technology (Nashua, NH). In addition, DMF was passed through an associated isocyanate "scrubbing" column to remove any amines. Water was purified by an Arium Pro from Sartorius (Goettingen, Germany).

5.5.3 Volatile removal

The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials with a rotary evaporator at water aspirator pressure (<20 torr) while maintaining a water bath below 40 °C. The phrase "volatiles were removed under reduced pressure" refers to the removal of solvents and other volatile materials with a rotary evaporator at vacuum-pump pressure (~2 torr) while maintaining a water bath below 40 °C.

5.5.4 Small molecule synthesis

Fmoc-6-Aminohexanoic acid: 6-Aminohexanoic acid (1.00 g, 7.62 mmol) was dissolved in a saturated aqueous solution of $NaHCO_3$ (50 mL). In a separate flask, Fmoc-OSu (2.82 g, 8.38 mmol) was dissolved in dioxane (50 mL). The two solutions were combined, and the mixture

became cloudy, and was stirred for 16 h. The reaction mixture was then concentrated under reduced pressure. The residue was dissolved in EtOAc and washed with aqueous 1 N HCl and brine. The organic layer was dried over Na_2SO_4(s), decanted, and concentrated under reduced pressure. Crude product was purified by silica gel chromatography, eluting with 1% v/v acetic acid and 40% v/v EtOAc in hexanes to yield Fmoc-6-aminohexanoic acid (2.556 g, 95%) as a white solid. HRMS–ESI (*m*/*z*): $[M + 1]^+$ calcd, 354.17; found, 354.17. ^{1}H NMR (400 MHz, MeOD, δ): 7.78 (d, J = 7.5 Hz, 2H), 7.63 (d, J = 7.5 Hz, 2H), 7.37 (t, J = 7.4 Hz, 2H), 7.33–7.26 (m, 2H), 4.33 (d, J = 6.8 Hz, 2H), 4.18 (t, J = 6.9 Hz, 1H), 3.08 (t, J = 7.0 Hz, 2H), 2.27 (t, J = 7.4 Hz, 2H), 1.60 (p, J = 7.5 Hz, 2H), 1.49 (p, J = 7.1 Hz, 2H), 1.33 (p, J = 10.1, 6.0 Hz, 2H). ^{13}C NMR (101 MHz, MeOD, δ): 157.49, 143.95, 141.20, 127.34, 126.71, 124.75, 119.50, 66.11, 47.13, 40.17, 33.71, 29.18, 25.97, 24.45.

Synthesis of Boc-Flp-OBn: Boc-Flp-OH (2.5 g, 10.7 mmol) was dissolved in DMF. Solid Cs_2CO_3 (1.5 g, 10.7 mmol) was added, and the reaction mixture was stirred for 10 min. Benzyl bromide (1.27 mL, 10.7 mmol) was added dropwise, and the reaction mixture was stirred for 16 h. The reaction mixture was then concentrated under reduced pressure. Crude product was purified by silica gel chromatography, eluting with 10% v/v EtOAc in hexanes to yield (2.26 g, 65 %). HRMS–ESI (*m*/*z*): $[M + 1]^+$ calcd, 324.15; found, 324.16. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.35 (d, J = 4.5 Hz, 5H), 5.32–5.19 (m, 1H), 5.19–5.14 (m, 1H), 5.14–5.04 (m, 1H), 4.58–4.38 (m, 1H), 3.99–3.75 (m, 1H), 3.61 (ddt, J = 36.1, 13.0, 3.9 Hz, 1H), 2.70–2.47 (m, 1H), 2.20–1.97 (m, 1H), 1.53–1.31 (m, 9H).

Synthesis of Boc-flp-Flp-OBn: Boc-Flp-OBn (2.26 g, 6.99 mmol) was dissolved in 4 N HCl (8.0 mL), and the reaction mixture was left to stir for 30 min. The reaction mixture was then concentrated under reduced pressure. The residue was dissolved in DMF. DIEA (4.87 mL, 27.96

mmol) was added dropwise. Solid PyBrOP (3.91 g, 8.39 mmol) and Boc-flp-OH (1.79 g, 7.69 mmol) was added, and the reaction mixture was stirred for 16 h. Volatiles were removed under reduced pressure. The residue was taken up in EtOAc and washed successively with 1 M HCl (2x), saturated aqueous $NaHCO_3$ (2x), and brine (2x).The organic layer was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure, yielding crude product (3.96 g, 129 %). HRMS–ESI (*m*/*z*): $[M + 1]^+$ calcd, 439.20; found, 439.20. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.35 (d, J = 4.0 Hz, 6H), 5.32–5.10 (m, 4H), 4.55 (dt, J = 20.6, 8.4 Hz, 1H), 4.06–3.77 (m, 4H), 3.66 (ddt, J = 36.6, 13.0, 3.5 Hz, 1H), 2.62 (dddt, J = 28.3, 16.6, 8.0, 1.8 Hz, 1H), 2.22–1.99 (m, 2H), 1.85 (ddt, J = 71.3, 13.3, 6.7 Hz, 1H), 0.94 (d, J = 6.7 Hz, 3H), 0.85 (dd, J = 6.7, 2.3 Hz, 4H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 172.16, 172.04, 155.00, 154.56, 135.47, 135.28, 128.65, 128.58, 128.52, 128.42, 128.33, 128.16, 92.65, 91.86, 90.86, 90.08, 77.25, 71.88, 71.83, 67.09, 67.02, 60.40, 57.72, 57.48, 53.58, 53.35, 53.15, 52.93, 37.77, 37.54, 36.71, 36.48, 28.36, 28.17, 27.98, 27.86, 19.03, 18.96, 18.92, 14.22.

Synthesis of Fmoc-Gly-flp-Flp-OBn: Crude Boc-flp-Flp-OBn (3.96 g, 9.02 mmol) was dissolved in 4 N HCl (8.0 mL) and reaction was left to stir for 30 min. The reaction mixture was then concentrated under reduced pressure. The residue was dissolved in DCM. DIEA (3.65 mL, 20.97 mmol) was added dropwise to the resulting solution. Solid Fmoc-Gly-OPfp (3.24 g, 6.99 mmol) was added, and the reaction mixture was stirred for 16 h. Volatiles were removed under reduced pressure. The residue was taken up in EtOAc and washed successively with 1 M HCl (2x), saturated aqueous $NaHCO_3$ (2x), and brine (2x).The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure. Crude product was purified via silica gel chromatography, eluting with dichloromethane followed by a methanol flush to yield product (3.70 g, 86 %). HRMS–ESI (*m*/*z*): $[M + NH_4]^+$ calcd, 635.23; found, 635.27 ^{1}H NMR (400

MHz, $CDCl_3$, δ): 7.75 (d, J = 7.5 Hz, 2H), 7.59 (dd, J = 7.5, 3.2 Hz, 2H), 7.43–7.26 (m, 9H), 5.86–5.48 (m, 1H), 5.44–5.02 (m, 4H), 4.85–4.67 (m, 2H), 4.45–4.02 (m, 4H), 4.03–3.60 (m, 3H), 2.62 (ddd, J = 20.3, 14.1, 8.1 Hz, 1H), 2.56–2.22 (m, 2H), 2.08 (dddd, J = 40.9, 18.3, 9.1, 4.8 Hz, 1H), 1.82 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 171.14, 168.75, 167.23, 156.27, 143.88, 141.26, 135.40, 128.58, 128.39, 128.29, 127.69, 127.10, 125.21, 119.95, 91.95, 77.25, 67.21, 57.96, 57.13, 47.08, 43.41, 35.31, 34.48, 34.26.

Synthesis of Fmoc-Gly-flp-Flp-OH: (3.70 g, 5.98 mmol) was dissolved in methanol (25 mL). The head space was purged with N_2(g). Pd/C (10% w/w, 0.64 g) was added, and the flask was capped with a septum. H_2(g) was added via a balloon. The reaction was monitored by TLC and observed to be complete at 6 h. The reaction mixture was filtered through celite and concentrated under reduced pressure. Crude product was purified by silica gel chromatography, eluting with 1% v/v acetic acid and 20% v/v methanol in EtOAc to yield product (2.95 g, 93 %). HRMS–ESI (m/z): $[M - 1]^-$ calcd, 526.52; found, 526.18 ^{1}H NMR (400 MHz, MeOD, δ): 7.81 (d, J = 7.6 Hz, 2H), 7.74–7.64 (m, 2H), 7.48–7.28 (m, 4H), 5.49–5.03 (m, 3H), 4.72–4.20 (m, 4H), 4.20–3.43 (m, 5H), 2.85–2.27 (m, 3H), 2.26–2.06 (m, 1H). ^{13}C NMR (101 MHz, $CDCl_3$, δ): 177.16, 173.49, 172.00, 160.94, 147.71, 145.13, 132.21, 131.59, 131.14, 130.97, 130.78, 128.99, 123.79, 96.63, 96.29, 94.84, 94.48, 71.07, 68.30, 61.94, 61.35, 50.96, 46.99, 39.48, 39.26, 38.24, 38.03.

5.5.5 Peptide synthesis

All peptides were cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptides were then purified by preparative HPLC using a gradient of B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA).

Nest-o: The open nest was synthesized by first doing a 0.05-mmol coupling of Fmoc-Lys(Fmoc)-OH to TGT S RAM resin (0.22 mmol/g). Next, Fmoc-6-aminohexanoic acid was single coupled at a 0.10-mmol scale. All further amino acid and small-molecule carboxylic acid additions were at a 0.10-mmol scale and double-coupled. Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acids and small-molecule carboxylic acids (5 equiv) were converted to an active ester using DIC and HOBt. The nest-o peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 5757.05; found, 5757.07. A 0.05-mmol scale synthesis afforded 7.6 mg (2.7%) of nest-o after purification. Its purity was assessed to be >95% by UPLC.

Nest-c: Following the same synthetic route as for the open nest, the closed nest was obtained by preforming olefin metathesis on the N-terminal 4-butenoic acids while the peptide was still on the resin, following a procedure similar to that of peptide stapling.[268] Peptide bound to resin was added to a Schlenk flask. The resin was then dried for at least 3 h on a high-vacuum manifold. The flask was then rigorously purged with N_2(g) at room temperature. While under N_2(g), the resin was pre-swelled in 2.5 mL of dry CH_2Cl_2 for at least 15 min prior to the addition 0.5 mL of a 2.5 mM solution of Grubbs GII in dry CH_2Cl_2 using standard Schlenk technique. The reaction flask was quickly equipped with an oven-dried reflux condenser, purged with N_2(g), and heated in an oil bath at 40 °C for 36 h under N_2(g). After 36 h, another 0.5-mL aliquot of Grubbs GII solution was added, and the flask was heated at 40 °C for another 36 h under N_2(g). Occasional additions of dry solvent were added over the course of the reaction to maintain at least 3 mL of CH_2Cl_2. The reaction mixture was then allowed to cool to room temperature and filtered. The resin was washed with DCM to remove any remaining catalyst. Peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 5729.02; found,

5728.55. A 0.05-mmol scale synthesis afforded 7.0 mg (2.4%) of nest-c after purification. UPLC showed many over laying peaks.

Nest-r: Following olefin metathesis and filtering the resin to remove catalyst. The resin was returned to a Schlenk flask and suspended in DCE. Again following literature precedent,[271] Umicore M2 (46.5 mg, 0.05 mmol) and Et_3SiH (0.80 mL, 9.4 mmol) were added to the flask. The vessel was capped with a septum and the reaction mixture was heated to 60 °C for 72 h. The resin was washed with DCM to remove any remaining catalyst. The peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 5731.03; found, 5731.67 A 0.05-mmol scale synthesis afforded 9.8 mg (3.4%) of nest-r after purification. Its purity was assessed to be >95% by UPLC.

Biotin–nest: For biotin conjugates of the open nest and reduced nest, Fmoc-Lys(Boc)-OH was coupled first to TGT S RAM resin (0.22 mmol/g). Following the initial Lys a $(Gly\text{-}Ser)_3$ sequence was incorporated. The rest of the biotin-nest synthesis followed as described above, stopping after peptide synthesis and reduction for the open nest and reduced nest, respectively. The peptide was cleaved from the resin.

Biotin–nest-o: The crude open nest with the additional lysine residue (5.3 mg, 0.84 μmol) was dissolved in 500 μL of DMSO. A solution of Biotin-NHS ester (4.0 mg, 11.7 μmol) and DIEA (0.1 mL, 0.57 mmol) in 500 μL of DMSO was added to the peptide solution. The mixture was allowed to react for 12 h. The peptide was then purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 6543.38; found, 6541.63. A 0.05-mmol scale synthesis afforded 0.4 mg (0.52%) of biotin-nest-o after purification. Its purity was assessed to be >95% by UPLC.

Biotin–nest-r: Crude reduced nest with an additional lysine residue (4.3 mg mg, 0.68 mmol) was dissolved in 500 μL of DMSO. A solution of Biotin-NHS ester (6.7 g, 19.6 μmol) and DIEA (0.1 mL, 0.57 mmol) in 500 μL of DMSO was added to the peptide solution. The mixture was allowed to react for 12 h. The peptide was then purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 6517.37; found, 6519.48. A 0.05-mmol scale synthesis afforded 0.2 mg (0.14%) of biotin-nest-r after purification. Its purity was assessed to be >90% by UPLC.

Biotin–fluorescein: Fmoc-Lys(Boc)-OH was coupled first to TGT S RAM resin (0.22 mmol/g) from Chem-Impex International (Wood Dale, IL), followed by a $(Gly\text{-}Ser)_3$ sequence that was synthesized by single amino acid couplings. The amino acid monomers (5 equiv) were converted to an active ester using DIC and HOBt. Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. While still on resin, biotin-NHS ester (5 equiv) was used to attach biotin to the N-terminus. The peptide was then cleaved from the resin. A solution made of 5(6)-carboxyfluorescein (9.5 mg, 25.2 μmol), HATU (8.7 mg, 22.8 μmol), and DIEA (100 μL, 0.57 mmol) in 500 μL DMSO was allowed to react for 15 min. Crude biotin-$(Gly\text{-}Ser)_3$-Lys-NH_2 (12.3 mg) was dissolved in 500 μL of DMSO, and the resulting solution was again allowed to react for 12 h. Peptide biotin-$(Gly\text{-}Ser)_3$-Lys(fluorescein)-NH_2 is then purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 1162.41; found, 1162.52. A 0.05-mmol synthesis afforded 1.5 mg (2.5%) of biotin-$(Gly\text{-}Ser)_3$-Lys(fluorescein)-NH_2 after purification. Its purity was assessed to be >95% by UPLC.

$(flp\text{-}Flp\text{-}Gly)_7$: $(flp\text{-}Flp\text{-}Gly)_7$ was synthesized by using the Fmoc-Gly-flp-Flp-OH tripeptide and Fmoc-flp-OH and Fmoc-Flp-OH monomers. First, six segment condensations of tripeptide were followed by two additions of monomer, by each addition being double-coupled

on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was converted to an active ester using DIC and HOBt. The peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 2028.77; found, 2028.75. A 0.05-mmol scale synthesis afforded 8.3 mg (8.14%) of $(flp\text{-}Flp\text{-}Gly)_7$ after purification. Its purity was assessed to be >95% by UPLC.

$(flp\text{-}Hyp\text{-}Gly)_7$: $(flp\text{-}Hyp\text{-}Gly)_7$ was synthesized by double-coupling each amino acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was converted to an active ester using DIC and HOBt. The peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 2015.80; found, 2015.10. A 0.05-mmol scale synthesis afforded 6.4 mg (6.4%) of $(flp\text{-}Hyp\text{-}Gly)_7$ after purification. The purity was assessed to be >95% by UPLC.

$(Pro\text{-}Ile\text{-}Gly)_7$: $(Pro\text{-}Ile\text{-}Gly)_7$ was synthesized by amino acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was converted to an active ester using DIC and HOBt. The peptide was then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 1890.31; found, 1890.30. A 0.05-mmol scale synthesis afforded 9.6 mg (10.2%) of $(Pro\text{-}Ile\text{-}Gly)_7$ after purification. Purity was assessed as >95% by UPLC.

$(Pro\text{-}Pro\text{-}Gly)_7$: $(Pro\text{-}Pro\text{-}Gly)_7$ was synthesized by amino acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-

methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was converted to an active ester using DIC and HOBt. The peptide is then cleaved from the resin and purified by HPLC. MALDI (*m*/*z*): $[M + H]^+$ calcd, 1777.02; found, 1776.87. A 0.05-mmol scale synthesis afforded 18.0 mg (20.2%) of $(Pro\text{-}Pro\text{-}Gly)_7$ after purification. The purity was assessed to be >95% by UPLC.

$(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$: $(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$ were synthesized and purified as described previously.[272] For $(Pro\text{-}Hyp\text{-}Gly)_7$ and $(Pro\text{-}Flp\text{-}Gly)_7$, purity was assessed as >95% by UPLC.

D-CMP–fluorescein: Ac-Lys-$(Ser\text{-}Gly)_3$-$(D\text{-}Pro\text{-}D\text{-}Pro\text{-}Gly)_7$ was synthesized by amino acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was converted to an active ester by using DIC and HOBt. The peptide is then cleaved from the resin. A solution made of 5(6)-carboxyfluorescein (9.7 mg, 25.7 μmol), HATU (9.1 mg, 23.9 μmol) and DIEA (100 μL, 0.57 mmol) in 500 μL DMSO was allowed to react for 15 min. A solution of crude Ac-Lys-$(Ser\text{-}Gly)_3$-$(D\text{-}Pro\text{-}D\text{-}Pro\text{-}Gly)_7$ (36.0 mg) dissolved in 500 μL of DMSO was added to the solution and allowed to react for 12 h. The peptide was then purified by HPLC. MALDI (*m*/*z*): $[M + Na]^+$ calcd, 2759.20; found, 2760.47. A 0.05-mmol synthesis afforded 2.5 mg (1.8%) of Ac-Lys(fluorescein)-$(Ser\text{-}Gly)_3$-$(D\text{-}Pro\text{-}D\text{-}Pro\text{-}Gly)_7$ after purification. The purity was assessed to be >95% by UPLC.

L-CMP–fluorescein: Ac-Lys-$(Ser\text{-}Gly)_3$-$(Pro\text{-}Pro\text{-}Gly)_7$ was synthesized by amino acid addition on preloaded Fmoc-Gly-Wang resin (0.65 mmol/g). Fmoc-deprotection was achieved by treatment with 4-methyl-piperidine (20% v/v) in DMF. The amino acid monomer (5 equiv) was

converted to an active ester by using DIC and HOBt. The peptide was then cleaved from the resin. A solution made of 5(6)-carboxyfluorescein (112.4 mg, 0.30 mmol), HATU (104.8 mg, 0.28 mmol) and DIEA (100 μL, 0.57 mmol) in 500 μL of DMSO was allowed to react for 15 min. A solution of crude Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ (74.2 mg) dissolved in 500 μL of DMSO was added and allowed to react for 12 h. The peptide was then purified by HPLC. MALDI (*m*/*z*): $[M + Na]^+$ calcd, 2759.20; found, 2760.47. A 0.05-mmol scale synthesis afforded 3.5 mg (2.5%) of Ac-Lys(fluorescein)-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ after purification. The purity was assessed to be >95% by UPLC.

5.5.6 Circular dichroism spectroscopy

Peptide solutions were prepared in 50 mM HOAc. Many of the critical experiments in this study involve mixtures of peptides and linked dimers, and preparation of equimolar mixtures of the components is of vital importance. To ensure equimolar mixtures, relative concentrations were determined by integrating the absorbance of peptides at 218 nm during UPLC, and concentrations of all components in interrogated samples were matched to that of 180 μM (Pro-Pro-Gly)$_{10}$. The extinction coefficients of all Xaa-Yaa-Gly repeats were assumed to be identical at 218 nm and used to scale integrations according to peptide length. To facilitate the formation of the thermodynamic assembly product, peptide solutions were heated to 65 °C, then cooled to 4 °C at a rate of −12 °C per hour. Samples were left at 4 °C for at least 48 h before data acquisition.

CD spectra of peptides (180 μM) were recorded at 4 °C with a 1-nm band-pass filter and an averaging time of 3 s in a 0.1-cm path-length quartz cuvette. For thermal denaturation experiments, the CD signal was monitored at 226 nm as the sample was heated at a rate of 12 °C per hour in 3 °C steps. The value of T_m, which is the temperature at the midpoint of the thermal

transition, was calculated by fitting the data to a Boltzmann sigmoidal curve with the program Prism from GraphPad (La Jolla, CA).

5.5.7 Analytical ultracentrifugation

Samples prepared at 180 μM were diluted to 90 μM prior to the experiments. The sample (100 μL) and matching buffer (110 μL) were placed in a cell with an Epon 12 mm double-sector charcoal-filled centrepiece from Beckman Coulter. Experiments were run at 4 °C for more than seven days at speeds of 20,000, 26,000, 34,000, and 42,000 revolutions per minute (r.p.m.), and gradients recorded at 235 nm were monitored until superimposable 4 h apart. Equilibrium gradients at 4 °C were modelled as single and multiple non-interacting species through nonlinear least-squares fits to the gradient data. Non-sedimenting baselines attenuance were applied for all the samples during the analysis, which was performed with programs written by D. R. McCaslin (UW BIF) for IGOR PRO software from WaveMetrics (Lake Oswego, OR).

5.5.8 Nest-coated beads for confocal and flow cytometry

A 200-μL suspension of streptavidin-coated fluorescent blue particles was added to 200 μL of 58 μM either open nest or reduced nest solution in H_2O. The mixture was agitated for 9 h at room temperature. Beads were pelleted by centrifugation at 12,000 rpm for 2 min. Supernatant was removed and beads were resuspended in 200 μL of H_2O and stored at 4 °C.

A 10-μL aliquot of suspended beads was treated with either Ac-Lys(fluorescein)-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ or Ac-Lys(fluorescein)-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ so that the final concentrations were 60 μM. Mixtures were allowed to anneal by heating the samples to 65 °C and cooling slowly to 4 °C at a rate of −12 °C per hour. Keeping samples at 4 °C, beads were pelleted and the supernatant removed and resuspended in water. This process was repeated three

times to wash the beads. Finally, beads were resuspended in 300 μL of water. For confocal images, beads (3 μL) were spotted on a microscope slides and allowed to dry. For flow cytometry, resuspended beads were imaged with 488 nm and 640 nm lasers with 530/30 nm and 675/25 nm filters respectively.

5.6 Acknowledgements

We are grateful to Dr. Martha M. Vestling and Dr. Darrell R. McCaslin for facility support and contributive discussions. Thank you to the Landis group for use of their glove box. This work was supported by Grant R01 AR044276 (NIH) and Instrumentation Grant GM044783 (NIH). Aubrey J. Ellison was supported by Chemistry-Biology Interface (CBI) Training Grant T32 GM008505 (NIH). MALDI–TOF mass spectrometer was obtained by the chemistry department with funding by the Bender gift to the University Wisconsin–Madison.

Chapter 6

Future directions

6.1 New substituted prolines

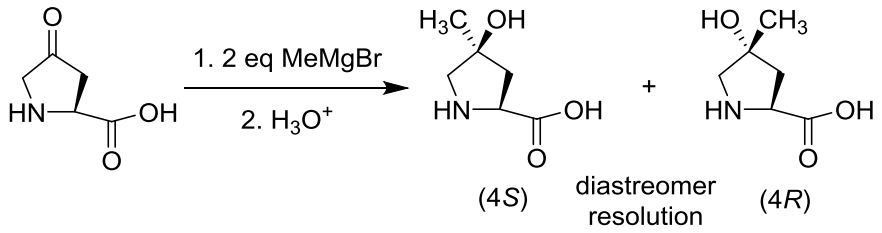
Many substituted prolines have served to tease out understanding of the stability of the collagen triple helix. These substituted prolines have been utilized for creating invasive collagen mimetic peptides (CMPs). Design of invasive CMPs relies on the disruption of their homo-triple helix but at the same time promotion of their contribution to a hetero-triple helix. This attribute allows for the binding of a CMP to a collagen surface.

Electronegative groups, such as hydroxy and fluoro, on the 4-position influence the ring pucker of proline through the gauche effect. But when substituted with bulkier groups such as a methyl, the ring pucker is now dictated by steric interactions. These effects result in (4*R)* –OH/F and (4*S*) $–CH_3$ substituted prolines all favoring a C^{γ}-exo ring pucker.

As the electronegative and steric directing substitutions direct from opposite faces of proline, they could be combined on the same proline. It would be expected that the directing of –OH/F and $–CH_3$ would prove to be synergistic further influencing the C^{γ}-exo equilibrium preference. (4*R*)-4-Fluoro-4-methyl-L-proline is in development by OmegaChem (Lévis, Québec) and soon will be commercially available and ready for study. In contrast, (4*R*)-4-hydroxy-4-methyl-L-proline is not in commercial development. It could be accessed through the reaction of 4-keto-L-proline and a methyl Grignard (Scheme 6.1), following separation of the diastereomeric products.

With these new substituted prolines, the C^{γ}-exo/C^{γ}-endo pucker equilibrium of the proline could be evaluated by ^{13}C NMR ratios. Additionally, (4*R*)-4-hydroxy-4-methyl-L-proline and (4*R*)-4-fluoro-4-methyl-L-proline could be incorporated in a CMP in the Yaa position of the Xaa-Yaa-Gly repeat. Using CD, the CMPs could be evaluated for tolerance of the new prolines in the

collagen triple helix. If the two groups work synergistically, as expected, then the resulting CMPs should form a highly stable triple helix. These proline modifications would add to the growing repertoire of additions that can tune the stability of a triple helix.



Scheme 6.1. Synthesis of (4*S*)-4-hydroxy-4-methyl-L-proline and (4*R*)-4-hydroxy-4-methyl-L-proline.

6.2 Invasive CMPs with controlled release of cargo

Invasive CMPs have been shown to bind to collagen surfaces and, through this adherence, coat the surface with a dye, growth factor, or sunscreen. In the use of growth factors or small-molecule drugs that are tethered to the CMP, the activity is currently limited to the outside of the cell. Incorporating a release mechanism into the peptide could expand the utility of invasive CMPs to deliver more intracellularly active drugs.

CMPs are currently utilized for slow drug release. In these manifestations, CMPs form a hydrogel that contains the drug within the matrix.[273] The drug is then released from the matrix diffusional control. Incorporating a chemical release mechanism into the design of invasive CMPs could improve upon drug delivery, as release could be controlled by relevant biological signals.

An example of a release mechanism that could be implemented employs protease cleavage. Many proteases are active during different parts of the wound-healing process.[274] One such protease is thrombin. Thrombin has a consensus sequence of Leu-Val-Pro-Arg-Gly-Ser, cleaving between Arg and Gly. This sequence could be included in the CMP through solid-phase peptide synthesis placing it between the collagen sequence and the cargo. Using the natural biological process, drug release could be implemented exactly when needed.

Before implementation *in vivo*, the release rate should be thoroughly studied *in vitro*. Modification of the consensus sequence could allow for tuning of the rate of release. Adhering a dye or drug to the surface that would be released only when needed could allow for CMPs to act as biosensors.

6.3 Nest studies

The creation and use of a double strand duplex was presented earlier in Chapter 5. This double-strand duplex, tethered at both ends, was dubbed nest-r. Nest-r showed utility in understanding the binding nature of invasive CMPs through the formation of a hetero-triple helix. Nest-r is a novel collagen structure that still requires more study as it could hold further insights into how the collagen triple helix folds.

As was stated previously, nest-r shows triple helix character even when alone in solution. AUC confirms that nest-r remains primarily monomeric, indicating that the "triple helix" CD signal arises from the structure of the nest alone. Structural analysis by X-ray crystallography or by solution NMR spectroscopy would allow for further insight in to the seemingly triple-helical nature of nest-r. Another way to probe the structure would be through more indirect methods such as enzymatic recognition. CNA, an adhesin from *Staphylococcus aureus*, or other proteins

that recognize a triple helix could be used in the presence of nest-r. Binding of CNA-fluorophore conjugate to a nest-r labeled surface could, for example, provide evidence for a "double helix" that maintains the pitch of a triple helix. The nature of the structure could provide insight into how denatured collagen may represent itself in host tissue.

The nest also provides a unique opportunity to study the kinetics of CMP association and hetero-triple helix formation. This analysis could be accomplished through Förster resonance energy transfer (FRET). Nest-r can be modified through the addition of a lysine on the C-terminus before branching of the two collagen mimetic peptides (Figure 6.1). The amino group on the side chain of lysine enables attachment of a coumarin 343 as part of a FRET pair. The corresponding FRET pair of 5(6)-carboxy-2′,7′-dichlorofluorescein could be attached to a CMP, again using an additional lysine residue as an attachment point. Initial tests contained a $(Ser\text{-}Gly)_3$ spacer between the C-terminal lysine residue and the collagen sequence for nest-r and CMP. This spacer could have led to a false negative result as the FRET pair was not within the necessary Förster's radius. Exclusion or shortening of the $(Ser\text{-}Gly)_3$ spacer could alleviate this issue, bringing the FRET pair to the proper distance when bound in a hetero-triple helix. FRET could allow for the real-time analysis of the rate of hybridization of the invasive CMPs to nest-r when mixed together (Figure 6.2). In the reverse these studies could also be utilized to show equilibrium dynamics as the CMPs are released from nest-r.

One of the constant goals in developing invasive CMPs is longer-lasting binding to a collagen surface. This goal could be accomplished through incorporation of substituted prolines that stabilize hetero-triple helix formation. There is, however, a balancing act to play between stability of a hetero-triple helix and disruption of a homo-triple helix.

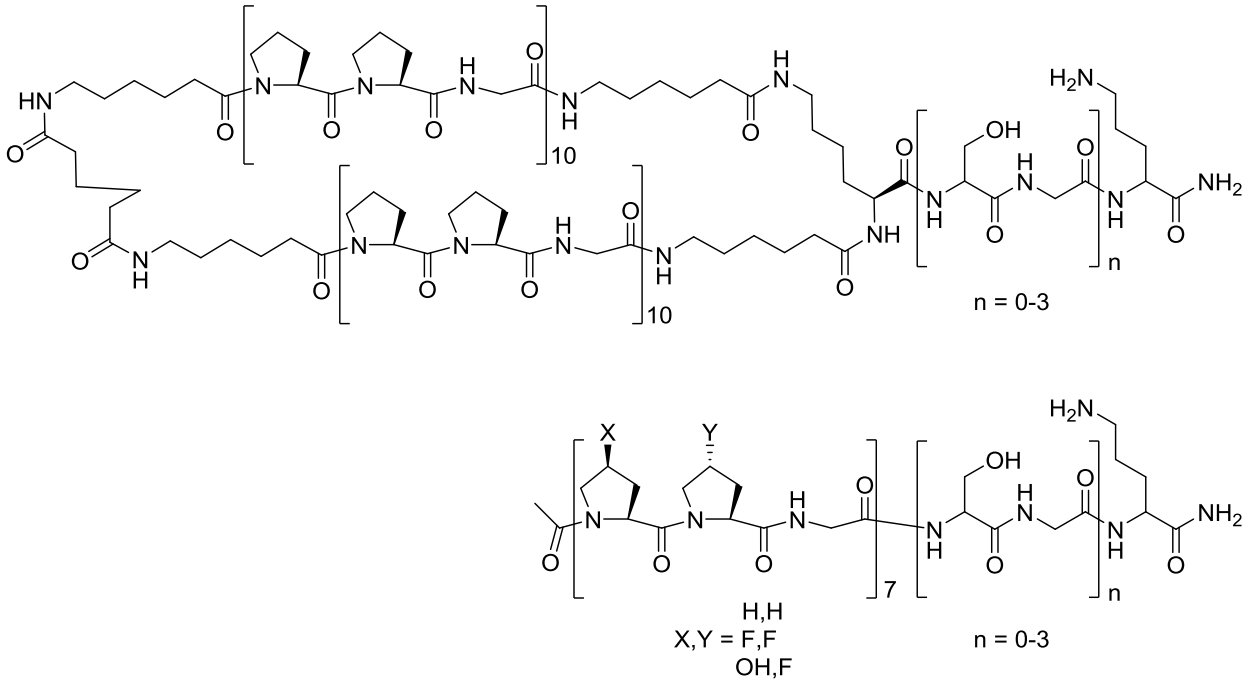


Figure 6.1. Structure of nest-r and CMP for FRET Studies. Each has a lysine side chain available for modification with a fluorophore.

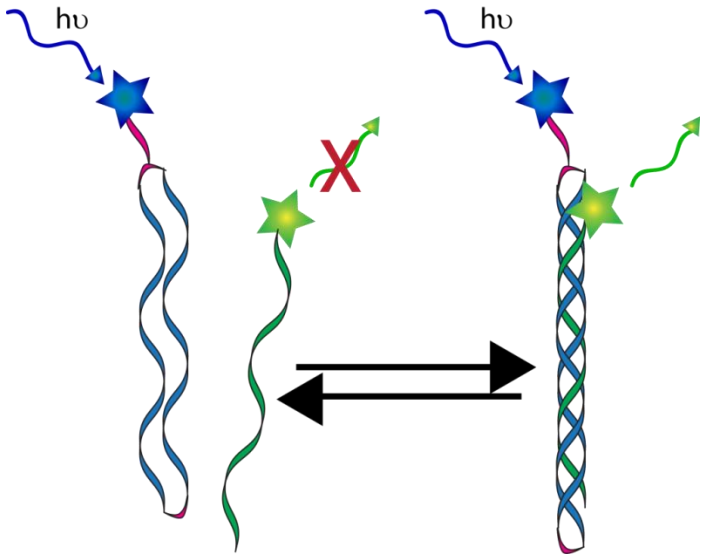


Figure 6.2. FRET occurring upon hetero-triple helix formation.

Nest-r primarily exists in solution as a monomer. The addition of CMPs induces hetero-triple helix formation. It could be reasoned that nest-r could also hybridize with a collagen surface, creating a hetero-triple helix and binding in a similar fashion as invasive CMPs. Because the linkers on the N-terminus and C-terminus restrain nest-r from forming complexes with itself, there is no longer the same concern of balancing hetero-triple helix formation and disruption of self-complexation through homo-triple helix formation. Thus, optimized factors for triple helix stability can be incorporated into the collagen sequence of nest-r. As nest-r can be modified with the addition of a C-terminal lysine residue, the nest can be chemically ligated to dyes, growth factors, or other small molecules for delivery to a collagen surface. Nest-r presents itself as a new avenue to create hyper-strong binding to a collagen surface, allowing for long-lasting adherence.

6.4 Targeting interactions for bacterial collagen

With the rise of antibiotic resistance, there is a need to develop new ways to combat infection. Discussed in Chapter 4, *Streptococcus pyogenes* displays collagen-like proteins Scl1 and Scl2 on its cell surface. These proteins are known adhesins and play an important role in virulence. We showed that these collagen-like proteins are able to bind and interact with mammalian collagen. Could disruption of this interaction diminish virulence of the bacteria?

A collagen duplex like the nests described in Chapter 5 could be used as a way to capture the collagen-like proteins on the bacterial surface. The nest could act as a "glove" in which the Scl1 and Scl2 can fit antagonizing their interaction with the host (Figure 6.3). As the nests are designed now, they are capable of binding mammalian collagen as well. The challenge then is to create a selective interaction that would only bind the bacterium. The collagen sequences of Sc1 and Scl2 are rich in charged amino acids. Delsuc and coworkers were able to create selective collagen duplexes through the utilization of charge–charge interactions for use in protein

recognition.[266] Accordingly, the collagen sequence of the nests could be made to be complementary to the charged residues of Scl1 and Scl2. The nests would then be deterred by charge repulsion from folding into a triple helix unless bound to collagen-like proteins of the bacterium (Figure 6.4). Once a selective system is established for the collagen-like proteins, virulence could be tested with treated bacteria. Virulence could be measured by studies of binding to a collagen surface and to human cells.

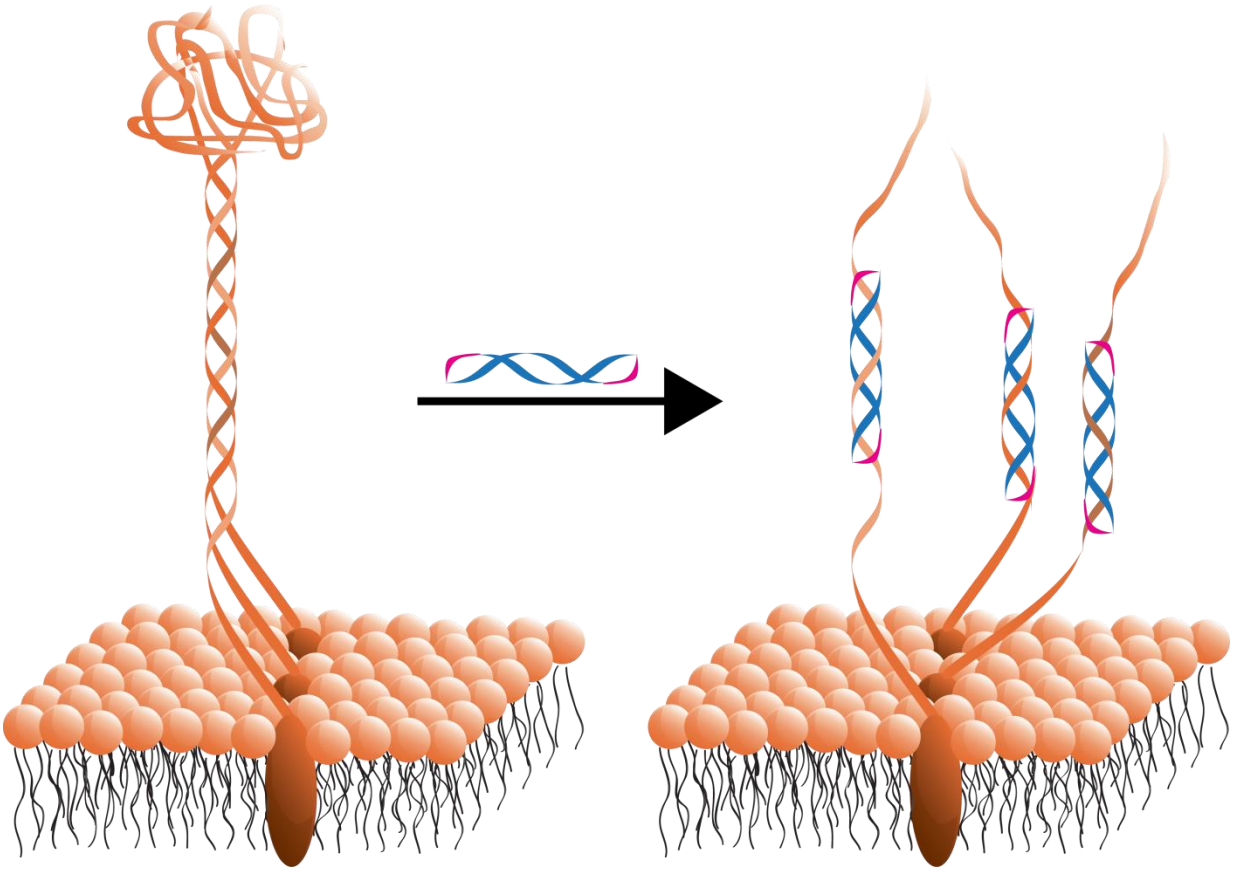


Figure 6.3. Proposed binding of Scl1 and Scl2 to disrupt virulence.

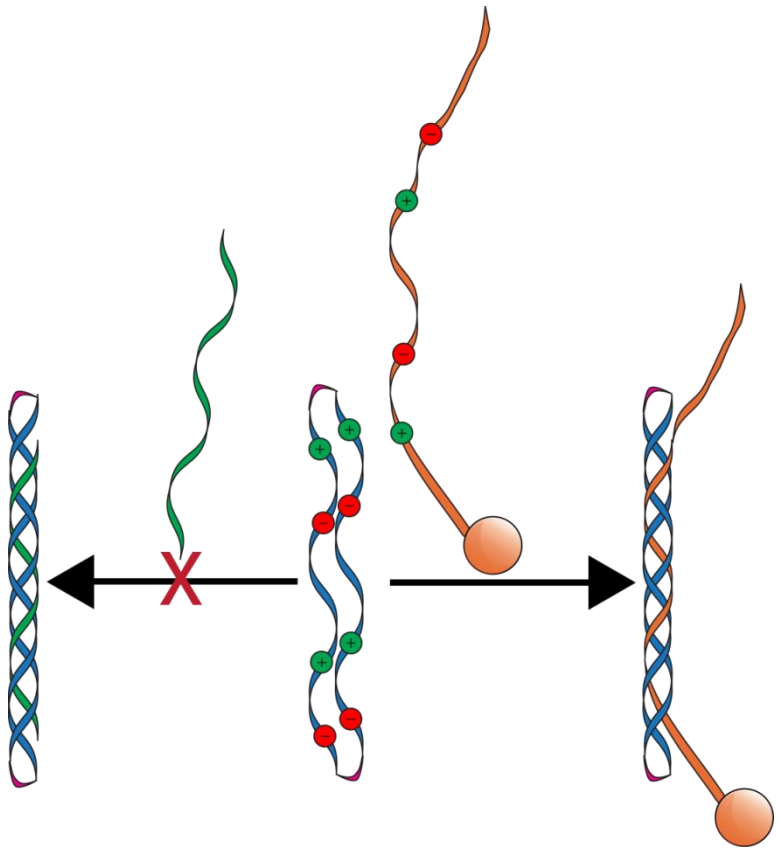


Figure 6.4. Selective formation of hetero-triple helix with Scl1 and Scl2 with a nest through charge–charge interactions.

Appendix A

LL37 conjugated to a CMP for antimicrobial attachment to a wound bed

Contributions

Professor Ronald T. Raines proposed creation of CMP–antimicrobial conjugate. Aubrey J. Ellison synthesized and characterized CMP-LL37. Professor Jonathan McAnulty and Dr. Patricia Kierski helped plan and perform mouse experiments. Aubrey J. Ellison and Ronald T. Raines planned experiments and analyzed data. Aubrey J. Ellison performed all other experiments and drafted the appendix and figures.

7.1 Introduction

Skin acts as a protective barrier to the outside world and any break must be quickly and efficiently repaired. Successful wound healing involves three stages: inflammatory response, tissue repair, and reepithelialization.[275,276] Wound healing is a complex process that requires collaboration of many working parts. If one of the cells, proteins, or signaling molecules fails to do its intended function at the needed time, the wound healing process can be severely stunted and lead to chronic wounds.[275] Slow healing wounds can result from prolonged bed rest, diabetes, vascular problems, and burns.[277] Chronic wounds are highly susceptible to infection that can further slow the healing process.[278,279]

The current approaches for treatment of pathological wounds work to provide protection from invading microbes.[277,280] Treatments to prevent infection employ physical barriers such as bandages and the introduction of antimicrobials either orally or topically. Because chronic wounds are slow healing, oral antimicrobials would need to be taken for the duration of the healing process. This has negative consequences for the patient's natural biome, leading to possible gastral intestinal problems as well as spreading of antimicrobial resistance.[277,281] Topical application localizes antimicrobials to the site of possible infection limiting the spread of antimicrobial resistance.[281,282] However, topical treatments need to be reapplied often.[277] The frequent application risks damaging tissue, prolonging the healing time. One route toward providing better protection from infection would be a long lasting topical application that lessens the risk of tissue damage and is located to the wounded area.[283]

Collagen is a fundamental component of connective tissue. As such, the deposition of new collagen a key component in the wound healing process through.[275,276] Collagen fibers are comprised of bundles of triple helices made of three parallel, left handed polyproline type II

strands.[11,284-286] During wound formation, the collagen triple helix becomes broken, denatured, and frayed. Work previously done in the Raines group has been directed toward the creation of collagen mimetic peptide (CMP) strands that are preorganized to invade and bind with the host collagen tissue.[156,157] The CMP strands are theorized to invade by forming a triple helix with the denatured collagen in the wound. To accelerate wound healing, CMP strands tethered to growth factors were introduced to wounds in mice to bind to the frayed and newly forming collagen.[157] This same invasive strand strategy can be utilized to immobilize and localize an antimicrobial to the endogenous tissue of a wound. Adhering an antimicrobial to a wound would work to stem the spread of antimicrobial resistance and prolong the potency against bacteria and other microbes while leaving the healing process uninterrupted.

7.2 Results and discussion

7.2.1 LL37 conjugated to an invasive CMP

The mechanism of antimicrobial action is crucial when selecting an antimicrobial to tether to the CMP. Because the CMP strand will tether the antimicrobial to a wound surface, the antimicrobial needs to be active without endocytosis. Therefore, antimicrobial activity should be directed toward the membrane or surface proteins. Additionally, the antimicrobial will need to have a synthetic handle which will not disrupt the mode of activity when covalently attached to a CMP.

LL37 is a 37 amino acid long human polypeptide. Endogenous to humans and produced by keratinocytes, LL37 is from the cathelicidin family of antimicrobial peptides (AMPs).[161,287,288] Cathelicidins are positively charged peptides that are attracted to the negatively charged cell

surface.[287] At the cell surface, the AMP invades the bacterial lipid membrane and forms pores, bursting the cell.[289]

In studies to parse out the amino acids necessary for antimicrobial activity in LL37, it was found that the first 12 amino acids are not required for antimicrobial activity.[290,291] These residues could be synthetically modified for attachment to a CMP. LL37 has been tethered before, in order to create an antimicrobial surface. A modified titanium surface had LL37 bound to it using a maleimide linker and a N-terminal added cysteine.[292] The surface-bound LL37 retained antimicrobial activity. Utilizing similar chemistry, I attached a CMP strand to LL37 with a maleimide linker. A Lys residue followed by a tri (Ser-Gly) spacer was added to the N-terminus. $(Ser\text{-}Gly)_3$ enhances solubility and the Lys provides a site for chemical ligation. An NHS ester-maleimide linker was added to the CMP strand and tethered via amide bond formation. N-terminal Cys-LL37 was chemically linked via Michael addition (Figure 7.1).

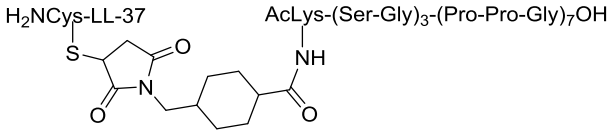


Figure 7.1. Sequence of CMP-LL37

7.2.2 Antimicrobial activity in solution

To test that LL37 maintained antimicrobial activity after attachment, I looked to evaluate any change in minimal inhibitory concentration (MIC). Cultures of *E. coli* (RP437) and *B. subtilis* (Ol 1085) were grown in a 96 well plate with varying concentrations of LL37, CMP-

LL37, and ampicillin, acting as a positive control. Optical density of the cultures showed the minimum concentration necessary to inhibit growth. LL37 had MICs of 10 μM and 2.5 μM against *E. coli* and *B.subtilis* respectively (Table 7.1). These values are comparable with literature values of 5-25 μM or up to 40 μM for *E. coli* strains D21 and K12, respectively and 0.59 μM for *B.subtilis*.[288,291] CMP-LL37 also had comparable MICs to the untethered antimicrobial at 5.0 μM and 1.56 μM for *E. coli* and *B.subtilis*, respectively.

Table 7.1. MICs of LL37, CMP-LL37, and CMP for *E. coli* and *B. subtilis*.

	Antimicrobial	**Experimental**	**LL37 (Literature)**[288,291]
E. coli (RP437)	LL37	10 μM	5-25 μM (D21) 40 μM (K12) 32 μM (atcc 25922)
	CMP-LL37	5 μM	
	CMP	No inhibition	
B. subtilis (Ol 1085)	LL37	2.5 μM	0.59 μM
	CMP-LL37	1.56 μM	
	CMP	No inhibition	

The retention of antimicrobial activity in solution of the tethered CMP-LL37 shows that the CMP and the tether do not interfere with the activity of LL37. The CMP strand was also evaluated for antimicrobial activity. There was no difference in inhibition of growth between the CMP strand and the negative control of phosphate buffered saline (PBS).

7.2.3 Hemolysis assay

The mode of activity for LL37 and other cathelicidins is to intercalate into lipid membranes. As such, cathelicidins can interact not just with bacterial cells, but also with human

cells. Human cells, however, have reduced susceptibility to the toxicity of LL37 compared to bacteria.[293] CMP-LL37 was evaluated for toxicity towards human blood cells by hemolysis (Figure 7.2). As expected, CMP-LL37 showed a similar toxicity curve to LL37 alone. Based on the data, the IC_{50} values for CMP-LL37 and LL37 were found to be 28.3 μM and 49.9 μM, respectively. This cytotoxicity test on human cells provides a therapeutic window for the use of CMP-LL37 as the IC50 values are an order of magnitude larger than the MIC found (2–5 μM).

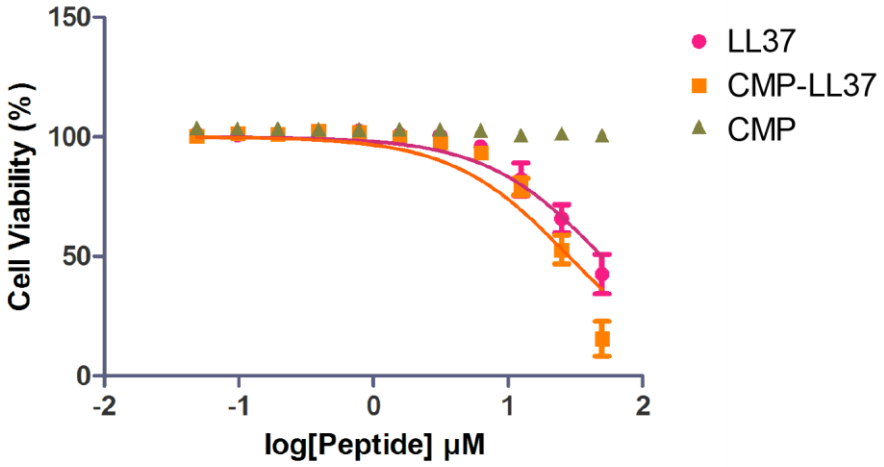


Figure 7.2. Cytotoxicity of LL37, CMP-LL37, and CMP towards human red blood cells determined by hemolysis.

7.2.4 Surface antimicrobial activity

While LL37 activity is maintained in solution, the antimicrobial activity on a collagen surface will better represent how the conjugate will behave in a wound. Collagen-coated wells

can act as a mimic to host tissue collagen. Collagen-coated wells were treated with LL37, CMP-LL37, and CMP. After incubation, the wells were washed of the solutions and *E. coli* cultures were inoculated and added to the wells. The OD of the bacteria in the wells was monitored for 9 h to obtain growth curves (Figure 7.3). The LL37 and CMP treated wells showed no inhibition in the growth curve. This lack of inhibition is expected as CMPs do not have antimicrobial activity and LL37 would be washed away. CMP-LL37, however, did stunt culture growth. This delayed growth is attributed to the CMP-LL37 conjugate remaining adherent to the collagen well and thus was not washed away.

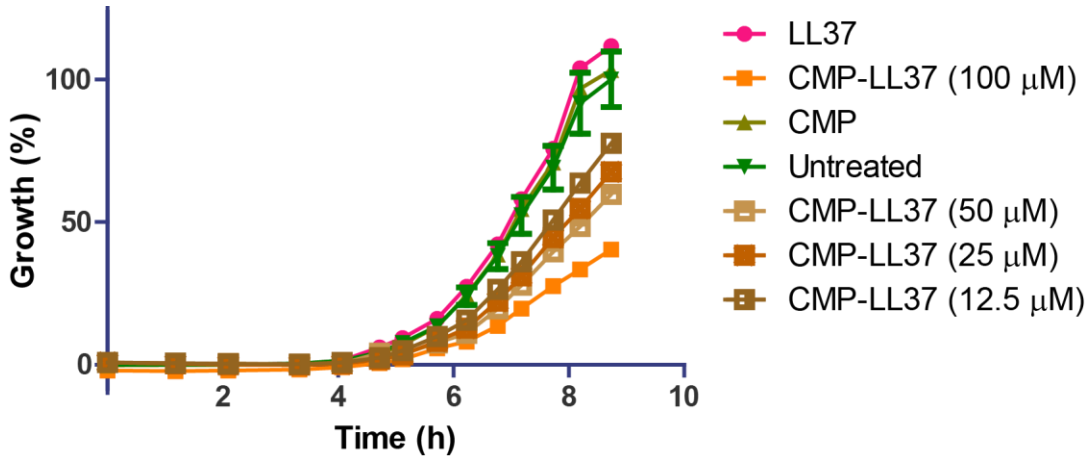


Figure 7.3. Bacterial growth curves of *E. coli* in collagen-coated wells treated with LL37, CMP-LL37, and CMP.

By removing the cultures and replacing with fresh media, a second growth curve was obtained (Figure 7.4). The bacterial growth was once again stunted. The slower growth showed that there was CMP-LL37 that was still adhered to the collagen-coated well. With this second

growth curve, however, the delay in growth is less prevalent then the first culture grown. The difference between the first and second growth curves could mean that CMP-LL37 is in equilibrium between being bound to the surface and being in solution. The antimicrobial activity that is observed in the stunted *E. coli* growth, could then be attributed to solution activity rather than interaction between the bacteria and the CMP-LL37 bound to the walls of the collagen-coated wells.

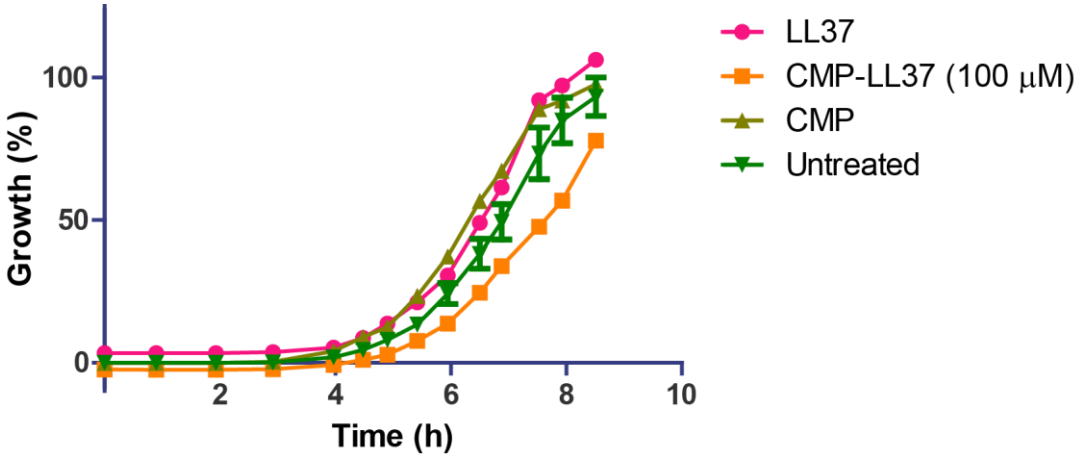


Figure 7.4. Bacterial growth curves of the second inoculation of *E. coli* in collagen-coated wells treated with LL37, CMP-LL37, and CMP.

In an effort to develop a different model to test the surface activity, I turned to evaluating cultures grown on an agar plate. These surface grown cultures would no longer be suspended in solution and thus any potential antimicrobial activity could then be attributed to surface activity. Biobrane was used to maintain a collagen surface to which the invasive CMP can bind. Biobrane is a biocomposite wound dressing that contains porcine dermal collagen. To make sure that biobrane would be a viable collagen surface, a CMP fluorophore conjugate was tested. A D-CMP

fluorescent conjugate was used as control to ensure that any binding is due to triple helix formation with the collagen surface and not some other attractive force. Only the L-CMP is able to form a triple helix with natural collagen strands. By imaging the fluorescence of the material remaining on Biobrane after a series of washes, I found that the L-CMP did indeed bind while the D-CMP was washed away (Figure 7.5).

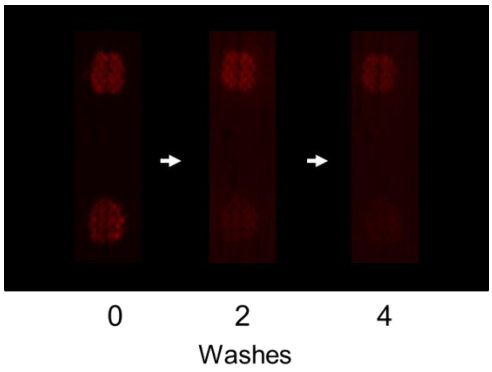


Figure 7.5. Adherence of L-CMP-Red (top) compared to D-CMP-Red (bottom) to Biobrane over a series of washes.

Biobrane was spotted with CMP-LL37 and positive control, kanamycin. L-CMP-Red was co-spotted at each location for visual conformation as to the origin of application. The treated Biobrane was then applied to an agar plate that had been inoculated with GFP expressing *E. coli*. After a bacterial lawn had formed, the plate was evaluated for areas of growth inhibition that coincided with the location of L-CMP-Red (Figure 7.6). Kanamycin, as the positive control, showed clear rings of inhibited growth around the spot of origin. CMP-LL37, however, did not have any discernable inhibition to the GFP expressing *E. coli*.

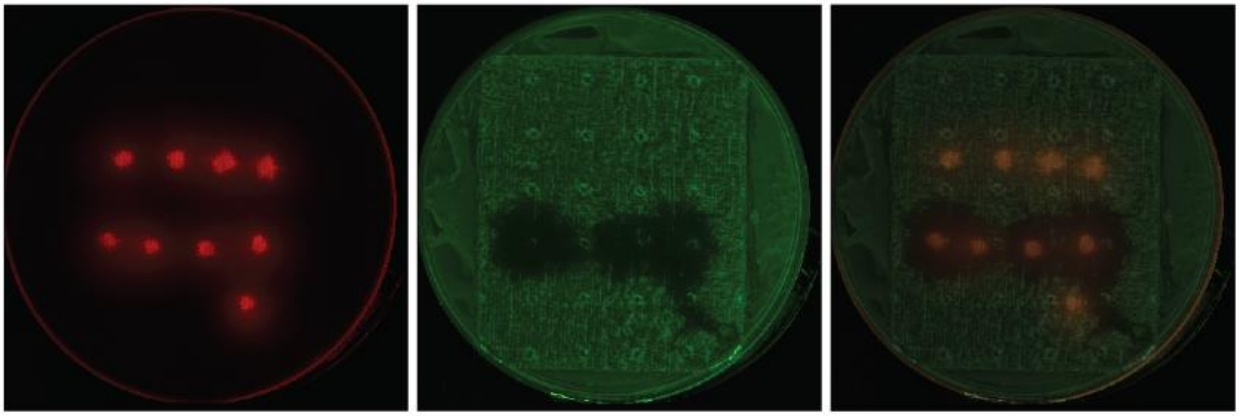


Figure 7.6. Surface antimicrobial activity against GFP expressing *E. coli* of 25 μM CMP-LL37 (top) compared to a positive control (bottom, 5 mg/mL kanamycin). Center of origin is indicated by L-CMP-Red co-spot.

7.2.5 Mouse pilot study

As a wound contains some moisture, the environment could be somewhere between the two *in vitro* models of bacteria grown in collagen-coated wells and grown on an agar plate. CMP-LL37 could again form an equilibrium between the surface and solution. To test CMP-LL37 in the context of a wound a mouse model was used. A wound was created using a 6 mm punch biopsy on the flank of the mice. The mouse wound was then treated with CMP-LL37, washed, and inoculated with *Pseudomonas aeruginosa* (27853). After 3 days the wounds were harvested and bacteria present in the wounds was determined by colony forming units (CFU). There was no observable difference between bacterial growth in untreated wounds and CMP-LL37 treated wounds (Figure 7.7).

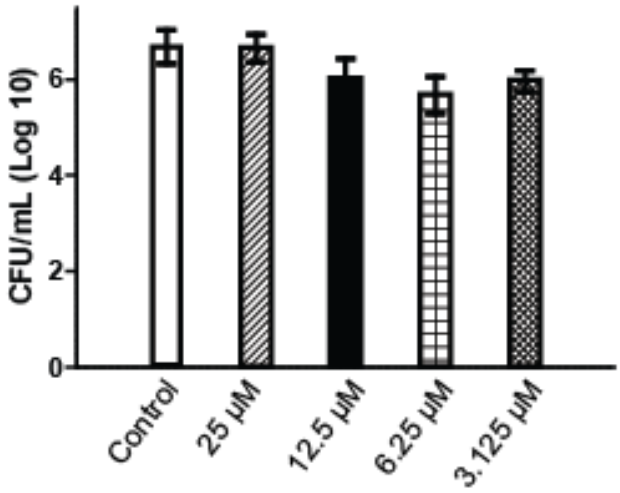


Figure 7.7. *P. aeruginosa* growth (expressed as CFU) after 3 days in mouse wounds treated with CMP-LL37.

There are many possible experimental reasons for the lack of antimicrobial activity. Because of the limitation of how long the mice can be anesthetized, the CMP-LL37 is given only 15 min of incubation before washing the wound. For *in vitro* models, CMP-LL37 was incubated for 1 h prior to washing. The shorter incubation time for the mouse wounds may not be enough for CMP adherence to the host surface.

While the concentrations of the antimicrobial conjugate are well over the MIC for solution activity, another possibility for the lack of antimicrobial activity is that the amount of CMP-LL37 applied does not create an appropriate surface density. Assuming a flat surface and that all the peptide applied at 25 μM attached to the surface, the calculated surface density would be 1.77×10^{-9} mol/cm^2. This surface density is greater than the previously reported surface density of 1.47×10^{-10} mol/cm^2 which proved to create an antimicrobial surface.[292] The actual surface density achieved after CMP-LL37 application may not be close to the calculated value.

This deviation could arise from washing of the wound (*vide supra*). Additionally, the assumption of a flat wound surface could lead to a misleading calculated surface density. The wound surface could be far from flat, with many wrinkles and crevices making the surface area much greater than assumed.

7.3 Conclusions and future directions

With LL37's mode of activity impacting the cell membrane and a portion of the peptide sequence available for being covalently linked to CMP strands, LL37 was an ideal candidate for invasive CMP delivery. While LL37 conjugated to a CMP retained antimicrobial activity in solution, antimicrobial activity on an *in vitro* and *in vivo* collagen surface was not observed.

Further surface tests for antimicrobial activity from CMP-LL37 could yet provide evidence that this conjugate could be used successfully for wound treatment. Experimental conditions similar to the mouse model could be optimized on an *ex vivo* skin model.

In addition to experimental modifications, changes to the CMP conjugate could improve detection of surface antimicrobial activity. Activity of surface-bound cathelicidins is subject to the length of the linker that holds it to a surface. Modulation of the linker between CMP and LL37 may show better antimicrobial activity. Other antimicrobials could be explored for linkage to a CMP. Vancomycin, for example, has been used to create antimicrobial surfaces.[294]

7.4 Materials and methods

7.4.1 Instrumentation

Solid-phase peptide synthesis was performed at the University of Wisconsin–Madison Biotechnology Center with a Prelude peptide synthesizer from Protein Technologies (Tucson, AZ). Synthetic peptide was purified by HPLC with Shimadzu Prominence instrument from

Shimadzu (Kyoto, Japan) equipped with a VarioPrep 250/21 C18 column from Macherey–Nagel (Düren, Germany). Molecular mass was determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid matrix with a Voyager DE-Pro instrument from Thermo Fischer Scientific (Waltham, MA) at the Biophysics Instrumentation Facility at the University of Wisconsin–Madison. Colorimetric reads were done with a Tecan Infinite M1000 plate reader (Männedorf, Switzerland). GE Healthcare LAS 4010 Imaging System and GE Typhoon FLA 9000 Gel Imaging Scanner (Marlborough, MA) were used for image capture.

7.4.2 General

Amino acid derivatives and HOBt were from Chem-Impex International (Wood Dale, IL). Fmoc-Gly-2-chlorotrityl resin was from EMD Millipore (La Jolla, CA). Rhodamine Red™-X, succinimidyl ester, 5-isomer was obtained from Thermo Fischer Scientific (Waltham, MA). Cys-LL37 made and obtained from Biomatik (Wilmington, Delaware). *N*,*N*-dimethylformamide (DMF) was dried with a Glass Contour system from Pure Process Technology (Nashua, NH). In addition, DMF was passed through an associated isocyanate "scrubbing" column to remove any amines. Water was purified by an Arium Pro from Sartorius (Goettingen, Germany). All other chemical reagents were from Sigma–Aldrich (St. Louis, MO) and were used without further purification.

7.4.3 Peptide synthesis

Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Using the Fmoc-D-Pro-D-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-D-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19

mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each residue was double-coupled between Fmoc-deprotections. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): $[M + H]^+$ calcd, 2380.6; found, 2380.0. A 0.05-mmol scale synthesis afforded 18.2 mg (15%) of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ after purification.

Ac-Lys(Red)-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$: Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$ (3.4 mg, 1.43 μmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (199.9 μL, 1.30 μmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (200 μL, 1.14 mmol) was added dropwise. The reaction mixture was allowed to stir for 8 h, and then diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 0.5 mg of peptide (11%). MALDI $[M + Na]^+$: calcd, 3056.42; found, 3056.36. Purity was assessed as >95% by UPLC.

Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$: Using the Fmoc-Pro-Pro-Gly-OH tripeptide synthesized in solution without chromatography[232] and an Fmoc-Pro-OH monomer, Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ was synthesized by two additions of monomer followed by six segment condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The tripeptide or amino acid monomer (4 equiv) was converted to an active ester using HATU and NMM. Each

residue was double-coupled between Fmoc-deprotections. Peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H_2O/TIPSH (5 mL), precipitated from diethyl ether at –80 °C, and isolated by centrifugation. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 50 min (A: H_2O containing 0.1% v/v TFA; B: acetonitrile containing 0.1% v/v TFA). MALDI (*m*/*z*): MALDI $[M + Na]^+$: calcd, 2402.61; found, 2403.15. A 0.05-mmol scale synthesis afforded 12.0 mg (10%) of Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ after purification.

Ac-Lys(Red)-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$: Ac-Lys-(Ser-Gly)$_3$-(Pro-Pro-Gly)$_7$ (6.4 mg, 2.68 µmol) was dissolved in 2.0 mL of DMSO. Rhodamine Red™-X, succinimidyl ester, 5-isomer (400 µL, 2.60 µmol) was added as a 5.0 mg/mL solution in DMSO. DIEA (2.2 µL, 13.0 µmol) was added dropwise. The reaction mixture was allowed to stir for 8 h, and then was diluted with 7 mL of H_2O, frozen, and lyophilized. The peptide was purified by preparative HPLC using a gradient of 65–95% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 1.0 mg of peptide (12%). MALDI $[M + Na]^+$: calcd, 3056.42; found, 3056.16. Purity was assessed as >95% by UPLC.

Ac-Lys(maleimide)-(Ser-Gly)$_3$(Pro-Pro-Gly)$_7$-: Ac-Lys-(Ser-Gly)$_3$(Pro-Pro-Gly)$_7$ (8.4 mg, 3.53 µmol) and 4-(*N*-maleimidomethyl)cyclohexanecarboxylic acid *N*-hydroxysuccinimide ester (3.0 mg, 8.97 µmol) was dissolved in 1 mL DMF. DIEA (200 µL, 1.14 mmol) was added and reaction was stirred for 16 h. Volatiles were removed under reduced pressure. The peptide was purified by preparative HPLC using a gradient of 10–50% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 1.2 mg of peptide (13%). MALDI $[M+Na]^+$: calcd 2621.85, found 2620.13. Analytical HPLC showed peptide as one peak >80% pure. [Buffer A: H_2O with 0.1% v/v TFA; BufferB: AcCN with 0.1% v/v TFA].

Ac-Lys(maleimide-LL37)-(Ser-Gly)$_3$(Pro-Pro-Gly)$_7$: Cys-LL37 (8.0 mg, 1.74 μmol) dissolved in 1 mL PBS buffer. A few crystals of *tris*(2-carboxyethyl)phosphine were added. After 20 min, 100 μL of Ac-Lys-(Ser-Gly)$_3$(Pro-Pro-Gly)$_7$-maleimide (1.2 mg, 0.46 μmol) dissolved in 1:1 H_2O/ACN was added. Reaction agitated for 16 h. The peptide was purified by preparative HPLC using a gradient of 10–70% v/v B over 55 min (A: H_2O containing 0.1% v/v TFA; B: methanol containing 0.1% v/v TFA) to yield 1.6 mg of peptide (48%). MALDI $[M+H]^+$: calcd 7196.31, found 7195.23. Analytical HPLC showed peptide as one peak >95% pure. [Buffer A: H_2O with 0.1% v/v TFA; BufferB: AcCN with 0.1% v/v TFA].

7.4.4 MIC assay

The MICs of LL37, CMP-LL37 and AcK(SG)$_3$(PPG)$_7$ were determined using microbroth dilution method in a 96-well microplate for *E. coli* (RP437) and *B.subtilis* (Ol 1085). The test strains were grown overnight in LB medium. 3 μL of the cultures were used to inoculate 30 mL of LB broth that was then distributed in 100 μL volumes into a 96-well microplate. The wells on the perimeter were filled with H_2O to prevent dehydration of cultures during growth. Antimicrobial solutions were added to a well containing LB medium. Contents of the well were thoroughly mixed and half of the solution was added to the adjacent well. This continued, to provide two fold dilutions ranging from μM to low nM. A positive control of ampicillin and a negative control of PBS buffer were also included. Cells were grown at 37 °C overnight with shaking. Growth was analyzed periodically by optically density at 600 nm. The lowest concentration of antimicrobial at which no turbidity was seen was regarded as the MIC.

7.4.5 Hemolysis Assay

Hemolysis assays were conducted as previously described using human red blood cells. Expired human blood was obtained from the University Wisconsin–Madison hospital. The

human red blood cells were washed with Tris-buffered saline (pH 7.2, 10 mM Tris-HCl, 155 mM NaCl) and centrifuged at 3500 rpm until the supernatant was clear (3 washes). Two-fold serial dilutions of the peptide mixture in Tris buffered saline were added to wells of a sterile 96-well plate with conical bottoms, for a total volume of 50 μL in each well. A 2% (v/v) red blood cells suspension (50 μL in Tris buffer) was added to each well. The plate was incubated at 37 °C for 1 h, and then the cells were pelleted in the plates by centrifugation at 3500 rpm for 5 min. Aliquots of the supernatants (80 μL) were transferred to a fresh plate, and hemoglobin was detected by measuring the OD at 405 nm. The average OD after incubation with TX-100 at 4 μg/mL defines 100% hemolysis; the average OD after incubation in Tris buffer defines 0%. IC_{50} is the concentration of a peptide mixture necessary to cause 50% hemolysis. IC_{50} values are reported as the median value of at least two assays run in duplicate.

7.4.6 Bacteria growth in collagen-coated wells

Collagen in the wells of a 24-well plate were treated with solutions of LL37, CMP-LL37, CMP, or left untreated. Wells were agitated at 37 °C for 1 h. Solutions were then removed, the wells were washed with sterile water, and 2 mL of *E. coli* inoculated LB was added. The growing cultures OD were monitored by a plate reader. After culture growth the medium was removed, fresh LB was added and growth was again monitored by a plate reader.

7.4.7 CMP labeling on Biobrane

Biobrane was spotted with 5 μL of a solution of L-CMP-red and D-CMP-red in water. The spots were allowed to dry and were imaged. The Biobrane was washed with water. After washing, the Biobrane was once again imaged. The same procedure was repeated for each subsequent wash.

7.4.8 Bacterial growth on Biobrane

Biobrane was spotted with 5 μL of 25 μM CMP-LL37 and 5 g/mL kanamycin. The spots were incubated at 37 °C for 1 h, until spots were dry. An LB agar plate was inoculated with 500 μL of GFP expressing *E. coli* (BL21) culture and the Biobrane was then place on top. The plate was incubated at 37 °C until a bacterial lawn formed Following bacterial lawn growth the plate was imaged.

7.4.9 Mouse pilot study

15 BalbC mice (Jackson Labs # 000651) 16 – 20 week, were anesthesized with 2% Isoflurane gas. Mice were shaved on their cranial half of dorsum and back nails clipped, 0.001mg buprenorphine SubQ was administered prior to wounding. All surgical instruments are autoclaved before use and cleaned between mice with ethanol as necessary. Shaved area was prepped with Betasept Scrub (4% chlorhexidine gluconate) and sterile saline (0.9%) 3 times each using sterile CTAs. Sterile O-rings (splint) (11 mm x 1.75 mm silicone O-ring, O-ring Warehouse #0568-013) are glued with CrazyGlue gel on each side of prepped dorsum, 6 interrupted sutures using 5-0 Monosof or 5-0 Ethilon are placed equidistant around O-ring. 2–6 mm wounds are made with 6 mm biopsy punch on prepped dorsum in center of each splint. Wounds on 3 mice were treated with 20 μL of 25 μM, 12.5 μM, 6.25 μM, 3.125 μM or 0 μM CMP-LL37 and allowed to dry for up to 15 min. After drying, wounds were rinsed 4 times with 200 μl of sterile PBS. All mice wounds were inoculated with 10^5 cfu/wound *P. aeruginosa* (27853) (10 μL of 10^7 cfu/mL *P. aeruginosa*). Wounds are covered with tegaderm and mouse weighed. The wounds/splints/coverings are evaluated daily and repairs made as needed.

All mice were be euthanized on day 3. All wounds were completely excised, to include a deep tissue layer, using aseptic technique with a 6-mm biopsy punch and homogenized for 15 min with 1 mL of phosphate-buffered saline. Homogenates were serially diluted and plated on blood agar plates. Plates were incubated overnight at 37 ºC and colony counts were performed.

7.5 Acknowledgements

We are grateful to the Professor Bernard Weisblum who allowed us to work in his lab to perform the hemolysis assay. This work was supported by Grant R01 AR044276 (NIH) and Instrumentation Grant GM044783 (NIH). A. J. Ellison was supported by Chemistry-Biology Interface (CBI) Training Grant T32 GM008505 (NIH).

Appendix B

Spectra and Supporting Information

8.1 Chapter 2

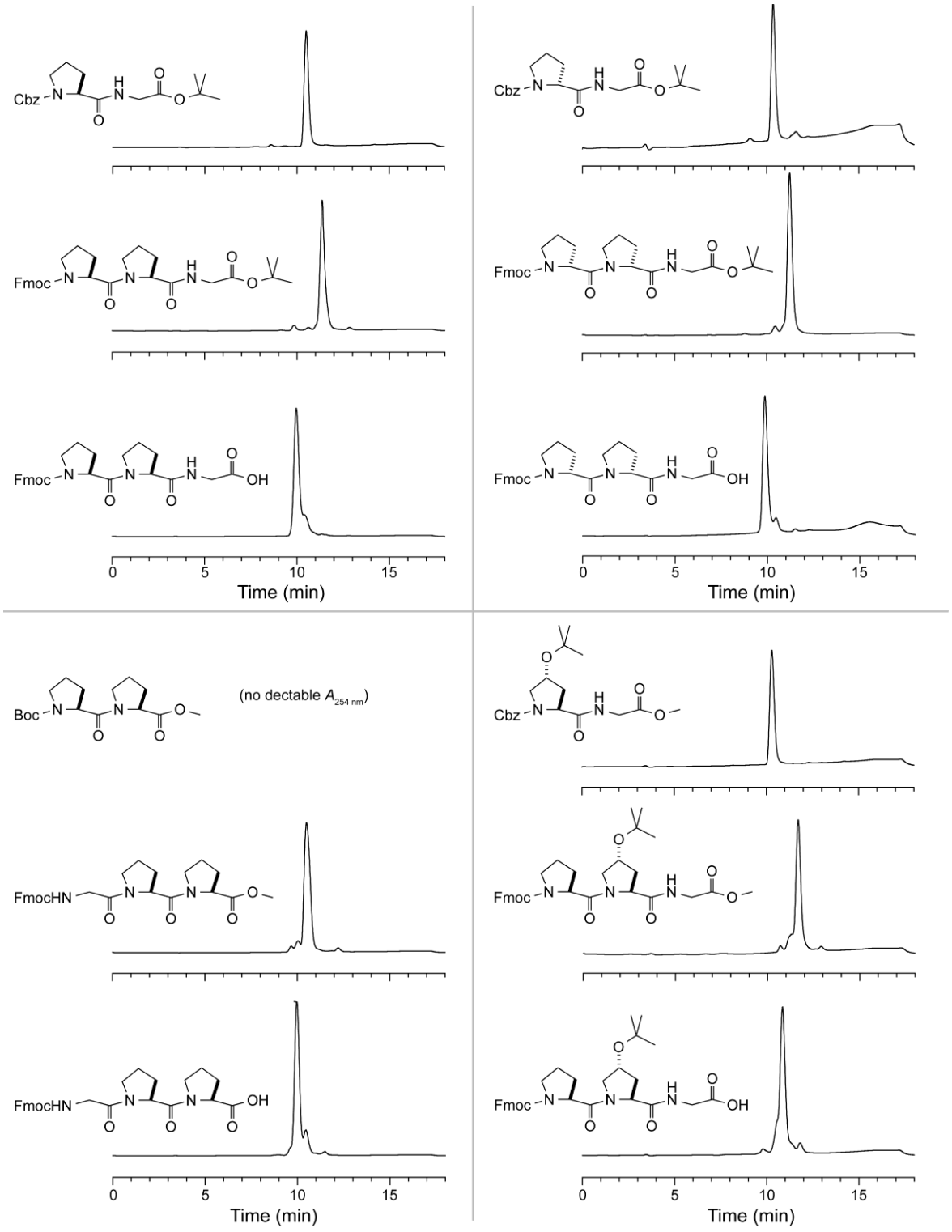


Figure 8.1. LCMS traces after each step in the chromatography-free synthesis of tripeptides. The traces for Fmoc-Pro-Pro-Gly-OH are also shown in Figure 1 of the main text.

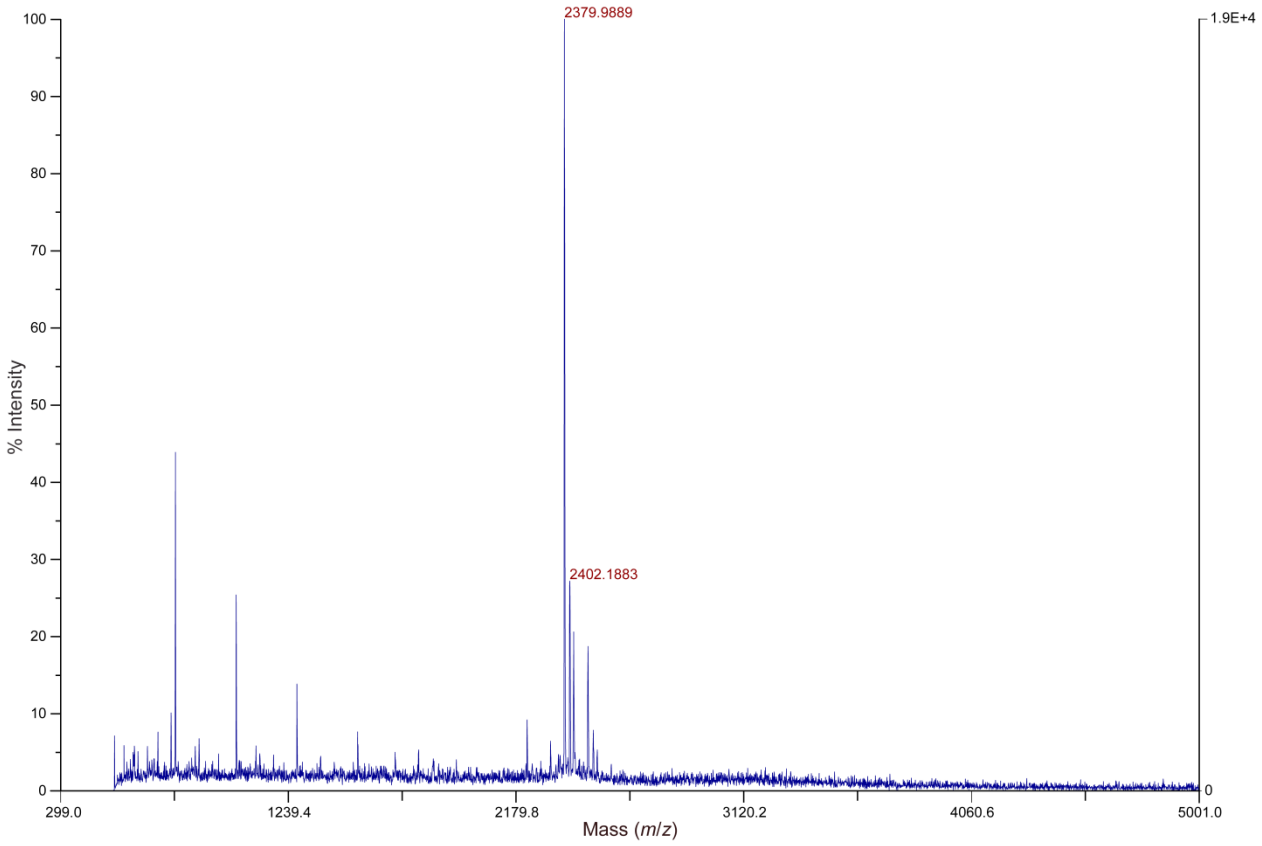
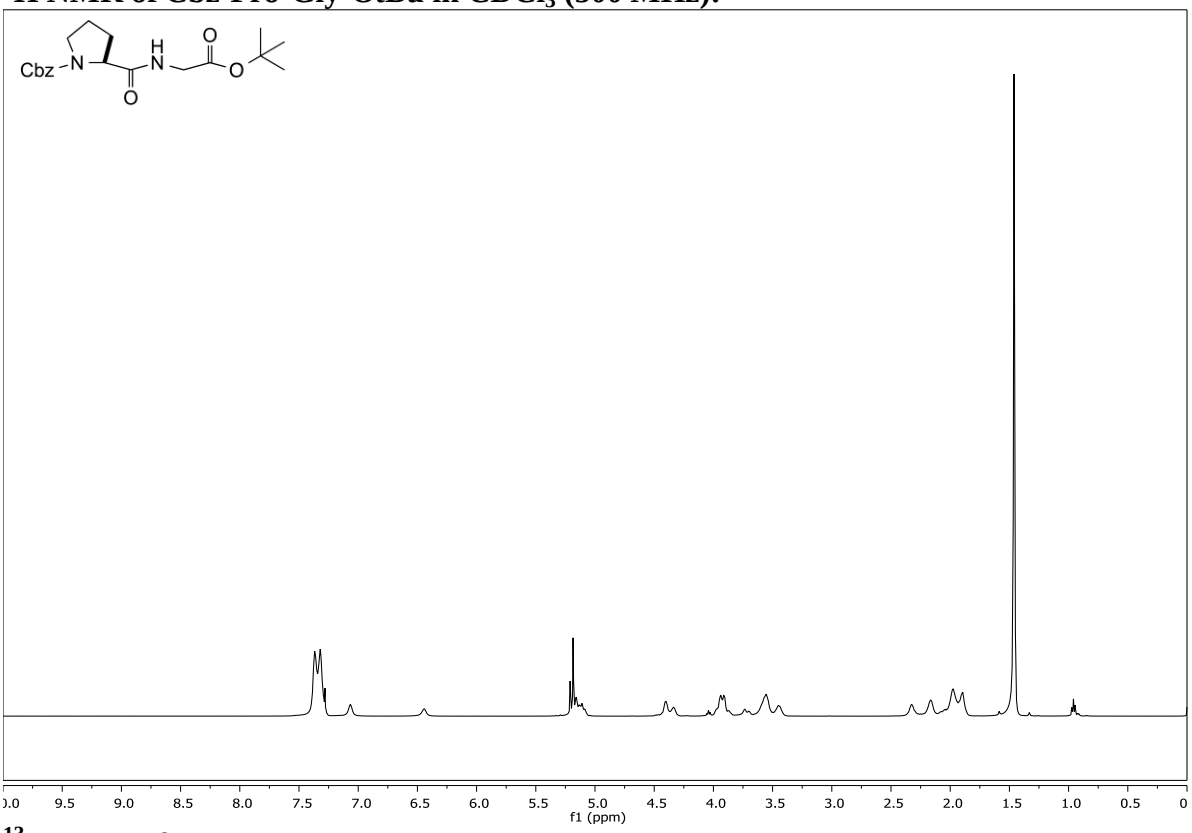
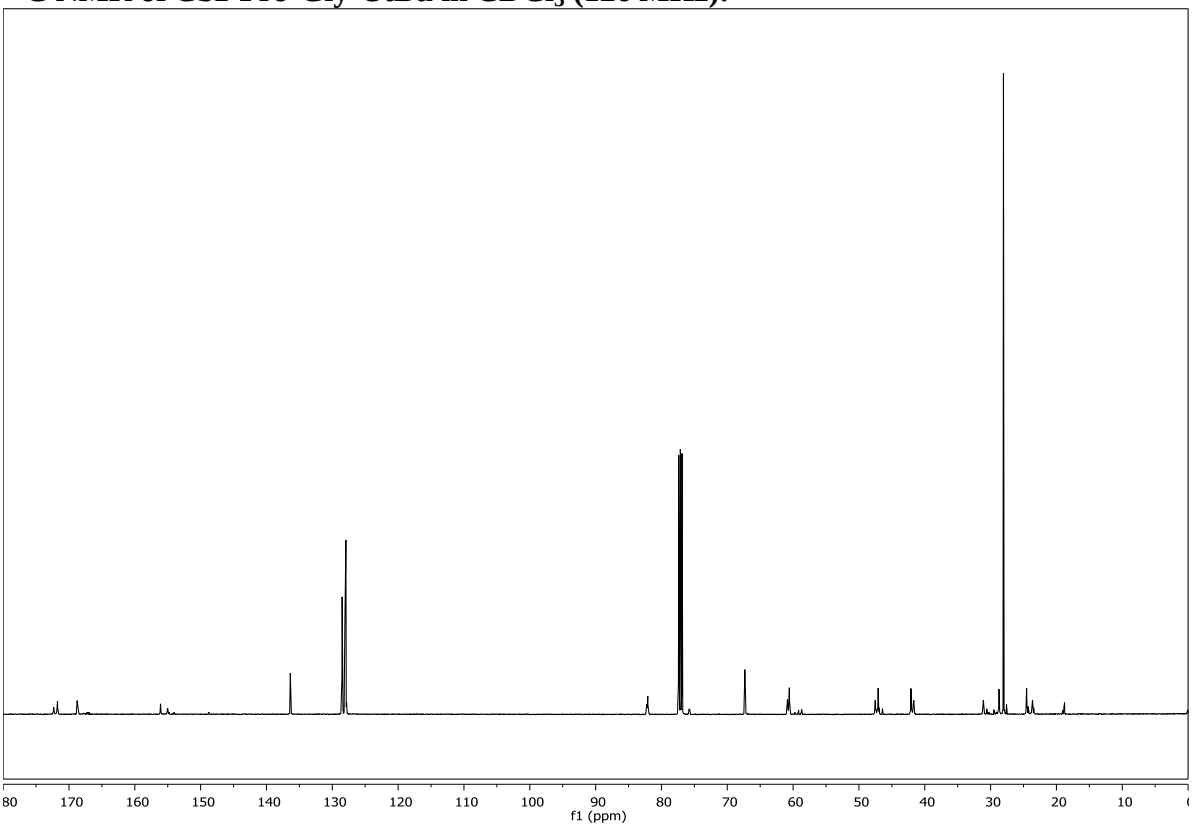


Figure 8.2. MALDI–TOF mass spectrum of Ac-Lys-(Ser-Gly)$_3$-(D-Pro-D-Pro-Gly)$_7$-OH. This 28-mer peptide was synthesized on a solid support by segment condensation of Fmoc-D-Pro-D-Pro-Gly-OH units synthesized in solution without chromatography. (*m*/*z*): $[M + H]^+$ calcd 2380.6, found 2380.0

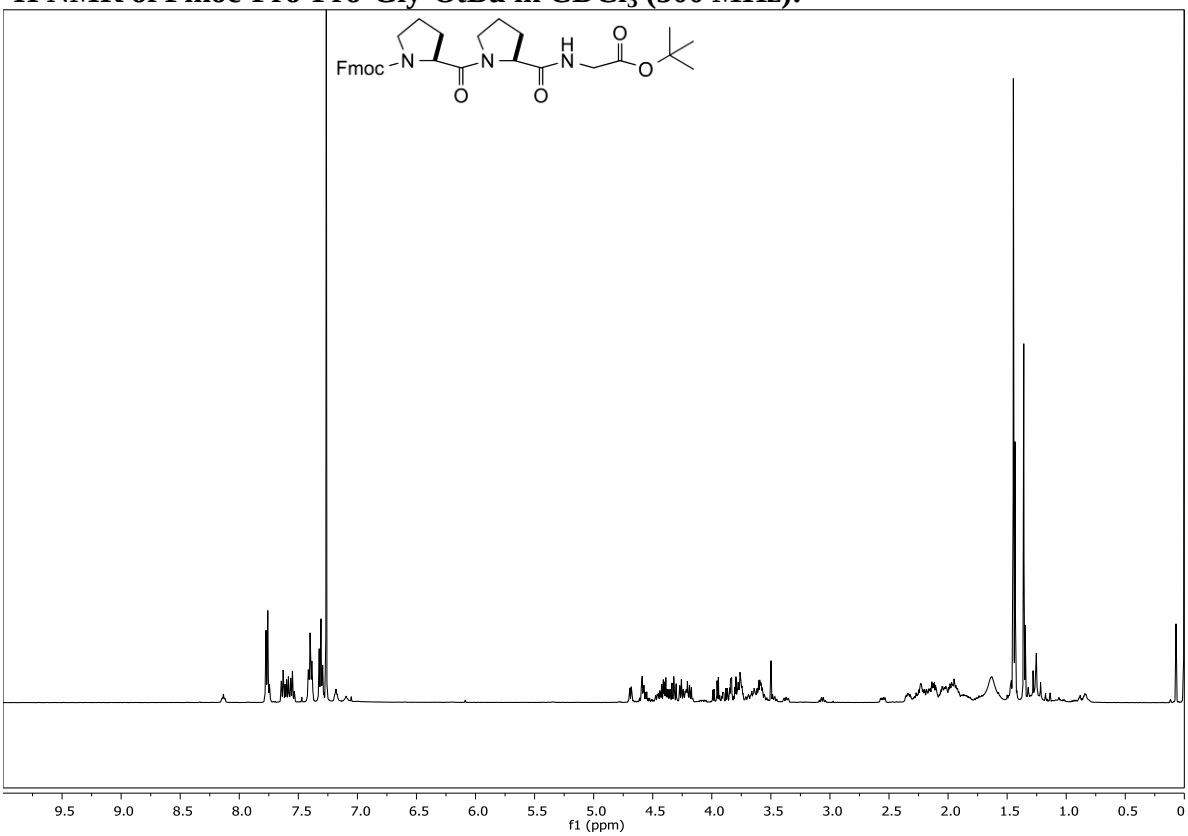
^{1}H NMR of Cbz-Pro-Gly-O*t*Bu in $CDCl_3$ (500 MHz):



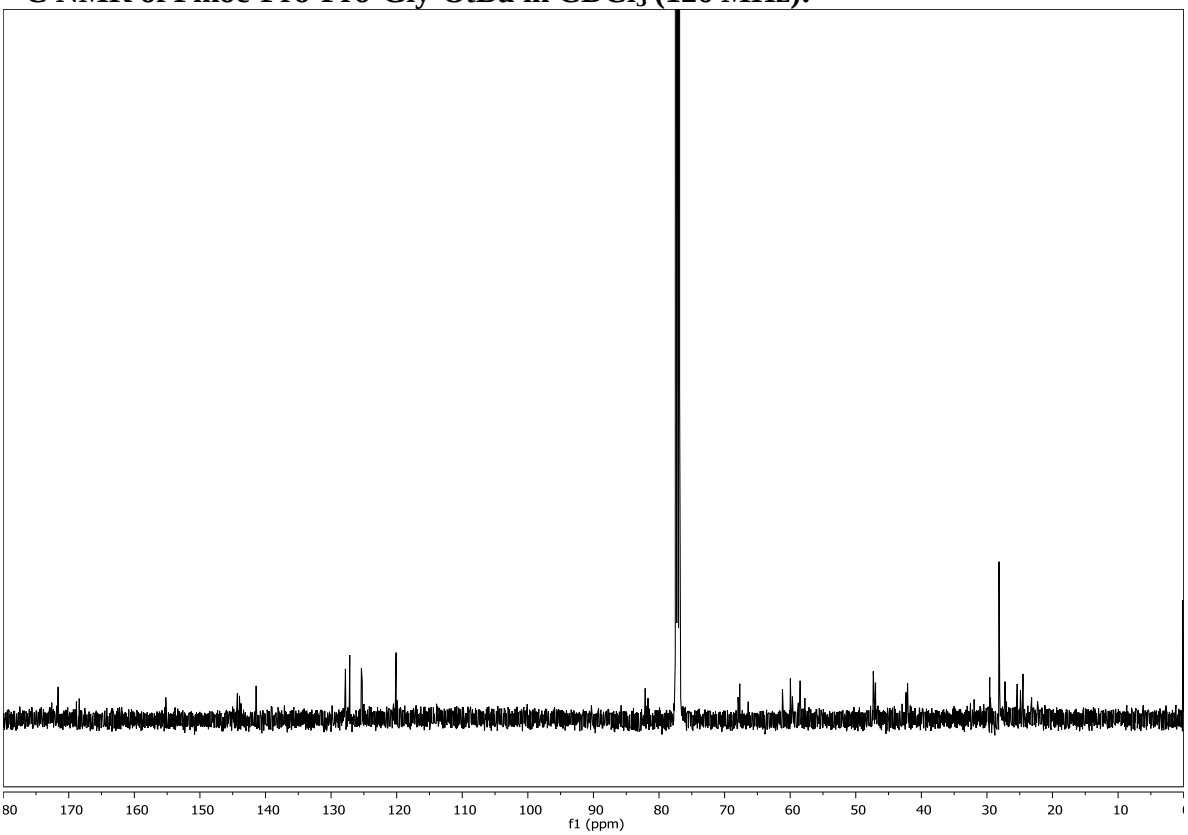
^{13}C NMR of Cbz-Pro-Gly-O*t*Bu in $CDCl_3$ (126 MHz):



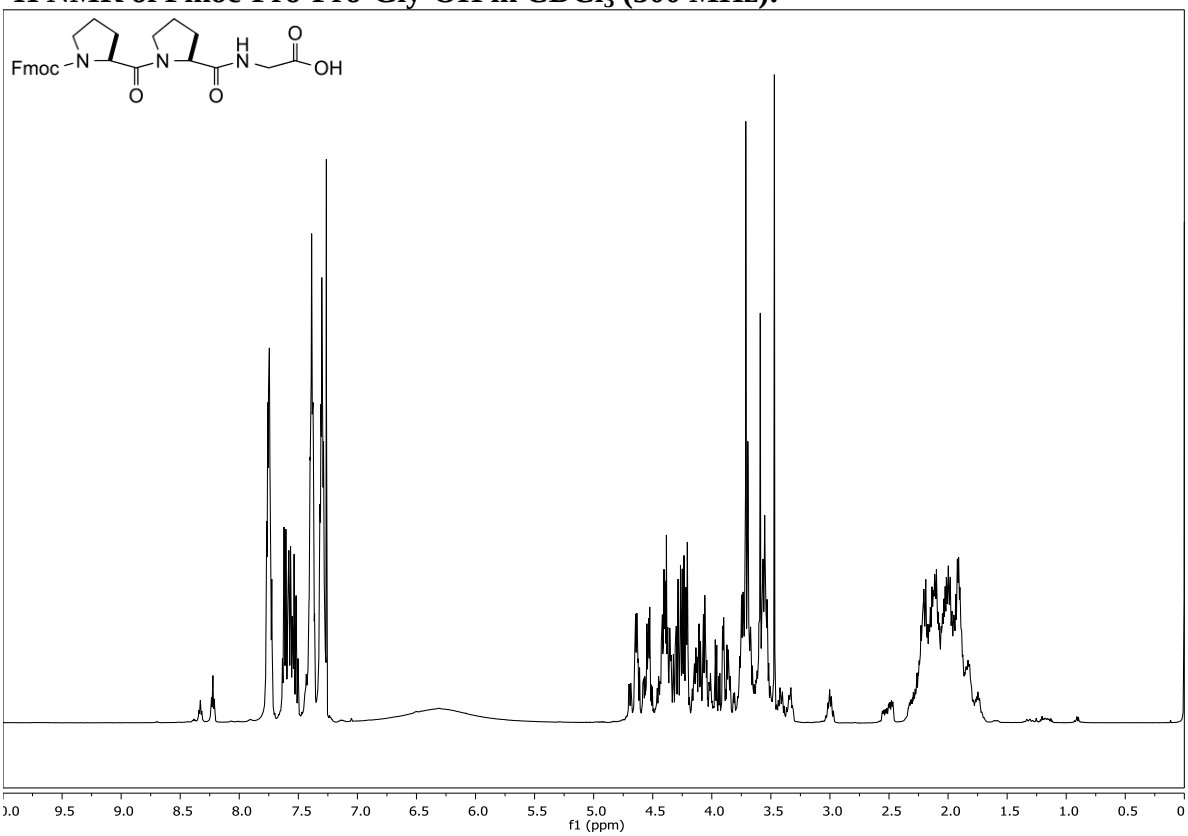
^{1}H NMR of Fmoc-Pro-Pro-Gly-O*t*Bu in $CDCl_3$ (500 MHz):



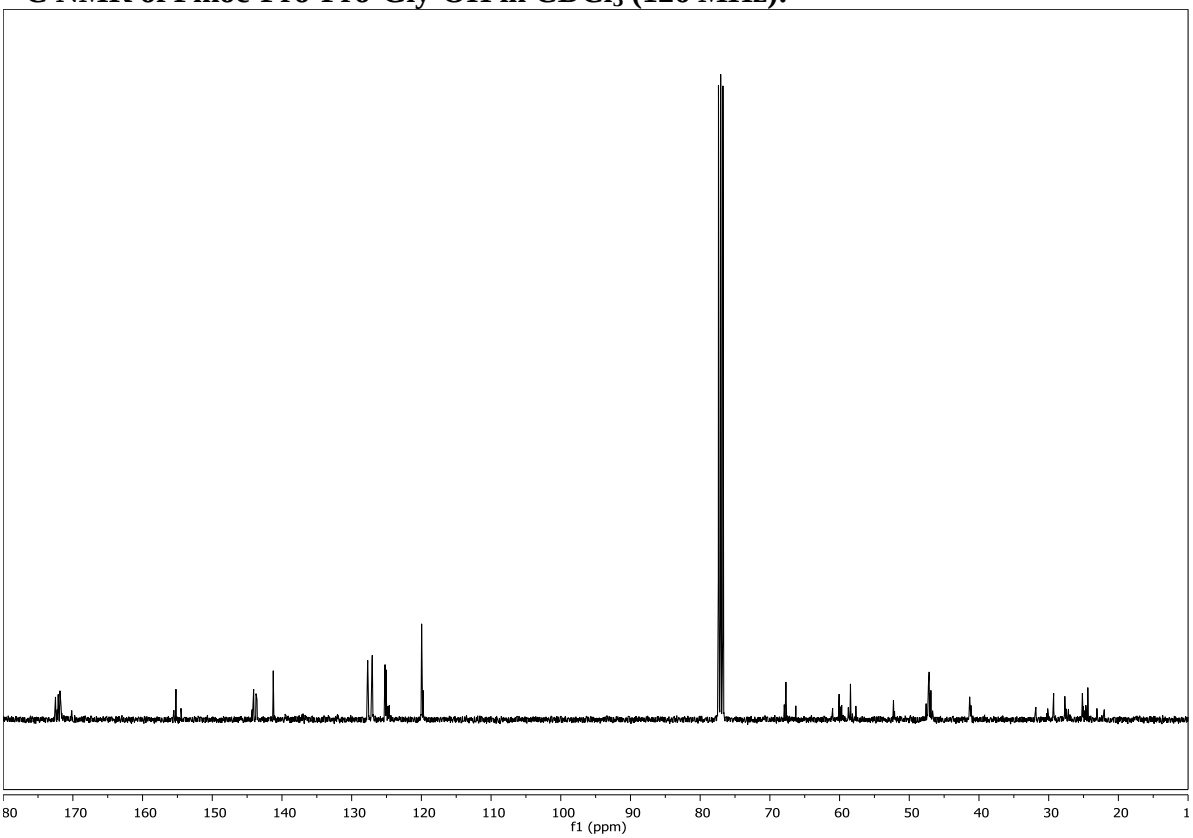
^{13}C NMR of Fmoc-Pro-Pro-Gly-O*t*Bu in $CDCl_3$ (126 MHz):



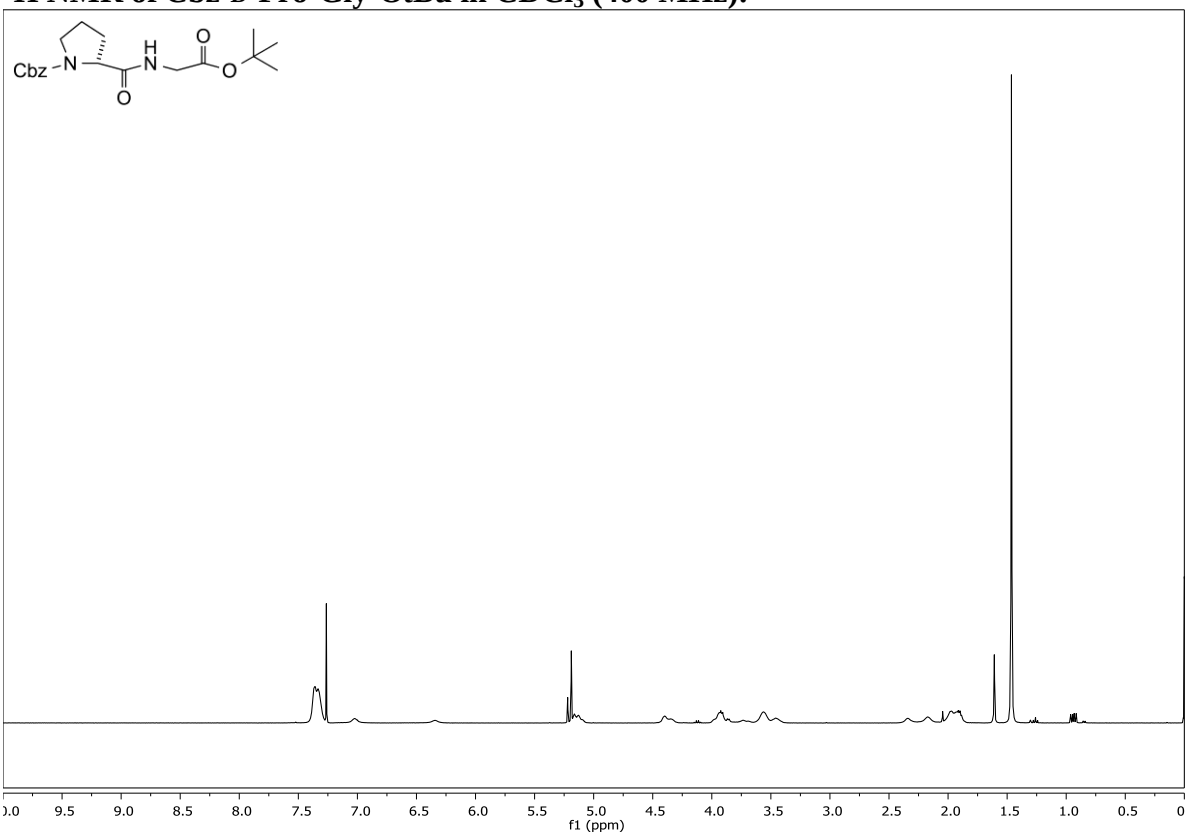
^{1}H NMR of Fmoc-Pro-Pro-Gly-OH in $CDCl_3$ (500 MHz):



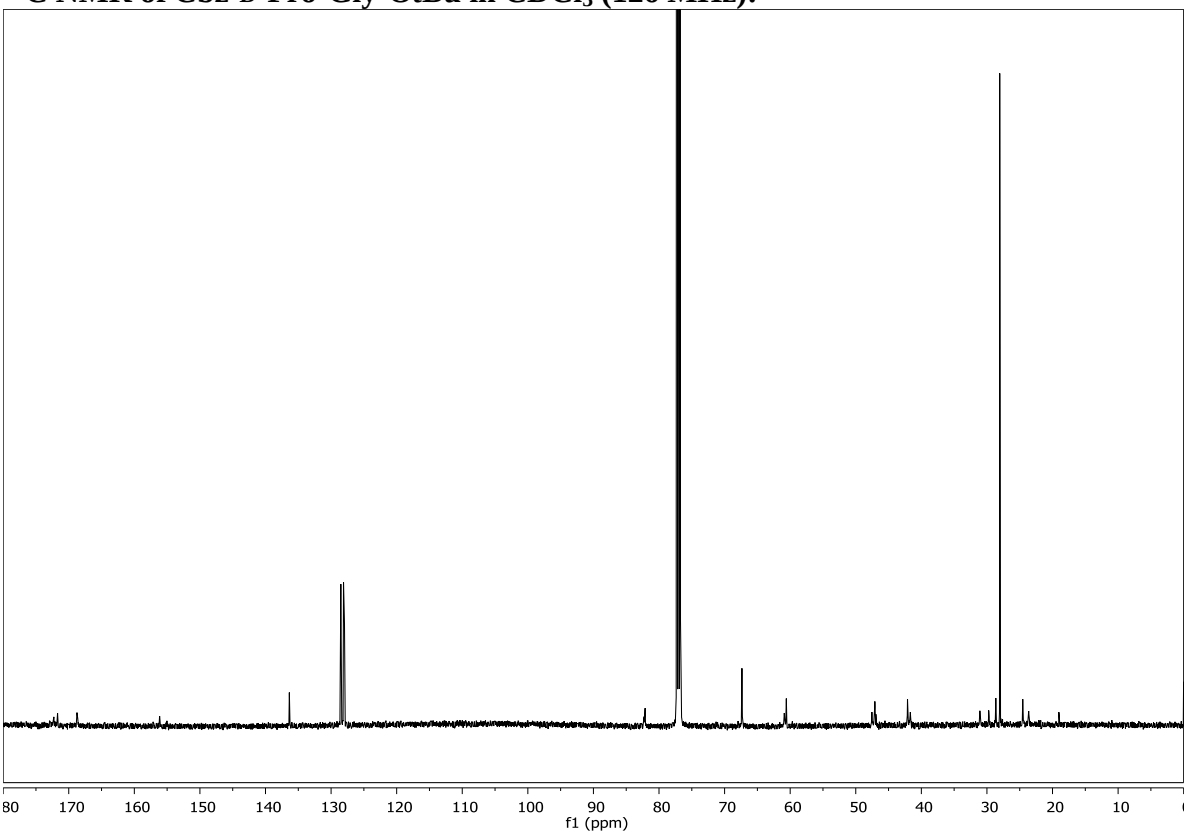
^{13}C NMR of Fmoc-Pro-Pro-Gly-OH in $CDCl_3$ (126 MHz):



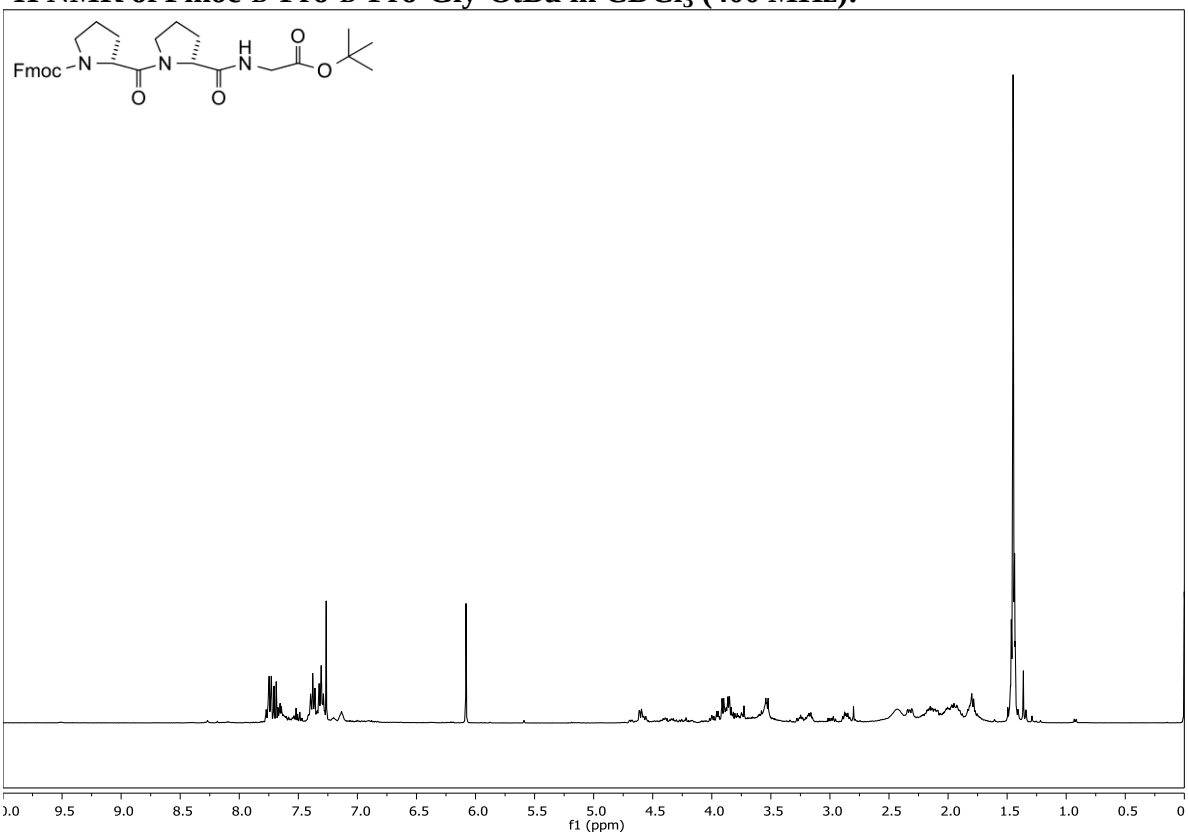
^{1}H NMR of Cbz-D-Pro-Gly-O*t*Bu in $CDCl_3$ (400 MHz):



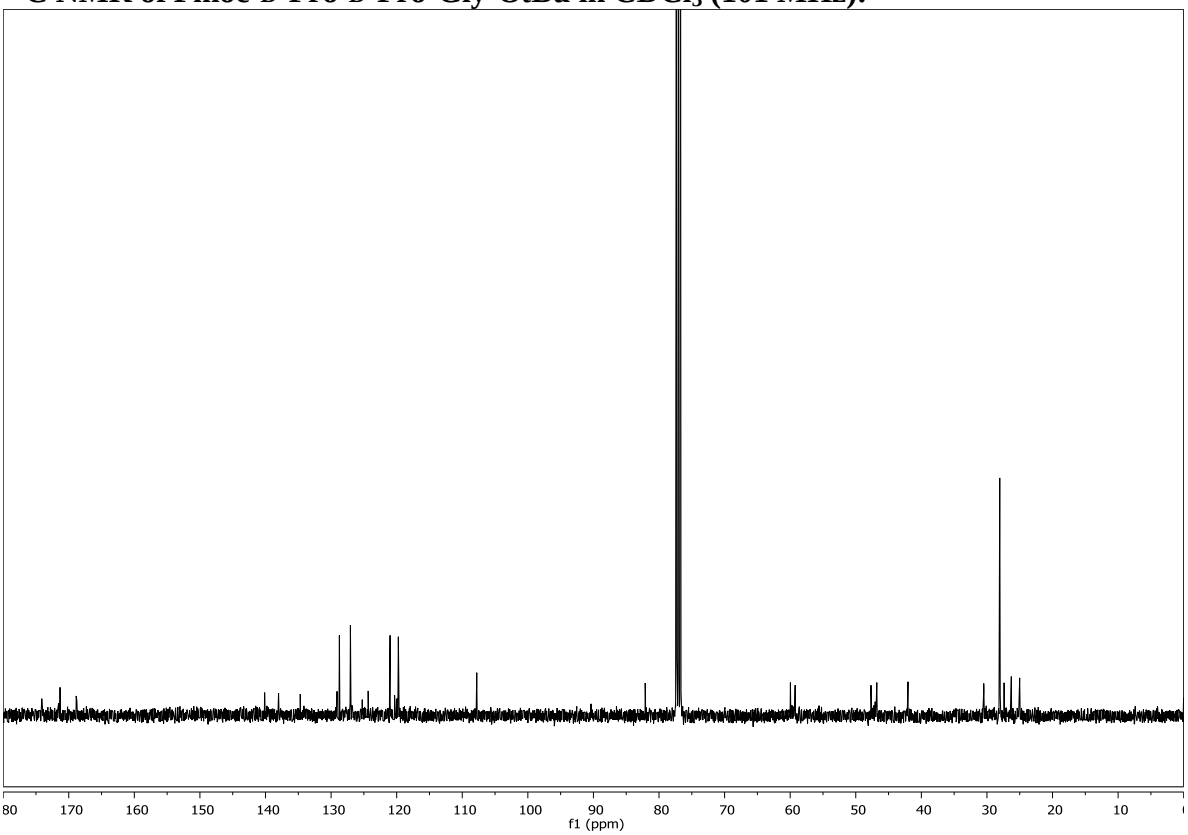
^{13}C NMR of Cbz-D-Pro-Gly-O*t*Bu in $CDCl_3$ (126 MHz):



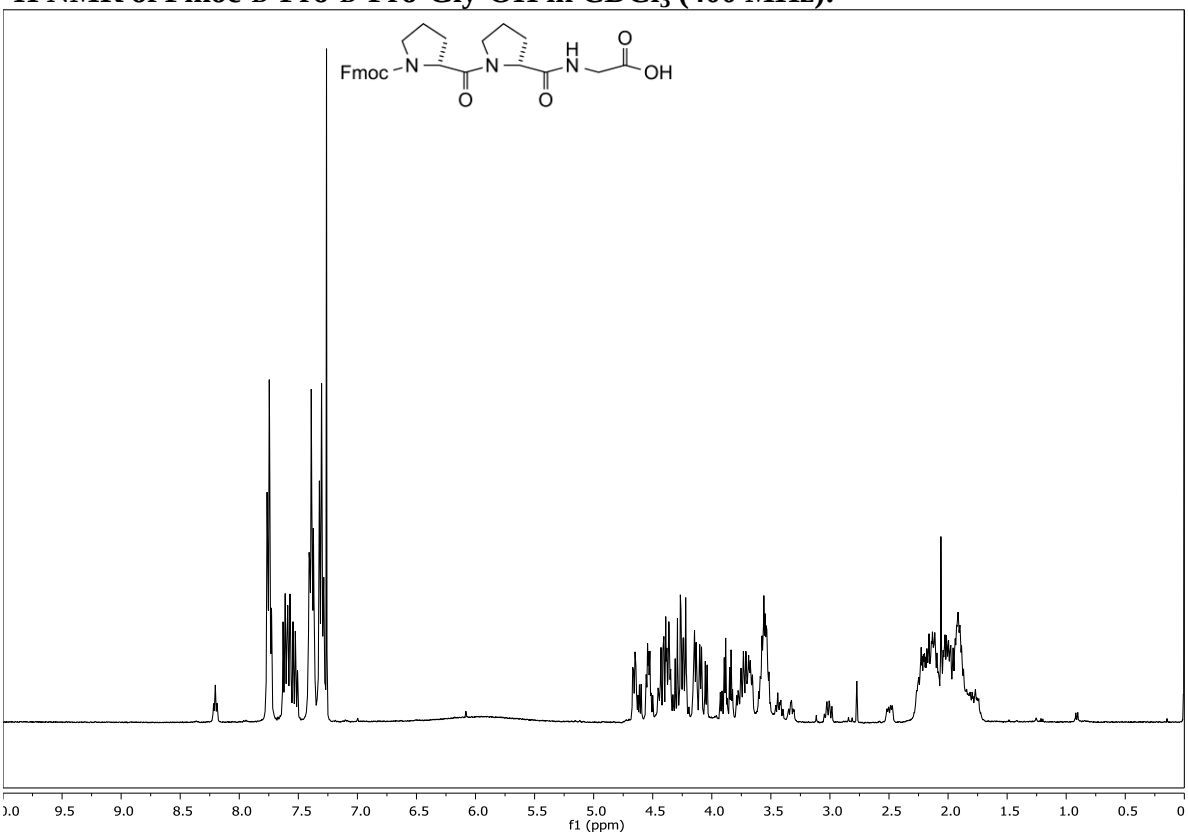
^{1}H NMR of Fmoc-D-Pro-D-Pro-Gly-O*t*Bu in $CDCl_3$ (400 MHz):



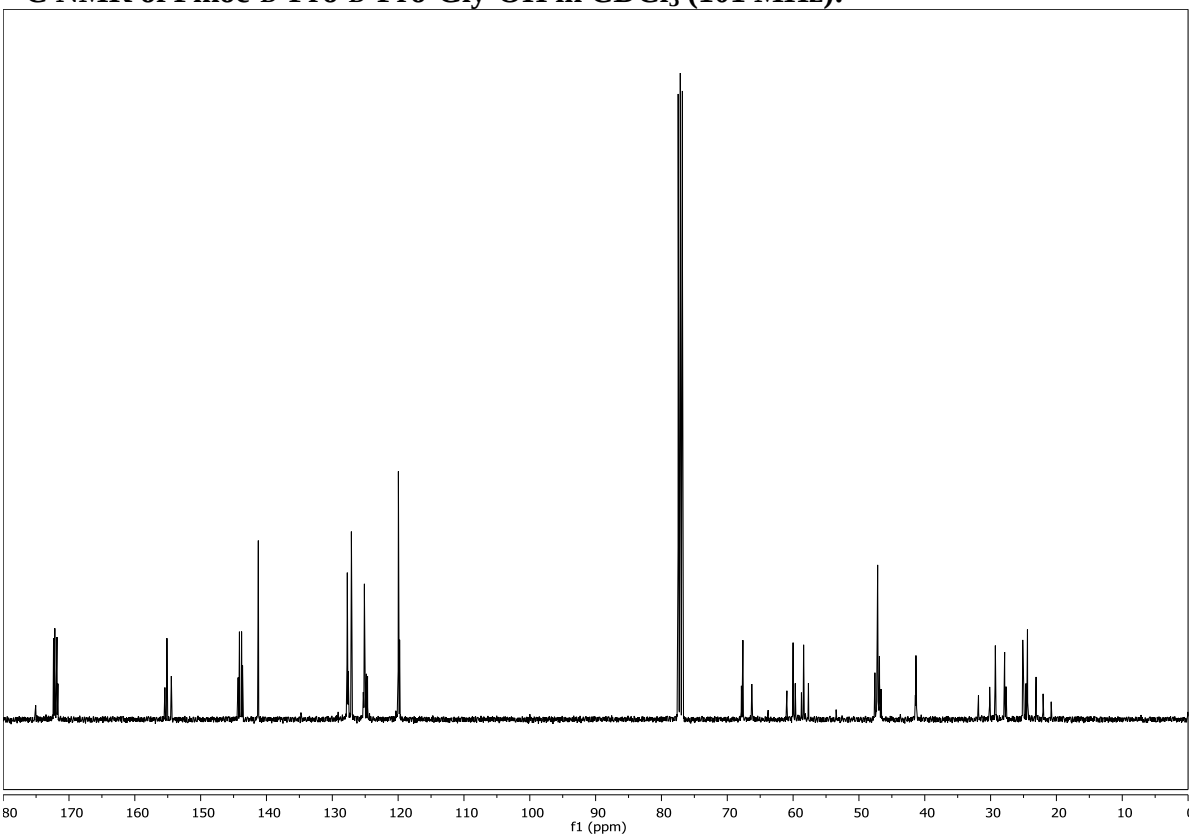
^{13}C NMR of Fmoc-D-Pro-D-Pro-Gly-O*t*Bu in $CDCl_3$ (101 MHz):



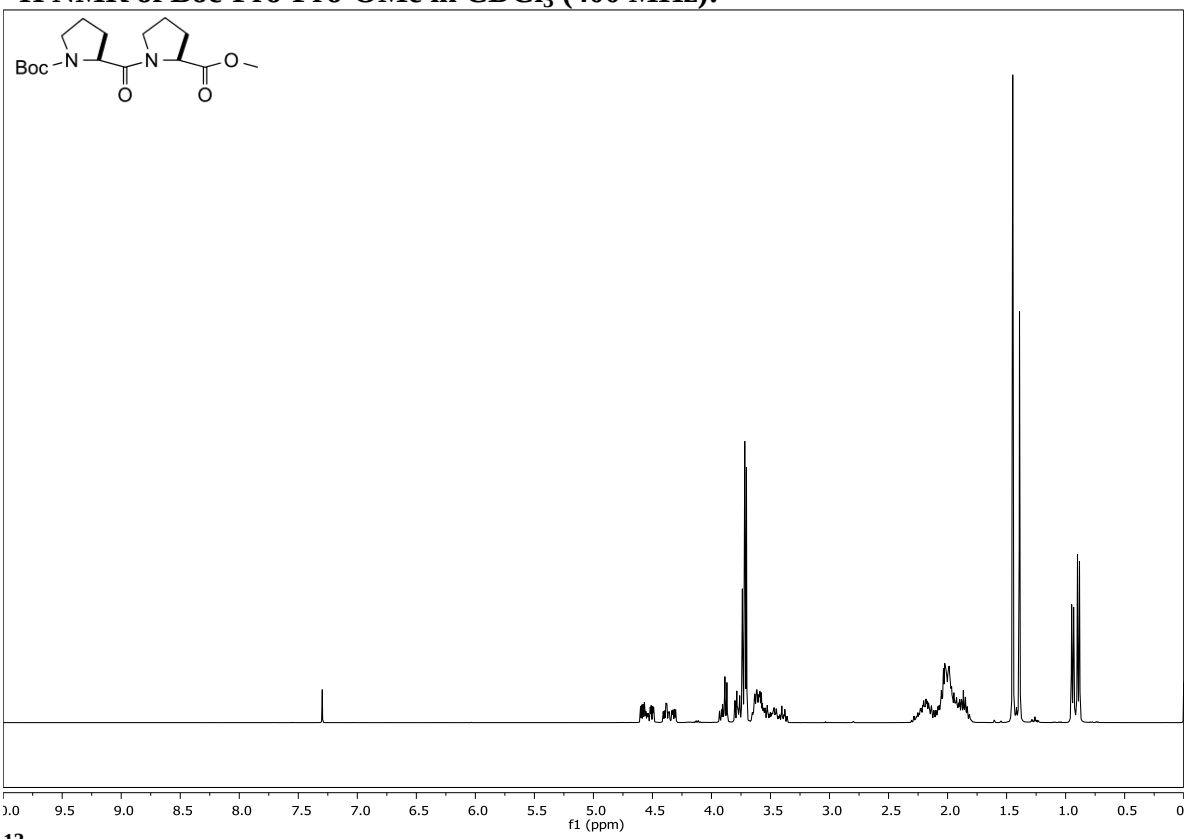
^{1}H NMR of Fmoc-D-Pro-D-Pro-Gly-OH in $CDCl_3$ (400 MHz):



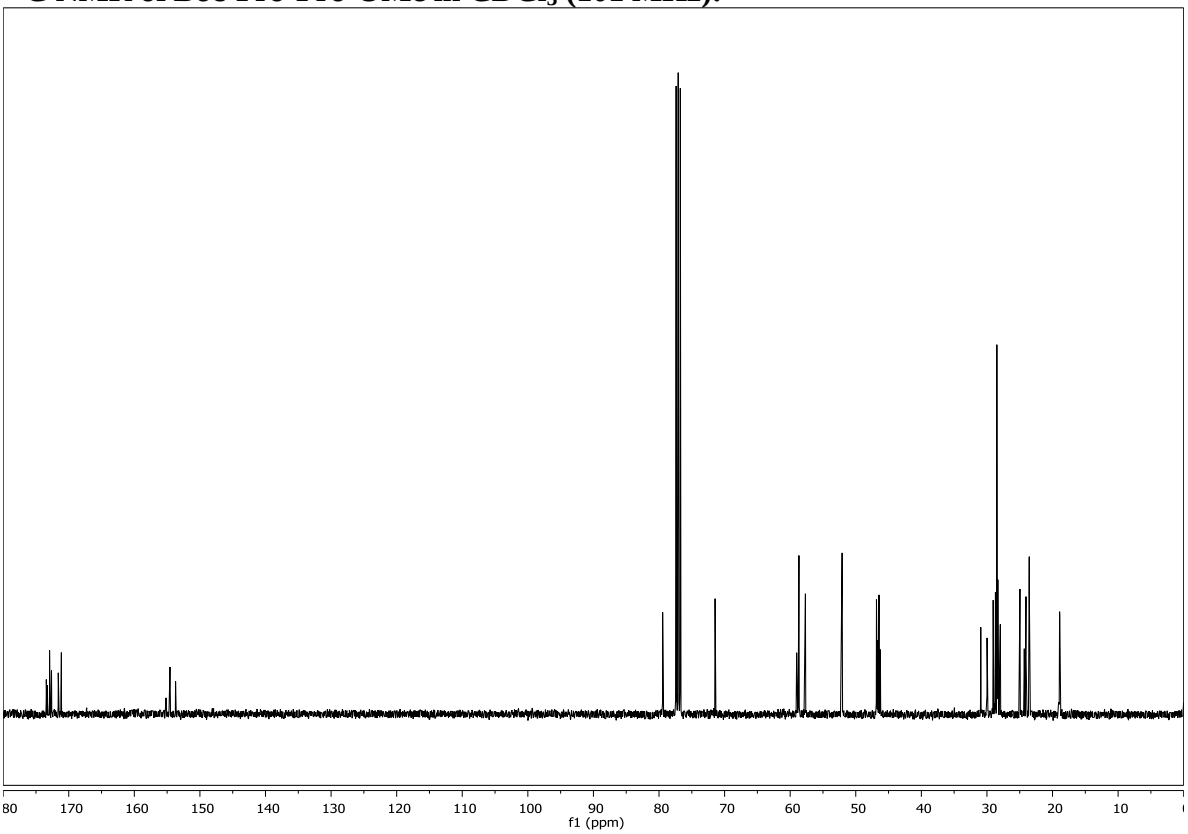
^{13}C NMR of Fmoc-D-Pro-D-Pro-Gly-OH in $CDCl_3$ (101 MHz):



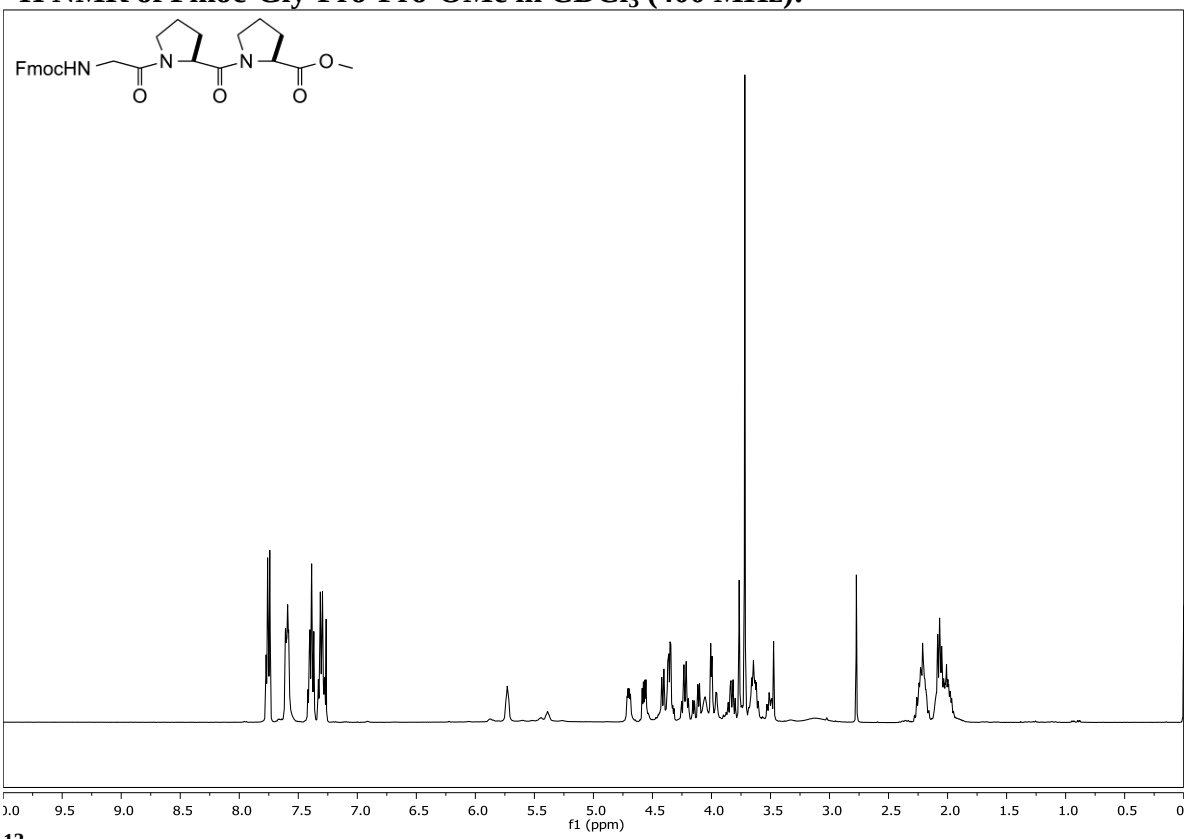
^{1}H NMR of Boc-Pro-Pro-OMe in $CDCl_3$ (400 MHz):



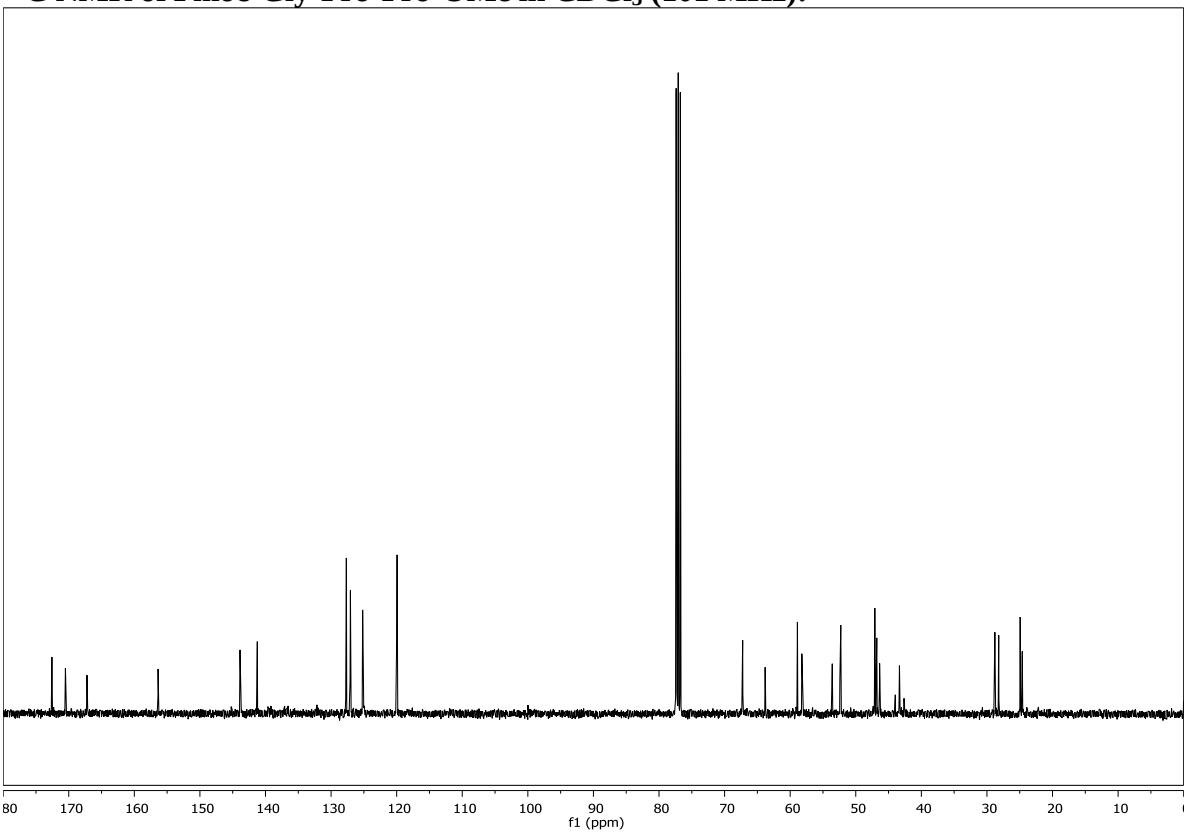
^{13}C NMR of Boc-Pro-Pro-OMe in $CDCl_3$ (101 MHz):



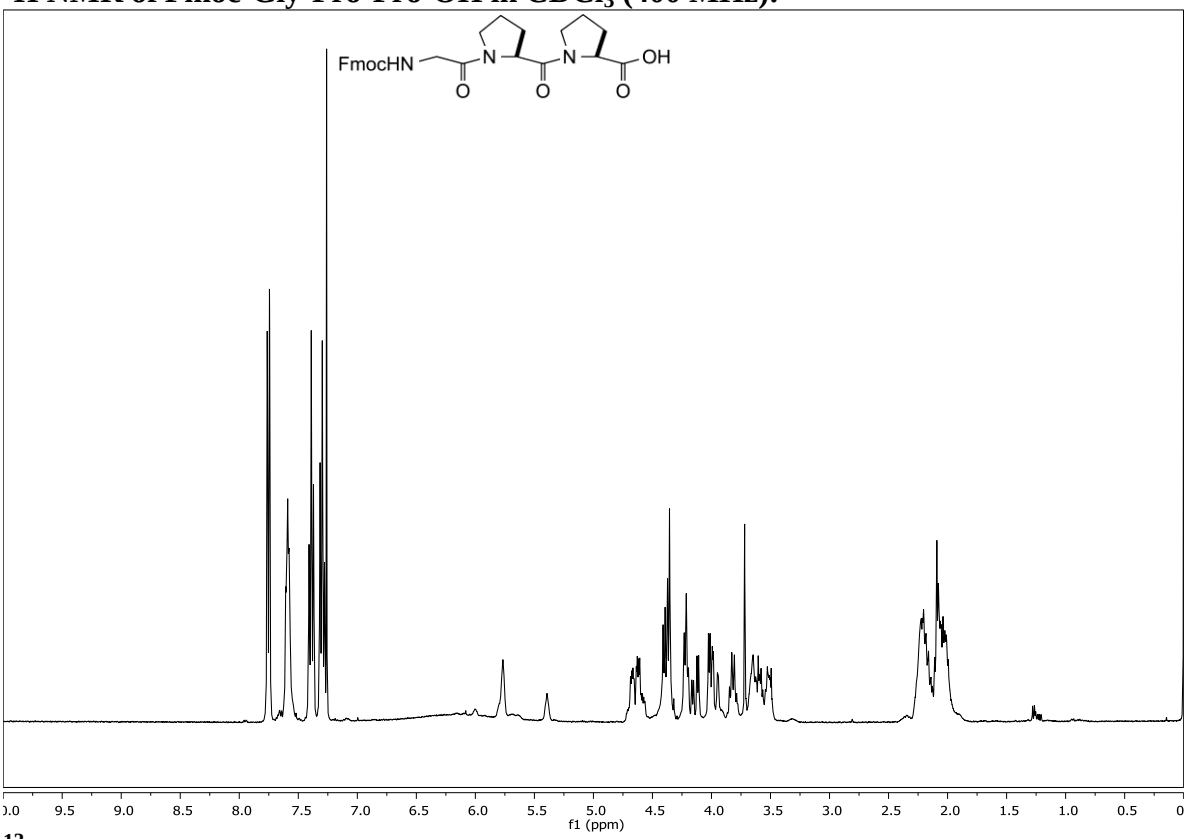
^{1}H NMR of Fmoc-Gly-Pro-Pro-OMe in $CDCl_3$ (400 MHz):



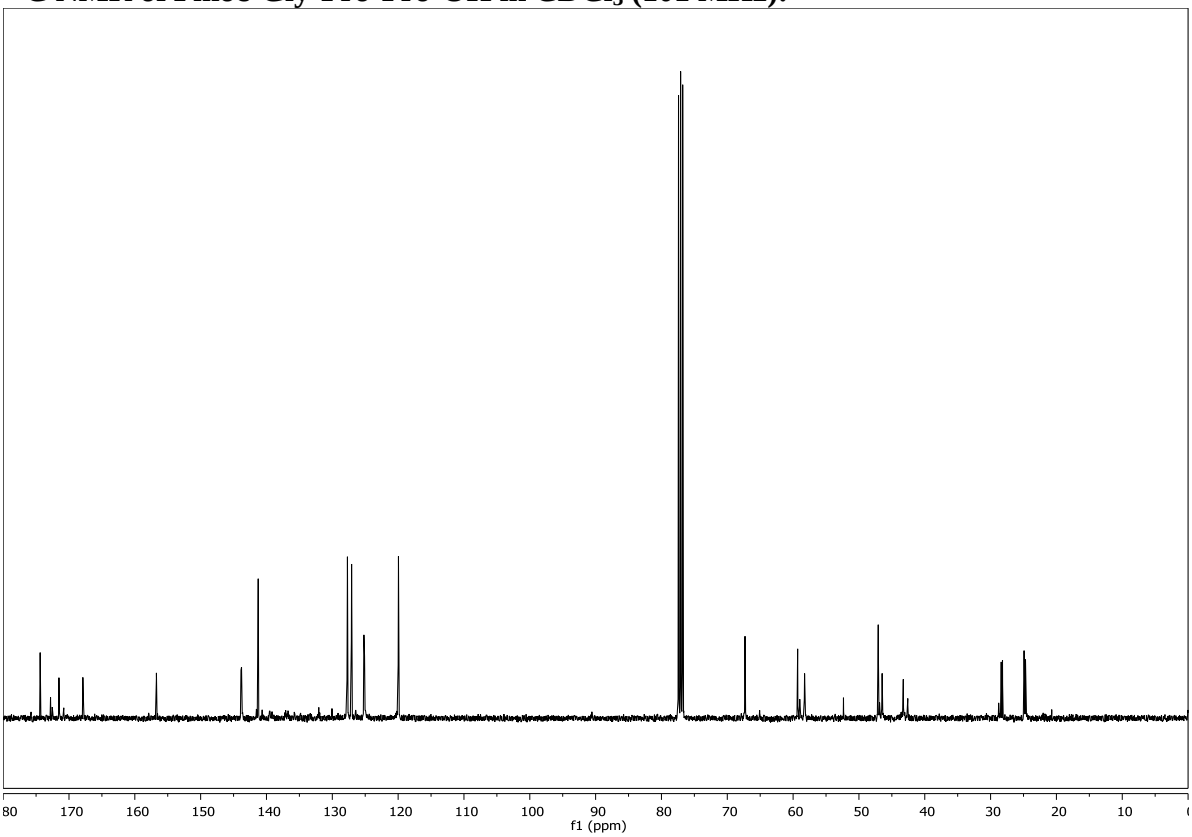
^{13}C NMR of Fmoc-Gly-Pro-Pro-OMe in $CDCl_3$ (101 MHz):



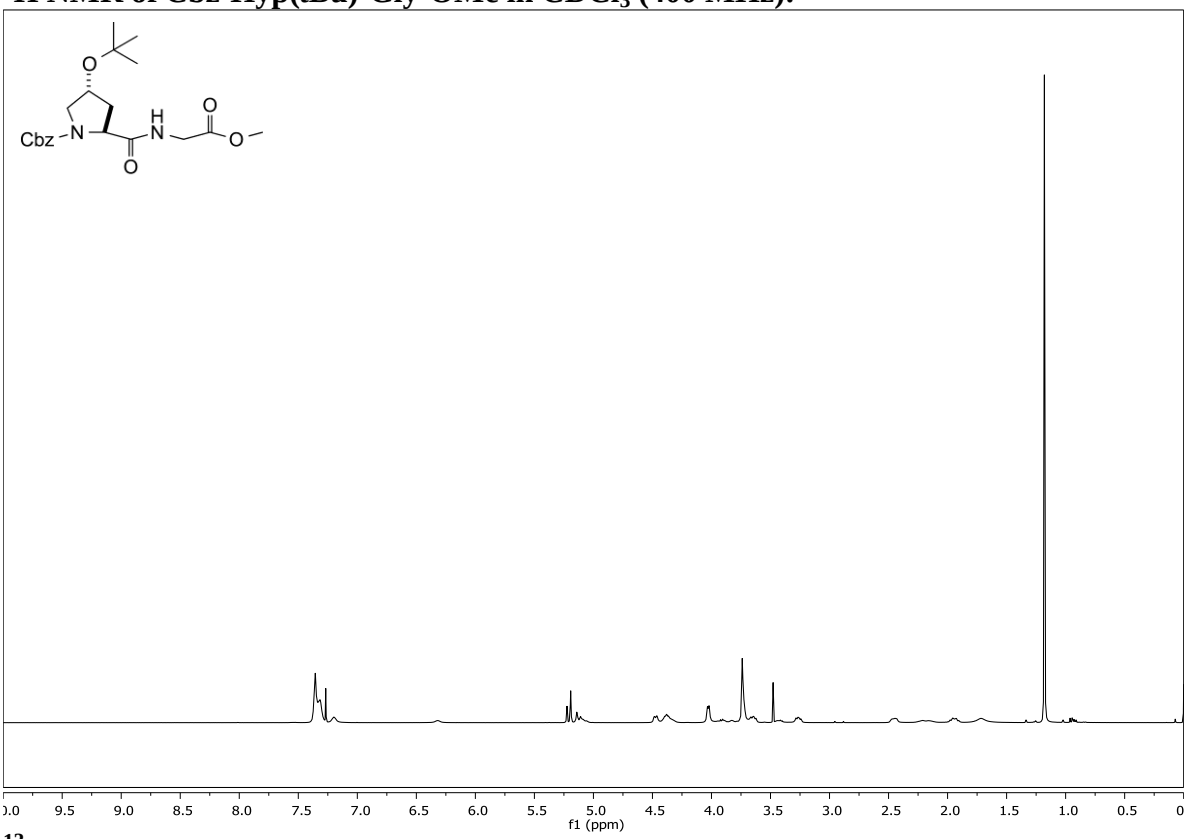
^{1}H NMR of Fmoc-Gly-Pro-Pro-OH in $CDCl_3$ (400 MHz):



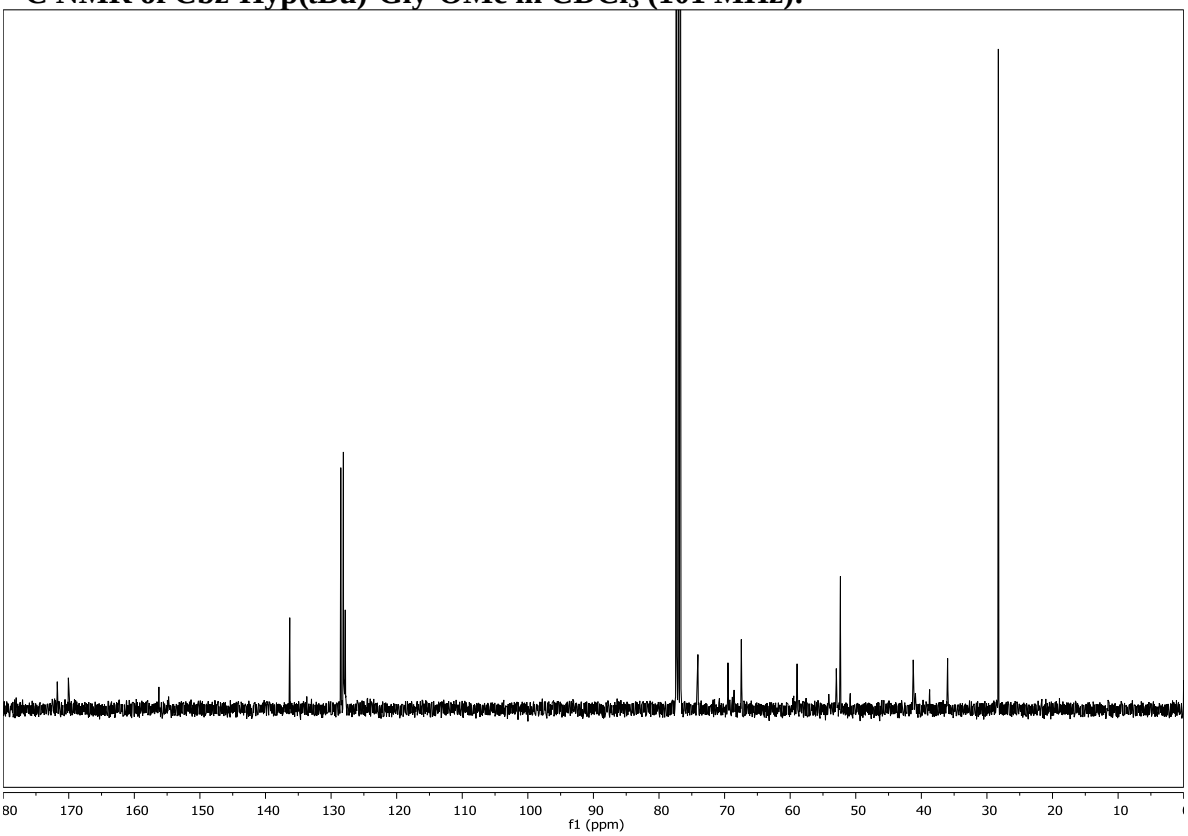
^{13}C NMR of Fmoc-Gly-Pro-Pro-OH in $CDCl_3$ (101 MHz):



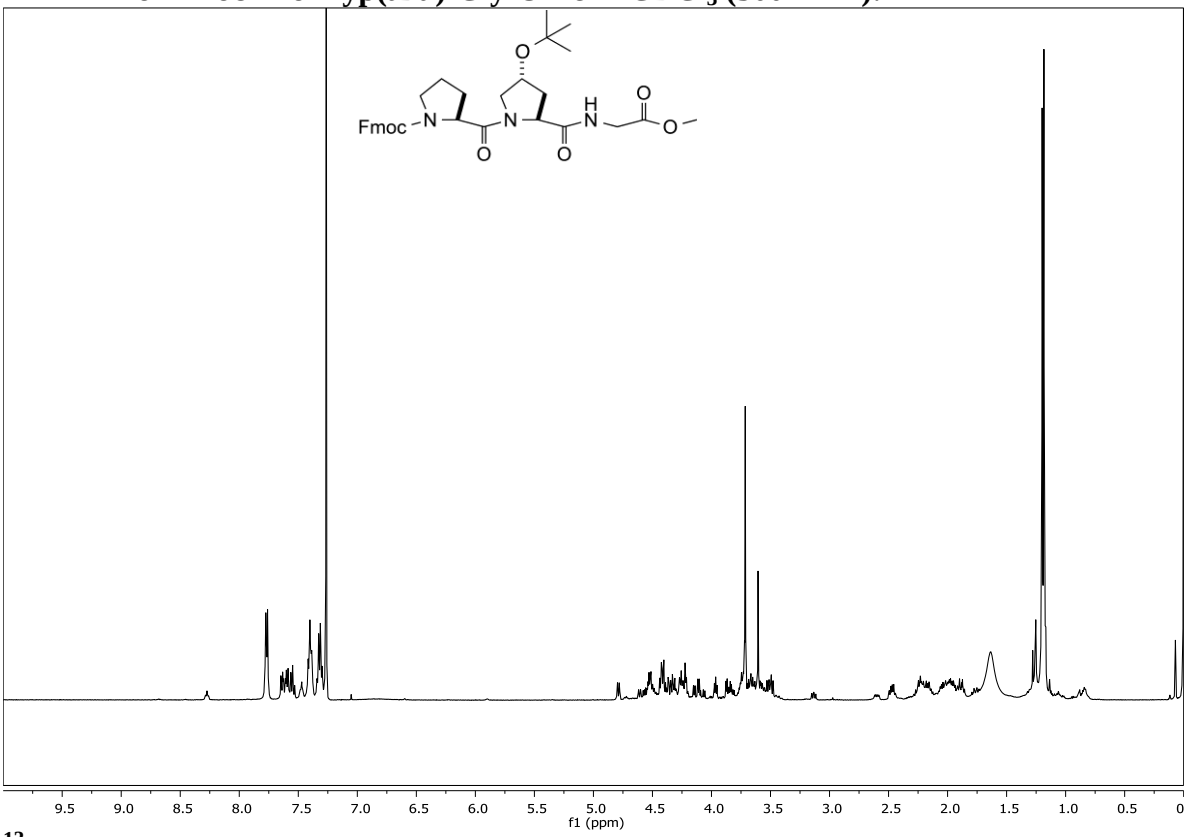
^{1}H NMR of Cbz-Hyp(*t*Bu)-Gly-OMe in $CDCl_3$ (400 MHz):



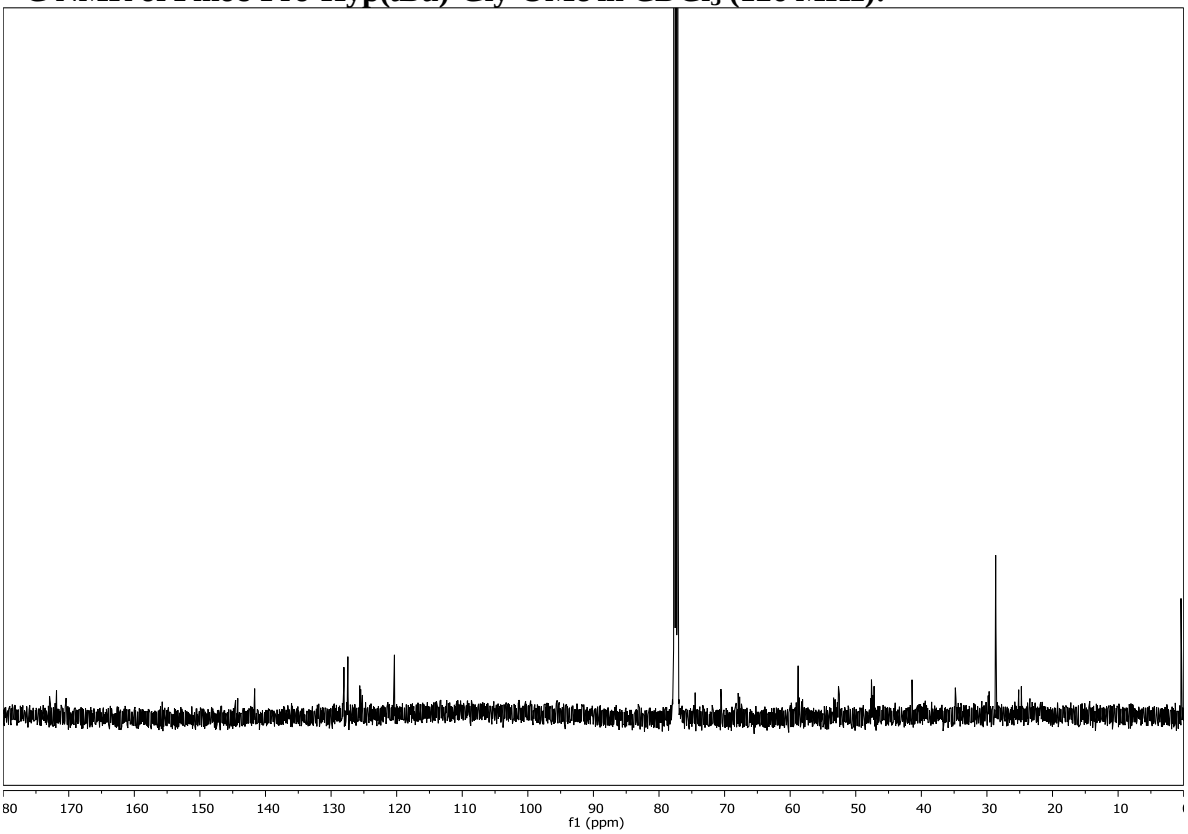
^{13}C NMR of Cbz-Hyp(*t*Bu)-Gly-OMe in $CDCl_3$ (101 MHz):



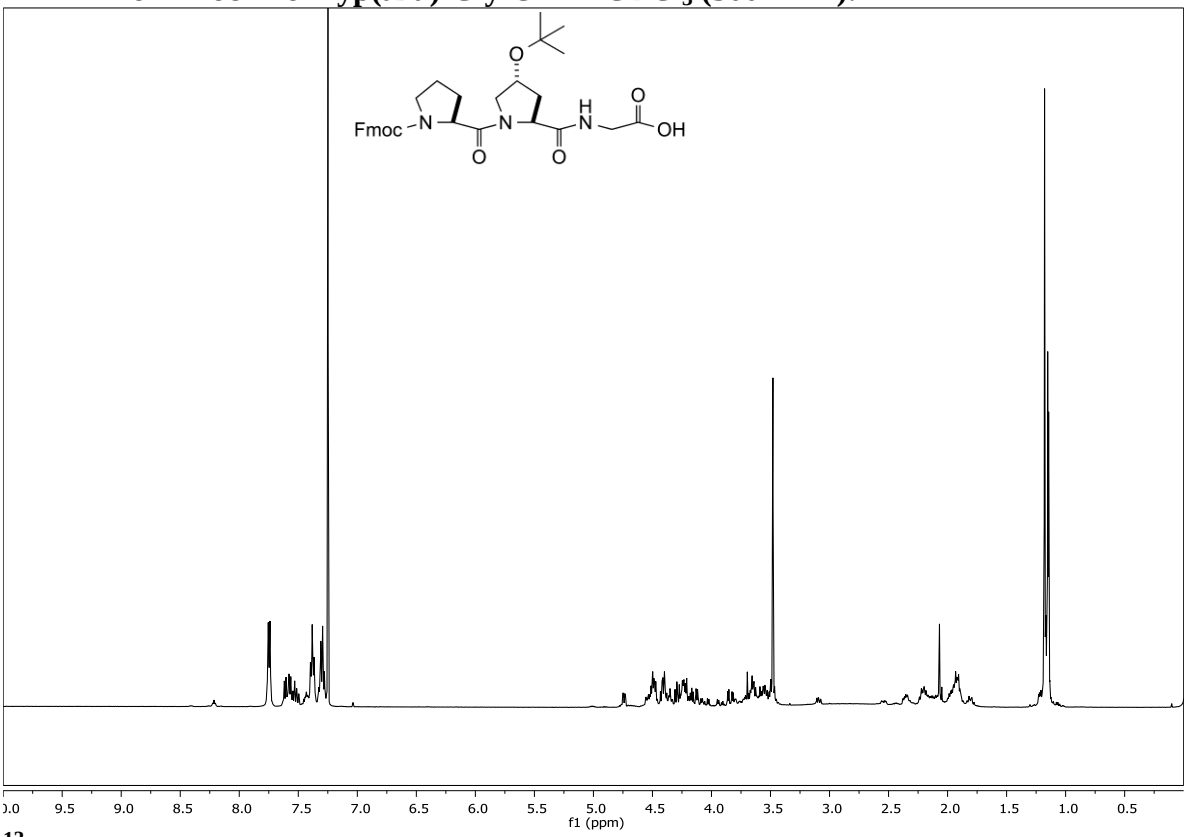
^{1}H NMR of Fmoc-Pro-Hyp(*t*Bu)-Gly-OMe in $CDCl_3$ (500 MHz):



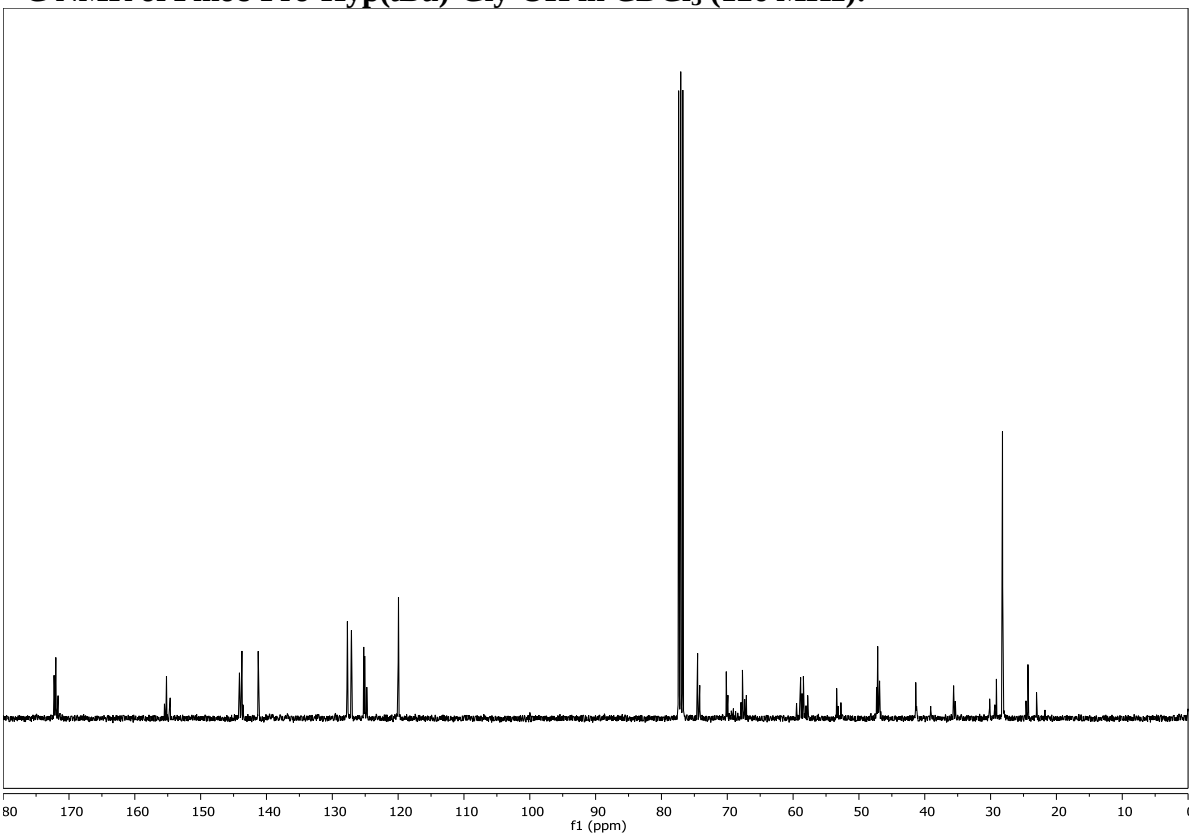
^{13}C NMR of Fmoc-Pro-Hyp(*t*Bu)-Gly-OMe in $CDCl_3$ (126 MHz):



^{1}H NMR of Fmoc-Pro-Hyp(*t*Bu)-Gly-OH in $CDCl_3$ (500 MHz):

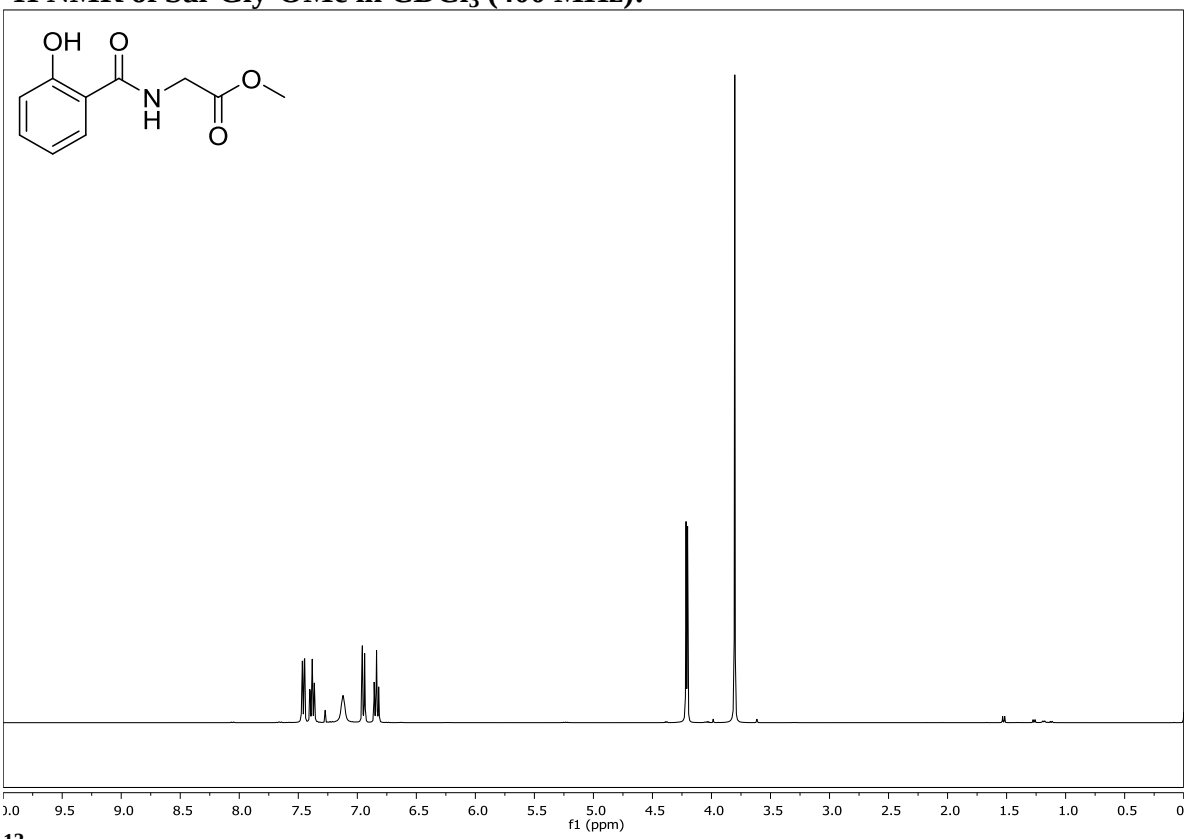


^{13}C NMR of Fmoc-Pro-Hyp(*t*Bu)-Gly-OH in $CDCl_3$ (126 MHz):

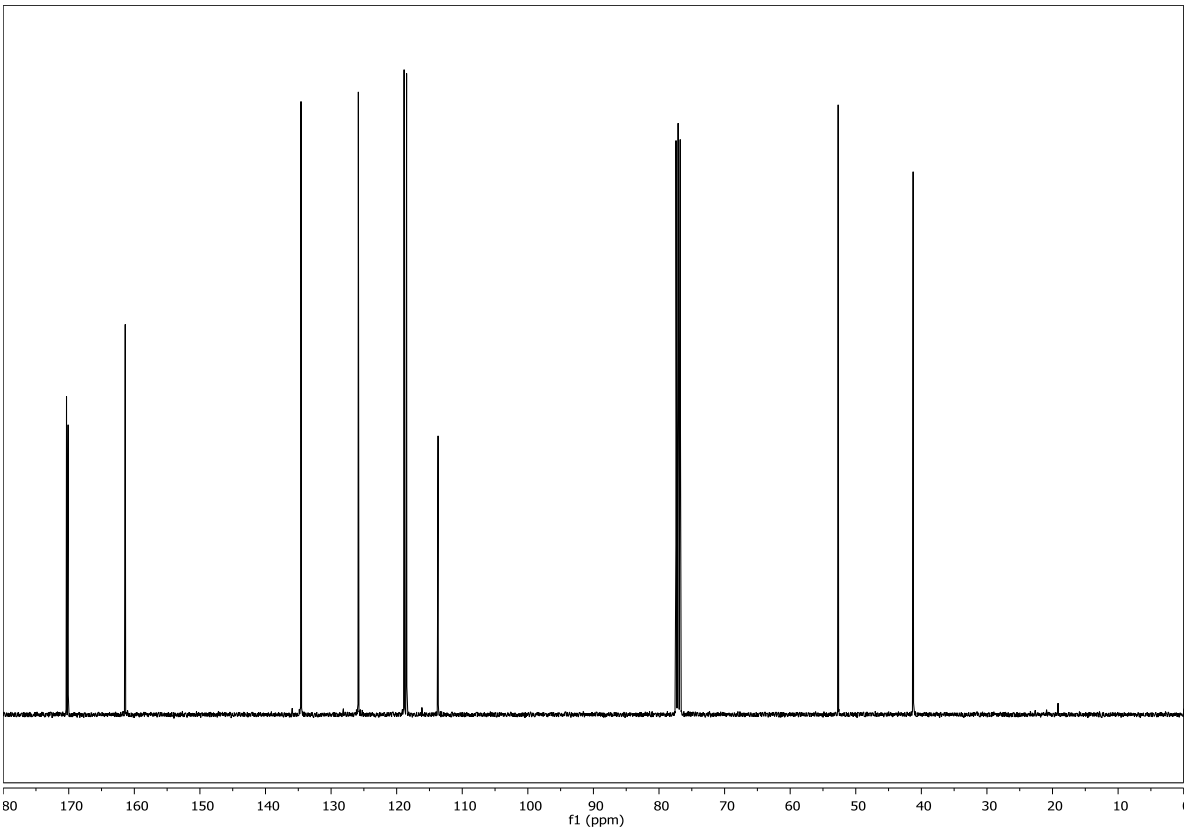


8.2 Chapter 3

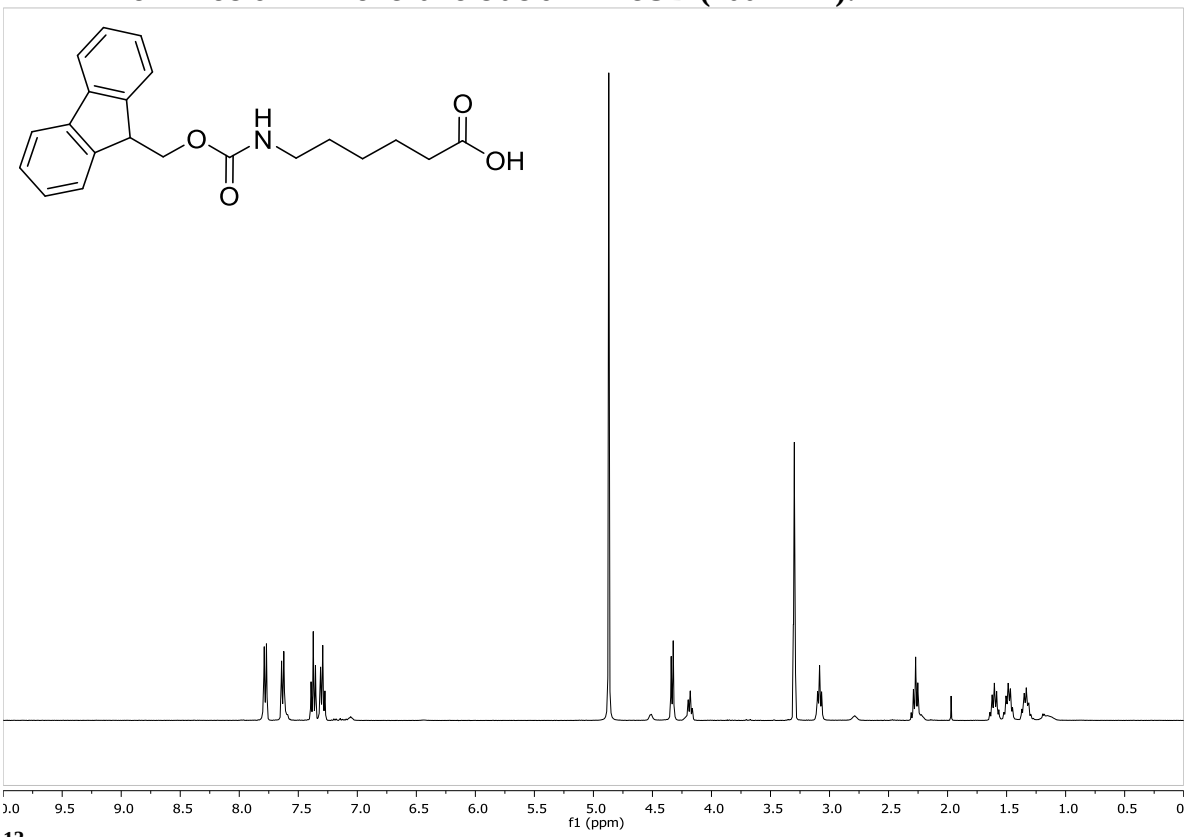
^{1}H NMR of Sal-Gly-OMe in $CDCl_3$ (400 MHz):



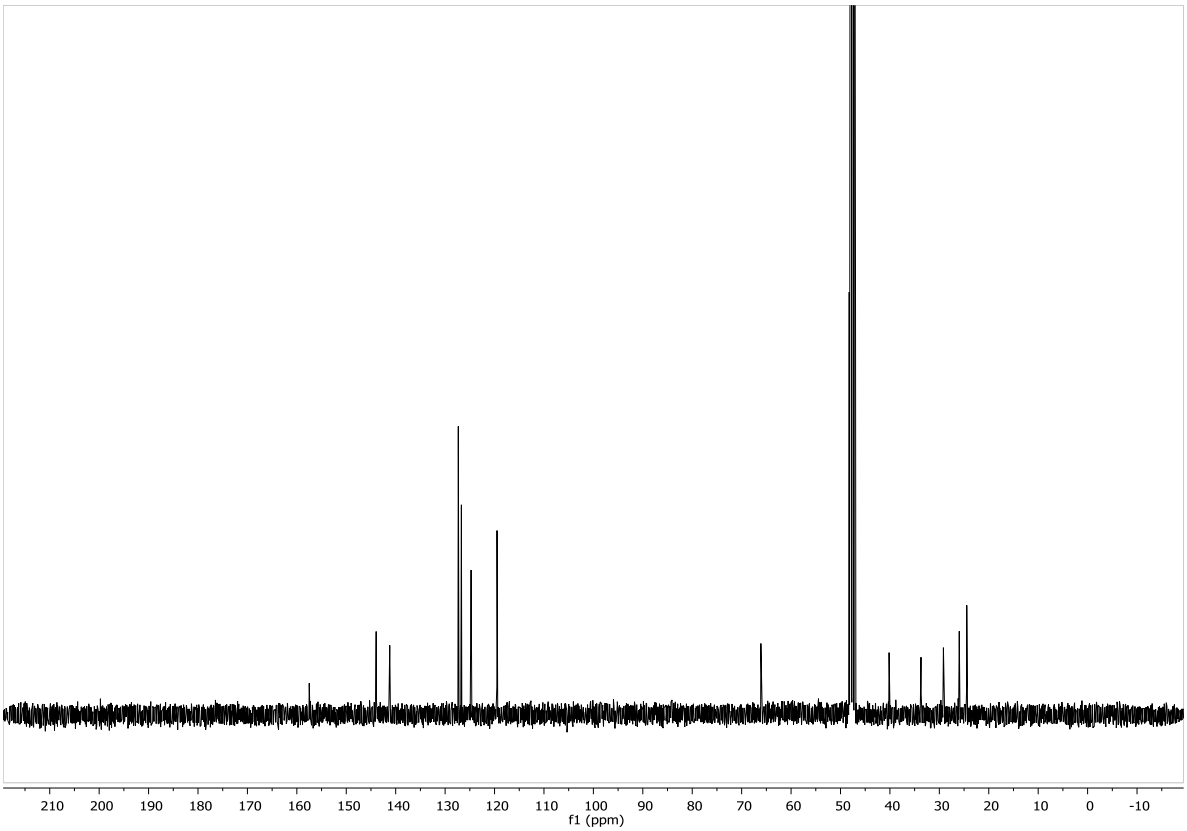
^{13}C NMR of Sal-Gly-OMe in $CDCl_3$ (101 MHz):



^{1}H NMR of Fmoc-6-Aminohexanoic acid in MeOD (400 MHz):

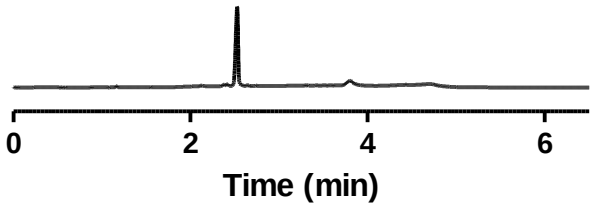


^{13}C NMR of Fmoc-6-Aminohexanoic acid in MeOD (101 MHz):

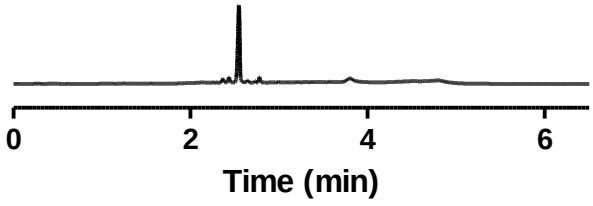


UPLC Traces

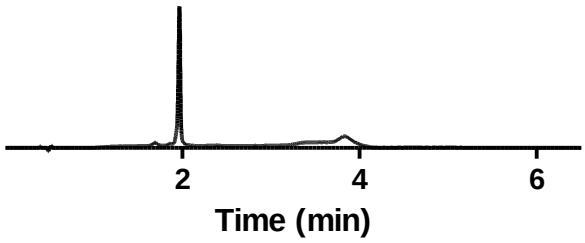
D-CMP-red



L-CMP-red



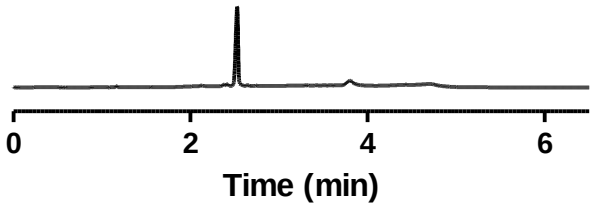
CMP-sal



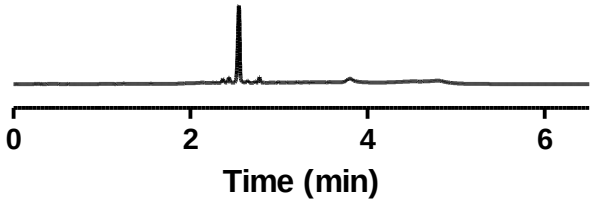
8.3 Chapter 4

UPLC Traces

D-CMP-red

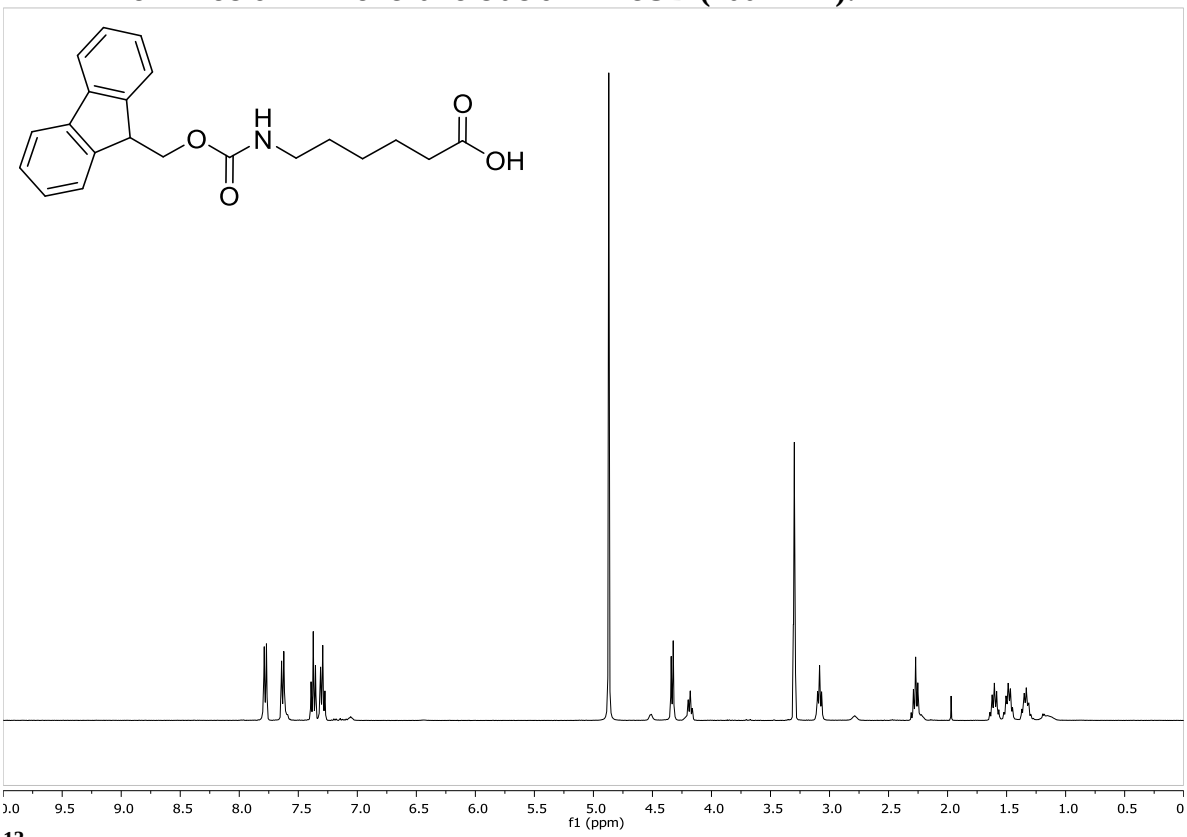


L-CMP-red

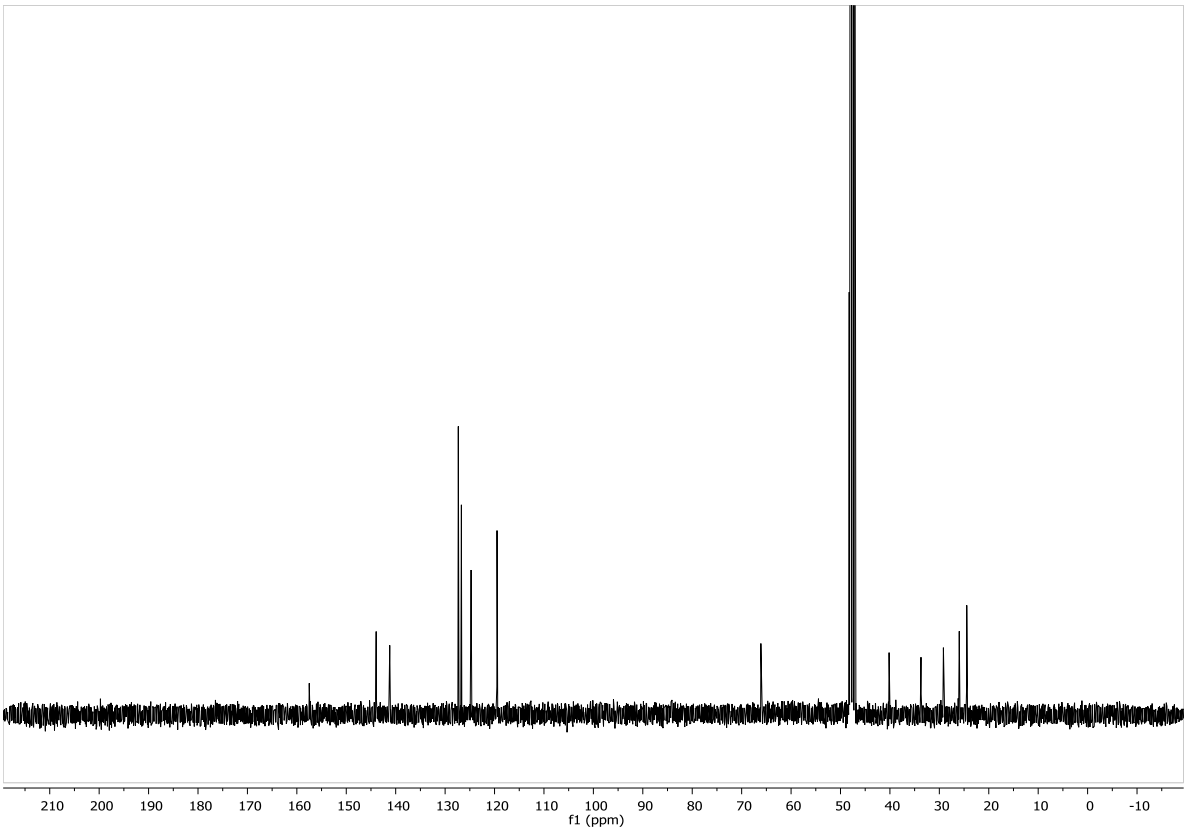


8.4 Chapter 5

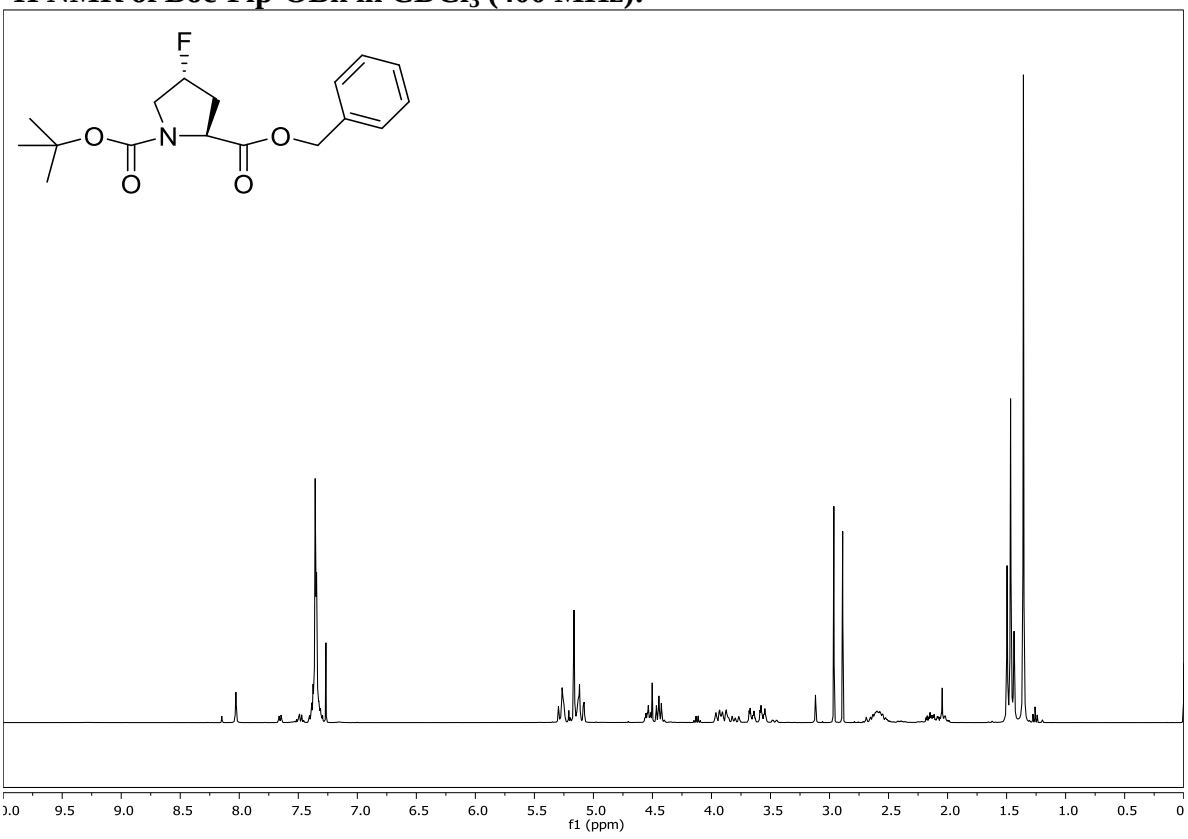
^{1}H NMR of Fmoc-6-Aminohexanoic acid in MeOD (400 MHz):



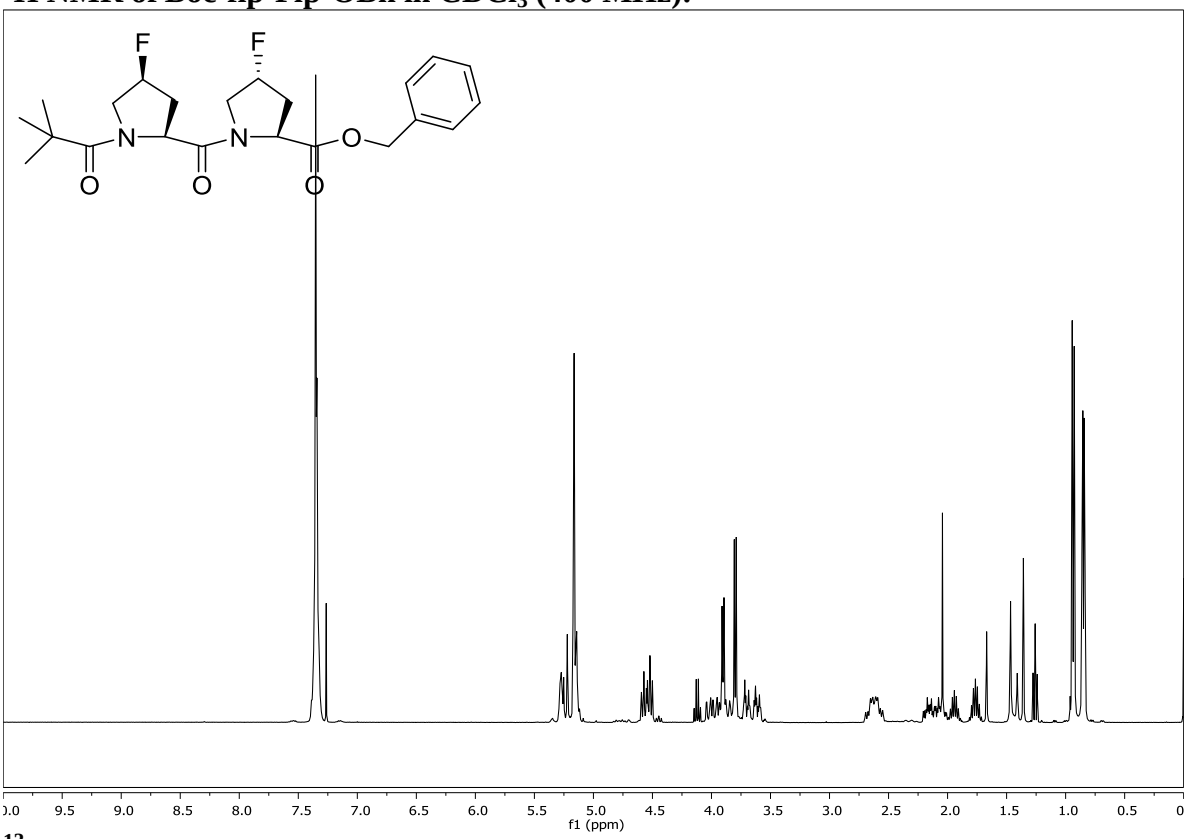
^{13}C NMR of Fmoc-6-Aminohexanoic acid in MeOD (101 MHz):



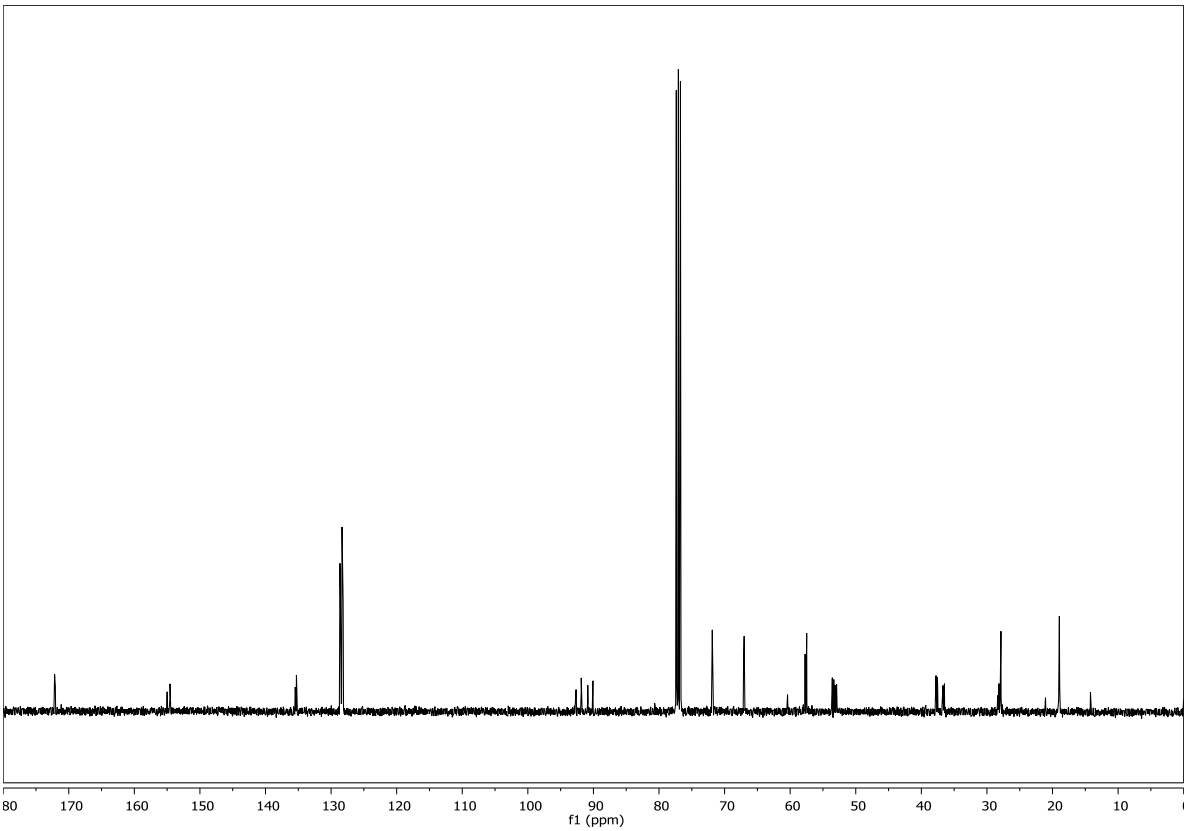
^{1}H NMR of Boc-Flp-OBn in $CDCl_3$ (400 MHz):



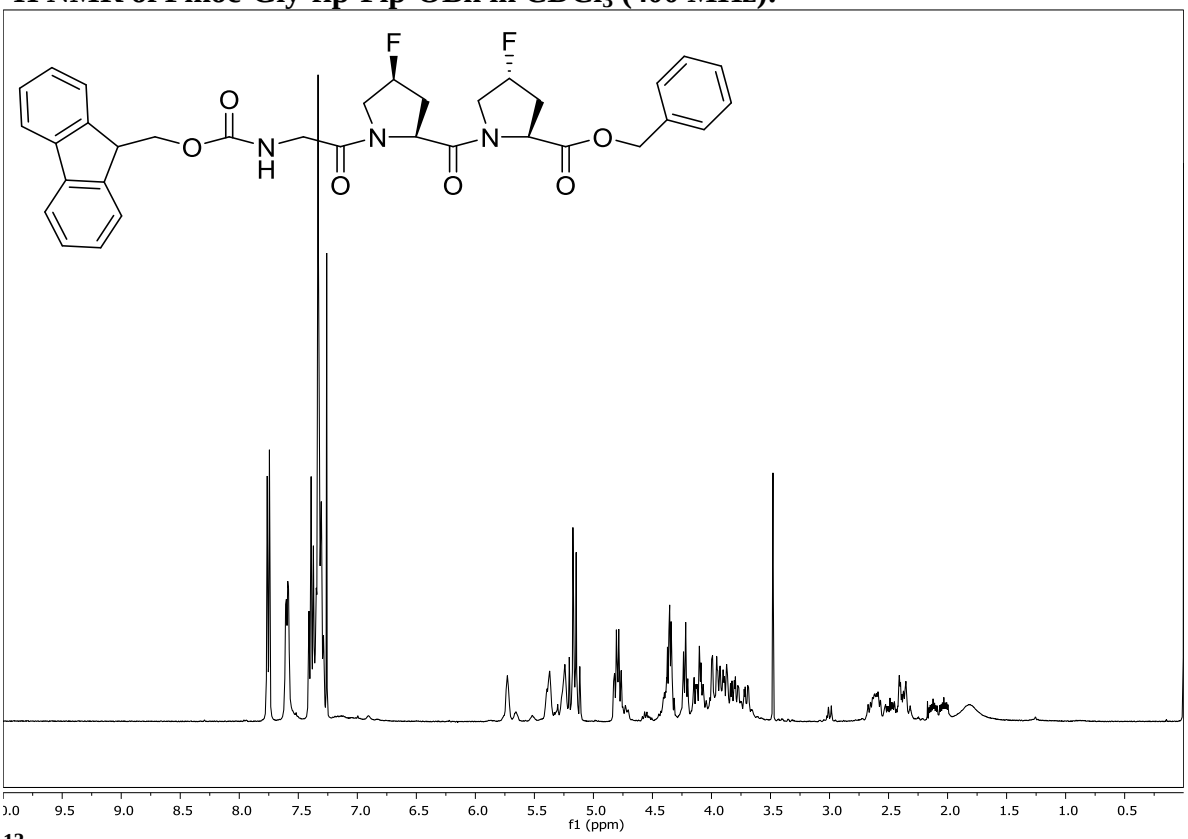
^{1}H NMR of Boc-flp-Flp-OBn in $CDCl_3$ (400 MHz):



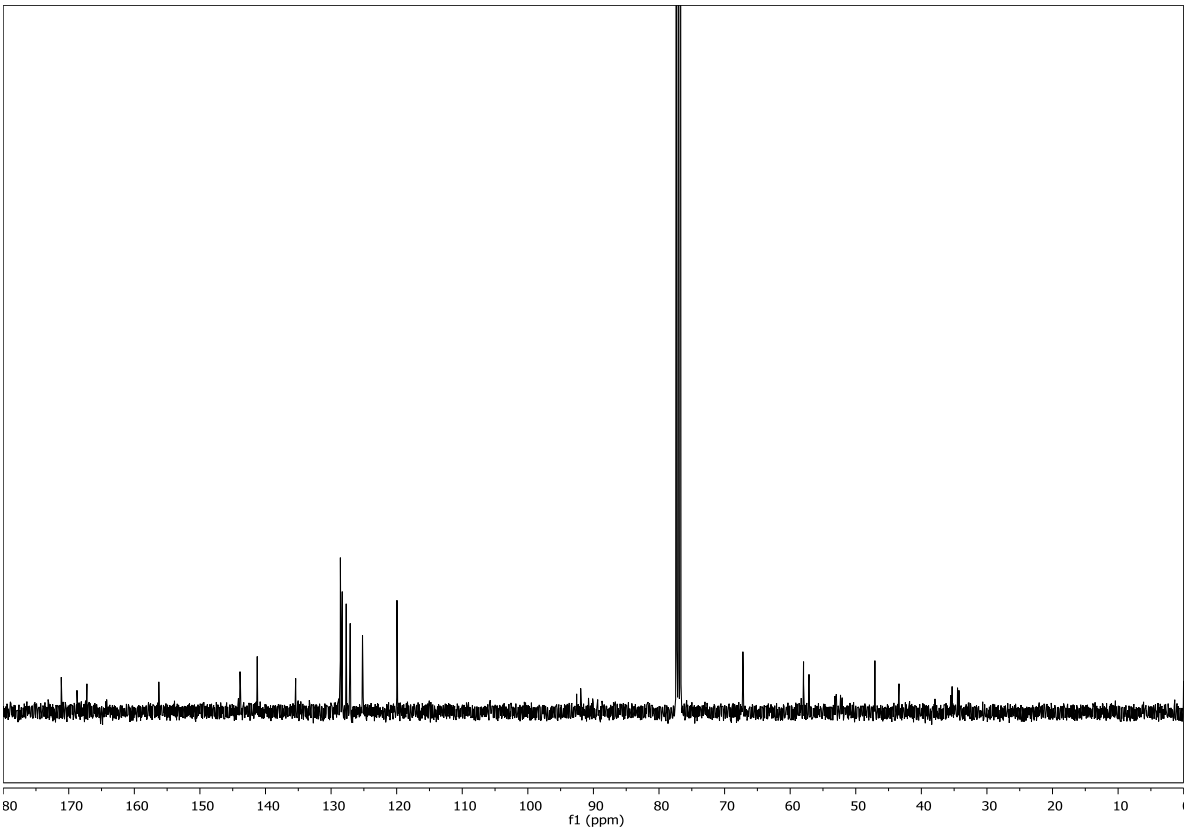
^{13}C NMR of Boc-flp-Flp-OBn in $CDCl_3$ (101 MHz):



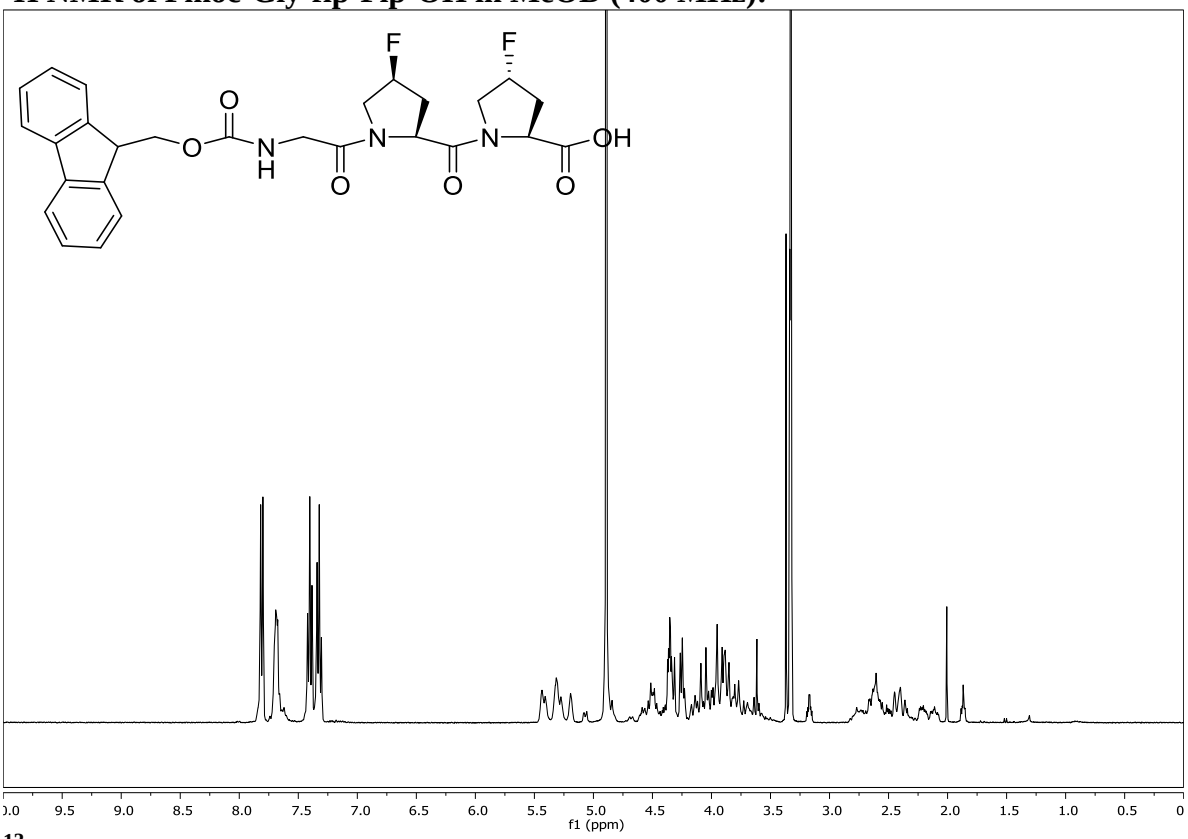
^{1}H NMR of Fmoc-Gly-flp-Flp-OBn in $CDCl_3$ (400 MHz):



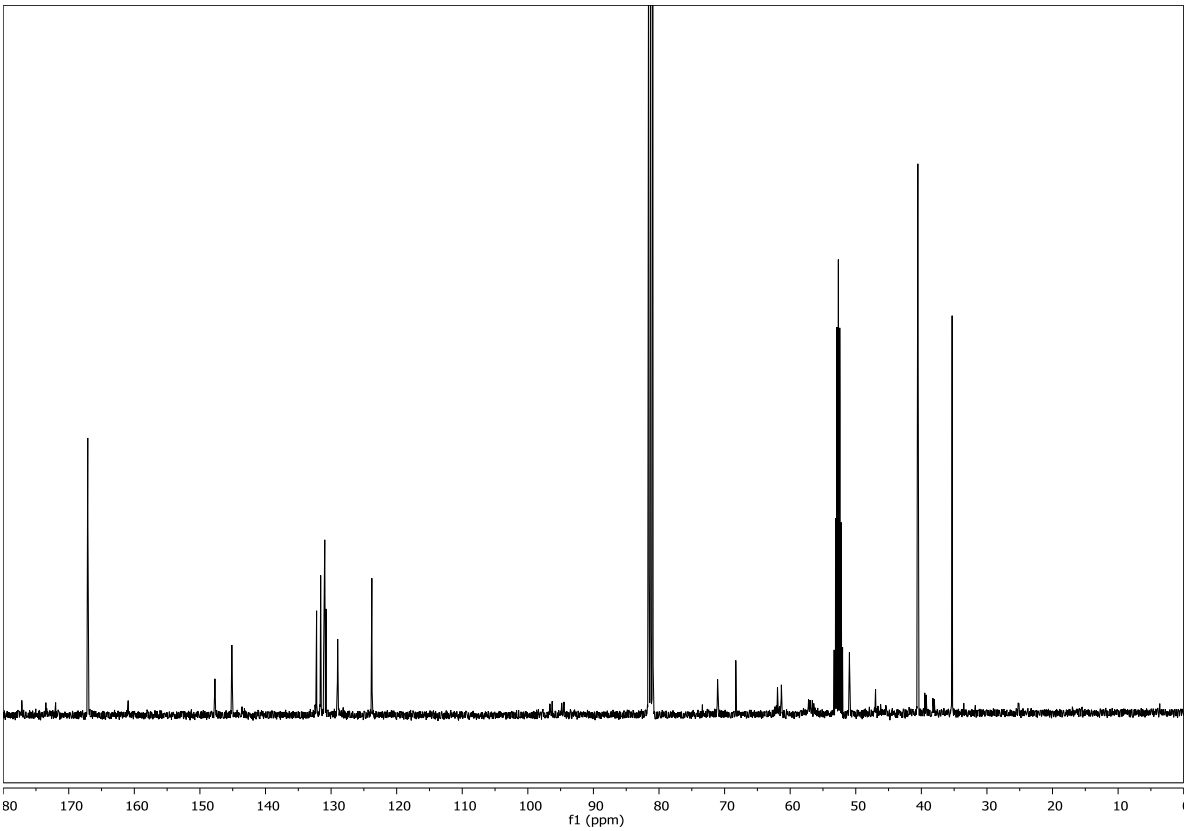
^{13}C NMR of Fmoc-Gly-flp-Flp-OBn in $CDCl_3$ (101 MHz):



^{1}H NMR of Fmoc-Gly-flp-Flp-OH in MeOD (400 MHz):

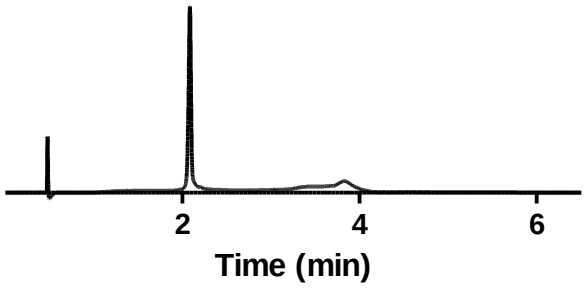


^{13}C NMR of Fmoc-Gly-flp-Flp-OBn in $CDCl_3$ (101 MHz):

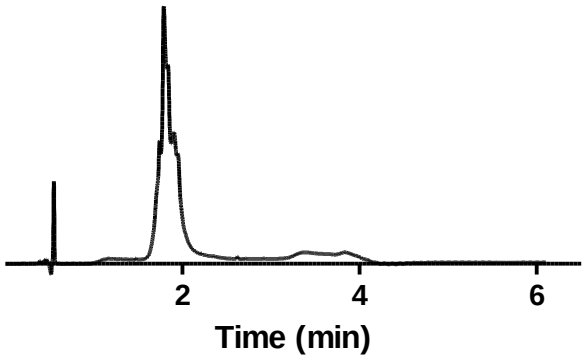


UPLC Traces

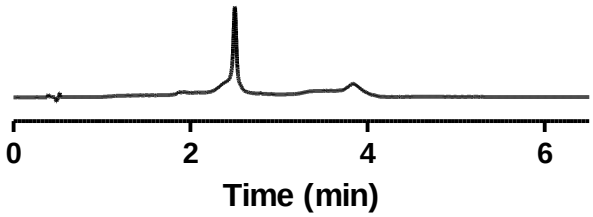
nest-o



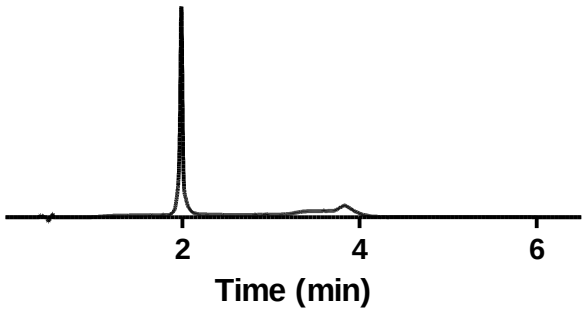
nest-c



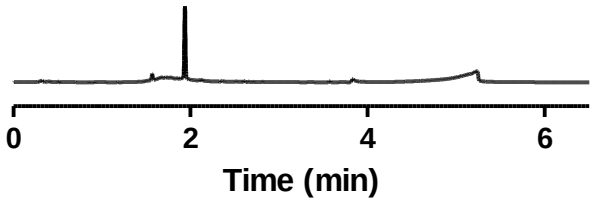
nest-r



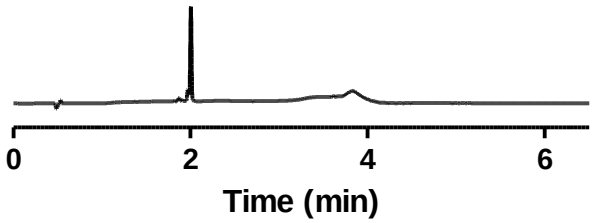
biotin-nest-o



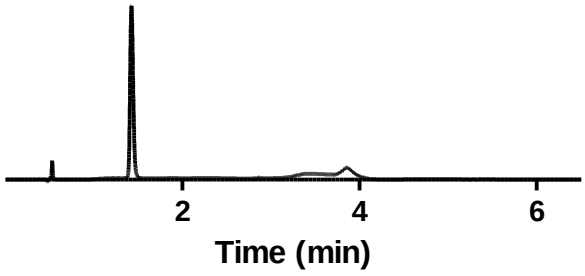
biotin-nest-r



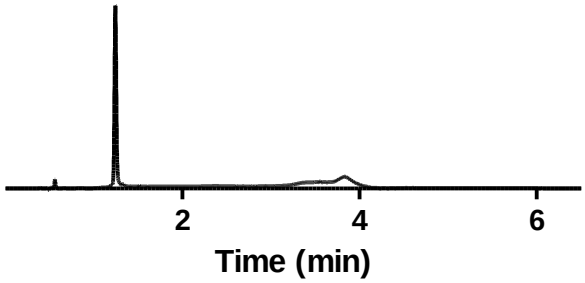
biotin-fluorescein



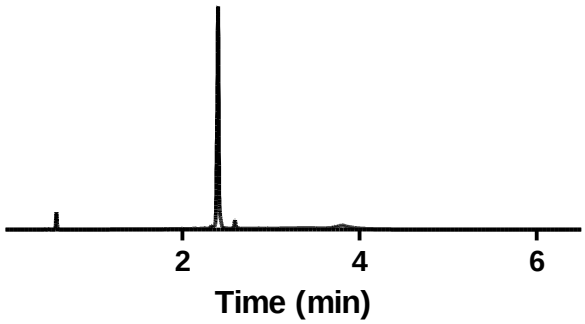
(flp-Flp-Gly)$_7$



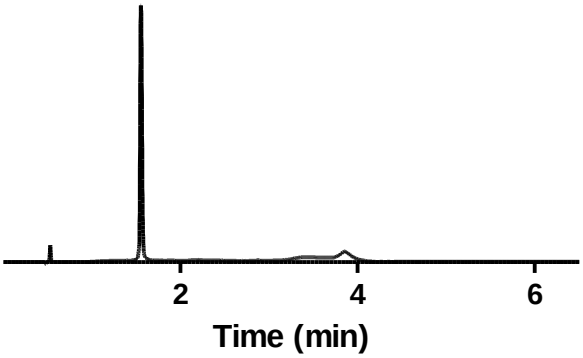
(flp-Hyp-Gly)$_7$



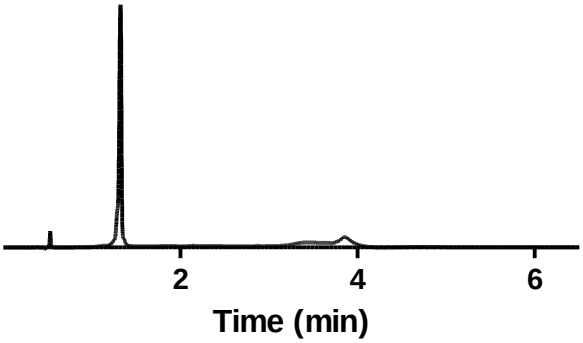
(Pro-Ile-Gly)$_7$



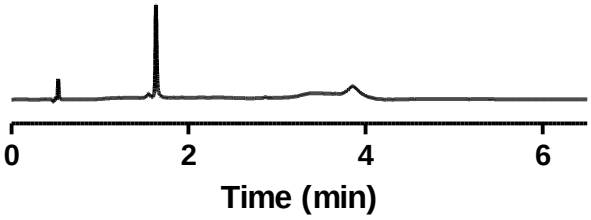
(Pro-Pro-Gly)$_7$



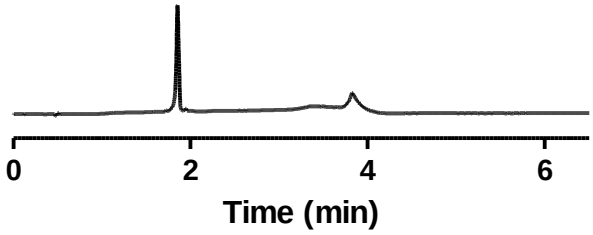
(Pro-Hyp-Gly)$_7$



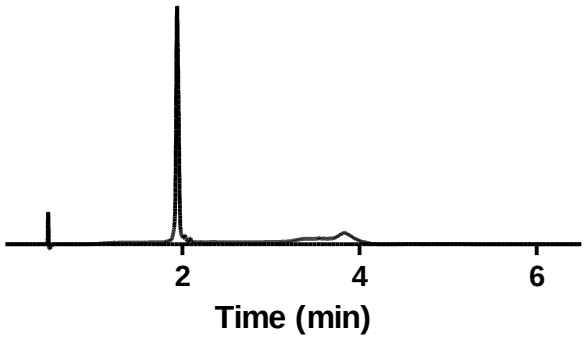
(Pro-Flp-Gly)$_7$



D-CMP-fluorescein



L-CMP-fluorescein



References

(1) Coelho, N. M.; McCulloch, C. A. Contribution of collagen adhesion receptors to tissue fibrosis. *Cell Tissue Res* **2016**, *365*, (3), 521-538.

(2) Manon-Jensen, T.; Kjeld, N. G.; Karsdal, M. A. Collagen-mediated hemostasis. *J Thromb Haemost* **2016**, *14*, (3), 438-448.

(3) Hollander, A. P.; Heathfield, T. F.; Webber, C.; Iwata, Y.; Bourne, R.; Rorabeck, C.; Poole, A. R. Increased damage to type II collagen in osteoarthritic articular cartilage detected by a new immunoassay. *J Clin Invest* **1994**, *93*, (4), 1722-1732.

(4) Conklin, M. W.; Eickhoff, J. C.; Riching, K. M.; Pehlke, C. A.; Eliceiri, K. W.; Provenzano, P. P.; Friedl, A.; Keely, P. J. Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am J Pathol* **2011**, *178*, (3), 1221-1232.

(5) Provenzano, P. P.; Inman, D. R.; Eliceiri, K. W.; Knittel, J. G.; Yan, L.; Rueden, C. T.; White, J. G.; Keely, P. J. Collagen density promotes mammary tumor initiation and progression. *BMC Med* **2008**, *6*, 11.

(6) Provenzano, P. P.; Eliceiri, K. W.; Campbell, J. M.; Inman, D. R.; White, J. G.; Keely, P. J. Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med* **2006**, *4*, (1), 38.

(7) Garvican, E. R.; Vaughan-Thomas, A.; Innes, J. F.; Clegg, P. D. Biomarkers of cartilage turnover. Part 1: Markers of collagen degradation and synthesis. *Vet J* **2010**, *185*, (1), 36-42.

(8) Theocharis, A. D.; Skandalis, S. S.; Gialeli, C.; Karamanos, N. K. Extracellular matrix structure. *Adv Drug Deliv Rev* **2016**, *97*, 4-27.

(9) Adzhubei, A. A.; Sternberg, M. J.; Makarov, A. A. Polyproline-II helix in proteins: Structure and function. *J Mol Biol* **2013**, *425*, (12), 2100-2132.

(10) Ramshaw, J. A. M.; Shah, N. K.; Brodsky, B. Gly-X-Y tripeptide frequencies in collagen: A context for host-guest triple-helical peptides. *J Struct Biol* **1998**, *122*, 86-91.

(11) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* **1994**, *266*, (5182), 75-81.

(12) Bella, J.; Berman, H. M. Crystallographic evidence for C alpha-H...O=C hydrogen bonds in a collagen triple helix. *J Mol Biol* **1996**, *264*, (4), 734-742.

(13) Shoulders, M. D.; Raines, R. T. Collagen structure and stability. *Annu Rev Biochem* **2009**, *78*, 929-958.

(14) Hollingsworth, S. A.; Karplus, P. A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol Concepts* **2010**, *1*, (3-4), 271-283.

(15) Berg, R.; Prockop, D. The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabiliz. *Biochem Biophys Res Commun* **1973**, *52*, (1), 115-120.

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

(17) Suzuki, E.; Fraser, R. D. B.; MacRae, T. P. Role of hydroxyproline in the stabilization of the collagen molecule via water molecules. *Int J Biol Macromol* **1980**, *2*, (1), 54-56.

(18) Bella, J.; Brodsky, B.; Berman, H. M. Hydration structure of a collagen peptide. *Structure* **1995**, *3*, (9), 893-906.

(19) Kramer, R. Z.; Berman, H. M. Patterns of hydration in crystalline collagen peptides. *J Biomol Struct Dyn* **1998**, *16*, (2), 367-380.
(20) Ramachandran, G. N.; Bansal, M. Stereochemistry of the pyrrolidine rings in the collagen structure. *Curr Sci* **1976**, *45*, (18), 647-649.
(21) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. Crystal structure of the collagen triple helix model $[(\text{Pro-Pro-Gly})_{10}]_3$. *Protein Sci* **2002**, *11*, (2), 262-270.
(22) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. Structural bases of collagen stabilization induced by proline hydroxylation. *Biopolymers* **2001**, *58*, (5), 459-464.
(23) Holmgren, S.; Taylor, K.; Bretscher, L.; Raines, R. T. Code for collagen's stability deciphered. *Nature* **1998**, *392*, 666-667.
(24) Holmgren, S.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. A hyperstable collagen mimic. *Chem Biol* **1999**, *6*, (2), 67-70.
(25) Than, M. E.; Henrich, S.; Huber, R.; Ries, A.; Mann, K.; Kuhn, K.; Timpl, R.; Bourenkov, G. P.; Bartunik, H. D.; Bode, W. The 1.9-A crystal structure of the noncollagenous (NC1) domain of human placenta collagen IV shows stabilization via a novel type of covalent Met-Lys cross-link. *Proc Natl Acad Sci* **2002**, *99*, (10), 6607-6612.
(26) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **2000**, *29*, 14960-14967.
(27) Yang, W.; Chan, V. C.; Kirkpatrick, A.; Ramshaw, J. A. M.; Brodsky, B. Gly-Pro-Arg confers stability similar to Gly-Pro-Hyp in the collagen triple-helix of host-guest peptides. *J Biol Chem* **1997**, *272*, (46), 28837-28840.
(28) Shah, N. K.; Ramshaw, J. A. M.; Kirkpatrick, A.; Shah, C.; Brodsky, B. A host-guest set of triple-helical peptides: Stability of Gly-X-Y triplets containing common nonpolar residues. *Biochemistry* **1996**, *35*, 10262-10268.
(29) Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. Prediction of collagen stability from amino acid sequence. *J Biol Chem* **2005**, *280*, (19), 19343-19349.
(30) Bretscher, L.; Jenkins, C. L.; Taylor, K. M.; Raines, R. T. Conformational stability of collagen relies on a stereoelectronic effect. *J Am Chem Soc* **2001**, *123*, 777-778.
(31) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on polyl ring conformations. *J Am Chem Soc* **2002**, *124*, 2497-2505.
(32) Hodges, J. A.; Raines, R. T. Stereoelectronic effects on collagen stability: The dichotomy of 4-fluoroproline diastereomers. *J Am Chem Soc* **2003**, *125*, 9263-9264.
(33) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. Characterization of collagen model peptides containing 4-fluoroproline; $(4(S)\text{-fluoroproline-Pro-Gly})_{10}$ forms a triple helix, but $(4(R)\text{-fluoroproline-Pro-Gly})_{10}$ does not. *J Am Chem Soc* **2003**, *125*, 9922-9923.
(34) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. Reciprocity of steric and stereoelectronic effects in the collagen triple helix. *J Am Chem Soc* **2006**, *128*, 8112-8113.
(35) Zhang, Y.; Malamakal, R. M.; Chenoweth, D. M. Aza-glycine induces collagen hyperstability. *J Am Chem Soc* **2015**, *137*, (39), 12422-12425.
(36) Zhang, Y.; Herling, M.; Chenoweth, D. M. General solution for stabilizing triple helical collagen. *J Am Chem Soc* **2016**, *138*, (31), 9751-9754.
(37) Kasznel, A. J.; Zhang, Y.; Hai, Y.; Chenoweth, D. M. Structural basis for aza-glycine stabilization of collagen. *J Am Chem Soc* **2017**, *139*, (28), 9427-9430.

(38) Heino, J.; Huhtala, M.; Kapyla, J.; Johnson, M. S. Evolution of collagen-based adhesion systems. *Int J Biochem Cell Biol* **2009**, *41*, (2), 341-348.
(39) Exposito, J. Y.; Valcourt, U.; Cluzel, C.; Lethias, C. The fibrillar collagen family. *Int J Mol Sci* **2010**, *11*, (2), 407-426.
(40) Lukomski, S.; Bachert, B. A.; Squeglia, F.; Berisio, R. Collagen-like proteins of pathogenic *streptococci*. *Mol Microbiol* **2017**, *103*, (6), 919-930.
(41) Rasmussen, M.; Jacobsson, M.; Björck, L. Genome-based identification and analysis of collagen-related structural motifs in bacterial and viral proteins. *J Biol Chem* **2003**, *278*, (34), 32313-32316.
(42) Doxey, A. C.; McConkey, B. J. Prediction of molecular mimicry candidates in human pathogenic bacteria. *Virulence* **2013**, *4*, (6), 453-466.
(43) Berisio, R.; Vitagliano, L. Polyproline and triple helix motifs in host-pathogen recognition. *Curr Protein Pept Sci* **2012**, *13*, (8), 855-865.
(44) Ghosh, N.; McKillop, T. J.; Jowitt, T. A.; Howard, M.; Davies, H.; Holmes, D. F.; Roberts, I. S.; Bella, J. Collagen-like proteins in pathogenic *E. coli* strains. *PLoS One* **2012**, *7*, (6), e37872.
(45) Lukomski, S.; Nakashima, K.; Abdi, I.; Cipriano, V. J.; Ireland, R. M.; Reid, S. D.; Adams, G. G.; Musser, J. M. Identification and characterization of the scl gene encoding a group A *Streptococcus* extracellular protein virulence factor with similarity to human collagen. *Infect Immun* **2000**, *68*, (12), 6542–6553.
(46) Mohs, A.; Silva, T.; Yoshida, T.; Amin, R.; Lukomski, S.; Inouye, M.; Brodsky, B. Mechanism of stabilization of a bacterial collagen triple helix in the absence of hydroxyproline. *J Biol Chem* **2007**, *282*, (41), 29757-29765.
(47) Xu, C.; Yu, Z.; Inouye, M.; Brodsky, B.; Mirochnitchenko, O. Expanding the family of collagen proteins: recombinant bacterial collagens of varying composition form triple-helices of similar stability. *Biomacromolecules* **2010**, *11*, (2), 348–356.
(48) Seo, N.; Russell, B. H.; Rivera, J. J.; Liang, X.; Xu, X.; Afshar-Kharghan, V.; Hook, M. An engineered α1 integrin-binding collagenous sequence. *J Biol Chem* **2010**, *285*, (40), 31046-31054.
(49) An, B.; DesRochers, T. M.; Qin, G.; Xia, X.; Thiagarajan, G.; Brodsky, B.; Kaplan, D. L. The influence of specific binding of collagen-silk chimeras to silk biomaterials on hMSC behavior. *Biomaterials* **2013**, *34*, (2), 402-412.
(50) An, B.; Abbonante, V.; Yigit, S.; Balduini, A.; Kaplan, D. L.; Brodsky, B. Definition of the native and denatured type II collagen binding site for fibronectin using a recombinant collagen system. *J Biol Chem* **2014**, *289*, (8), 4941-4951.
(51) Peng, Y. Y.; Stoichevska, V.; Schacht, K.; Werkmeister, J. A.; Ramshaw, J. A. Engineering multiple biological functional motifs into a blank collagen-like protein template from *Streptococcus pyogenes*. *J Biomed Mater Res A* **2014**, *102*, (7), 2189-2196.
(52) Yu, Z.; Visse, R.; Inouye, M.; Nagase, H.; Brodsky, B. Defining requirements for collagenase cleavage in collagen type III using a bacterial collagen system. *J Biol Chem* **2012**, *287*, (27), 22988-22997.
(53) Chattopadhyay, S.; Raines, R. T. Review collagen-based biomaterials for wound healing. *Biopolymers* **2014**, *101*, (8), 821-833.
(54) Han, R.; Zwiefka, A.; Caswell, C. C.; Xu, Y.; Keene, D. R.; Lukomska, E.; Zhao, Z.; Hook, M.; Lukomski, S. Assessment of prokaryotic collagen-like sequences derived from

Streptococcal Scl1 and Scl2 proteins as a source of recombinant GXY polymers. *Appl Microbiol Biotechnol* **2006**, *72*, (1), 109-115.

(55) Ramshaw, J. A.; Werkmeister, J. A.; Dumsday, G. J. Bioengineered collagens: emerging directions for biomedical materials. *Bioengineered* **2014**, *5*, (4), 227-233.

(56) Werkmeister, J. A.; Ramshaw, J. A. Recombinant protein scaffolds for tissue engineering. *Biomed Mater* **2012**, *7*, (1), 012002.

(57) Sylvestre, P.; Couture-Tosi, E.; Mock, M. A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. *Mol Microbiol* **2002**, *45*, (1), 169-178.

(58) Duncan, C.; Prashar, A.; So, J.; Tang, P.; Low, D. E.; Terebiznik, M.; Guyard, C. Lcl of *Legionella pneumophila* is an immunogenic GAG binding adhesin that promotes interactions with lung epithelial cells and plays a crucial role in biofilm formation. *Infect Immun* **2011**, *79*, (6), 2168-2181.

(59) Bachert, B. A.; Choi, S. J.; Snyder, A. K.; Rio, R. V.; Durney, B. C.; Holland, L. A.; Amemiya, K.; Welkos, S. L.; Bozue, J. A.; Cote, C. K.; Berisio, R.; Lukomski, S. A unique set of the *Burkholderia* collagen-like proteins provides insight into pathogenesis, genome evolution and niche adaptation, and infection detection. *PLoS One* **2015**, *10*, (9), e0137578.

(60) Lukomski, S.; Nakashima, K.; Abdi, I.; Cipriano, V. J.; Shelvin, B. J.; Graviss, E. A.; Musser, J. M. Identification and characterization of a second extracellular collagen-like protein made by group A *Streptococcus*: control of production at the level of translation. *Infect Immun* **2001**, *69*, (3), 1729-1738.

(61) Caswell, C. C.; Oliver-Kozup, H.; Han, R.; Lukomska, E.; Lukomski, S. Scl1, the multifunctional adhesin of group A *Streptococcus*, selectively binds cellular fibronectin and laminin, and mediates pathogen internalization by human cells. *FEMS Microbiol Lett* **2010**, *303*, (1), 61-68.

(62) Humtsoe, J. O.; Kim, J. K.; Xu, Y.; Keene, D. R.; Hook, M.; Lukomski, S.; Wary, K. K. A *streptococcal* collagen-like protein interacts with the $\alpha_2\beta_1$ integrin and induces intracellular signaling. *J Biol Chem* **2005**, *280*, (14), 13848-13857.

(63) Caswell, C. C.; Lukomska, E.; Seo, N. S.; Hook, M.; Lukomski, S. Scl1-dependent internalization of group A *Streptococcus* via direct interactions with the $\alpha_2\beta_1$ integrin enhances pathogen survival and re-emergence. *Mol Microbiol* **2007**, *64*, (5), 1319-1331.

(64) Caswell, C. C.; Barczyk, M.; Keene, D. R.; Lukomska, E.; Gullberg, D. E.; Lukomski, S. Identification of the first prokaryotic collagen sequence motif that mediates binding to human collagen receptors, integrins $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$. *J Biol Chem* **2008**, *283*, (52), 36168-36175.

(65) Chen, S.-M.; Tsai, Y.-S.; Wu, C.-M.; Liao, S.-K.; Wu, L.-C.; Chang, C.-S.; Liu, Y.-H.; Tsai, P.-J. *Streptococcal* collagen-like surface protein 1 promotes adhesion to the respiratory epithelial cell. *BMC Microbiol* **2010**, *10*, (1), 320.

(66) Han, R.; Caswell, C. C.; Lukomska, E.; Keene, D. R.; Pawlowski, M.; Bujnicki, J. M.; Kim, J. K.; Lukomski, S. Binding of the low-density lipoprotein by *Streptococcal* collagen-like protein Scl1 of *Streptococcus pyogenes*. *Mol Microbiol* **2006**, *61*, (2), 351-367.

(67) Caswell, C. C.; Han, R.; Hovis, K. M.; Ciborowski, P.; Keene, D. R.; Marconi, R. T.; Lukomski, S. The Scl1 protein of M6-type group A *Streptococcus* binds the human

complement regulatory protein, factor H, and inhibits the alternative pathway of complement. *Mol Microbiol* **2008**, *67*, (3), 584-596.

(68) Reuter, M.; Caswell, C. C.; Lukomski, S.; Zipfel, P. F. Binding of the human complement regulators CFHR1 and factor H by *streptococcal* collagen-like protein 1 (Scl1) via their conserved C termini allows control of the complement cascade at multiple levels. *J Biol Chem* **2010**, *285*, (49), 38473-38485.

(69) Pahlman, L. I.; Marx, P. F.; Morgelin, M.; Lukomski, S.; Meijers, J. C.; Herwald, H. Thrombin-activatable fibrinolysis inhibitor binds to *Streptococcus pyogenes* by interacting with collagen-like proteins A and B. *J Biol Chem* **2007**, *282*, (34), 24873-24881.

(70) Döhrmann, S.; Anik, S.; Olson, J.; Anderson, E. L.; Etesami, N.; No, H.; Snipper, J.; Nizet, V.; Okumura, C. Y. M. Role for *streptococcal* collagen-like protein 1 in M1T1 group A *Streptococcus* resistance to neutrophil extracellular traps. *Infect Immun* **2014**, *82*, (10), 4011-4020.

(71) LaRock, C. N.; Nizet, V. Cationic antimicrobial peptide resistance mechanisms of *streptococcal* pathogens. *Biochim Biophys Acta* **2015**, *1848*, (11 Pt B), 3047-3054.

(72) Dinkla, K.; Rohde, M.; Jansen, W. T. M.; Carapetis, J. R.; Chhatwal, G. S.; Talay, S. R. *Streptococcus pyogenes* recruits collagen via surface-bound fibronectin: a novel colonization and immune evasion mechanism. *Mol Microbiol* **2003**, *47*, (3), 861–869.

(73) Xu, Y.; Keene, D. R.; Bujnicki, J. M.; Hook, M.; Lukomski, S. *Streptococcal* Scl1 and Scl2 proteins form collagen-like triple helices. *J Biol Chem* **2002**, *277*, (30), 27312-27318.

(74) Lannergard, J.; Flock, M.; Johansson, S.; Flock, J. I.; Guss, B. Studies of fibronectin-binding proteins of *Streptococcus equi*. *Infect Immun* **2005**, *73*, (11), 7243-7251.

(75) Tiouajni, M.; Durand, D.; Blondeau, K.; Graille, M.; Urvoas, A.; Valerio-Lepiniec, M.; Guellouz, A.; Aumont-Nicaise, M.; Minard, P.; van Tilbeurgh, H. Structural and functional analysis of the fibronectin-binding protein FNE from *Streptococcus equi* spp. *equi*. *FEBS J* **2014**, *281*, (24), 5513-5531.

(76) Enjel, J.; Bachinger, H. P. Structure, stability and folding of the collagen triple helix. *Top Curr Chem* **2005**, *247*, 7-33.

(77) Legay, C. Why so many forms of acetylcholinesterase? *Microsc Res Tech* **2000**, *49*, (1), 56-72.

(78) Okamoto, Y.; Arita, Y.; Nishida, M.; Muraguchi, M.; Ouchi, N.; Takahashi, M.; Igura, T.; Inui, Y.; Kihara, S.; Nakamura, T.; Yamashita, S.; Miyagawa, J.; Funahashi, T.; Matsuzawa, Y. An adipocyte-derived plasma protein, adiponectin, adheres to injured vascular walls. *Horm Metab Res* **2000**, *32*, (2), 47-50.

(79) Gupta, G.; Surolia, A. Collectins: sentinels of innate immunity. *Bioessays* **2007**, *29*, (5), 452-464.

(80) Kouser, L.; Madhukaran, S. P.; Shastri, A.; Saraon, A.; Ferluga, J.; Al-Mozaini, M.; Kishore, U. Emerging and novel functions of complement protein C1q. *Front Immunol* **2015**, *6*, 317.

(81) Bowdish, D. M.; Gordon, S. Conserved domains of the class A scavenger receptors: Evolution and function. *Immunol Rev* **2009**, *227*, (1), 19-31.

(82) Ricard-Blum, S. The collagen family. *Cold Spring Harb Perspect Biol* **2011**, *3*, (1), a004978.

(83) Cox, G.; Kable, E. Secound-harmonic imaging of collagen. *Methods Mol Biol* **2006**, *319*, 15-35.
(84) Lacomb, R.; Nadiarnykh, O.; Campagnola, P. J. Quantitative second harmonic generation imaging of the diseased state osteogenesis imperfecta: experiment and simulation. *Biophys J* **2008**, *94*, (11), 4504-4514.
(85) Zipfel, W. R.; Williams, R. M.; Christie, R.; Nikitin, A. Y.; Hyman, B. T.; Webb, W. W. Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proc Natl Acad Sci* **2003**, *100*, (12), 7075-7080.
(86) Pavithra, V.; Sowmya, S. V.; Rao, R. S.; Patil, S.; Augustine, D.; Haragannavar, V. C.; Nambiar, S. Tumor-associated collagen signatures: An insight. *WJD* **2017**, *8*, (3), 224-230.
(87) Sun, W.; Chang, S.; Tai, D. C.; Tan, N.; Xiao, G.; Tang, H.; Yu, H. Nonlinear optical microscopy: use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. *J Biomed Opt* **2008**, *13*, (6), 064010.
(88) Lin, S.-J.; Wu, R.-J.; Tan, H.-Y.; Lo, W.; Lin, W.-C.; Young, T.-H.; Hsu, C.-J.; Chen, J.-S.; Jee, S.-H.; Dong, C.-Y. Evaluating cutaneous photoaging by use of multiphoton fluorescence and second-harmonic generation microscopy. *Optics Lett* **2005**, *30*, (17), 2275-2277.
(89) Mostaco-Guidolin, L.; Rosin, N. L.; Hackett, T. L. Imaging collagen in scar tissue: developments in second harmonic generation microscopy for biomedical applications. *Int J Mol Sci* **2017**, *18*, (8), 1772.
(90) Heino, J. The collagen family members as cell adhesion proteins. *BioEssays* **2007**, *29*, (10), 1001-1010.
(91) Svensson, L.; Oldberg, Å.; Heinegård, D. Collagen binding proteins. *Osteoarthr Cartil* **2001**, *9*, S23-S28.
(92) Tuckwell, D.; Humphries, M. Integrin-collagen binding. *Semin Cell Dev Biol* **1996**, *7*, (5), 649-657.
(93) Kitajima, T.; Terai, H.; Ito, Y. A fusion protein of hepatocyte growth factor for immobilization to collagen. *Biomaterials* **2007**, *28*, (11), 1989-1997.
(94) Ishikawa, T.; Eguchi, M.; Wada, M.; Iwami, Y.; Tono, K.; Iwaguro, H.; Masuda, H.; Tamaki, T.; Asahara, T. Establishment of a functionally active collagen-binding vascular endothelial growth factor fusion protein in situ. *Arterioscler Thromb Vasc Biol* **2006**, *26*, (9), 1998-2004.
(95) Ishikawa, T.; Terai, H.; Kitajima, T. Production of a biologically active epidermal growth factor fusion protein with high collagen affinity. *J Biochem* **2001**, *129*, (4), 627-633.
(96) Ishikawa, T.; Terai, H.; Yamanoto, T.; Harada, K.; Kitajima, T. Delivery of a growth factor fusion protein having collagen-binding activity to wound tissues. *Artif Organs* **2003**, *27*, (2), 147-154.
(97) Okiyama, N.; Kitajima, T.; Ito, Y.; Yokozeki, H.; Miyasaka, N.; Kohsaka, H. Addition of the collagen binding domain of fibronectin potentiates the biochemical availability of hepatocyte growth factor for cutaneous wound healing. *J Dermatol Sci* **2011**, *61*, (3), 215-217.
(98) Addi, C.; Murschel, F.; De Crescenzo, G. Design and use of chimeric proteins containing a collagen-binding domain for wound healing and bone regeneration. *Tissue Eng Part B Rev* **2017**, *23*, (2), 163-182.

(99) Ribet, D.; Cossart, P. How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes Infect* **2015**, *17*, (3), 173-183.
(100) Ross, C. L.; Liang, X.; Liu, Q.; Murray, B. E.; Hook, M.; Ganesh, V. K. Targeted protein engineering provides insights into binding mechanism and affinities of bacterial collagen adhesins. *J Biol Chem* **2012**, *287*, (41), 34856-34865.
(101) Patti, J. M.; Boles, J. O.; Hook, M. Identification and biochemical characterization of the ligand binding domain of the collagen adhesin from *Staphylococcus aureus*. *Biochemistry* **1993**, *32*, (42), 11428-11435.
(102) Rich, R. L.; Kreikemeyer, B.; Owens, R. T.; LaBrenz, S.; Narayana, S. V. L.; Weinstock, G. M.; Murray, B. E.; Hook, M. Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*. *J Biol Chem* **1999**, *274*, (38), 26939–26945.
(103) Nallapareddy, S. R.; Weinstock, G. M.; Murray, B. E. Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Mol Microbiol* **2003**, *47*, (6), 1733-1747.
(104) Lannergard, J.; Frykberg, L.; Guss, B. CNE, a collagen-binding protein of *Streptococcus equi*. *FEMS Microbiol Lett* **2003**, *222*, (1), 69-74.
(105) Sato, Y.; Okamoto, K.; Kagami, A.; Yamanoto, Y.; Igarashi, T.; Kizaki, H. *Streptococcus mutans* strains harboring collagen-binding adhesin. *J Dent Res* **2004**, *83*, (7).
(106) Kreikemeyer, B.; Nakata, M.; Oehmcke, S.; Gschwendtner, C.; Normann, J.; Podbielski, A. *Streptococcus pyogenes* collagen type I-binding Cpa surface protein. Expression profile, binding characteristics, biological functions, and potential clinical impact. *J Biol Chem* **2005**, *280*, (39), 33228-33239.
(107) Shimoji, Y.; Ogawa, Y.; Osaki, M.; Kabeya, H.; Maruyama, S.; Mikami, T.; Sekizaki, T. Adhesive surface proteins of *Erysipelothrix rhusiopathiae* bind to polystyrene, fibronectin, and type I and IV collagens. *J Bacteriol* **2003**, *185*, (9), 2739-2748.
(108) Xu, Y.; Liang, X.; Chen, Y.; Koehler, T. M.; Hook, M. Identification and biochemical characterization of two novel collagen binding MSCRAMMs of *Bacillus anthracis*. *J Biol Chem* **2004**, *279*, (50), 51760-51768.
(109) Zong, Y.; Xu, Y.; Liang, X.; Keene, D. R.; Hook, A.; Gurusiddappa, S.; Hook, M.; Narayana, S. V. A 'collagen hug' model for *Staphylococcus aures* CNA binding to collagen. *EMBO J* **2005**, *24*, (4224-4236).
(110) Boerboom, R. A.; Krahn, K. N.; Megens, R. T.; van Zandvoort, M. A.; Merkx, M.; Bouten, C. V. High resolution imaging of collagen organisation and synthesis using a versatile collagen specific probe. *J Struct Biol* **2007**, *159*, (3), 392-399.
(111) Krahn, K. N.; Bouten, C. V.; van Tuijl, S.; van Zandvoort, M. A.; Merkx, M. Fluorescently labeled collagen binding proteins allow specific visualization of collagen in tissues and live cell culture. *Anal Biochem* **2006**, *350*, (2), 177-185.
(112) Chen, J.; Lee, S. K.; Abd-Elgaliel, W. R.; Liang, L.; Galende, E. Y.; Hajjar, R. J.; Tung, C. H. Assessment of cardiovascular fibrosis using novel fluorescent probes. *PLoS One* **2011**, *6*, (4), e19097.
(113) Megens, R. T. A.; Oude Egbrink, M. G. A.; Cleutjens, J. P. M.; Kuijpers, M. J. E.; Schiffers, P. H. M.; Merkx, M.; Slaaf, D. W.; Van Zandvoort, M. A. M. J. Imaging collagen in intact viable healthy and atherosclerotic arteries using fluorescently labeled CNA35 and two-photon laser scanning microscopy. *Mol Imag* **2007**, *6*, (4), 247-260.

(114) Reulen, S. W. A.; Dankers, P. Y. W.; Bomans, P. H. H.; Meijer, E. W.; Merkx, M. Collagen targeting using protein-functionalized micelles: The strength of multiple weak interactions. *J Am Chem Soc* **2009**, *131*, (21), 7304-7312.
(115) Madani, A.; Garakani, K.; Mofrad, M. R. K. Molecular mechanics of *Staphylococcus aureus* adhesin, CNA, and the inhibition of bacterial adhesion by stretching collagen. *PLoS One* **2017**, *12*, (6), e0179601.
(116) Caravan, P.; Das, B.; Dumas, S.; Epstein, F. H.; Helm, P. A.; Jacques, V.; Koerner, S.; Kolodziej, A.; Shen, L.; Sun, W. C.; Zhang, Z. Collagen-targeted MRI contrast agent for molecular imaging of fibrosis. *Angew Chem Int Ed Engl* **2007**, *46*, (43), 8171-8173.
(117) Helm, P. A.; Caravan, P.; French, B. A.; Jacques, V.; Shen, L. H.; Xu, Y.; Beyers, R. J.; Kramer, C. M.; Epstein, F. H. Postinfarction myocardial scarring in mice: Molecular MR imaging with use of a collagen-targeting contrast agent. *Radiology* **2008**, *247*, (3), 788-796.
(118) Federico, S.; Pierce, B. F.; Piluso, S.; Wischke, C.; Lendlein, A.; Neffe, A. T. Design of decorin-based peptides that bind to collagen I and their potential as adhesion moieties in biomaterials. *Angew Chem Int Ed Engl* **2015**, *54*, (37), 10980-10984.
(119) Paderi, J. E.; Panitch, A. Design of a synthetic collagen-binding peptidoglycan that modulates collagen fibrillogenesis. *Biomacromolecules* **2008**, *9*, (9), 2562-2566.
(120) Stuart, K.; Paderi, J.; Snyder, P. W.; Freeman, L.; Panitch, A. Collagen-binding peptidoglycans inhibit MMP mediated collagen degradation and reduce dermal scarring. *PLoS One* **2011**, *6*, (7), e22139.
(121) Helms, B. A.; Reulen, S. W. A.; Nijhuis, S.; de Graaf-Heuvelmans, P. T. H. M.; Merkx, M.; Meijer, E. W. High-affinity peptide-based collagen targeting using synthetic phage mimics: From phage display to dendrimer display. *J Am Chem Soc* **2009**, *131*, (33), 11683-11685.
(122) Timpi, R. Antibodies to collagen and procollagens. *Methods Enzymol* **1982**, *82*, (472-498).
(123) Glattauer, V.; Werkmeister, J. A.; Kirkpatrick, A.; Ramshaw, J. A. M. Identification of the epitope for a monoclonal antibody that blocks platelet aggregation induced by type III collagen. *Biochem J* **1997**, *323*, 45-49.
(124) Hashimoto, N.; Jin, H.; Liu, T.; Chensue, S. W.; Phan, S. H. Bone marrow–derived progenitor cells in pulmonary fibrosis. *J Clin Invest* **2004**, *113*, (2), 243-252.
(125) Kobielarz, M.; Szotek, S.; Glowacki, M.; Dawidowicz, J.; Pezowicz, C. Qualitative and quantitative assessment of collagen and elastin in annulus fibrosus of the physiologic and scoliotic intervertebral discs. *J Mech Behav Biomed Mater* **2016**, *62*, 45-56.
(126) Nemirovskiy, O. V.; Dufield, D. R.; Sunyer, T.; Aggarwal, P.; Welsch, D. J.; Mathews, W. R. Discovery and development of a type II collagen neoepitope (TIINE) biomarker for matrix metalloproteinase activity: From in vitro to in vivo. *Anal Biochem* **2007**, *361*, (1), 93-101.
(127) Breurken, M.; Lempens, E. H.; Merkx, M. Protease-activatable collagen targeting based on protein cyclization. *ChemBioChem* **2010**, *11*, (12), 1665-1668.
(128) Breurken, M.; Lempens, E. H.; Meijer, E. W.; Merkx, M. Semi-synthesis of a protease-activatable collagen targeting probe. *Chem Commun* **2011**, *47*, (28), 7998-8000.
(129) Xu, J.; Rodriguez, D.; Kim, J. J.; Brooks, P. C. Generation of monoclonal antibodies to cryptic collagen sites by using subtractive immunization *Hybridoma* **2004**, *19*, (5), 375-385.

(130) Freimark, B.; Clark, D.; Pernasetti, F.; Nickel, J.; Myszka, D.; Baeuerle, P. A.; Van Epps, D. Targeting of humanized antibody D93 to sites of angiogenesis and tumor growth by binding to multiple epitopes on denatured collagens. *Mol Immunol* **2007**, *44*, (15), 3741-3750.
(131) Pernasetti, F.; Nickel, J.; Clark, D.; Baeuerle, P. A.; Van Epps, D.; Freimark, B. Novel anti-denatured collagen humanized antibody D93 inhibits angiogenesis and tumor growth: An extracellular matrix-based therapeutic approach. *Int J Oncol* **2006**, *29*, (6), 1371-1379.
(132) Mueller, J.; Gaertner, F. C.; Blechert, B.; Janssen, K. P.; Essler, M. Targeting of tumor blood vessels: a phage-displayed tumor-homing peptide specifically binds to matrix metalloproteinase-2-processed collagen IV and blocks angiogenesis in vivo. *Mol Cancer Res* **2009**, *7*, (7), 1078-1085.
(133) Paterlini, M. G.; Nemethy, G.; Scheraga, H. A. The energy of formation of internal loops in triple-helical collagen polypeptides. *Biopolymers* **1995**, *35*, (6), 607-619.
(134) Long, C. G.; Thomas, M.; Brodsky, B. Atypical Gly-X-Y sequences surround interruptions in the repeating tripeptide pattern of basement membrane collagen. *Biopolymers* **1995**, *35*, (6), 621-628.
(135) Leikina, E.; Mertts, M. V.; Kuznetsova, N.; Leikin, S. Type I collagen is thermally unstable at body temperature. *Proc Natl Acad Sci* **2002**, *99*, (3), 1314-1318.
(136) Van Der Rest, M.; Garrone, R. Collagen family of proteins. *FASEB J* **1991**, *5*, (13), 2814-2823.
(137) Paul, D. W.; Ghassemi, P.; Ramella-Roman, J. C.; Prindeze, N. J.; Moffatt, L. T.; Alkhalil, A.; Shupp, J. W. Noninvasive imaging technologies for cutaneous wound assessment: A review. *Wound Repair Regen* **2015**, *23*, (2), 149-162.
(138) Miles, C. A.; Bailey, A. J. Thermally labile domains in the collagen molecule. *Micron* **2001**, *32*, (3), 325-332.
(139) Mo, X.; An, Y.; Yun, C. S.; Yu, S. M. Nanoparticle-assisted visualization of binding interactions between collagen mimetic peptide and collagen fibers. *Angew Chem Int Ed* **2006**, *45*, (14), 2267-2270.
(140) San, B. H.; Li, Y.; Tarbet, E. B.; Yu, S. M. Nanoparticle assembly and gelatin binding mediated by triple helical collagen mimetic peptide. *ACS Appl Mater Interfaces* **2016**, *8*, (31), 19907-19915.
(141) Santos, J. L.; Li, Y.; Culver, H. R.; Yu, M. S.; Herrera-Alonso, M. Conducting polymer nanoparticles decorated with collagen mimetic peptides for collagen targeting. *Chem Commun* **2014**, *50*, (95), 15045-15048.
(142) Wang, A. Y.; Foss, C. A.; Leong, S.; Mo, X.; Pomper, M. G.; Yu, S. M. Spatio-temporal modification of collagen scaffolds mediated by triple helical propensity. *Biomacromolecules* **2008**, *9*, (7), 1755-1763.
(143) Hwang, J.; San, B. H.; Turner, N. J.; White, L. J.; Faulk, D. M.; Badylak, S. F.; Li, Y.; Yu, S. M. Molecular assessment of collagen denaturation in decellularized tissues using a collagen hybridizing peptide. *Acta Biomater* **2017**, *53*, 268-278.
(144) Zitnay, J. L.; Li, Y.; Qin, Z.; San, B. H.; Depalle, B.; Reese, S. P.; Buehler, M. J.; Yu, S. M.; Weiss, J. A. Molecular level detection and localization of mechanical damage in collagen enabled by collagen hybridizing peptides. *Nat Commun* **2017**, *8*, 14913.
(145) Wang, A. Y.; Leong, S.; Liang, Y.-C.; Haung, R. C. C.; Chen, C. S.; Yu, S. M. Immobilization of growth factors on collagen scaffolds mediated by polyanionic collagen

mimetic peptides and its effect on endothelial cell morphogenesis. *Biomacromolecules* **2008**, *9*, (10), 2929-2936.
(146) Stahl, P. J.; Chan, T. R.; Shen, Y. I.; Sun, G.; Gerecht, S.; Yu, S. M. Capillary network-like organization of endothelial cells in PEGDA scaffolds encoded with angiogenic signals via triple helical hybridization. *Adv Funct Mater* **2014**, *24*, (21), 3213-3225.
(147) Chan, T. R.; Stahl, P. J.; Li, Y.; Yu, S. M. Collagen-gelatin mixtures as wound model, and substrates for VEGF-mimetic peptide binding and endothelial cell activation. *Acta Biomater* **2015**, *15*, 164-172.
(148) Chan, T. R.; Stahl, P. J.; Yu, S. M. Matrix-bound VEGF mimetic peptides: design and endothelial cell activation in collagen scaffolds. *Adv Funct Mater* **2011**, *21*, (22), 4252-4262.
(149) Li, Y.; Foss, C. A.; Summerfield, D. D.; Doyle, J. J.; Torok, C. M.; Dietz, H. C.; Pomper, M. G.; Yu, S. M. Targeting collagen strands by photo-triggered triple-helix hybridization. *Proc Natl Acad Sci* **2012**, *109*, (37), 14767-14772.
(150) Li, Y.; San, B. H.; Kessler, J. L.; Kim, J. H.; Xu, Q.; Hanes, J.; Yu, S. M. Non-covalent photo-patterning of gelatin matrices using caged collagen mimetic peptides. *Macromol Biosci* **2015**, *15*, (1), 52-62.
(151) Li, Y.; Ho, D.; Meng, H.; Chan, T. R.; An, B.; Yu, H.; Brodsky, B.; Jun, A. S.; Michael Yu, S. Direct detection of collagenous proteins by fluorescently labeled collagen mimetic peptides. *Bioconjug Chem* **2013**, *24*, (1), 9-16.
(152) Hwang, J.; Huang, Y.; Burwell, T. J.; Peterson, N. C.; Connor, J.; Weiss, S. J.; Yu, S. M.; Li, Y. In situ imaging of tissue remodeling with collagen hybridizing peptides. *ACS Nano* **2017**, *11*, (10), 9825-9835.
(153) Bennink, L. L.; Smith, D. J.; Foss, C. A.; Pomper, M. G.; Li, Y.; Yu, S. M. High serum stability of collagen hybridizing peptides and their fluorophore conjugates. *Mol Pharm* **2017**, *14*, (6), 1906-1915.
(154) Hodges, J. A.; Raines, R. T. Stereoelectronic and steric effects in the collagen triple helix: toward a code for strand association. *J Am Chem Soc* **2005**, *127*, 15923-15932.
(155) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nishio, H.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. Collagen-like triple helix formation of synthetic (Pro-Pro-Gly)$_{10}$ analogues: (4(*S*)-hydroxyprolyl-4(*R*)-hydroxyprolyl-Gly)$_{10}$, (4(*R*)-hydroxyprolyl-4(*R*)-hydroxyprolyl-Gly)$_{10}$ and (4(*S*)-fluoroprolyl-4(*R*)-fluoroprolyl-Gly)$_{10}$. *J Pept Sci* **2005**, *11*, (10), 609-616.
(156) Chattopadhyay, S.; Murphy, C. J.; McAnulty, J. F.; Raines, R. T. Peptides that anneal to natural collagen in vitro and ex vivo. *Org Biomol Chem* **2012**, *10*, (30), 5892-5897.
(157) Chattopadhyay, S.; Guthrie, K. M.; Teixeira, L.; Murphy, C. J.; Dubielzig, R. R.; McAnulty, J. F.; Raines, R. T. Anchoring a cytoactive factor in a wound bed promotes healing. *J Tissue Eng Regen Med* **2014**, *10*, (12), 1012-1020.
(158) Jin, H. E.; Farr, R.; Lee, S. W. Collagen mimetic peptide engineered M13 bacteriophage for collagen targeting and imaging in cancer. *Biomaterials* **2014**, *35*, (33), 9236-9245.
(159) Egli, J.; Siebler, C.; Maryasin, B.; Erdmann, R. S.; Bergande, C.; Ochsenfeld, C.; Wennemers, H. pH-responsive aminoproline-containing collagen triple helices. *Chem Eur J* **2017**, *23*, (33), 7938-7944.
(160) Siebler, C.; Erdmann, R. S.; Wennemers, H. Switchable proline derivatives: Tuning the conformational stability of the collagen triple helix by pH changes. *Angew Chem Int Ed Engl* **2014**, *53*, (39), 10340-10344.

(161) Bahar, A. A.; Ren, D. Antimicrobial peptides. *Pharmaceuticals* **2013**, *6*, (12), 1543-1575.
(162) Colby Davie, E. A.; Mennen, S. M.; Xu, Y.; Miller, S. J. Asymmetric catalysis mediated by synthetic peptides. *Chem Rev* **2007**, *107*, (12), 5759-5812.
(163) Grove, T. Z.; Regan, L. New materials from proteins and peptides. *Curr Opin Struct Biol* **2012**, *22*, (4), 451-456.
(164) Kent, S. B. The critical role of peptide chemistry in the life sciences. *J Pept Sci* **2015**, *21*, (3), 136-138.
(165) Pepe-Mooney, B. J.; Fairman, R. Peptides as materials. *Curr Opin Struct Biol* **2009**, *19*, (4), 483-494.
(166) Pipkorn, R.; Braun, K.; Wiessler, M.; Waldeck, W.; Schrenk, H. H.; Koch, M.; Semmler, W.; Komljenovic, D. A peptide & peptide nucleic acid synthesis technology for transporter molecules and theranostics—the SPPS. *Int J Med Sci* **2014**, *11*, (7), 697-706.
(167) Lewandowski, B.; Wennemers, H. Asymmetric catalysis with short-chain peptides. *Curr Opin Chem Biol* **2014**, *22*, 40-46.
(168) Tsomaia, N. Peptide therapeutics: Targeting the undruggable space. *Eur J Med Chem* **2015**, *94*, 459-470.
(169) Fischer, E.; Fourneau, E. Ueber einige Derivate des Glykocolls. *Ber Deutsch Chem Ges* **1901**, *34*, (August), 2868-2877.
(170) Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J Am Chem Soc* **1963**, *85*, (14), 2149-2154.
(171) Solution-Phase Peptide Synthesis. Tsuda, Y.; Okada, Y. In *Amino Acids, Peptides and Proteins in Organic Chemistry*; Hughes, A. B., Ed.; WILEY-VCH Verlag & Co. KGaA: Weinheim, Germany, 2011; Vol. 3, p 203-251.
(172) Chaturvedi, N. C.; Park, W. K.; Smeby, R. R.; Bumpus, F. M. Analogs of angiotensin II. I. Solid phase synthesis. *J Med Chem* **1970**, *13*, 177-181.
(173) Stawikowski, M. F., G. B. Introduction to peptide synthesis. *Curr Protoc Protein Sci* **2002**, *26*, (18.1), 1-17.
(174) Abu-Baker, S.; Garber, P.; Hina, B.; Reed, T.; Shahrokh, G.; Al-Saghir, M.; Lorigan, G. Microwave assisted peptide synthesis as a new gold standard in solid phase peptide synthesis: Phospholamban as an example. *Open J Synthesis Theory and Appl* **2014**, *3*, (1), 1-4.
(175) Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S. High-efficiency solid phase peptide synthesis (HE-SPPS). *Org Lett* **2014**, *16*, (3), 940-943.
(176) Albericio, F. Developments in peptide and amide synthesis. *Curr Opin Chem Biol* **2004**, *8*, (3), 211-221.
(177) Palomo, J. M. Solid-phase peptide synthesis: An overview focused on the preparation of biologically relevant peptides. *RSC Adv* **2014**, *4*, (62), 32658-32672.
(178) Shelton, P. T.; Jensen, K. J. Linkers, resins, and general procedures for solid-phase peptide synthesis. *Methods Mol Biol* **2013**, *1047*, 23-41.
(179) Solid-Phase Peptide Synthesis: Historical Aspects. Marshall, G. R. In *Amino Acids, Peptides and Proteins in Organic Chemistry*; Hughes, A. B., Ed.; Wiley–VCH Verlag GmbH: Weinheim, Germany, 2011; Vol. 3, p 253-272.
(180) Difficult peptides. Machini Miranda, M. T.; Liria, C. W.; Remuzgo, C. In *Amino Acids, Peptides and Proteins in Organic Chemistry*; Hughes, A. B., Ed.; Wiley–VCH Verlag GmbH: Weinheim, Germany, 2011; Vol. 3, p 549-569.

(181) Gisin, B. F.; Merrifield, R. B.; Tosteson, D. C. Solid-phase synthesis of the cyclododecadepsipeptide valinomycin. *J Am Chem Soc* **1969**, *91*, (10), 2691-2695.
(182) Gisin, B. F.; Dhundale, A. R. Synthesis of lysine-valinomycin by solid-phase segment condensation. *International journal of peptide and protein research* **1979**, *14*, (4), 356-363.
(183) Fields, G. B. Synthesis and biological applications of collagen-model triple-helical peptides. *Org Biomol Chem* **2010**, *8*, (6), 1237-1258.
(184) Fields, G. B.; Prockop, D. J. Perspective on the synthesis and application of triple-helical collagen-model peptides. *Biopolymers* **1996**, *40*, (4), 345-357.
(185) Boulègue, C.; Musiol, H. J.; Götz, M. G.; Renner, C.; Moroder, L. Natural and artificial cystine knots for assembly of homo- and heterotrimeric collagen models. *Antioxid Redox Signal* **2008**, *10*, 113-125.
(186) Yu, S. M.; Li, Y.; Kim, D. Collagen mimetic peptides: Progress towards functional applications. *Soft Matter* **2011**, *7*, (18), 7927-7938.
(187) Luo, T.; Kiick, K. L. Collagen-like peptides and peptide-polymer conjugates in the design of assembled materials. *Eur Polym J* **2013**, *49*, (10), 2998-3009.
(188) Renner, C.; Saccà, B.; Moroder, L. Synthetic heterotrimeric collagen peptides as mimics of cell adhesion sites of the basement membrane. *Biopolymers* **2004**, *76*, 34-47.
(189) Siebler, C.; Erdmann, R. S.; Wennemers, H. From azidoproline to functionalizable collagen. *Chimia* **2013**, *67*, 891-895.
(190) Fallas, J. A.; O'Leary, L. E.; Hartgerink, J. D. Synthetic collagen mimics: Self-assembly of homotrimers, heterotrimers and higher order structures. *Chem Soc Rev* **2010**, *39*, (9), 3510-3527.
(191) Przybyla, D. E.; Chmielewski, J. Higher-order assembly of collagen peptides into nano- and microscale materials. *Biochemistry* **2010**, *49*, (21), 4411-4419.
(192) Erdmann, R. S.; Wennemers, H. Synthesis of Fmoc-Pro-Hyp(TBDPS)-Gly-OH and its application as a versatile building block for the preparation of collagen model peptides. *Synthesis* **2009**, 143-149.
(193) Inouye, K.; Sakakibara, S.; Prockop, D. J. Effects of the stereo-configuration of the hydroxyl group in 4-hydroxyproline on the triple-helical structures formed by homogenous peptides resembling collagen. *Biochim Biophys Acta* **1976**, *420*, 133-141.
(194) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. Synthesis of (Pro–Hyp–Gly)n of defined molecular weights. Evidence for the stabilization of collagen triple helix by hydroxyproline. *Biochim Biophys Acta* **1973**, *303*, 198-202.
(195) Ottl, J.; Musiol, H. J.; Moroder, L. Heterotrimeric collagen peptides containing functional epitopes. Synthesis of single-stranded collagen type I peptides related to the collagenase cleavage site. *J Pept Sci* **1999**, *5*, (2), 103-110.
(196) Jenkins, C. L.; Vasbinder, M. M.; Miller, S. J.; Raines, R. T. Peptide bond isosteres: Ester or (*E*)-alkene in the backbone of the collagen triple helix. *Org Lett* **2005**, *7*, (13), 2619-2622.
(197) Kotch, F. W.; Raines, R. T. Self-assembly of synthetic collagen triple helices. *Proc Natl Acad Sci* **2006**, *103*, (9), 3028-3033.
(198) Tanrikulu, I. C.; Raines, R. T. Optimal interstrand bridges for collagen-like biomaterials. *J Am Chem Soc* **2014**, *136*, (39), 13490-13493.

(199) Kovacs, J.; Kisfaludy, L.; Ceprini, M. Q. On the optical purity of peptide active esters prepared by *N*,*N*'-dicyclohexylcarbodiimide and "complexes" of *N*,*N*'-dicyclohexylcarbodiimide–pentachlorophenol and *N*,*N*'-dicyclohexylcarbodiimide–pentafluorophenol. *J Am Chem Soc* **1967**, *89*, 183-184.
(200) Vaughan, J. R., Jr. . Acylalkylcarbonates as acylating agents for the synthesis of peptides. *J Am Chem Soc* **1951**, *73*, 3547.
(201) Albericio, F.; Carpino, L. A. Coupling reagents and activation. *Methods Enzymol* **1997**, *289*, (109-126).
(202) McDermott, J. R.; Benoiton, N. L. *N*-Methylamino acids in peptide synthesis. III. Racemization during deprotection by saponification and acidolysis. *Can J Chem* **1973**, *51*, 2555-2561.
(203) McDermott, J. R.; Benoiton, N. L. *N*-Methylamino acids in peptide synthesis. IV. Racemization and yields in peptide-bond formation. *Can J Chem* **1973**, *51*, 2562-2570.
(204) Racemization in peptide synthesis. Kemp, D. S. In *The Peptides: Analysis, Synthesis, Biology*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, NY, 1979; Vol. 1, p 315-383.
(205) Benoiton, N. L.; Lee, Y.; Chen, F. M. F. Isopropyl chloroformate as a superior reagent for mixed anhydride generation and couplings in peptide synthesis. *Int J Pept Protein Res* **1988**, *31*, (3), 577-580.
(206) Hall, H. K., Jr. Correlation of the base strengths of amines. *J Am Chem Soc* **1954**, *79*, 5441-5444.
(207) Capicciotti, C. J.; Trant, J. F.; Leclere, M.; Ben, R. N. Synthesis of C-linked triazole-containing AFGP analogues and their ability to inhibit ice recrystallization. *Bioconjug Chem* **2011**, *22*, (4), 605-616.
(208) Rogers, H. W.; Weinstock, M. A.; Feldman, S. R.; Coldiron, B. M. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol* **2015**, *151*, (10), 1081-1086.
(209) Armstrong, B. K.; Kricker, A. The epidemiology of UV induced skin cancer. *J Photochem Photobiol B* **2001**, *63*, 8-18.
(210) Shafie Pour, N.; Saeedi, M.; Morteza Semnani, K.; Akbari, J. Sun protection for children: A review. *Pediatr Rev* **2015**, *3*, (1), 1-7.
(211) Gasparro, F. P. Sunscreens, skin photobiology, and skin cancer: The need for UVA protection and evaluation of efficacy. *Environ Health Perspect* **2000**, *108*, (suppl 1), 71-78.
(212) Dawes, J. M.; Antunes-Martins, A.; Perkins, J. R.; Paterson, K. J.; Sisignano, M.; Schmid, R.; Rust, W.; Hildebrandt, T.; Geisslinger, G.; Orengo, C.; Bennett, D. L.; McMahon, S. B. Genome-wide transcriptional profiling of skin and dorsal root ganglia after ultraviolet-B-induced inflammation. *PLoS One* **2014**, *9*, (4), e93338.
(213) Darvin, M. E.; Richter, H.; Ahlberg, S.; Haag, S. F.; Meinke, M. C.; Le Quintrec, D.; Doucet, O.; Lademann, J. Influence of sun exposure on the cutaneous collagen/elastin fibers and carotenoids: Negative effects can be reduced by application of sunscreen. *J Biophotonics* **2014**, *7*, (9), 735-743.
(214) Quist, S. R.; Wiswedel, I.; Quist, J.; Gollnick, H. P. Kinetic profile of inflammation markers in human skin in vivo following exposure to ultraviolet B indicates synchronic release of cytokines and prostanoids. *Acta Derm Venereol* **2016**, *96*, (7), 910-916.

(215) Green, A. C.; Williams, G. M.; Logan, V.; Strutton, G. M. Reduced melanoma after regular sunscreen use: Randomized trial follow-up. *J Clin Oncol* **2011**, *29*, (3), 257-263.
(216) van der Pols, J. C.; Williams, G. M.; Pandeya, N.; Logan, V.; Green, A. C. Prolonged prevention of squamous cell carcinoma of the skin by regular sunscreen use. *Cancer Epidemiol Biomarkers Prev* **2006**, *15*, (12), 2546-2548.
(217) Hughes, M. C.; Williams, G. M.; Baker, P.; Green, A. C. Sunscreen and prevention of skin aging: A randomized trial. *Ann Intern Med* **2013**, *158*, (11), 781-790.
(218) Moyal, D. D.; Fourtanier, A. M. Broad-spectrum sunscreens provide better protection from solar ultraviolet-simulated radiation and natural sunlight-induced immunosuppression in human beings. *J Am Acad Dermatol* **2008**, *58*, (5 Suppl 2), S149-S154.
(219) Green, A.; Williams, G.; Nèale, R.; Hart, V.; Leslie, D.; Parsons, P.; Marks, G. C.; Gaffney, P.; Battistutta, D.; Frost, C.; Lang, C.; Russell, A. Daily sunscreen application and betacarotene supplementation in prevention of basal-cell and squamous-cell carcinomas of the skin: A randomised controlled trial. *Lancet* **1999**, *354*, (9180), 723-729.
(220) Society, A. C. Cancer facts & figures 2017. **2017**, 1-71.
(221) Seite, S.; Fourtanier, A. M. The benefit of daily photoprotection. *J Am Acad Dermatol* **2008**, *58*, (5 Suppl 2), S160-S166.
(222) Forestier, S. Rationale for sunscreen development. *J Am Acad Dermatol* **2008**, *58*, (5 Suppl 2), S133-S138.
(223) Jansen, R.; Osterwalder, U.; Wang, S. Q.; Burnett, M.; Lim, H. W. Photoprotection: Part II. Sunscreen: Development, efficacy, and controversies. *J Am Acad Dermatol* **2013**, *69*, (6), 867.e861-867.e814.
(224) Mancuso, J. B.; Maruthi, R.; Wang, S. Q.; Lim, H. W. Sunscreens: An update. *Am J Clin Dermatol* **2017**, *18*, 643-650.
(225) Diffey, B. Sunscreens: Expectation and realization. *Photodermatol Photoimmunol Photomed* **2009**, *25*, 233-236.
(226) Diffey, B.; Osterwalder, U. Labelled sunscreen SPFs may overestimate protection in natural sunlight. *Photochem Photobiol Sci* **2017**, *16*, (10), 1519-1523.
(227) Solky, B. A.; Phillips, P. K.; Christenson, L. J.; Weaver, A. L.; Roenigk, R. K.; Otley, C. C. Patient preferences for facial sunscreens: A split-face, randomized, blinded trial. *J Am Acad Dermatol* **2007**, *57*, (1), 67-72.
(228) Lawrence, G. D.; PFishelson, S. UV catalysis, cyanotype photography, and sunscreens. *J Chem Ed* **1999**, *76*, (9), 1199-1200.
(229) Stanfield, J.; Osterwalder, U.; Herzog, B. In vitro measurements of sunscreen protection. *Photochem Photobiol Sci* **2010**, *9*, (4), 489-494.
(230) Diffey, B. L.; Tanner, P. R.; Matts, P. J.; Nash, J. F. In vitro assessment of the broad-spectrum ultraviolet protection of sunscreen products. *J Am Acad Dermatol* **2000**, *43*, (6), 1024-1035.
(231) Bruge, F.; Tiano, L.; Astolfi, P.; Emanuelli, M.; Damiani, E. Prevention of UVA-induced oxidative damage in human dermal fibroblasts by new UV filters, assessed using a novel in vitro experimental system. *PLoS One* **2014**, *9*, (1), e83401.
(232) Ellison, A. J.; VanVeller, B.; Raines, R. T. Convenient synthesis of collagen-related tripeptides for segment condensation. *Peptide Sci* **2015**, *104*, (6), 674-681.

(233) Walker, M. J.; Barnett, T. C.; McArthur, J. D.; Cole, J. N.; Gillen, C. M.; Henningham, A.; Sriprakash, K. S.; Sanderson-Smith, M. L.; Nizet, V. Disease manifestations and pathogenic mechanisms of Group A *Streptococcus*. *Clin Microbiol Rev* **2014**, *27*, (2), 264-301.
(234) Cunningham, M. W. Pathogenesis of group A *Streptococcal* infections. *Clin Microbiol Rev* **2000**, *13*, (3), 470–511.
(235) Finlay, B. B.; Cossart, P. Exploitation of mammalian host cell functions by bacterial pathogens. *Science* **1997**, *276*, (5313), 718-725.
(236) Pizarro-Cerda, J.; Cossart, P. Bacterial adhesion and entry into host cells. *Cell* **2006**, *124*, (4), 715-727.
(237) Dickey, S. W.; Cheung, G. Y. C.; Otto, M. Different drugs for bad bugs: Antivirulence strategies in the age of antibiotic resistance. *Nat Rev Drug Discov* **2017**, *16*, (7), 457-471.
(238) Kline, K. A.; Falker, S.; Dahlberg, S.; Normark, S.; Henriques-Normark, B. Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* **2009**, *5*, (6), 580-592.
(239) Rasmussen, M.; Edén, A.; Björck, L. SclA, a novel collagen-like surface protein of *Streptococcus pyogenes*. *Infect Immun* **2000**, *68*, (11), 6370–6377.
(240) Yu, Z.; An, B.; Ramshaw, J. A.; Brodsky, B. Bacterial collagen-like proteins that form triple-helical structures. *J Struct Biol* **2014**, *186*, (3), 451-461.
(241) Squeglia, F.; Bachert, B.; De Simone, A.; Lukomski, S.; Berisio, R. The crystal structure of the streptococcal collagen-like protein 2 globular domain from invasive M3-type group A *Streptococcus* shows significant similarity to immunomodulatory HIV protein gp41. *J Biol Chem* **2014**, *289*, (8), 5122-5133.
(242) Bober, M.; Morgelin, M.; Olin, A. I.; von Pawel-Rammingen, U.; Collin, M. The membrane bound LRR lipoprotein Slr, and the cell wall-anchored M1 protein from *Streptococcus pyogenes* both interact with type I collagen. *PLoS One* **2011**, *6*, (5), e20345.
(243) Dinkla, K.; Rohde, M.; Jansen, W. T.; Kaplan, E. L.; Chhatwal, G. S.; Talay, S. R. Rheumatic fever-associated *Streptococcus pyogenes* isolates aggregate collagen. *J Clin Invest* **2003**, *111*, (12), 1905-1912.
(244) Reissmann, S.; Gillen, C. M.; Fulde, M.; Bergmann, R.; Nerlich, A.; Rajkumari, R.; Brahmadathan, K. N.; Chhatwal, G. S.; Nitsche-Schmitz, D. P. Region specific and worldwide distribution of collagen-binding M proteins with PARF motifs among human pathogenic *streptococcal* isolates. *PLoS One* **2012**, *7*, (1), e30122.
(245) Bohinc, K.; Dražić, G.; Fink, R.; Oder, M.; Jevšnik, M.; Nipič, D.; Godič-Torkar, K.; Raspor, P. Available surface dictates microbial adhesion capacity. *Int J Adhes Adhes* **2014**, *50*, 265-272.
(246) Livak, K. J.; Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **2001**, *25*, (4), 402-408.
(247) http://bioinfo.ut.ee/primer3-0.4.0/primer3/, accessed on 04/03/2017.
(248) Jenkins, C. L.; Raines, R. T. Insights on the conformational stability of collagen. *Nat Prod Rep* **2002**, *19*, (1), 49-59.
(249) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Triple-helix propensity of hydroxyproline and fluoroproline: comparison of host-guest and repeating tripeptide collagen models. *J Am Chem Soc* **2003**, *125*, 11500-11501.

(250) Chan, V. C.; Ramshaw, J. A. M.; Kirkpatrick, A.; Beck, K.; Brodsky, B. Positional preferences of ionizable residues in Gly-X-Y triplets of the collagen triple-helix. *J Biol Chem* **1997**, *272*, (50), 31441-31446.
(251) Shah, N. K.; Sharma, M.; Kirkpatrick, A.; Ramshaw, J. A. M.; Brodsky, B. Gly-Gly-containing triplets of low stability adjacent to a type III collagen epitope. *Biochemistry* **1997**, *36*, 5878-5883.
(252) Heidemann, E.; Germann, H. P. A synthetic model of collagen: An experimental investigation of the triple-helix stability. *Biopolymers* **1988**, *27*, (1), 157-163.
(253) Thakur, S.; Vadolas, D.; Germann, H. P.; Heidemann, E. Influence of different tripeptides on the stability of the collagen triple helix. II. An experimental approach with appropriate variations of a trimer model oligotripeptide. *Biopolymers* **1986**, *25*, (6), 1081-1086.
(254) Fields, C. G.; Lovdahl, C. M.; Miles, A. J.; Hageini, V. L. M.; Fields, G. B. Solid-phase synthesis and stability of triple-helical peptides incorporating native collagen sequences. *Biopolymers* **1993**, *33*, (11), 1695-1707.
(255) Fields, C. G.; Mickelson, D. J.; Drake, S. L.; McCarthy, J. B.; Fields, G. B. Melanoma cell adhesion and spreading activities of a synthetic 124-residue triple-helical "mini-collagen". *J Biol Chem* **1993**, *19*, 14153-14160.
(256) Tanaka, T.; Wada, Y.; Nakamura, H.; Doi, T.; Imanishi, T. A synthetic model of collagen structure taken from bovine macrophage scavenger receptor. *FEBS Lett* **1993**, *334*, (3), 272-276.
(257) Tanaka, Y.; Suzuki, K.; Tanaka, T. Synthesis and stabilization of amino and carboxy terminal constrained collagenous peptides. *J Pept Res* **1998**, *51*, (6), 413-419.
(258) Ottl, J.; Bttistuta, R.; Pieper, M.; Tschesche, H.; Bode, W.; Kuhn, K.; Moroder, L. Design and synthesis of heterotrimeric collagen peptides with a built-in cystine-knot Models for collagen catabolism by matrix-metalloproteases. *FEBS Lett* **1996**, *398*, (1), 31-36.
(259) Ottl, J.; Moroder, L. Disulfide-bridged heterotrimeric collagen peptides containing the collagenase cleavage site of collagen type I. Synthesis and conformational properties. *J Am Chem Soc* **1999**, *121*, (4), 653-661.
(260) Constanze, J.; Muller, D.; Ottl, J.; Moroder, L. Heterotrimeric collagen peptides as fluorogenic collagenase substrates: Synthesis, conformational properties, and enzymatic digestion. *Biochemistry* **2000**, *39*, (17), 5111-5116.
(261) Ottl, J.; Gabriel, D.; Murphy, G.; Knauper, V.; Tominaga, Y.; Nagase, H.; Kroger, M.; Tschesche, H.; Bode, W.; Moroder, L. Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chem Biol* **2000**, *7*, (2), 119-132.
(262) Goodman, M.; Melacini, G.; Feng, Y. Collagen-like triple helices incorporating peptoid residues. *J Am Chem Soc* **1996**, *118*, (44), 10928-10929.
(263) Lebruin, L. T.; Banerjee, S.; O'Rourke, B. D.; Case, M. A. Metal ion-assembled micro-collagen heterotrimers. *Biopolymers* **2011**, *95*, (11), 792-800.
(264) Fields, G. B. Induction of protein-like molecular architecture by self-assembly processes. *Bioorganic Med Chem* **1999**, *7*, (1), 75-81.
(265) Frank, S.; Kammerer, R. A.; Mechling, D.; Schulthess, T.; Landwehr, R.; Bann, J.; Guo, Y.; Lustig, A.; Bachinger, H. P.; Engel, J. Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol* **2001**, *308*, (5), 1081-1089.

(266) Delsuc, N.; Uchinomiya, S.; Ojida, A.; Hamachi, I. A host-guest system based on collagen-like triple-helix hybridization. *Chem Commun* **2017**, *53*, (51), 6856-6859.
(267) Sardi, F.; Manta, B.; Portillo-Ledesma, S.; Knoops, B.; Comini, M. A.; Ferrer-Sueta, G. Determination of acidity and nucleophilicity in thiols by reaction with monobromobimane and fluorescence detection. *Anal Biochem* **2013**, *435*, (1), 74-82.
(268) Blackwell, H. E.; Clemons, P. A.; Schreiber, S. L. Exploiting site-site interactions on solid support to generate dimeric molecules. *Org Lett* **2001**, *3*, (8), 1185-1188.
(269) Ramachandran, G. N. Stereochemistry of collagen. *Int J Pept Protein Res* **1988**, *31*, (1), 1-16.
(270) Tamburro, A. M.; Guantieri, V.; Cabrol, D.; Broch, H.; Vasilescu, D. Experimental and theoretical conformational studies on polypeptide models of collagen. *Int J Pept Protein Res* **1984**, *24*, (6), 627-635.
(271) Jida, M.; Betti, C.; Schiller, P. W.; Tourwe, D.; Ballet, S. One-pot isomerization-cross metathesis-reduction (ICMR) synthesis of lipophilic tetrapeptides. *ACS Comb Sci* **2014**, *16*, (7), 342-351.
(272) Shoulders, M. D.; Raines, R. T. Interstrand dipole-dipole interactions can stabilize the collagen triple helix. *J Biol Chem* **2011**, *286*, (26), 22905-22912.
(273) Luo, T.; Kiick, K. L. Collagen-like peptide bioconjugates. *Bioconjug Chem* **2017**, *28*, (3), 816-827.
(274) Iuonut, A. M.; Dindelegan, G. C.; Ciuce, C. Proteases as biomarkers in wound healing. *TMJ* **2011**, *61*, 65-73.
(275) Martin, P. Wound healing--aiming for perfect skin regeneration. *Science* **1997**, *276*, 75-81.
(276) Singer, A. Cutaneous wound healing. *N Engl J Med* **1999**, *341*, 738-746.
(277) Church, D.; Elsayed, S.; Reid, O.; Winston, B.; Lindsay, R. Burn wound infections. *Clin Microbiol Rev* **2006**, *19*, (2), 403-434.
(278) Robson, M. C. Wound infection: A failure of wound healing caused by an imbalance of bacteria. *Surg Clin North Am* **1997**, *77*, 637-650.
(279) Bowler, P. G.; Duerden, B. I.; Armstrong, D. G. Wound microbiology and associated approaches to wound management. *Clin Microbiol Rev* **2001**, *14*, (2), 244-269.
(280) Wounds: Biology, pathology, and management. Lorenz, H. P.; Longaker, M. T. In *In: Norton J.A. et al. (eds) Surgery*; Springer: New York, NY, 2003, p 77-88.
(281) Dantas, G.; Sommer, M. O. A. How to fight back against atribiotic resistance. *Am Sci* **2014**, *102*, 42-51.
(282) Neely, A. N.; Gardner, J.; Durkee, P.; Warden, G. D.; Greenhalgh, D. G.; Gallagher, J. J.; Herndon, D. N.; Tompkins, R. G.; Kagan, R. J. Are topical antimicrobials effective against bacteria that are highly resistant to systemic antibiotics? *J Burn Care Res* **2008**, *30*, 19-29.
(283) O'Neal, P. B.; Itani, K. M. Antimicrobial formulation and delivery in the prevention of surgical site infection. *Surg Infect* **2016**, *17*, (3), 275-285.
(284) Ramachandran, G.; Kartha, G. Structure of collagen. *Nature* **1954**, *176*, 593-595.
(285) Ramachandran, G.; Kartha, G. Structure of collagen. *Nature* **1954**, *174*, (269-270).
(286) Rich, A.; Crick, F. The structure of collagen. *Nature* **1955**, *176*, 915-916.
(287) Brown, K. L.; Hancock, R. E. Cationic host defense (antimicrobial) peptides. *Curr Opin Immunol* **2006**, *18*, (1), 24-30.

(288) Durr, U. H.; Sudheendra, U. S.; Ramamoorthy, A. LL-37, the only human member of the cathelicidin family of antimicrobial peptides. *Biochim Biophys Acta* **2006**, *1758*, (9), 1408-1425.
(289) Fjell, C. D.; Hiss, J. A.; Hancock, R. E.; Schneider, G. Designing antimicrobial peptides: Form follows function. *Nat Rev Drug Discov* **2012**, *11*, (1), 37-51.
(290) Wu, W. K.; Wang, G.; Coffelt, S. B.; Betancourt, A. M.; Lee, C. W.; Fan, D.; Wu, K.; Yu, J.; Sung, J. J.; Cho, C. H. Emerging roles of the host defense peptide LL-37 in human cancer and its potential therapeutic applications. *Int J Cancer* **2010**, *127*, (8), 1741-1747.
(291) Li, X.; Li, Y.; Han, H.; Miller, D. W.; Wang, G. Solution structures of human LL-37 fragments and NMR-based identification of a minimal membrane-targeting antimicrobial and anticancer region. *J Am Chem Soc* **2006**, *128*, 5776-5785.
(292) Gabriel, M.; Nazmi, K.; Veerman, E. C.; Amerongen, A. N.; Zenter, A. Preparation of LL-37-grafted titanium surfaces with bactericidal activity. *Bioconjug Chem* **2006**, *17*, 548-550.
(293) Nagant, C.; Pitts, B.; Nazmi, K.; Vandenbranden, M.; Bolscher, J. G.; Stewart, P. S.; Dehaye, J. P. Identification of peptides derived from the human antimicrobial peptide LL-37 active against biofilms formed by Pseudomonas aeruginosa using a library of truncated fragments. *Antimicrob Agents Chemother* **2012**, *56*, (11), 5698-5708.
(294) Stewart, S.; Barr, S.; Engiles, J.; Hickok, N. J.; Shapiro, I. M.; Richardson, D. W.; Parvizi, J.; Schaer, T. P. Vancomycin-modified implant surface inhibits biofilm formation and supports bone-healing in an infected osteotomy model in sheep: A proof-of-concept study. *J Bone Joint Surg Am* **2012**, *94*, (15), 1406-1415.