EXPANDING THE SCOPE OF HETEROGENEOUS BACKBONE DESIGNS TO INCLUDE SINGLE- AND MULTI-DOMAIN ZINC FINGER PROTEINS

by

Kelly Lynette George

B.S. Biology, James Madison University, 2011

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kelly Lynette George

It was defended on

August 31st, 2018

and approved by

Stephen G. Weber, Professor, Department of Chemistry

Kabirul Islam, Assistant Professor, Department of Chemistry

Andrea J. Berman, Assistant Professor, Department of Biological Sciences

Dissertation Advisor: W. Seth Horne, Associate Professor, Department of Chemistry

Expanding the Scope of Heterogeneous Backbone Designs to Include Single- and Multi-Domain Zinc Finger Proteins

Kelly L. George, PhD

University of Pittsburgh, 2018

Polymers with specific folded conformations — termed "foldamers" — lend themselves to a range of potential applications in therapeutic development and biomedical and material sciences. One promising approach to foldamer design is to incorporate a combination of unnatural building blocks that retain a peptide's natural side chains while altering the backbone upon which the side chains are displayed. This creates a peptide variant with a heterogenous backbone and native-like folds and functions.

The focus of this dissertation is on recent work towards expanding the scope of heterogeneous backbone design strategies and consists of two main goals. The first goal is to mimic the tertiary structure of a single zinc finger domain (the third finger of Specificity protein 1, Sp1-3). This expands design strategies past simple tertiary structures and presents dynamic challenges such as retaining a precise metal coordination site. Data shown here indicate that appropriately designed alterations can allow for a tertiary structure that is superior to the native backbone in the context of folded stability and ligand binding.

The second goal discussed in this work is two part: 1) to investigate if the increase in folded stability seen with single domain Sp1-3 backbone variants is generalizable to other Cys_2His_2 zinc finger sequences and 2) to synthesize a multi-domain zinc finger protein through native chemical ligation. Each of these aims attempts to push towards the longer-term goal of gaining a better understanding of how backbone alterations impact DNA binding in multi-domain zinc fingers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATIONS

Aib … 2-aminoisobutyric acid

CD … Circular Dichroism

COSY … Correlation Spectrosocy

DIEA … *N,N*-diisopropylethylamine

Fmoc … 9-Fluorenylmethoxycarbonyl

HCTU … 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate

HPLC … High Performance Liquid Chromatography

ITC … Isothermal Titration Calorimetry

MALDI-TOF MS … Matrix-assisted laser desorption/ionization time of flight mass spectrometry

NMP … *N*-methylpyrrolidinone

NMR … Nuclear Magnetic Resonance

NOE … Nuclear Overhauser effect

NOESY … Nuclear Overhause effect Spectroscopy

PDB … Protein Data Bank

PPI … Protein-Protein Interaction

SPPS … Solid Phase Peptide Synthesis

TFA … Trifluoroacetic acid

1.0   INTRODUCTION

A portion of this chapter has been published in:

George, K. L.; Horne, W. S. Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Acc. Chem. Res.* **2018,** 51(5):1220-1228

and is used with permission of the publisher.

1.1   PROTEIN DESIGN BACKGROUND

The intricacy of protein structure belies a remarkable simplicity in covalent connectivity. From an alphabet of just 20 α-amino acids arranged in chains of varying sequence and length, countless complex folds and functions emerge. Nature accomplishes this feat by placing protein covalent structure (primary sequence) at the root of a hierarchy of higher-order structure that results from defined non-covalent intra- and inter-chain contacts. This hierarchy consists of local folding motifs (secondary structure), ordered arrangements of such motifs in larger domains (tertiary structure), and associations of folded chains (quaternary structure). Chemists have long sought to recreate protein-like structural hierarchies in oligomers with backbones that differ from natural α-peptides, termed "foldamers".[1] Research on this subject dating back >20 years has shown that a variety of artificial backbones can support ordered folds.[1-3] Progress toward secondary

structure in short chains gave way to quaternary assemblies based on these motifs. In looking toward the goal of complex protein-like tertiary folding patterns, a key barrier became apparent. How does one design, from scratch, both backbone connectivity and side-chain sequence that will support multiple different secondary structures and their defined packing in a single chain?

Two developments were key to overcoming the above barrier. First was the recognition that, while proteins are homogeneous-backbone polymers with variable side chains, a foldamer can have a "heterogeneous backbone" that blends monomers differing in both backbone connectivity and side-chain identity.[4] Second, when one or more α-residues in a biological sequence are replaced by an artificial backbone unit compatible with the native fold, as detailed further below, the resulting artificial oligomer can mimic the fold and function of the native sequence. Taken together, these ideas lead to a view of a protein as having two orthogonal sequences: a sequence of side chains and a sequence of backbone units displaying those side chains.

In this dissertation, I describe my work leveraging the above precedent to develop design principles for constructing heterogeneous-backbone foldamers that adopt the complex tertiary structure and metal binding environment of zinc fingers. I review precedent that informed these efforts, survey key results obtained, and look to the future potential applications of peptidomimetics.

Chemists have long been interested in the macromolecules that underlie the most basic aspects of human biology. Attempts to understand and even improve upon the folding and binding of proteins has been at the forefront of chemistry for decades. Seeing the importance of peptide and protein structure, chemists also sought to design methods towards structural mimicry with both natural and unnatural moieties and an eye on potential applications as diagnostic and

therapeutic tools.[5] Designed polymers with specific folding conformations are termed "foldamers".[6] Since the term was defined in 1998, the pursuit of foldamer designs has grown into a large field with a common pursuit of mimicking biopolymer folds and functions as vast and complex as those found in nature.

Success in protein mimicry by foldamers began with the most common secondary structure in natural proteins:α helices.[7] Foldamers with helical folding propensity have been formed by homogeneous repeats of both α-peptoids (with side chains attached to nitrogen rather than the a carbon) and short β-peptides (with an additional carbon in the backbone).[8,9] Varying the substitution patterns of α- and β residues to create a heterogeneous backbone (Figure 1) along a helix has also led to a better understanding of helical stability and side chain projection.[10,11,12,13]
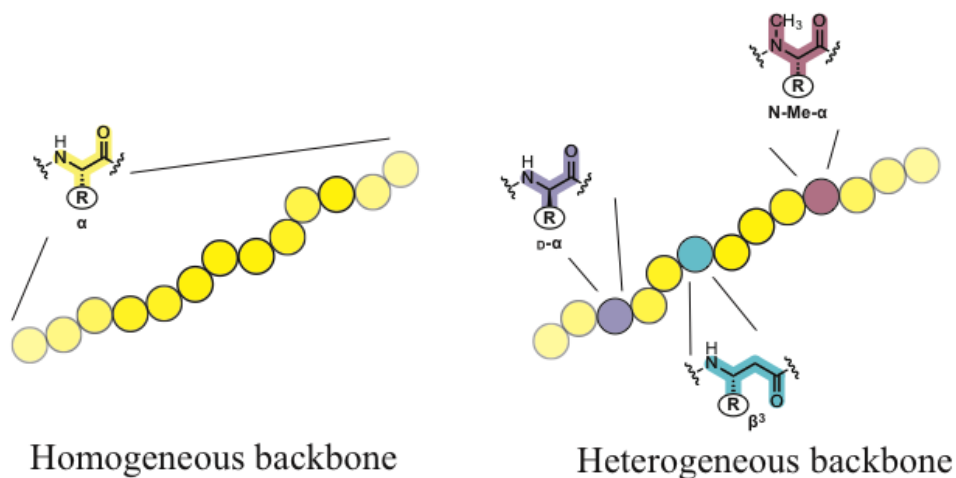
Figure 1. Key differences between a polypeptide consisting of a homogeneous backbone and heterogeneous backbone.

Backbone engineering of natural helical sequences has proven a powerful way to create protease-resistant analogues with native-like folds and molecular recognition capabilities.[14] Among the most widely used building blocks to this end are C_α-methylated α-residues such as Aib

and acyclic β-residues (e.g., $β^3$) (Figure 2). Incorporation of these monomers at one site per turn of an α-peptide helix leads to heterogeneous backbones that can fold to form quaternary assemblies encoded by the side-chain sequence of the prototype, including homomeric helix bundles[15,16] and heteromeric complexes with protein receptors.[17,18]
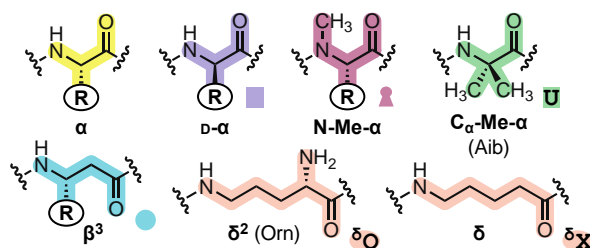
Figure 2. Native α residue alongside a selection of backbone modified analogs

To study β-sheet folding, β-hairpin model systems were used to identify monomers, such as such as D-Pro and Aib, that induce a stable turn in short peptides.[19,20] These modifications are particularly effective when paired with Gly at the $i+1$ position. Later studies showed the non-canonical amino acid Orn can promote a tight turn when the polypeptide chain is propagated from the side-chain amine (in this context, Orn is a δ-residue and replaces two α-residues in the native turn).[21] Of note, some of the above modifications as well as other more exotic moeties had been examined in protein tertiary fold contexts.[22-26]

As for backbone alterations to β-strands, many disrupt hydrogen bonding. One example is the introduction of a methyl group on the backbone amide nitrogen (α→N-Me-α substitution). This was used to engineer a variant of the protein interleukin 8 that retained a native-like fold but was unable to associate via the β-sheet interface found in the native dimer.[27] Another established strategy for β-sheet backbone modification was the replacement of extended peptide segments

with aromatic amino acids. The resulting monomer can display hydrogen bond donor and acceptor moieties such that one face of the strand is complementary to a native α-peptide, while the other is blocked. Such selective hydrogen bond disruption lends substantial credence to the utility of advancing heterogenous backbone designs as these modifications have proved instrumental in revealing structure and assembly mechanisms of amyloidogenic sequences[28] and developing inhibitors of amyloid formation.[29-31] C_α,C_α-dialkylated residues of increased steric bulk relative to Aib can also promote extended backbone conformations and cap growing sheet-based aggregates.[32]

A drawback to backbone modifications that disrupt hydrogen bonding in a β-strand is incompatibility with incorporation at an internal position of a larger β-sheet. Given the prevalence of such motifs in proteins, alternative substitution strategies were also evaluated by my previous lab mates. Initial efforts focused on β-residues, as these had proved so versatile in helices. Thus, a systematic comparison was completed of β-residues with varied backbone stereochemistry, substitution pattern, and ring constraint for the ability to promote β-sheet folding at cross-strand positions in a β-hairpin host.[33] The results showed that some β-residue classes (e.g., β^3) supported native-like folds; however, α→β substitution altered the display of side chains on residues flanking the modification. It was then established that the same monomer types could be used in alternate substitution patterns (αα→β, αα→ββ), resulting in heterogeneous backbone sheets that displayed side chains in a native-like fashion.[34] A cyclically constrained γ-residue, *cis*-1-amino-3-cyclohexanecarboxylic acid (ACC), was shown as a strong sheet stabilizer compatible with placement at internal β-strands.[35] Thus, α→γ^{cyc} substitution at two sites in a host hairpin led to a native-like folded state and a folding free energy superior to the natural backbone.[34] Although the above investigations were carried out in two-stranded hairpins, high-resolution structures obtained

suggested the modified backbones would support hydrogen bonding as part of β-sheets in larger proteins.

The precedents outlined above on the mimicry of individual secondary structures by backbones in which one or two α-residues is replaced by some artificial counterpart set the stage to examine combinations of such modifications in larger, more complex tertiary folds. When I joined the Horne Lab, colleagues had begun applying backbone alterations strategies to the tertiary structure of the B1 domain of *Streptococcal* protein G (GB1)(Figure 3).[36] This data set has led to improved understanding of how to design effective secondary structure mimics, an appreciation of the interplay among varied backbone modifications in a protein context, and insights into structural and thermodynamic consequences of backbone alteration. This work showed that the loops, helix, sheet, and turns of GB1 could be individually modified without abolishing tertiary folding, although stabilities of heterogeneous-backbone variants were slightly to somewhat compromised.[36] Individual modifications (α→β^3 substitutions in helix and loops, α→N-Me-α in the sheet, α→D-Pro/Aib in the turns) were then combined into a single variant. The resulting protein, in which ~20% of the residues in the prototype were replaced, retained a cooperative unfolding transition but was significantly less thermodynamically stable than the native.
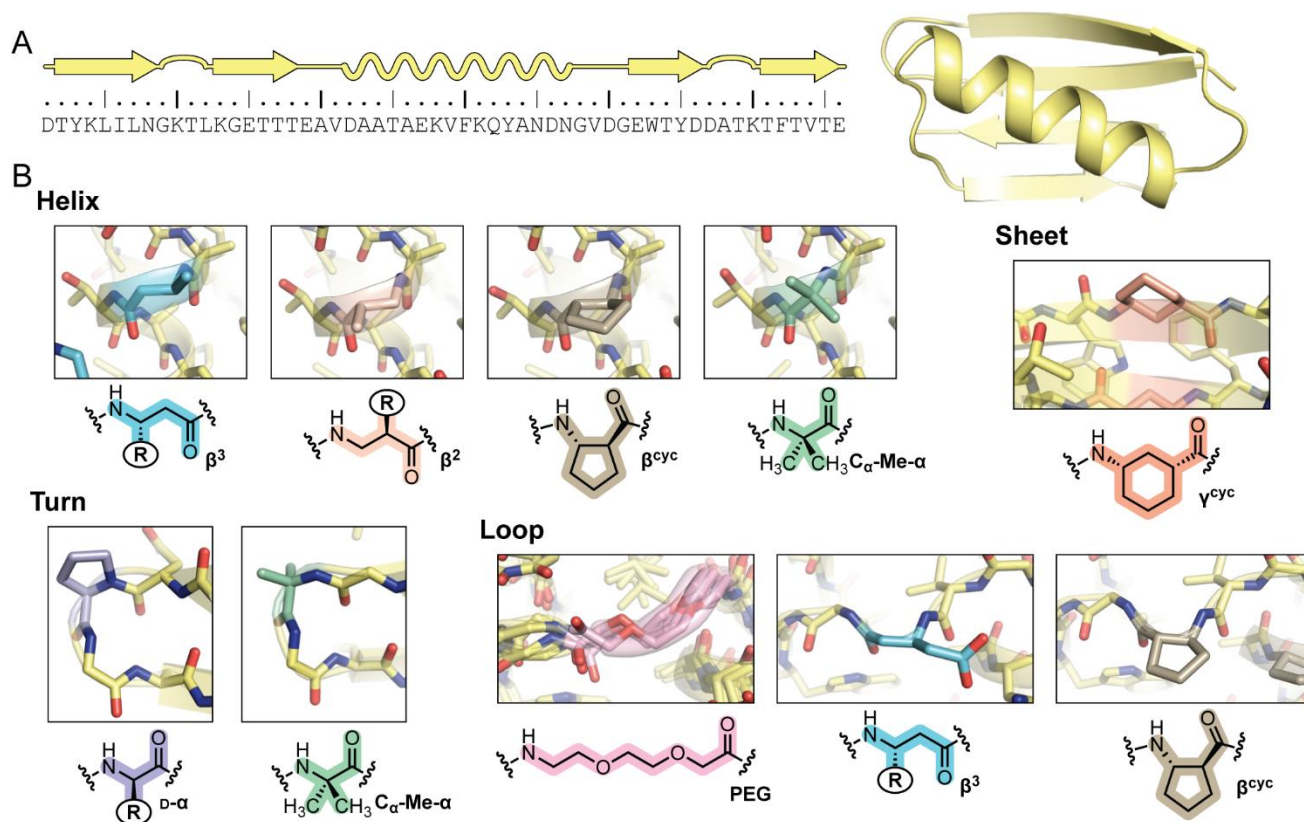
Figure 3. Mimicry of GB1 by heterogeneous backbones. (A) Sequence and crystal structure (PDB 2QMT) of wild-type GB1. (B) Some backbone modifications made to GB1, grouped by secondary structure context. All panels are from published crystal structures of GB1 variants (PDB: 5HFY, 5HI1, 4OZB, 4OZC, 4KGR, 4KGS, 4KGT), except data for the PEG-modified loop (snapshots from a molecular dynamics simulation) and the γ^{cyc}-modified sheet (NMR structure of a hairpin peptide derived from GB1). Reprinted with permission from George, K. L.; Horne, W. S. Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Acc. Chem. Res.* **2018, 51(5):1220-1228 Copyright 2018 American Chemical Society**

Collectively, a growing body of results by our lab and others spanning multiple protein systems suggests there is a great deal of plasticity with respect to the backbone chemical structures upon which sequence-encoded tertiary folds can manifest. Moreover, these efforts suggest

sequence-guided backbone alteration as a broadly applicable strategy for generating foldamers with complex tertiary folding patterns (Figure 4).
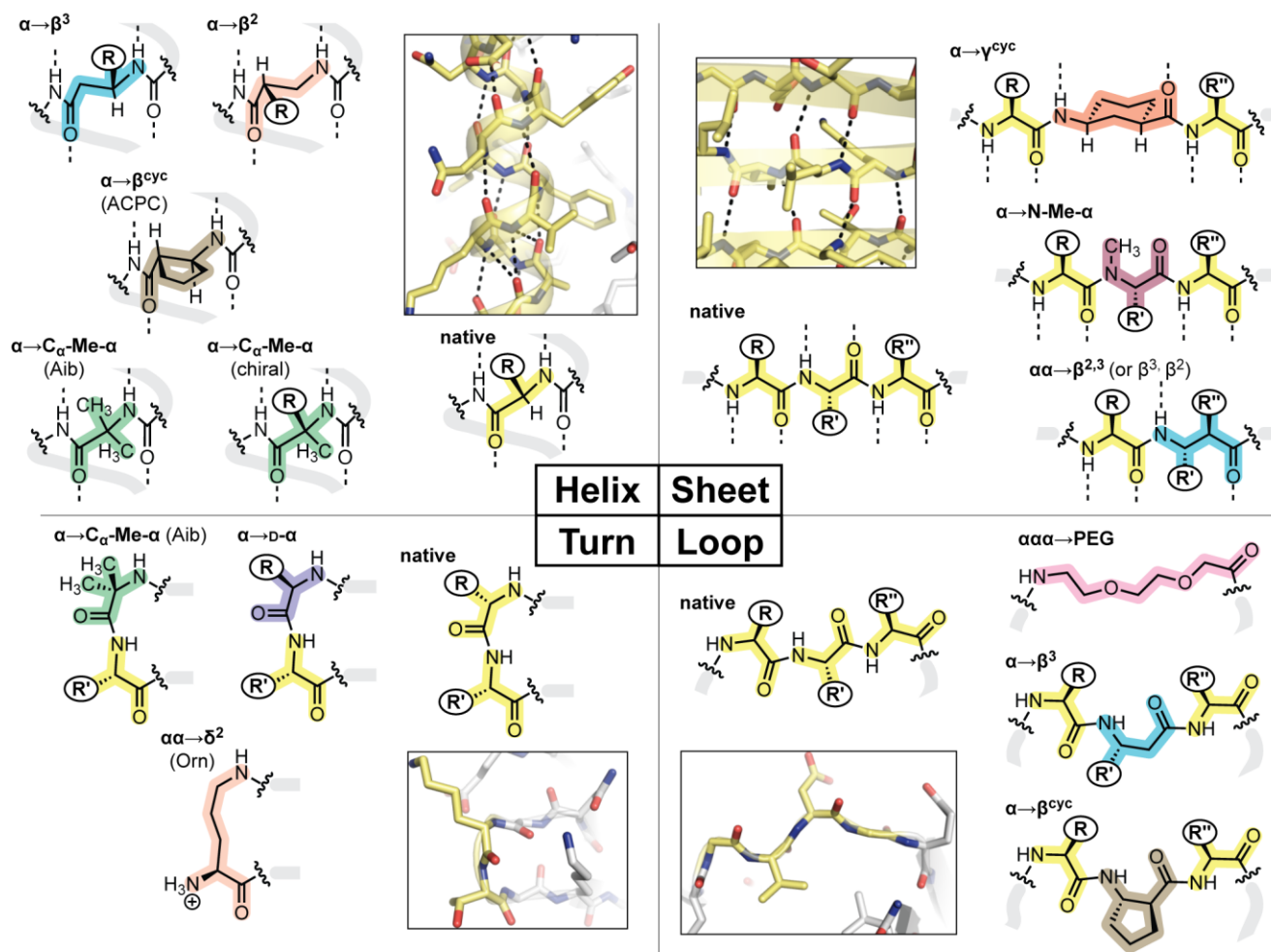
Figure 4. Overview of design principles for sequence-guided protein backbone alteration. Each quadrant presents a canonical secondary structure with a chemical representation and example from a folded protein (PDB: 2QMT) for the native alongside select replacement strategies shown viable in tertiary fold contexts. Adjusted with permission from George, K. L.; Horne, W. S. Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Acc. Chem. Res.* **2018, 51(5):1220-1228 Copyright 2018 American Chemical Society**

A majority of natural proteins require specific tertiary or quaternary structures for biological function.[37] Yet at the onset of my research, examples of unnatural oligomers that mimic higher order protein folds beyond those discussed above were limited as the complexity of the target fold increases. An important unmet aim in heterogeneous backbone designs of foldamer research still existed: mimicry of a tertiary structure with native ligand binding and potential biomolecular interaction.

1.1.1 DNA Recognition by Designed Proteins

With the growing understanding that a number of diseases, including cancer, originate with the misregulation of transcription factors, designed synthetic agents capable of reproducing DNA recognition have been developed with increasing sophistication. Miniaturized models of transcription factors (both peptidic and small molecule) have been shown to replicate DNA binding properties along the major groove but design with specific binding and high affinity have proved challenging.[38,39,40] Many designs are based on basic Leucine Zipper Proteins (bZIPs), which bind DNA through the N-terminal basic region of a leucine zipper dimer motif (**Error! Reference source not found.**).[41] It was shown that the leucine zipper portion (highlighted in green in **Error! Reference source not found.**) could be replaced with various linkers including disulfide bonds, an array of transition-metal complexes and many other moieties. From this work it was found that simply linking the two helices together was not sufficient to facilitate predictable DNA recognition and that the identity and geometry of the linker greatly influenced the binding affinity and sequence selectivity.[42-45]
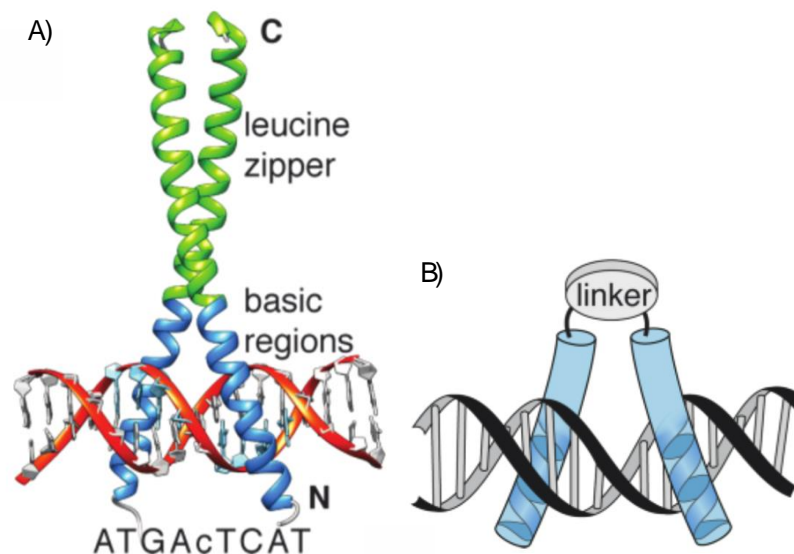
Figure 5. A) Structure of Leucine Zipper Proteins (bZIPs) bound to DNA. B) Designed variant of Leucine Zipper Protein with various linkers. Reprinted with permission from Sonia, B.; David, B.; Diego, G. P.; Miguel, V. L.; Eugenio, V. M., *European Journal of Organic Chemistry 2018* **(3), 249-261 Copyright 2018 Wiley Materials and with permission from Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C.,** *Cell* 1992, *71* **(7), 1223-1237. Copyright 1992 Elsevier**

Another and more recent approach includes designs that combine transcription factor motifs with short minor groove anchors. Linking together a short minor groove binding tail from the Serum Response Factor to the DNA helix of Elk1 was not only found to bind DNA but also was readily internalized into the nuclei of lung adenocarcinoma cells inducing a response and downregulating its targeted gene.[46]

1.1.2   DNA and RNA Recognition by Foldamers

Several foldamer designs have been made to mimic DNA and RNA recognition of natural motifs. For example, homogeneous β-peptides in the form of amphiphilic 14-helices have been

10

designed by Seebach to mimic the binding pattern of eukaryotic transcription factors along the major groove of DNA.[47] Additionally, foldamer designs with α-peptide/α-peptoid chimeras and homogeneous β-peptides have each been shown to mimic the HIV tat protein native recognition of RNA and even inhibit the infectivity of herpes simplex virus.[48-50]

1.2   ZINC FINGER BACKGROUND

1.2.1   Zinc Finger Structure and Function

First characterized in 1983, zinc fingers (ZFs) were originally identified through studies in the African clawed frog (*Xenopus laevis*) in which the interaction between the protein transcription factor IIIA (TFIIIA) and 5S rRNA necessitated zinc ions for specific DNA binding.[51] Six years later, studies showed that removal of zinc from TFIIIA resulted in a loss of both DNA binding and tertiary fold.[52] In the absence of zinc, the peptide returned to a random coil conformation. Previous to this, metal ions were known only to play functional roles — in enzyme catalysis, for example. The above studies confirmed that zinc could also play the structural role of maintaining a tertiary fold.[53]

This stabilization of the precise tertiary structure is crucial for transcription factors as it enables side chains to reside in the specific three dimensional space required for selective nucleic acid binding. In the case of zinc fingers, it is a tetrahedral coordination of metal with various combinations of cysteines and histidines (Cys_4, Cys_3His, Cys_2His_2) which stabilize the tertiary fold along with a hydrophobic core. Individual zinc fingers consist of around 30 residues and are linked

together in sets of three to nine to create tandem-binding motifs that can span along wide stretches of protein or nucleic acids. While the number and order of metal binding residues had traditionally been used as a classification system, more recently zinc fingers have begun to be classified by their overall folding motif.[54] Common folding motifs include the treble clef, zinc ribbon and the classical Cys_2His_2 motif (Figure 6).

Zinc Ribbon   Treble Cleft Cys_2His_2
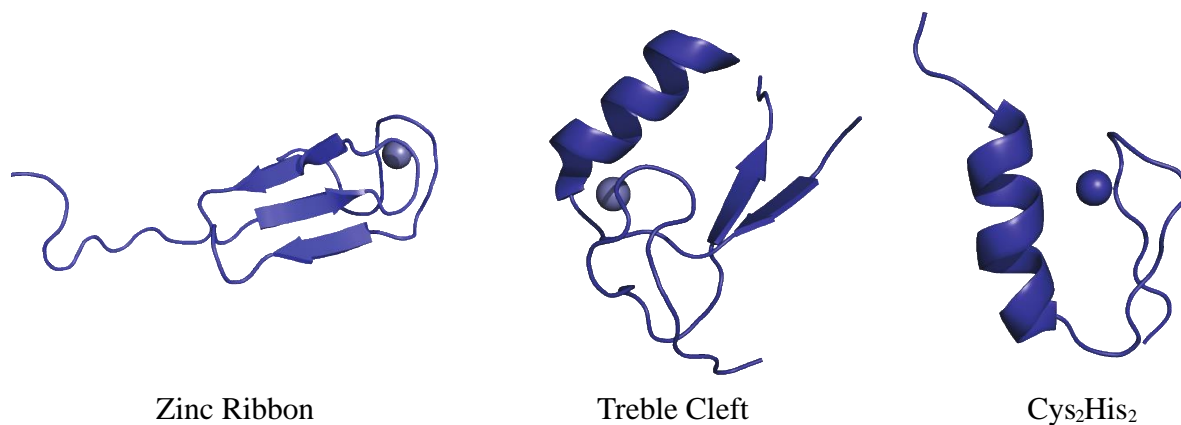
Figure 6. Main types of zinc finger motifs. Zinc Ribbon (PDB: 1PFT), Treble Cleft (PDB: 1HC7) and Cys_2His_2 (PDB: 1VA3)

In Cys_2His_2 zinc fingers, residues along the recognition helix are numbered starting with the C-terminal end of the helix at position 1. DNA contacts are made by residues at positions -1, +3 and +6 along the helix to the coding strand of DNA and by residue +2 to the noncoding strand of DNA (Figure 7).[55] Each finger contacts a 4-base pair site that overlaps with those of the adjacent fingers.
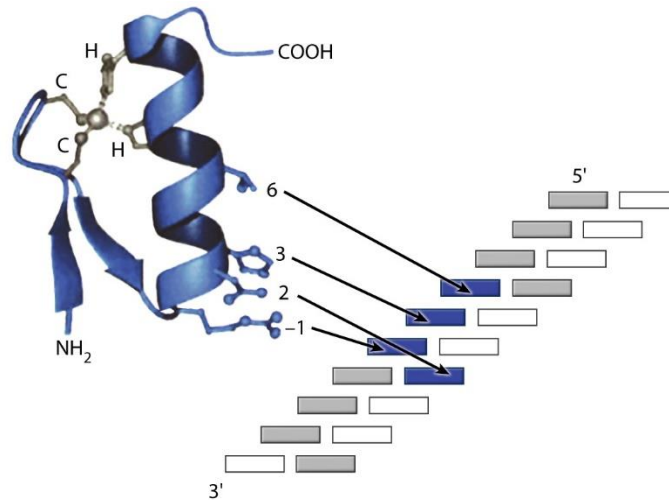
Figure 7. View of a Cys2His2 zinc finger and its contacts with DNA through residues at helical positions -1, 2, 3 and 6. Reprinted with permission from Klug, A., *The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation.* **Annual Review of Biochemistry, 2010.** 79(1): p. 213-231.

Interest in better understanding the structure and function of zinc fingers is rooted in the clinical association of mutations with developmental abnormalities and cancer.[56-59] Non aromatic substitutions within the hydrophobic core, for example, are known to impair DNA binding and transcriptional activation *in vitro* and have been associated with cancer and birth defects *in vivo*.[56] One study investigated the point mutation on the human tumor suppressor WT1 gene that produces a zinc finger with a Leu in place of the consensus Phe and results in Fraiser syndrome, a rare genetic disorder. Synthesis and analysis via circular dichrosim (CD) and isothermal titration calorimetry (ITC) demonstrated the consensus Phe at the site in question, typically considered a integral part of the hydrophobic core, is necessary for precise DNA interactions. Without it, aberrant DNA interaction results in phenotypes correlated to Fraiser syndrome.[60]

1.2.2 Zinc Finger Mimics and Engineering Precedents

The folding and DNA binding propensity of a zinc finger as well as its affinity for zinc are all correlated to the placement of consensus side chains; modification of the structure is generally a delicate endeavor.[61] Nonetheless, a diversity of approaches to mimicking the structure, metal binding and DNA binding of zinc fingers has been demonstrated since the first crystal structure was published in 1991.[62] The design of ZFs with specific DNA binding propensities was pioneered by Klug and Berg.[63,64] Various approaches were made to create ZF libraries exploring the coding relationship between zinc fingers and DNA.[65,66] Work focusing on zinc finger point mutations[67], sequence and length alterations to the linker[68] (residues connecting each zinc finger), and substitutions in the hydrophobic core[60] also contributed to understanding of base pair recognition. Other designed mimics of zinc fingers have taken the core $\beta\beta\alpha$ structure as a template that can be built upon to create a variant of the $\beta\beta\alpha$ motif in the absence of a metal ion.[69,70] Sequences identified by computational processes to resemble the structure of zinc fingers have also been synthesized and characterized.[71] Though, the computationally predefined stability does not always accurately predict the experimentally stability.[72]

1.2.3 Zinc Finger Nucleases

The combined knowledge of zinc finger engineering and the characterization of the dimerizing restriction endonuclease FokI, together, ushered in the era of targeted, site-specific gene editing with the pioneering design of Zinc Finger Nucleases (ZFNs).[70,73] Until the late 2000's, ZFNs were the only gene modification nucleases used in clinical trials (ClinicalTrials.gov identifier numbers: NCT02388594, NCT00842634, NCT01044654, NCT01252641,

14

NCT02225665, NCT02500849). Many clinical trials with ZFNs are still ongoing. Yet, often the success of ZFNs *in vitro* has been followed by a lack of genome modification *in vivo*.[74]

While RNA-guided nuclease gene editing with CRISPER/Cas9 has shifted attention to a simpler method that does not require protein engineering,[73] the gap between the excitement behind ZFNs and resulting level of success deserves consideration. The majority of engineered designs of ZFNs focused on the affinity and selection of DNA binding whereas the impacts of sequence alterations on critical aspects such as kinetics and thermodynamic stability were rarely considered.[75] Along with this, the design of sequences that bind to predicted DNA bases proved far more challenging than anticipated and interactions between zinc fingers had a greater impact than previously expected.[73] It is plausible that the engineered ZFNs failed to reach their target sequence because the balance between protein folding, protein scanning and protein binding was not entirely understood or taken fully into account.[75] Likewise, the complexity of the protein-protein, protein-DNA, and protein-zinc interactions were equally unappreciated. The need for a more thorough comprehension of the downstream impacts imbedded in sequence alteration for complex biological processes such as this underlies the importance of work within the foldamer field. Continued systematic studies of the impacts of protein sequence design with detailed analysis of kinetic and thermodynamic consequences could potentially allow future pursuits to avoid the pitfalls faced by ZFNs.

While work discussed in this dissertation does not specifically extend into a space where these pitfalls could be addressed in their entirety, the data and discussions herein do represent a step towards a better understanding of the complexities of natural and unnatural zinc fingers.

2.0   HETEROGENEOUS BACKBONE MIMICS OF A ZINC FINGER DOMAIN

A portion of this chapter has been published in:

George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* **2017,** *139*, 7931-7938.

and is used with permission of the publisher.

2.1   INTRODUCTION
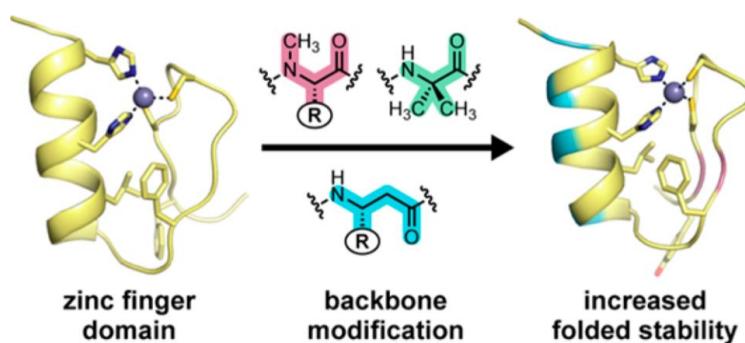
Figure 8. Backbone Alterations of Sp1-3 Zinc Finger

A variety of oligomeric backbones with compositions deviating from biomacromolecules can fold in defined ways. Termed "foldamers," these agents have diverse potential applications. A number of protein-inspired secondary structures (e.g., helices, sheets) have been produced from unnatural backbones, yet examples of tertiary folds combining several secondary structural elements in a single entity are rare. One promising strategy to address this challenge is the systematic backbone alteration of natural protein sequences, through which a subset of the native side chains is displayed on an unnatural building block to generate a

16

heterogeneous[1] backbone. A drawback to this approach is that substitution at more than one or two sites often comes at a significant energetic cost to fold stability. Here we report heterogeneous-backbone foldamers that mimic the zinc finger domain, a ubiquitous and biologically important metal-binding tertiary motif, and do so with a folded stability that is superior to the natural protein on which their design is based. A combination of UV−vis spectroscopy, isothermal titration calorimetry, and multidimensional NMR reveals that suitably designed oligomers with >20% modified backbones can form native-like tertiary folds with metal-binding environments identical to the prototype sequence (the third finger of specificity factor 1) and enhanced thermodynamic stability. These results expand the scope of heterogeneous-backbone foldamer design to a new tertiary structure class and show that judiciously applied backbone modification can be accompanied by improvement to fold stability.

The characteristic of proteins that enables their varied and sophisticated biological roles is the ability to fold into compact, well-ordered conformations encoded by amino acid sequence. Biomacromolecules are not unique in having the capacity for sequence-encoded folding, and diverse oligomeric backbones of alternate chemical compositions can show discrete folding propensities. The bulk of research on this class of molecules, termed "foldamers,"[6] has focused on secondary structures (helices, sheets, and turns) that are either inspired by motifs found in nature or completely abiotic in origin.[2,3,76-81] More recent work has sought to create unnatural backbones that mimic target folds with increased intricacy, such as multi-helix quaternary assemblies,[15,82-84] multi-stranded sheets,[85-87] and unimolecular tertiary structures.[36,88-90] The design of such molecules remains a significant challenge, the magnitude of which scales with the complexity of the desired fold.

17

While tertiary structure is crucial to protein function, so too in many cases is the ability to bind a cognate ligand. It has been estimated that a third of all proteins bind to some metal ion and half of enzymes require a specific metal to function.[91] Despite this ubiquity in nature, examples of metal binding in foldamers represent only a small fraction of the many unnatural folded units reported to date.[92] Most examples have been based on oligomers with metal-coordinating backbones,[93-96] while some have shown metal binding is possible in designed oligomers with peptide-based backbones closer to those found in nature.[97-106]

A powerful strategy for designing molecules that mimic sophisticated protein folds is the systematic backbone alteration of biologically derived sequences — inducing a natural arrangement of side chains to manifest a native-like fold on a backbone consisting of a blend of natural and unnatural building blocks.[107] Such heterogeneous-backbone oligomers can resist degradation by protease enzymes, depending on the nature of the modification,[108] and have potential applications in areas ranging from therapeutic[109] and diagnostic[110] to entirely novel functions.[111-113] Foundational work led to robust design principles for mimicry of α-helix and β-sheet secondary structures by this approach. More recent efforts have expanded to tertiary structure contexts, specifically a disulfide-cyclized helix-turn-helix motif derived from *Staphylococcal* protein A[110,114] and the α/β tertiary fold of the B1 domain of *Streptococcal* protein G (GB1).[36,87,115,116]

As noted above, protein backbone alteration can be accompanied by functional benefits; however, these benefits are almost always offset by compensating penalties to folded stability when more than one or two residues are altered.[117-121] The degree to which these observations may prove general vs. system dependent is not clear from precedent. Is it the case that natural side-chain sequences, finely honed by evolution to encode folding in an α-peptide context, will always

function best on a natural backbone? Alternatively, is it possible that certain significantly modified backbones may surpass nature in the capacity to manifest complex folds encoded by biologically-derived sequences? The best way to answer these questions is to expand heterogeneous-backbone foldamer design to a wider array of tertiary folds. Here, we report efforts to target the zinc finger, a ubiquitous motif in nature where folding and metal binding are intimately coupled. Examination of backbone-modified variants of a biologically derived domain found oligomers with >20% unnatural content that are identical to the native in folded structure and superior in folded stability.

Zinc fingers make up the largest class of eukaryotic proteins[122] and are estimated to be encoded by 3% of human genes.[55] The folding of these motifs is driven by and entirely dependent upon binding to a Zn^{2+} ion. Zinc finger domains are typically arrayed in tandem as part of larger DNA-binding proteins, giving rise to exquisite sequence-specific recognition of DNA through docking into the major groove of the double helix. As a modular protein domain, there has been a great deal of progress in designing and linking individual zinc fingers to engineer larger DNA-binding proteins able to modulate expression of specific genes.[123] Notable examples include designer zinc fingers that inhibit the replication of HIV[124] and others that stimulate angiogenesis.[125] Moreover, the inherent cell permeability of zinc fingers has paved the way for their use as a delivery system.[126]

We chose the third finger of specificity factor 1 (**Sp1-3**) as the prototype zinc finger in which to investigate sequence-guided backbone modification detailed below. **Sp1-3** is a classical example of the Cys_2His_2 zinc finger discussed in the introduction (Figure 9). In the presence of Zn^{2+}, the 29-residue **Sp1-3** adopts a ββα tertiary fold supporting a metal coordination site consisting of two Cys residues and two His residues (Cys_2His_2, Figure 9).[127] The fold is further stabilized by a conserved hydrophobic core consisting primarily of two Phe and one Leu residue.

19

Based on **Sp1-3**, peptide **1** differs from the native sequence only in a Phe^2→Tyr substitution made to facilitate spectroscopic concentration determination. Peptide **1** (Figure 9) served as the starting point for backbone modification and as a control for experiments aimed at characterization of folding and metal binding in the single domain zinc finger mimics.
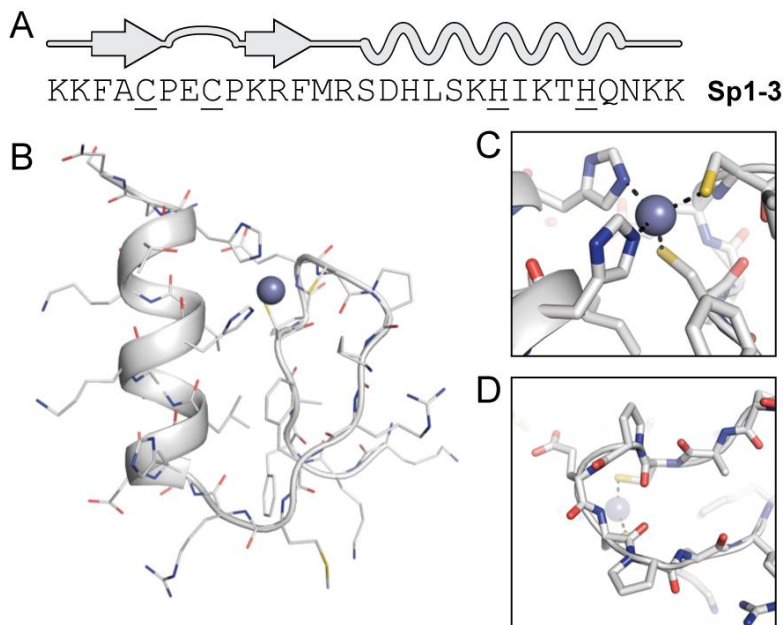
Figure 9. (A) Primary sequence and secondary structure map of the Sp1-3 zinc finger, with metal-coordinating residues underlined. (B) NMR structure of Sp1-3 (PDB 1SP1) with close-up views of the Cys$_2$His$_2$ metal-binding environment (C) and β-turn between the two strands (D). Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

To explore the relationship between backbone composition and folding, we prepared eight analogues (**2**-**9**) of the parent **Sp1-3** sequence (Figure 10). Collectively, these oligomers incorporate a variety of unnatural building blocks in place of a subset of the α-residues in **1**: D-α-, N-Me-α-, C$_\alpha$-Me-α-, β-, and δ-residues. In terms of composition, 21-25% of the backbone in each

sequence is altered. Design rationales behind each oligomer are detailed below. Relative to peptide **1**, an additional Met[12]→Nle mutation was made to reduce complications arising from oxidation of this residue during synthesis, purification, and characterization. Peptides **1**-**9** were synthesized by microwave-assisted Fmoc solid-phase methods and purified by reverse-phase HPLC; the identity and purity of each were confirmed by MALDI mass spectrometry and analytical HPLC.
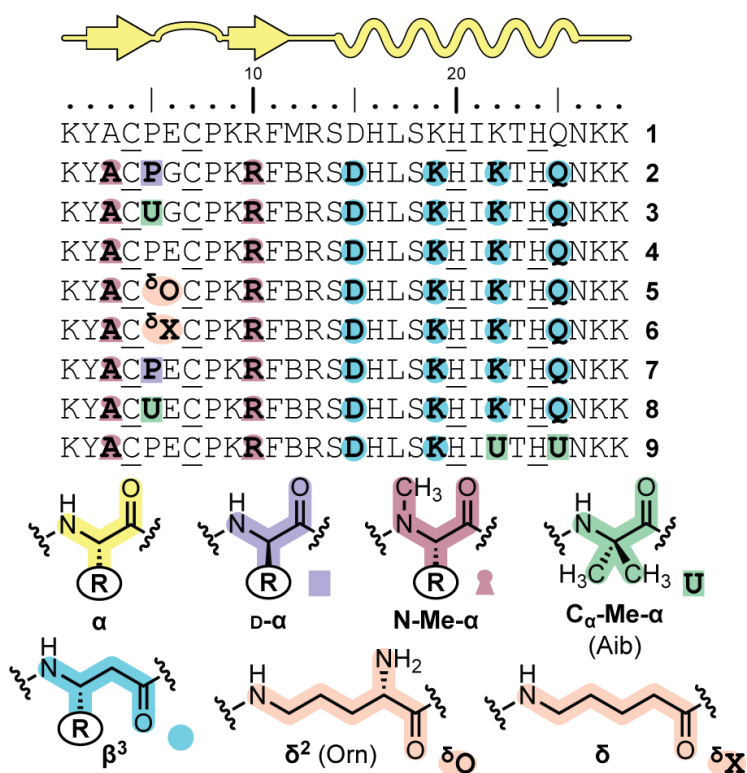
Figure 10. Primary sequence of peptide 1 and backbone-modified analogues 2-9. Bold residues indicate positions where the backbone is altered. For residue types with "R" groups, the identity of the side chain is that of the corresponding α-residue denoted by the single letter code found in the sequence. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

As a primary assay for folding and metal binding, we carried out UV-vis spectroscopy in the presence of Co^{2+}. Substitution of the native Zn^{2+} ligand in this way gives rise to spectroscopic features highly sensitive to the metal-binding environment and is a common tool in the study of zinc-binding proteins.[128] Thus, we acquired the spectrum of a 100 μM solution of peptide in 20 mM HEPES, pH 7.0, with a slight excess of Co^{2+} (Figure 11). The data obtained for **1** are consistent with prior published results for **Sp1-3**, indicating the small sequence change did not alter the fold.[129] Bands in the 300-350 nm range result from thiolate-to-Co^{2+} ligand-to-metal charge-transfer transitions with intensities correlating to the number of coordinating Cys residues.[130,131] Diagnostic d-d absorption bands in the 500-700 nm region arise from a tetrahedral cobalt(II) species, and the shape, intensity, and position of these peaks are diagnostic of coordination residue type and number.[132,133] Thus, the observation of peaks at 640 nm and 570 nm in a 3:1 ratio for **1** is consistent with the expected Cys_2His_2 tetrahedral metal binding site.
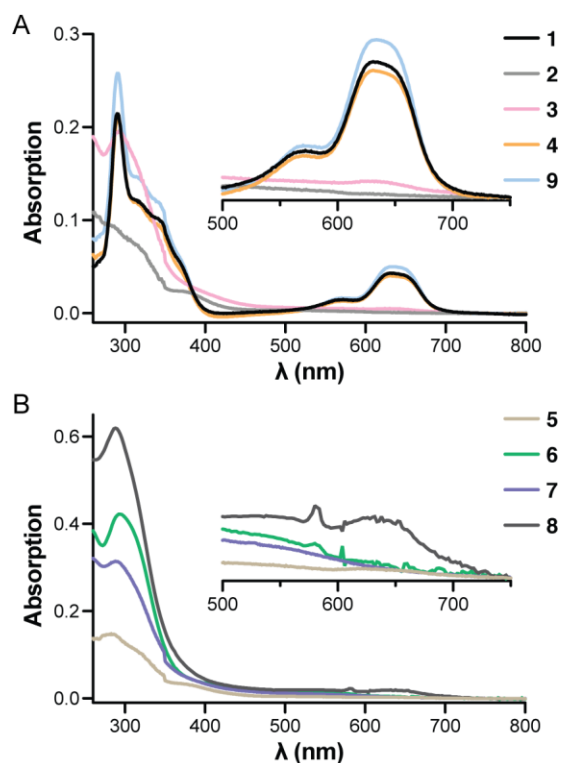
Figure 11. Interaction of 1-9 with Co^{2+} as determined by UV-vis spectroscopy. Measurements were carried out in a 1 cm path length cell at 100 μM peptide concentration in 20 mM HEPES pH 7.0 with 1.3 equiv of Co^{2+} (2 equiv Co^{2+} for peptide 5). Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* **2017,** *139***, 7931-7938. Copyright 2017 American Chemical Society**

The design of analogues **2** and **3** was based on our prior published observations for backbone modification strategies most effective at mimicry of the GB1 tertiary structure.[36,87,115,116] A single α→N-Me-α substitution was made in each strand of the hairpin segment at sites where the newly introduced methyl group should project toward solvent. α→$β^3$ substitutions were incorporated in each turn of the α-helix, such that added CH_2 groups in the backbone would reside opposite the hydrophobic core of the tertiary fold. For both the N-Me-α- and $β^3$-residues, the native **Sp1-3** side chain was retained at each point of modification. In place of the central residues of the

β-turn (Pro^{5}Glu^6), we incorporated one of two unnatural turn inducers in the form of an XxxGly sequence, where Xxx is either D-Pro or the C$_\alpha$-Me-α-residue Aib.

UV-vis spectra obtained for **2** and **3** in the presence of Co^{2+} showed that the signal at ~300 nm was retained in each case, suggesting interaction between Cys residues and the metal. However, the absence of peaks in the visible region indicated the well-defined metal coordination environment of native **Sp1-3**, and thus the tertiary fold, was abolished upon backbone modification.

In considering the inability of **2** and **3** to mimic the parent zinc finger fold, several considerations drew our attention to the turn region. First, the central turn is located directly between two metal-binding residues (Cys^4 and Cys^7); small changes in local conformation could alter alignment of these side chains and their propensity to bind metal in the folded state. Second, analysis of the NMR structure of **Sp1-3** indicates the backbone reversal is best classified as a "double turn,"[134] where a type-I turn encompassing Cys^4-Cys^7 overlaps with a type-VIII turn encompassing Pro^5-Pro^8 (Figure 12). The unnatural turn replacements employed in **2** and **3** are proven substitutions for canonical type-I and type-II turns in diverse systems;[19,20] however, they tend to promote mirror image type-I' and II' turn conformations that may be incompatible with the overlapping type-VIII turn directly following in **Sp1-3**. Finally, the side chain of Glu^6 is lost in the two analogues. This residue is found near the zinc ion in the NMR structure, where it may contribute as a second sphere ligand and/or stabilize the metal binding site through interaction with His.[24] Ideal phi and psi angles for each of the β turn types discussed are plotted in pale colors on the Ramachandran plot in Figure 12b along with the angles of residues Pro^5, Glu^6, and Cys^7 in Sp1-3.
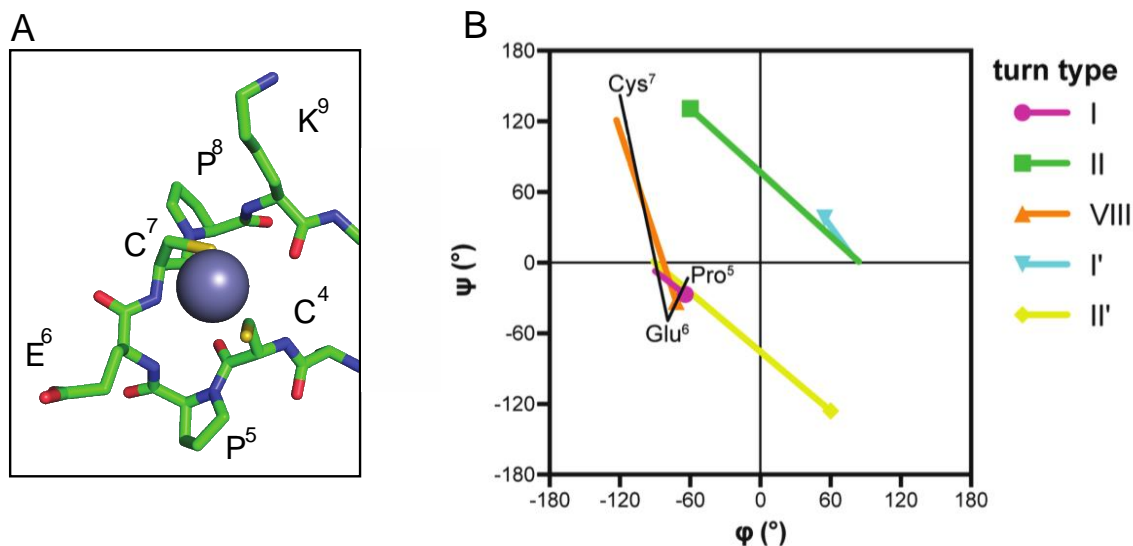
Figure 12. A. Sp1-3 Turn (PDB: 1VA3) B. Ramachandran plot depicting the tandem type-I, type-VIII double turn in Sp1-3. Idealized values are shown for residues *i*+1 (symbol) and *i*+2 (end of line) in the canonical turn types indicated, 1 alongside the conformation of residues Pro5, Glu6, and Cys7 in Sp1-3. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

To test the hypothesis that the modified turns were responsible for the inability of **2** and **3** to fold, we examined peptide **4**. This oligomer retains most of the backbone alterations present in **2** and **3** but restores the turn to its native composition. In the UV-vis assay with Co^{2+}, peptide **4** showed spectral features virtually identical to the parent domain **1**. This result, along with further experiments detailed below, provides compelling evidence for a native-like tertiary structure.

With evidence that the Aib-Gly and D-Pro-Gly segments in **2** and **3** were incompatible with the **Sp1-3** fold, we explored other modifications to the turn region. Peptides **5** and **6** were designed to examine whether the turn conformation might be the leading determinant against folding. Each incorporates a δ-residue as a dipeptide replacement in the center of the type-II turn encompassing Pro^5-Pro^8. Peptide **5** employs ornithine connected via its side chain to the preceding residue (^δO),

a well-established turn inducer.[21] Peptide **6** is an analogue of **5** with a less constrained, achiral δ-residue (**^{δ}X**). We reasoned the enhanced flexibility of **^{δ}X** relative to **^{δ}O** might help the two flanking Cys residues adopt an appropriate metal binding geometry. Peptides **7** and **8** restore the native Glu^6 side chain to **2** and **3**, respectively, and were designed to test the hypothesis that the loss of this side chain accompanying backbone modification was responsible for the inability to fold.

Analysis of the Co^{2+} binding characteristics of peptide **5-8** by UV-vis showed no evidence of a tetrahedral metal coordination environment and, thus, argues against any of these oligomers adopting a native-like tertiary fold. As with **2** and **3**, ligand-to-metal-charge-transfer transitions are observed in all cases; however, none of the modified backbones show peaks characteristic of Cys_2His_2 tetrahedral coordination. Peptide **8** with the Aib-Glu turn did show some signal ~640 nm, but the low intensity of this peak and broad background across the visible spectrum suggest significant heterogeneity in any interaction of the peptide with Co^{2+}.

Collectively, the inability of any of the turn modifications examined to be tolerated in **Sp1-3** highlights the challenges of chemical protein backbone alteration. Nevertheless, we were heartened by the promising results obtained for peptide **4** with modifications throughout the helix and hairpin. To further explore relationships between backbone composition and folding in the zinc finger motif, we prepared peptide **9**, an analogue of **4** where two of the β^3-residues in the helix are replaced by the $C_α$-Me-α residue Aib. We recently reported a thermodynamic comparison of the helical folding propensity of β^3- and $C_α$-Me-α-residues in the GB1 tertiary fold and were motivated to see to what degree those observations might hold in a fundamentally different structural context.[116] The data obtained in the Co^{2+} binding assay showed that peptide **9** supports a metal coordination environment indistinguishable from both analogue **4** and the native backbone

(**1**). Thus, we advanced peptides **1**, **4**, and **9** to experiments aimed at a more thorough characterization of folding.

2.2   FOLDING AND METAL BINDING THERMODYNAMICS

The thermodynamics behind protein folding has long been an area of intense interest. Both side-chain mutations and backbone modifications can often have dramatic thermodynamic impacts, and a solid understanding of the fundamentals driving these changes is an important precursor to understanding folding behavior.

For most protein folding events, thermodynamic contributions towards the overall free energy consist of enthalpic drive and entropic opposition.[135] Main contributors are the hydrophobic effect, protein solvation/desolvation, and the formation of hydrogen bonds. The hydrophobic effect describes the packing of nonpolar residues within the peptide core and in the process releases highly organized water molecules (desolvation effect), which contributes to the entropic drive toward the folded state. Counter to this is the entropic penalty conferred due to the loss of torsional freedom as the backbone transitions from a disordered unfolded ensemble to an organized folded state. The multitude of hydrogen bonds also greatly contributes enthaplically. The interplay of these factors creates the delicate balance described as enthalpy-entropy compensation (EEC).[136] In most cases, a folded conformation is reached which is more stable than the denatured state by around 5-15 kcal/mol.[135]

2.2.1 Thermodynamics of Folding in Heterogeneous Backbones

The folding thermodynamics of heterogeneous backbones is not nearly as well studied as that of natural peptides and proteins. One building block that has been studied to some extent is the unnatural residue amino isobutyric acid (Aib). Aib resides in restricted Ramachandran space and has been shown to enhance the thermodynamic stability of a folded protein when replacing an alanine residue.[137] This is likely due to both the decreased entropic penalty caused by rigidity in both denatured and folded state as well as an increased enthalpic drive through helix stabilization.

While assumptions can be made to predict potential impact of backbone modification on such events as peptide desolvation, conformational freedom and intramolecular interactions upon folding, caution must be taken, as many of the thermodynamic consequences of altering the backbone may not be apparent. For example, the increased conformational freedom that accompanies substitution with $\beta^{\square}$ residues might be expected to increase the entropy of the unfolded state and create a greater penalty for folding. Previous studies from our group in the protein GB1, however, have observed the opposite: $\beta^{\square}$ residue substitution incurs an entropic drive for folding.[115] It is hypothesized that this is due to the altered solvation properties of the unfolded peptide that contains more hydrophobic regions and thus forces additional order in the solvent. This ordered solvent lessens the entropy of the unfolded state (compared to solvent interacting with a native sequence) and thus creates an entropic advantage to dismantle the ordered cages of solvent and drive folding.

2.2.2   Thermodynamics of Zinc Finger Folding and Binding

An additional aspect to consider when analyzing the thermodynamics of zinc finger folding/binding, in particular, is the transition of zinc from a hydrated complex to a peptide-bound state (Figure 13). The enthalpy associated with zinc coordination by cysteine/histidine in comparison to the coordination in the zinc hydration shell is one of the largest overall contributions driving the zinc finger folding and binding event.
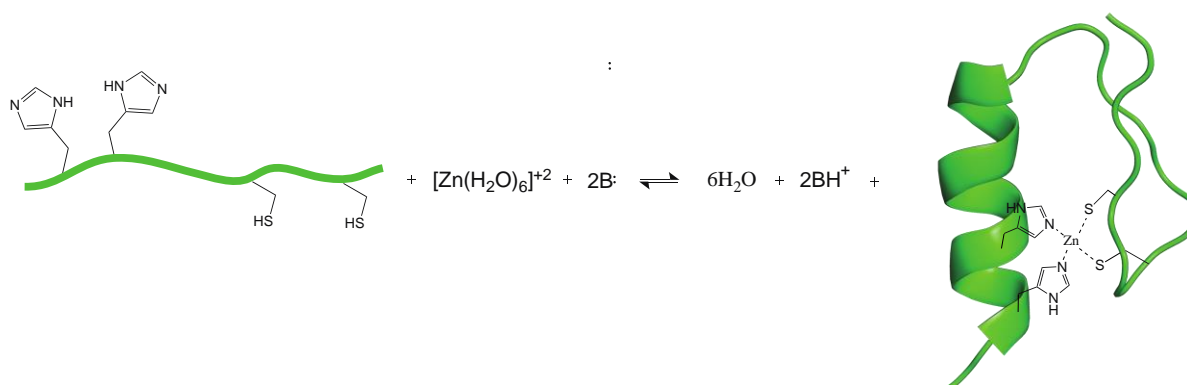
Figure 13. Overall reaction scheme of zinc - Sp1-3 binding. (PDB: 1VA3) At pH 7.4, both the cysteines and the histidines of Sp1-3 are mostly in their neutral form prior to folding (pKa = 8.2 and 6.15, respectively[138]) and zinc is in its hydrated state. Upon folding, thiols deprotonate and released protons are taken up by buffer.

In zinc fingers, folding and metal binding are closely coupled in a single overall process. Interestingly, one study aimed to better understand this interplay by uncoupling zinc finger folding and binding through the design of several minimalistic sequences consisting entirely of glycine apart from a binding site of either Cys4, Cys3His or Cys2His2.[139] These unstructured, glycine rich sequences teach us little about a true zinc finger system as they lack the torsional restraints that

accompany structure. Yet, the analysis of the complexities incurred when binding and folding are coupled is beneficial as we consider the thermodynamic comparison of variants to peptide **1** (Figure 14). Numerous other studies have looked into zinc finger binding and folding yet the relative thermodynamic contributions of ZF-metal verses ZF-ZF interactions is still not entirely understood.[138] What is well understood is the importance of side chain positioning and the formation of the hydrophobic core.
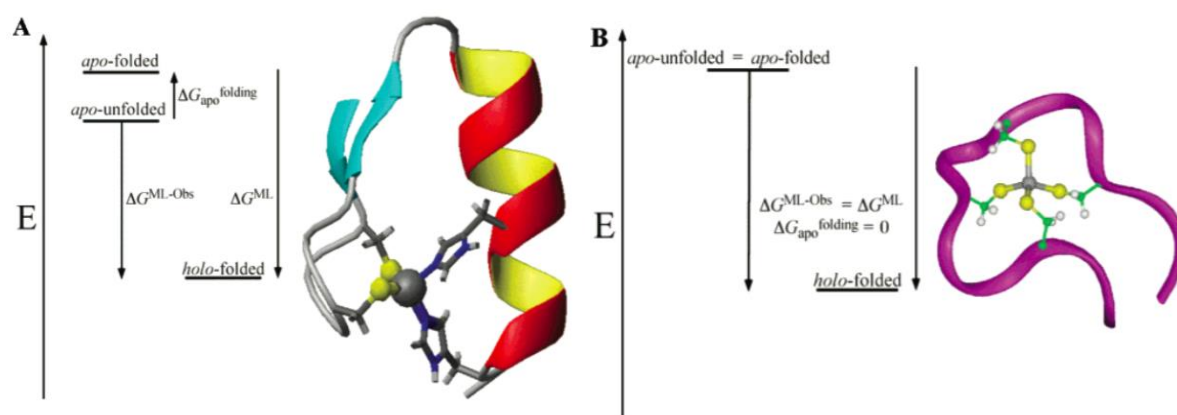
Figure 14. Folding events of A) a zinc finger with coupled binding and folding and B) a minimalistic zinc finger in which folding is not coupled to binding. ML-Obs = observed metal ligand binding free energy. Reprinted with permission from Reddi et al., *JACS* **2007** *129* **(42), 12815-12827. Copyright 2007 American Chemical Society.**

In a series of related sequences with consistent metal binding sites, contributions from peptide-metal interactions and accompanying proton transfer to buffer are similar; differences in observed energetics within the series are then dominated by differences in folding energetics.[61] Thus, comparison of metal binding thermodynamics between wild-type peptide **1** and analogues **4** and **9** has the capacity to shed light specifically on the energetic consequences of backbone modification in the zinc finger domain.

We first examined the binding affinity of each peptide for Co^{2+} by monitoring the intensity of the UV-vis peak at 620 nm as a function of added metal. The resulting binding isotherms (Figure 15) revealed dissociation constants (K_d) ~2 μM for wild-type sequence **1** as well as backbone-modified **4** and **9**[106] (Table 1). These values are comparable to literature precedent for **Sp1-3**,[129,138,140] indistinguishable within the uncertainty of the measurement, and at the limit of the affinity detectable under the conditions of the experiment. Some deviation from the expected curve for 1:1 binding stoichiometry was observed in the fitting. This behavior has been previously reported for other Cys_2His_2 zinc fingers and is attributed to a process in which multiple peptide chains bind metal competing with the formation of the 1:1 complex.[141,142] Based on the relative dissociation constants for this competing pathway (K_{d2}), the expected 1:1 complex dominates at all but the lowest peptide:Co^{2+} ratios.
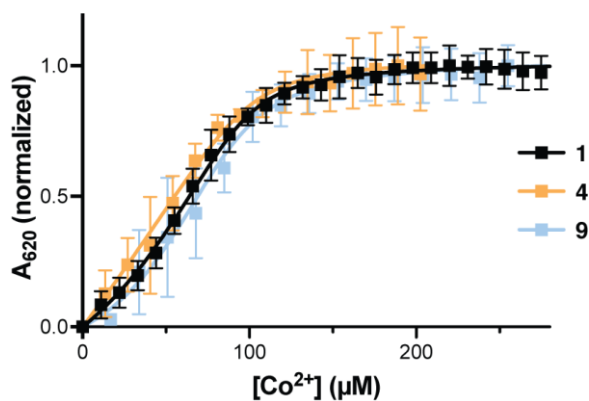
Figure 15. Co^{2+} binding isotherms for peptide 1 and backbone-modified analogues 4 and 9 determined by UV-Vis spectroscopy. Samples consisted of 100 μM peptide in 20 mM HEPES, pH 7.0. Error bars represent standard deviation from 3 independent measurements. Lines are the result of fitting to a 1:1 binding model with competing formation of a 2:1 peptide:metal complex (see methods for details). Reprinted with

We next assessed the interaction of peptides **1**, **4**, and **9** with Zn^{2+} by isothermal titration calorimetry (ITC). Besides confirming the ability of the modified backbones to fold in the presence of the native ligand, ITC experiments provide more accurate measurements of energetics for folding and metal binding compared to UV-vis titration as well as new information in the form of entropy/enthalpy components. ITC results (Figure 16, Table 1) show that the affinity of both backbone-modified analogues for Zn^{2+} is higher than that of parent sequence **1**. Consistent with prior studies on zinc finger domains, binding stoichiometries (*n*) were found to be slightly less than 1, which we attribute to partial oxidation of Cys residues during the measurement.[143] The observation that the metal binding affinities, and thus folding free energies, are so similar among **1**, **4**, and **9** is significant. In our previous work on analogous backbone modifications in the GB1 tertiary structure, a comparable fraction of backbone alteration led to a destabilization of the tertiary fold by ~8 kcal mol^{-1}.[36] Here, backbone modification has an opposite *favorable* effect on folding, with a magnitude as large as 0.5 kcal mol^{-1} in the case of peptide **4**. As the folded state of each analogue remains identical to that of the native sequence, we reason that the differences in fold stability originate elsewhere along the folding pathway, either in the denatured ensemble[115] or folding intermediates.[144]

Table 1. Thermodynamic Parameters for Metal-Binding Interactions of Peptides 1, 4, and 9 with Co^{2+} and Zn^2

	peptide		
parameter	**1**	**4**	**9**
Co^{2+}			
K_d (μM)^a	1.4 $\pm$ 0.2	2.7 $\pm$ 1.2	1.4 $\pm$ 0.5
K_{d2} (μM^2)^a	150 $\pm$ 30	1500 $\pm$ 1100	70 $\pm$ 40
Zn^{2+}			
K_d (μM)^b	2.2 $\pm$ 0.3	0.8 $\pm$ 0.1	1.7 $\pm$ 0.2
ΔG (kcal mol^{-1})^b	-7.72 ± 0.08	-8.30 ± 0.09	-7.90 ± 0.07
ΔH (kcal mol^{-1})^b	-13.5 ± 0.2	-9.8 ± 0.1	-11.0 ± 0.1
TΔS (kcal mol^{-1})^b	-5.8 ± 0.2	-1.5 ± 0.1	-3.3 ± 0.1
n^b	0.81	0.8	0.78

[a]Determined by UV-vis titration with Co^{2+}; K_d and K_{d2} are the dissociation constants for formation of the 1:1 and 2:1 ligand to metal complex, respectively. [b]Determined by ITC with Zn^{2+}. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society
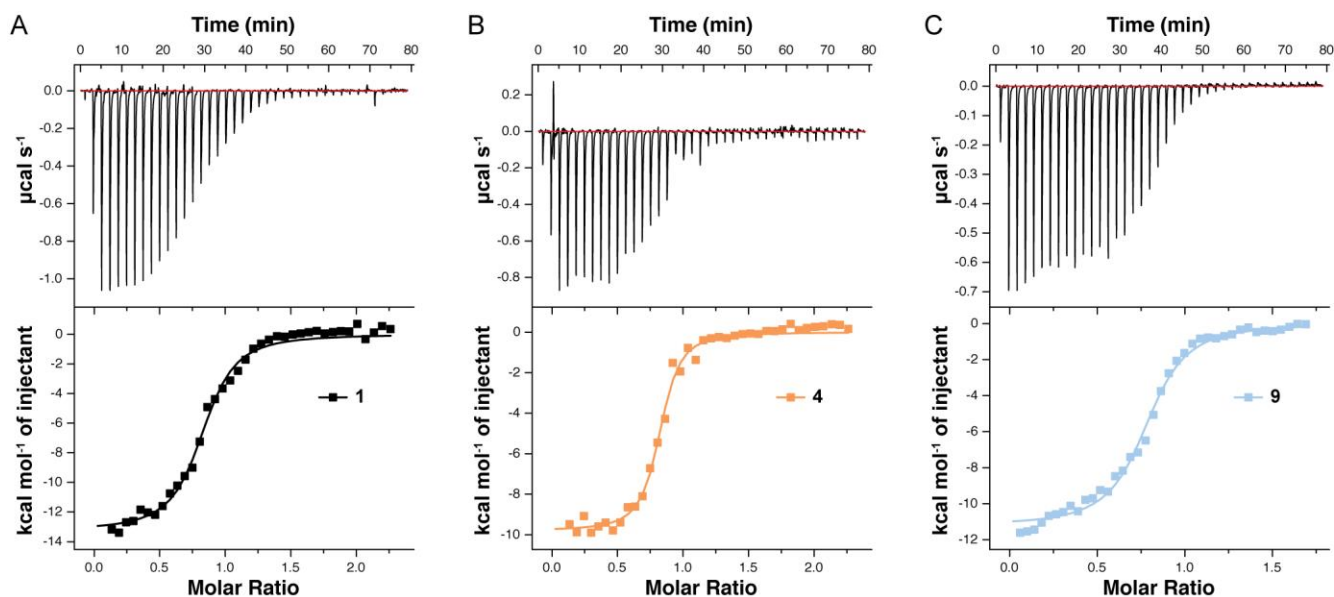
Figure 16. Isothermal titration calorimetry (ITC) data for the interaction of Zn^{2+} with peptide 1 and backbone modified analogues 4 and 9 in 50 mM HEPES, 100 mM NaCl, pH 7.4. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* **2017,** *139***, 7931-7938. Copyright 2017 American Chemical Society**

ITC data also provided insights into the thermodynamic trends underlying the observed stabilization of the fold upon backbone modification. Like the parent sequence **1**, folding and metal binding in analogues **4** and **9** are enthalpically driven and entropically opposed. Taking the energetic values for **1** as a baseline for comparison (Figure 17), the backbone modifications in **4** (2 α→N-Me-α; 4 α→β^3) have a moderately unfavorable effect on folding ΔH (~0.6 kcal mol^{-1} per substitution) that is more than compensated for by a favorable effect on folding TΔS (~0.7 kcal mol^{-1} per substitution). The above qualitative trend follows that seen upon α→β^3 modifications in the helix of GB1;[115] however, the quantitative balance between parameters differs between the two systems.[36] The magnitude of the unfavorable ΔΔH in **4** vs. **1** is comparable in GB1 and Sp1-3 yet the corresponding favorable ΔΔS is much greater in the zinc finger motif. Two key structural dissimilarities between GB1 and Sp1-3 may be responsible for this difference: chain length [56 residues in GB1 vs. 29 in Sp1-3] and fraction apolar side chains [59% in GB1 vs. 28% in Sp1-3]. The favorable entropic impact of additional methylene groups (from α→β^3 substitutions) is greater in the shorter, more hydrophilic Sp1-3.
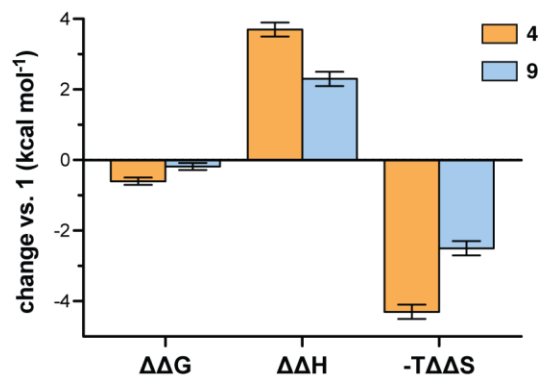
Figure 17. Summary of changes to the free energy, enthalpy, and entropy for the interaction between the indicated peptide and Zn^{2+} as determined by ITC in 50 HEPES, 100mM NaCl, pH 7.4. Changes are reported relative to corresponding parameters for natural backbone 1. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

In peptide **9**, which contains 2 α→C_α-Me-α replacements in place of 2 α→$β^3$ replacements in **4,** the enthalpic penalty and entropic boost accompanying backbone modification are both reduced in magnitude when compared to **4**. These data argue that the $β^3$-residues are the primary source of both the unfavorable ΔΔH as well as the favorable ΔΔS in the heterogeneous backbones. Interestingly, and counter to observations in GB1, the effect of replacing $β^3$-residues with Aib is overall unfavorable in the zinc finger. While the origins of this observation are unclear, the impacts of α→C_α-Me-α replacement in place of α→$β^3$ replacement in **Sp1-3** and GB1 are not solely due to a change in backbone composition but a loss of the native side chain as well. While Ala and Asn side chains were altered in the GB1 system, in **Sp1-3** Lys and Gln are modified.

2.3   HIGH-RESOLUTION STRUCTURAL CHARACTERIZATION OF THE
FOLDED STATE

The experiments detailed above all involve measurements related to the interaction between the zinc finger and metal. While folding and metal binding are highly interdependent, such analyses provide only indirect evidence bearing on the nature of the folded state. For a direct assessment of folded structure in the modified backbones and how it compares to the parent zinc finger, we subjected peptides **1**, **4**, and **9** to multidimensional NMR spectroscopy. We acquired ^{1}H/^{1}H COSY, TOCSY, and NOESY spectra for each peptide at 1.3-1.6 mM peptide concentration with 1.2 equiv. $ZnCl_2$ in 9:1 H_2O/D_2O at pH 7.0 (uncorrected). Comparison of backbone H_α chemical shifts for canonical α-residues present in all three sequences (Figure 18) showed significant deviations from random coil values and an identical pattern and magnitude of those

deviations along the chain. Collectively, these observations argue strongly for a similar folded state.
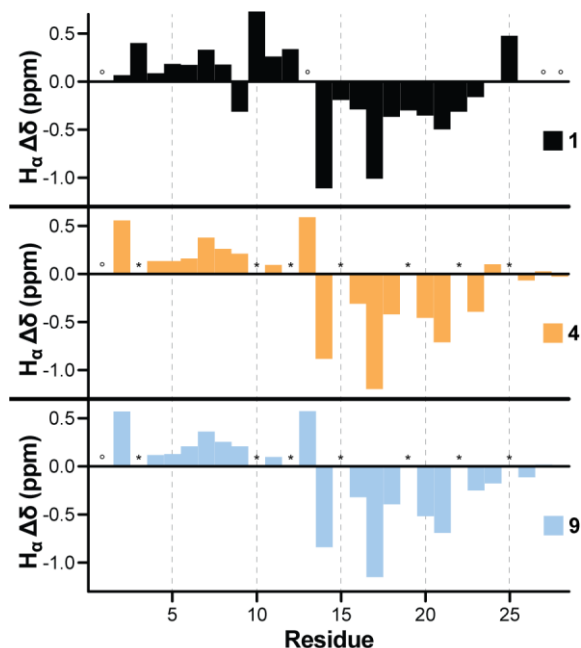
Figure 18. Backbone H$_\alpha$ chemical shift deviation from predicted random coil values for peptide 1 and backbone-modified analogues 4 and 9 as determined by NMR spectroscopy at pH 7.0 with 1.2 eq. Zn^{2+}. Asterisks indicate unnatural residues (no literature random coil chemical shift available for comparison) and open circles cases where H$_\alpha$ could not be unambiguously assigned. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

To evaluate the folds of **4** and **9** at higher resolution, we assigned the majority of backbone and side-chain resonances in each and analyzed both local and long-range NOEs. Select short-range correlations confirmed the secondary structure pattern of the parent domain was retained in the modified backbones (Figure 18). Strong sequential $i \rightarrow i+1$ H$_\alpha$-H$_N$ cross peaks were observed in each strand of the putative hairpin, along with a strong inter-strand H$_\alpha$-H$_\alpha$ correlation between

N-Me-Ala[3] and N-Me-Arg[10]. Near-continuous $i{\rightarrow}i{+}1$ H_N-H_N and $i{\rightarrow}i{+}3$ H_α-H_N correlations were seen in the expected helical region. In addition to short-range correlations indicative of secondary structure, a number of long-range NOEs were observed between the helix and hairpin that were consistent with the expected tertiary fold (Figure 19). Root mean square (RMSD) values comparing the C-α atoms of each of the nine lowest energy structures of peptide **4** to that of the previously reported tertiary folded structure for the native Sp1-3 domain are shown in Table 2. All root mean square deviations are below 1.5 Å suggesting that each lowest energy conformation of peptide **4** is highly similar to the previously reported structure. The ninth lowest energy structure shows the lowest RMSD of 1.205 Å while structure 5 shows the highest RMSD of 1.459 Å.
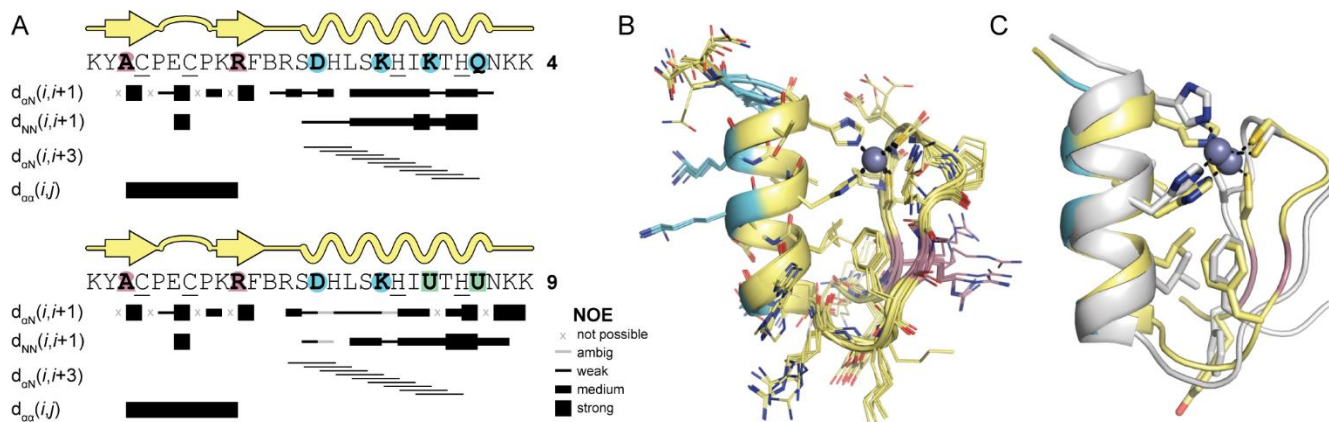
Figure 19. Analysis of the folded structure of 4 and 9 by NMR. (A) Secondary structure maps with select short- and medium-range NOE correlations supporting that assignment. (B) Overlay of the 9 lowest energy structures of peptide 4 as determined by simulated annealing with NOE distance restraints. Color code for carbons matches that in Figure 2. (C) Overlay of the NMR structure of peptide 4 (color) with that of Sp1-3 (white). Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

Table 2. RMSD values in comparison to previously reported tertiary structure of Sp13 (PDB: 1SP1).

	RMSD (Å)
5US3, target_state = 1	1.287
5US3, target_state = 2	1.325
5US3, target_state = 3	1.324
5US3, target_state = 4	1.352
5US3, target_state = 5	1.459
5US3, target_state = 6	1.342
5US3, target_state = 7	1.374
5US3, target_state = 8	1.252
5US3, target_state = 9	1.205

To capture a high-resolution picture of the folded state for one of the heterogeneous backbones, we tabulated a set of unambiguous inter-residue NOEs for peptide 4 and performed simulated annealing with distance restraints derived from NOE intensities. Preliminary calculations with just the NOE restraints supported the native-like pattern of secondary structures, overall tertiary folding topology, and proximity of Cys^4, Cys^7, His^{20} and His^{24} side chains around the expected binding site for Zn^{2+}. We therefore repeated the simulation with additional geometric restraints for hydrogen-bonds in the helix as well as the idealized tetrahedral coordination of the Zn^{2+} ion. The resulting ensemble of lowest energy structures showed excellent internal agreement and close homology to the reported folded structure for the native Sp1-3 domain (Figure 18).[127]
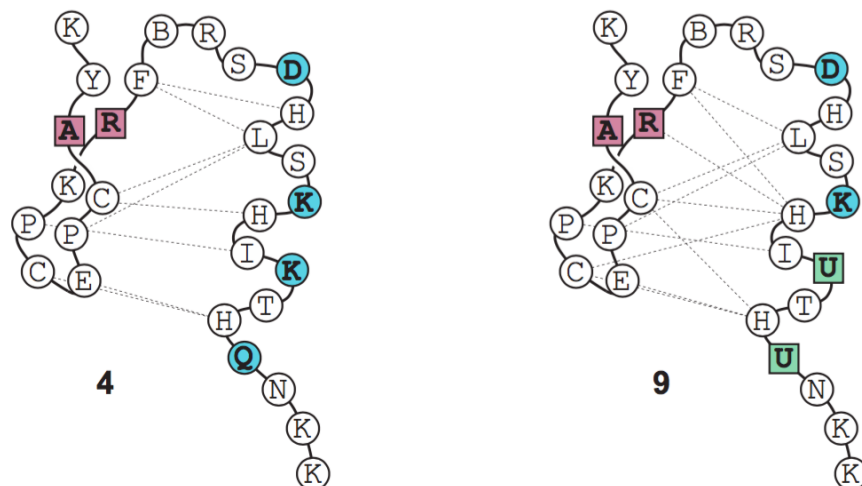
Figure 20. Selected long-range NOEs identified in 2D NMR analysis are shown as dotted lines for peptides 4 and 9. Reprinted with permission from George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* 2017, *139*, 7931-7938. Copyright 2017 American Chemical Society

2.4   PROTEOLYTIC DEGRADATION

The classical medicinal goal of peptidomimetic design was aimed toward drug development.[145] High levels of enzymatic degradation in systemic circulation can lead to such low half-life times that delivery of material to the target tissue is quantitatively insufficient. Many strategies have been developed to fortify the proteolytic stability of native sequences including cyclization, L- residue to D- residue replacement, amidation or acetylation of either the N- or C-terminus, and PEGylation.[146] Additionally, a systematic comparison of proteolytic protection incurred by selected unnatural residues has also been reported for the enzyme chymotrypsin.[108]

In an interest to determine the extent of proteolytic protection imbedded in the backbone modifications of peptide **4** and **9**, each was subjected to Trypsin digests alongside the native sequence in the presence of zinc. Trypsin cleaves at the C-terminal end of lysine and arginine residues.

Results of the trypsin digest indicate each of the three peptide variants of Sp1-3 have very short half lives in the range of only minutes (Figure 21). Interactions between Tris buffer and zinc may have sequestered the metal to such an extent that the peptide was left in an unbound, unfolded state which would be highly susceptible to degradation. To test this, digests were repeated in HEPES buffer with an increased concentration of zinc.
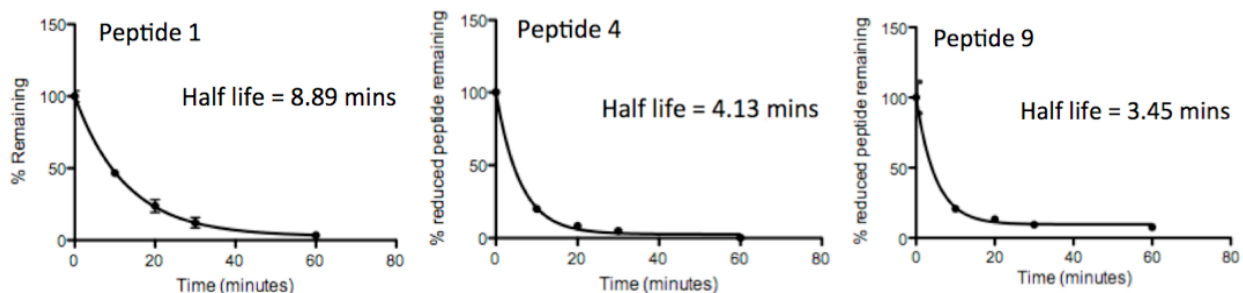
Figure 21. Trypsin digest of each peptide sequence with 1 μM Trypsin, 50 μM peptide, 1.3 eq ZnSO₄, 500 μM TCEP and 50 mM Tris at pH 7. Each digest was run in triplicate.

With the concentration of both peptide and trypsin remaining equivalent to the previous digests in TrisHCl, dramatically difference results were seen in HEPES buffer (Figure 22). At one-hour, each of the peptides remains undigested. The lower metal binding propensity of HEPES coupled with an increase in zinc concentration from 1.3 to 3.6 equivalents likely enabled each of the peptides to bind metal and form a stable tertiary structure. While satisfying to see the increase

in proteolytic stability, these results do not indicate the extent to which the backbone modifications impact proteolytic degradation. Future steps to better understand this relationship would include a digest at higher concentrations of Trypsin. Additionally, digests could be run to examine the impact of a wider array of proteases such as the more promiscuous protease K and the relatively selective chymotrypsin.
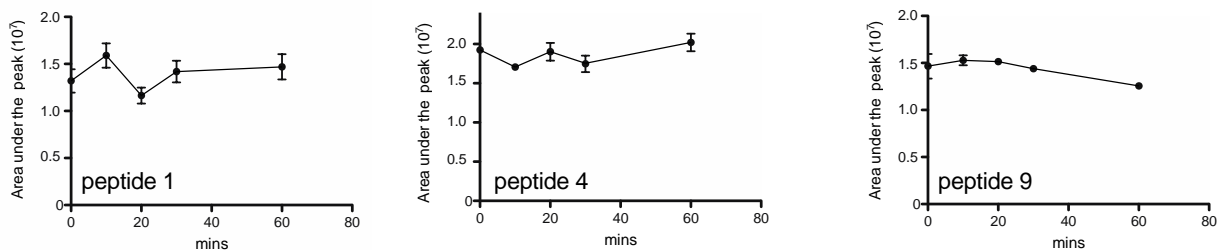
Figure 22. Trypsin digest of each peptide sequence with 1 μM Trypsin, 50 μM peptide, 3.6 eq ZnSO₄, 500 μM TCEP and 50 mM HEPES at pH 7. Each digest was run in triplicate.

2.5   CONCLUSIONS

In summary, we have demonstrated here that the systematic backbone alteration of a ubiquitous metal-binding tertiary structure — the zinc finger domain **Sp1-3** — can lead to heterogeneous-backbone mimics with native-like tertiary folded structures and metal binding environments along with superior folded stability. These findings are significant in the demonstration that a natural sequence of amino acid side chains can fold more effectively on a ~20% unnatural backbone than on the natural all-α-peptide backbone on which that sequence evolved. The successful mimicry of a new tertiary structure class through sequence-guided

backbone alteration broadens the scope of this approach and suggests it is capable of recreating a wide array of natural folding motifs. The reported results on the **Sp1-3** system also highlight the challenges and limitations of applying the method to complex folds through the extreme sensitivity of the metal-binding β-turn to alteration. Continued application to new prototype tertiary structures will lead to a greater understanding of the role of backbone chemical connectivity in protein folding and refinement of state-of-the-art design principles.

2.6   METHODS

Peptide Synthesis and Purification. All peptides were prepared by microwave-assisted Fmoc solid-phase synthesis on NovaPEG Rink Amide Resin. Microwave couplings were performed at 70 °C for 4 min with N-α-Fmoc-protected amino acid (6 equiv), HCTU (6 equiv), DIEA (10 equiv) in NMP unless otherwise specified. Coupling reactions of Fmoc-Cys(Trt)-OH, Fmoc-His(Trt)-OH, and Fmoc-Arg(Pbf)-OH were performed at room temperature for 30 min. Coupling reactions to N-Me-α-residues were performed in the microwave with 6 equiv PyAOP in place of HCTU. The pseudoproline dipeptide Fmoc-Leu-Ser-OH (Bachem) was utilized at Leu[17]Ser[18] in each sequence and was coupled with PyAOP (6 equiv) and DIEA (10 equiv) in NMP at room temperature for 1 h. Fmoc deprotections were performed with 20% v/v 4-methylpiperidine in DMF

at 80 °C for 2 min. Resin was washed 3 x 1 min with DMF after each coupling and deprotection step. Cleavage from resin was achieved by treatment with a solution of TFA/EDT/H_2O/TIS (94:2.5:2.5:1 by volume) over a period of 3 h followed by precipitation in cold ethyl ether. Pellets were collected via centrifugation and dissolved in solutions of 0.1% TFA in water and acetonitrile. Crude peptide was purified by preparative HPLC on C_{18} columns (100 Å pore size, 10 μm particle size using gradients between water and acetonitrile with 0.1% TFA. The identity of each peptide was confirmed by MALDI-TOF MS and the purity of each assessed by analytical HPLC (Figure 23, Table 1).

Preparation and Storage of Stock Solutions. Peptides were stored as lyophilized powders under high vacuum. Peptide stock solutions were prepared fresh prior to each measurement. Metal stock solutions were stored at room temperature for a maximum of 3 weeks. Solvents used for sample preparation were purged by freeze-pump-thaw cycles and/or argon bubbling. The absence of rigorous efforts to exclude oxygen in solution preparation and manipulation led to rapid peptide oxidation on a time scale that varied depending on sequence, pH and temperature in the range of minutes to hours. Concentration of peptide solutions used in biophysical experiments were quantified by DTNB assay (aka Ellman's assay) according to the procedure given by Thermo Scientific.[147] Cobalt and zinc stock solution concentrations were determined either by ICP-AES, ICP-MS, or spectrophometrically with Zincon monosodium salt according to published methods.[148]

UV-vis Spectroscopy. UV-vis spectra were acquired at 20 °C on a Cary 5000 Spectrophotometer in a 1 cm path length cuvette. Solutions contained 100 μM peptide, 20 mM HEPES (pH 7.0), 1.8 mM TCEP with 1.3 equivalents of Co^{2+} (2 equiv Co^{2+} for peptide **5**). Analytical HPLC prior to and post each experiment confirmed peptide was fully reduced.

UV-vis Titrations. CoCl$_2$ (1700 μM) was titrated into a 100 μl solution of 100 μM peptide, 20 mM HEPES (pH 7.0) with 1.8 mM TCEP in a 1 cm path length cuvette. Titrations were monitored at 20 °C on an HP Diode Array Spectrophotometer. Full spectra (200 - 800 nm) were recorded after each addition, following a 1 min equilibration. Spectra were corrected for dilution. Three replicate titrations were performed for each peptide. Analytical HPLC prior to and post each experiment confirmed peptide remained fully reduced. UV absorbance at 620 nm was plotted as a function of metal concentration and data were analyzed with Dynafit[149] (system input: P + M <--> MP; P + P + M <--> MP$_2$). The extinction coefficients of the two complexes and dissociation constants for their formation were allowed to float in the fit. The total concentration of metal and peptide were held constant based on experimental parameters. Uncertainties for the values are listed as the standard error output from the fits.

Isothermal Titration Calorimetry (ITC). ITC measurements were performed at 25 °C with a Malvern MicroCal iTC200. All solutions were purged with argon and adjusted to pH 7.4 ± 0.05 prior to titration. At the onset of the experiment, 300 μL of a solution containing 100 μM peptide, 50 mM HEPES, 100 mM NaCl was added to the cell. The syringe was filled with a solution of ZnCl$_2$ (0.9 - 1.6 mM) in 50 mM HEPES, 100 mM NaCl. Typical experiments consisted of 2 μL additions per injection with a total of 20 injections in an anaerobic environment. Equilibrium was reached between each injection. Data were analyzed using software supplied with the instrument. Heats of zinc, peptide and buffer dilution were each measured from corresponding control titrations and subtracted from the experimental data prior to fitting.

NMR Spectroscopy. For 2D NMR experiments, samples were prepared with 1.25 – 2.56 mM peptide, 1.3 equiv ZnCl$_2$, and 0.2 mM DSS in 9:1 H$_2$O:D$_2$O at pH 7.0 (uncorrected). Spectra were recorded at 4 °C on a Bruker Avance-700 spectrometer. 512 t_1-increments of 2048 points were

acquired for COSY, TOCSY and NOESY, with mixing times of 80 ms for the TOCSY and 400 ms for the NOESY. The free induction decay in both dimensions were multiplied by a phase-shifted sine bell apodization function, zero filled, and Fourier-transformed to yield 4096 by 4096 matrices. All spectra were processed using Topspin and referenced to internal DSS. Random coil chemical shift values used to generate the plots in Figure 7 were calculated for peptide **1** using the Poulsen IDP/IUP Random Coil Chemical Shift Server [https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/], which implements published methods.[150-152]

NMR Structure Calculations. Structure determination was carried out by simulated annealing using the Crystallography and NMR System software suite (CNS v1.3)[153,154] and based on published methods previously applied to isolated zinc finger domains.[127,155,156] For peptide **4**, 234 unique unambiguous inter-residue cross-peaks in the NOESY spectrum were assigned, integrated, and converted to interatomic distances based on calibration to $i{\rightarrow}i{+}1$ H_N-H_N correlations observed in the α-helix (set as 2.8 Å).[127] The resulting NOE-derived distances were sorted and conservatively classified as strong (≤2.7 Å), medium (≤3.5 Å), weak (≤4.5 Å), or very weak (≤5.5 Å). Topology and parameter files for the N-Me-α-, β^3-, and norleucine residues were manually prepared and included along with the implicit solvent force field for NMR structure determination included in CNS (protein-allhdg5-4). Simulated annealing was performed with 1000 steps of high-temperature dynamics, followed by a 2000-step slow-cooling in torsional space, a 6000-step cooling cycle in Cartesian space, and a final 200-step minimization. The NOE energy term scale factor was set to 150 during heating and the first cooling cycle and 75 for the Cartesian cooling and final minimization. Other parameters were kept as their default values. Initially, 100 trial structures were generated using just the NOE distance restraints. The resulting ensemble showed

evidence supporting a secondary structure distribution, overall tertiary folding topology, and metal-binding site matching the published native **Sp1-3** structure.[127] Thus, additional restraints were prepared corresponding to hydrogen-bonds in the helix (residue range determined by short-range NOE correlations) and tetrahedral coordination of Zn^{2+} by Cys^4, Cys^7, His^{20}, and His^{24} (2.3 Å for S^{γ}-Zn, 2.0 Å for N^{ε}-Zn).[157] 200 trial structures were generated by simulated annealing using NOE distances along with the above additional restraints. The resulting set of models was filtered for NOE violations > 0.5 Å, resulting in the 9 structures presented as the NMR ensemble. Ensemble coordinates and additional experimental data are deposited in the PDB under accession code 5US3.

Proteolysis Reactions Concentrations of active Trypsin were quantified by UV-vis spectroscopy with the substrate Nα-Benzoyl-L-arginine ethyl ester (BAEE). Reactions were each quenched with 0.5% TFA and analyzed with analytical RP-HPLC. Integration of the resulting chromatogram was then plotted against total reaction time.

Table 3. MALDI-TOF MS data for peptides 1-9

#	[M+H]+ m/z (average)	
	Calculated	**Observed**
1	3412.0	3411.7
2	3405.9	3405.0
3	3393.9	3392.4

4	3477.9	3478.7
5	3365.7	3364.6
6	3348.7	3351.0
7	3477.9	3476.3
8	3465.9	3465.1
9	3363.8	3363.5
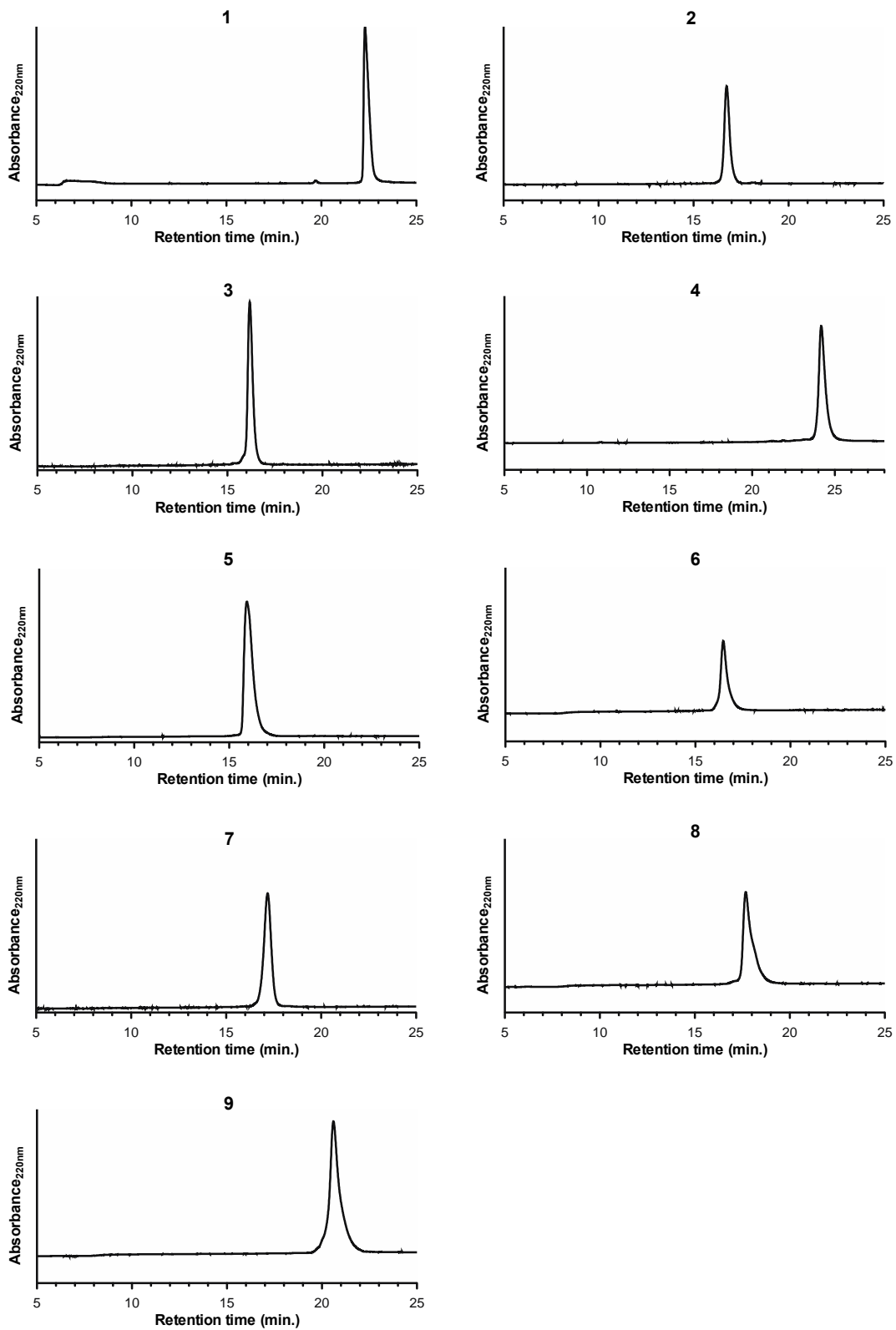
Figure 23. Analytical HPLC chromatograms of purified peptides 1-9.

49

Table 4. Backbone H$_N$ and H$_\alpha$ chemical shifts (ppm) for peptide 1, 4 and 9

1	H$_\alpha$	H$_N$	4	H$_\alpha$	H$_N$	9	H$_\alpha$	H$_N$
K	-[a]	-[a]	K	-[a]	-[a]	K	-[a]	-[a]
Y	4.662	8.845	Y	5.153	9.013	Y	5.166	8.981
A	4.707	8.893	NMeA	5.426	-[b]	NMeA	5.473	-[b]
C	4.824	8.653	C	4.871	8.647	C	4.856	8.642
P	4.654	-[b]	P	4.605	-[b]	P	4.597	-[b]
E	4.469	9.649	E	4.457	9.683	E	4.503	9.697
C	5.129	8.166	C	5.174	8.195	C	5.157	8.145
P	4.62	-[b]	P	4.702	-[b]	P	4.697	-[b]
K	3.893	8.391	K	4.419	8.335	K	4.415	8.321
R	4.985	7.883	NMeR	5.661	-[b]	NMeR	5.677	-[b]
F	4.902	8.55	F	4.733	8.719	F	4.738	8.703
M	4.796	7.376	Nle	4.345	-[a]	Nle	4.501	8.489
R	-[a]	-[a]	R	4.952	7.852	R	4.935	7.803
S	3.314	8.205	S	3.541	8.608	S	3.584	8.685
D	4.38	7.237	β³D	4.654	8.451	β³D	4.647	8.45
H	4.371	7.241	H	4.351	7.14	H	4.341	8.448
L	3.372	7.179	L	3.184	7.686	L	3.231	7.795
S	4.049	8.28	S	3.996	7.948	S	4.023	7.989
K	3.995	7.547	β³K	4.107	7.605	β³K	4.137	7.724
H	4.319	7.609	H	4.215	8.181	H	4.155	8.061
I	3.646	9.038	I	3.431	8.983	I	3.45	9.599
K	4.069	7.023	β³K	3.992	6.904	Aib	-[b]	7.223
T	4.146	7.782	T	3.913	7.57	T	4.057	7.806
H	4.686	7.279	H	4.781	7.778	H	4.505	7.733
Q	4.797	8.957	β³Q	4.242	7.498	Aib	-[b]	7.511
N	4.683	6.983	N	4.61	8.52	N	4.561	8.455
K	-[a]	-[a]	K	4.321	8.473	K	4.309	8.277
K	-[a]	-[a]	K	4.255	8.436	K	4.268	8.393

[a] Resonance unable to be assigned unambiguously. [b] Atom does not exist.

3.0 PROGRESS TOWARD MIMICRY OF A MULTI- DOMAIN ZINC FINGER PROTEIN

Work in the preceding chapter described the mimicry of the tertiary fold of a metalloprotein with heterogeneous-backbone variants that retained metal binding environments and dissociation constants indistinguishable from those of the native sequence. As prior precedents in heterogeneous-backbone modifications were limited for tertiary structures and absent in metal-binding tertiary structures, this work advances the heterogeneous foldamer field. Remarkably, these backbone-modified designs also display increased thermodynamic fold stability relative to the native domain. This demonstrates that systematically placed backbone alterations can produce protein variants with characteristics superior to the native backbone. I was motivated to continue pushing the boundaries of complexity and apply the lessons learned in the single-domain zinc finger mimics to a larger and more intricate system: a multi-domain zinc finger protein and its interaction with DNA.

The design and construction of peptide and protein mimics capable of defined interactions with biomacromolecular targets is an area of intensive research with a multitude of validated approaches. For example, many agents with structures reminiscent of secondary and tertiary motifs have been designed as modulators of protein-protein interactions (PPIs).[113,158] Similar approaches have been applied to the develop peptides with antimicrobial activity and as actors in specific disease related signally pathways.[17,159,160] Examples of this include designed α/β peptides with potent agonist activity for GPCRs such as the parathyroid hormone receptor and GLP-1R.[161,162]

Challenges and opportunities inherent in targeting a multi-domain zinc finger protein include 1) expanding beyond mimicking structure to mimicking biological recognition of a

consensus DNA sequence, 2) extending the size of proteins able to be targeted by heterogenous backbone sequence designs, and 3) probing the generality of the designs successful in a single finger domain in a more complex context. Further, testing the heterogeneous backbone design principals in a prototype of increased complexity also opens the possibility to better understand optimal strategies for the design of such agents.

From a practical standpoint, the goal of protein functional mimicry inherent in recreating the complex biological recognition of a consensus DNA sequence by an artificial agent would take the field of heterogeneous backbone designs one step closer to the long-term goal of real-world biomedical applications. Each increase in target complexity that can be met with a native-like folded mimic increases the possible sequences and structures for which backbone alterations could be applied expanding the scope of possible applications in such agents. Yet, as the complexity of the target fold or function increases, so too does the likelihood that a seemingly minor alteration will result in unintended consequences that potentially abolish or alter the fold or biological function (i.e. binding). Likewise, the more complex a prototype becomes the more opportunity we have to better understand the multitude of interactions which contribute to protein folding/binding and have been fine tuned in nature. Much like the lessons learned in prior work developing strategies for tertiary structure mimicry based on data accumulated related to secondary structure mimics, insights into the structural and thermodynamic consequences of ZF-ZF interactions and ZF-DNA interactions can be gained from both backbone designs that work as we envision as well those detrimental to function.

An additional challenge presented by the attempt to mimic a multi-domain zinc finger protein is the size of the target. The ~ 90 residue length was far larger than the 30 - 60 residues typical of artificial backbone sequences synthesized by solid-phase peptide synthesis (SPPS),

which had been the workhorse of the group up to this point. This increased length necessitated a different synthetic approach.

3.1   METHODS FOR PROTEIN CHEMICAL SYNTHESIS

The total synthesis of peptide and protein sequences of arbitrary size has long been considered a 'grand challenge' of synthetic organic chemistry. Emil Fischer referenced this in his Nobel Prize address in 1902:

"Nevertheless, the chemical enigma of Life will not be solved until organic chemistry has mastered another, even more difficult subject, the proteins, in the same way as it has mastered the carbohydrates."

The challenge posed by peptide and protein synthesis, both at the turn of last century as well as at the turn of this century, has been continually met with marked progress and intense interest. Particularly dramatic strides resulted from protecting groups and coupling reagents becoming more widely available in the 1950's and 1960's and when Merrifield published his protocol for SPPS.[163] Iterative optimizations of these protocols over the next decades led to improvements in efficiency; however the length of chains accessible remained limited to ~40 residues. Various ligation chemistries to enable the synthesis of longer sequences did become more widely available by the late 1980's, yet it was not until 1994 that the particularly sought-after characteristic of forming a native peptide bond was enabled.[164-166] This ability to generate a native amide bond at the ligation site, along with its chemoselectivity, mild, aqueous environment and a lack of additional protecting groups quickly drove native chemical ligation (NCL) to the forefront of protein synthesis methodology.[167]

NCL involves the reaction between two peptide fragments: one with a C-terminal thioester, the other with a N-terminal cysteine. Mechanistically, NCL proceeds via a thiol-thioester exchange through transthioesterification that generates a thioester-linked intermediate. A subsequent $S{\to}N$ acyl transfer then produces a stable native peptide bond (Figure 24). This reaction proceeds at near neutral pH in phosphate buffer with a chaotropic agent, such as guanidine, to aid in peptide dissolution and tris(2-carboxyethyl)phosphine (TCEP) to minimize oxidation of unprotected cysteines.
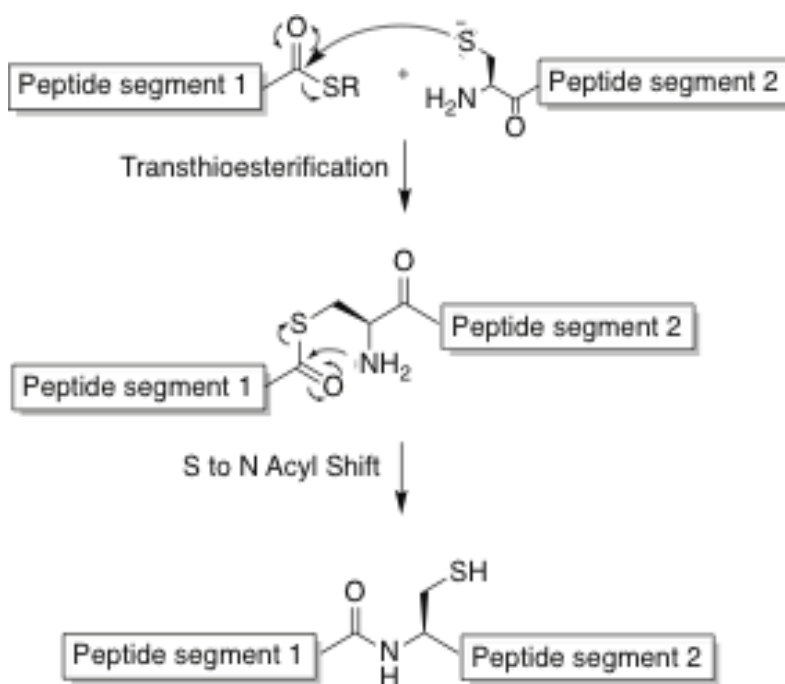
Figure 24. Scheme detailing the mechanism behind native chemical ligation.

A variant of NCL is expressed protein ligation, in which a chemically synthesized peptide is chemoselctively and regioselectively ligated to a recombinant protein.[168] In the context of our heterogenous backbone design work, expressed protein ligation opens the doors to introduce backbone-altered residues as necessary via SPPS in one segment while subsequently ligating this to a second segment, which contains all α residues and is produced through expression.

As discussed above, two fragments are necessary for NCL or expressed protein ligation: one bearing an N-terminal thiol and the other displaying a C-terminal thioester. Historically, C-terminal thioesters were prepared as derivatives of alkyl thioesters by in-situ neutralization driving Boc-SPPS. However, nearly half (43%) of the published ligations carried out since Kent's 1994 publication have made use of thioesters generated by Fmoc methods via either N-acylbenzimidazolinone (Nbz, 13%) or hydrazide[169] (30%) (Figure 25).[170] Each of these methods rely on a C-terminal functional moiety conveniently stable to Fmoc-SPPS conditions and readily able to convert to a thioester by a specific chemical reaction, while still attached to the resin in the case of Nbz or post cleavage in the case of hydrazide.
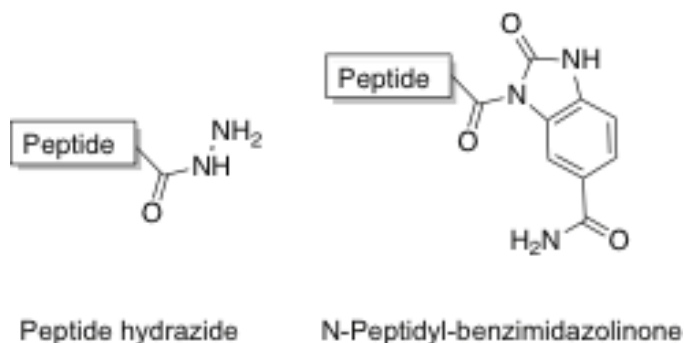
Figure 25. Structures of common thioester precursors prepared via hydrazide and N-acylbenzimidazolinone.

3.1.1   Zif268

Zif268 is a three finger, multidomain Cys_2His_2 zinc finger protein that regulates eukaryotic gene expression via recognition of specific DNA sequences, much like the Sp1 system in which the single domain studies from the previous chapter were carried out. Each finger adopts a $\beta\beta\alpha$ tertiary fold upon metal coordination and is stabilized by a conserved hydrophobic core. The crystal structure of Zif268 bound to a consensus DNA sequence is shown in Figure 26 with each

of the fingers extending along the major groove and primary base contacts residing along the exterior edge of each N-terminal helix.
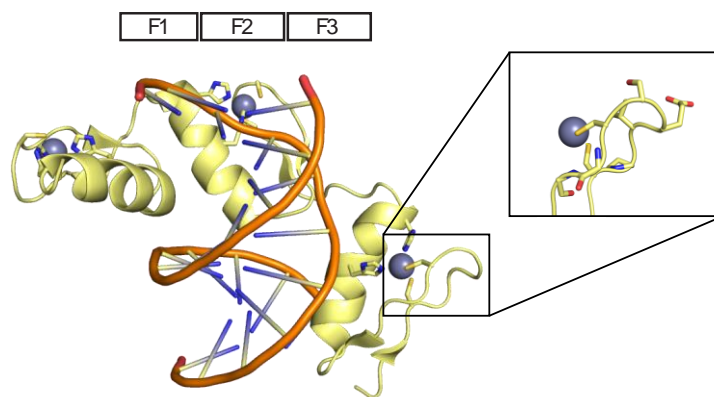
Figure 26. Zif268 bound to the consensus DNA sequence with a zoom of the finger 1 turn. (PDB: 1AAY)

Crucially, the first finger of Zif268 (Zif268-1) contains an extended 'irregular' type I CX_4C turn rather than the CX_2C turn in the Sp1-3 sequence (**Figure 27**). Given that each of the turn mimics applied in the single domain Sp1-3 prevented the formation of the $\beta\beta\alpha$ tertiary structure, the CX_4C turn in Zif268-1 presents an opportunity to examine the same turn alterations in a second structural context with a similar $\beta\beta\alpha$ tertiary fold.

Zif268-1
(PDB: 1AAY)
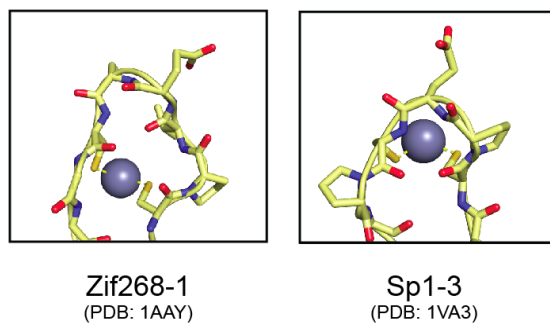
Sp1-3
(PDB: 1VA3)

Figure 27. Comparison of native sequence turns in Sp1-3 and Zif268-1.

Probing these turn alterations in Zif268-1 aids in determining if the lack of folding seen in turn variants of Sp1-3 can be correlated to the smaller (CX_2C) and more complex 'double turn' of

the Sp1-3 system specifically or if there is an additional aspect of the ββα zinc finger structure that has yet eluded us. It was our hope that the additional two residues buttressing the metal binding cysteines would allow sufficient space to ensure turn modifications do not perturb the precise tetrahedral coordination of metal by neighboring cysteines. The notion that the extended turn might allow for additional turn alterations is supported by work done by Thornton in which the CX₄C turn of zinc finger YY1-3 was successfully modified with a bicyclic beta turn mimetic.[171]

Previous work published toward total synthesis of Zif268 includes two approaches. In 1999, Dawson et al. demonstrated the utility of his recently published native chemical ligation methodology by ligating together three chemically prepared (via Boc SPPS) fragments of Zif268 at two Xaa-Cys ligation sites with thioesters derived from a trityl-associated mercaptopropionic acid-leucine (TAMPL) resin.[172] In later work, Fehr et al. took a semi-synthetic approach toward Zif268 with expressed protein ligation in which a chemically synthesized finger three of Zif268 was ligated to an expressed segment containing Zif268(1-64).[173] In each of these cases, the DNA binding capabilities of the final Zif268 product were assessed via CD spectroscopy and showed that both synthetically accessed Zif268 sequence did bind native DNA.

3.1.2   Design of a Synthetic Route to Heterogeneous-Backbone Zif268 Mimics

As Xaa-Cys ligations lack the need for any post ligation synthesis manipulations, the optimal potential ligation sites in the Zif268 sequence are: Gln^{35}-Cys^{36}, Ile^{38}-Cys^{39}, Ala^{63}-Cys^{64} and Ile^{66}-Cys^{67}. Given the known incompatibility of the hydrazine moiety with a C-terminal glutamine residue (due to an intramolecular cyclization that occurs between the hydrazide and the glutamine side-chain during SPPS)[169], along with the slow reaction time for beta branched residues such as isoleucine, the use of Nbz as the acyl donor for the Zif268 enables two more easily

57

obtainable ligation sites (Gln-Cys and Ala-Cys). Of note, the acyl hydrazide strategy would also be feasible at the Ala-Cys site.

Summarizing the proposed route to form the C-terminal thioester from an aminoanilide moitety, the Dbz residue is coupled to the resin prior to SPPS extension of the peptide sequence. After synthesis of the desired segment is complete, the entire resin-bound peptide is subjected to *p*-nitro- phenylchloroformate which, along with the subsequent basic conditions, results in a cyclic C-terminal N-acyl urea. The N-terminus of the peptide chain protected via a Boc group during this cyclization. Following cleavage from the resin and simultaneous Boc deprotection, the peptide-Nbz (N-acyl-benzimidazolinone) undergoes thiolysis to yield a thioester that is then available for ligation with a free thiol (Figure 28).
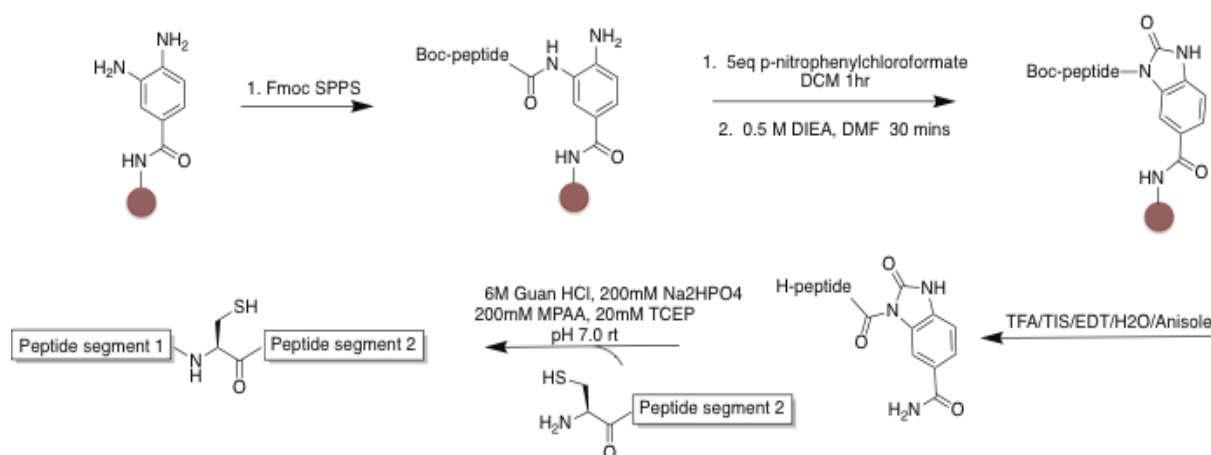
Figure 28. Generic scheme for activation and cyclization of Dbz residue followed by ligation.

While an N-terminal thiol is straightforward to generate via SPPS, in the context of expressed protein ligation additional considerations must be taken into account due to fact that the start codon for translation also codes for methionine. One method to produce a peptide segment

with an N-terminal thiol via expression takes advantage of the protease thrombin. This protease recognizes the consensus sequence LVPR|GS as well as similar sequences with a cysteine in place of the glycine. A sequence similar to this consensus sequence can be placed N-terminal to the cysteine that is intended to be ligated in NCL, thereby mimicking the consensus sequence. Incubation with thrombin, post expression, can effectively produce an N-terminal thiol (Figure 29).
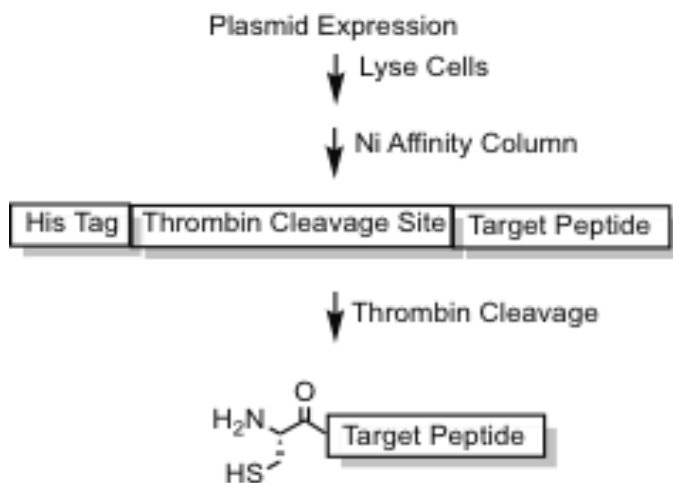
Figure 29. Scheme depicting protocol of post expression procedures to acquire the target sequence with an N-terminal cysteine.

3.2 HETEROGENEOUS BACKBONE VARIANTS OF ZIF268

Designs of heterogeneous-backbone variants of Zif268 focused on altering the backbone of the first finger while leaving the backbone along the second and third fingers undisturbed. Reminiscent of the Sp1-3 designs, two cross-strand N-methyl residues were placed along the β strand and four β^3-residues were placed along the exterior of the helix. Designing these alterations based on results from Sp1-3 allows us to test the hypothesis that the strategies for backbone

modification would be generally applicable. Four variants were designed with turn modifications Glu[9] and Ser[10] of the prototype sequence, D-ProGly, AibGly and Ornithine (**Error! Reference source not found.**). Each of the above alterations were placed with consideration of minimizing possible impacts related to 1) single domain folding, 2) ZF-ZF interactions and 3) ZF-DNA interactions.
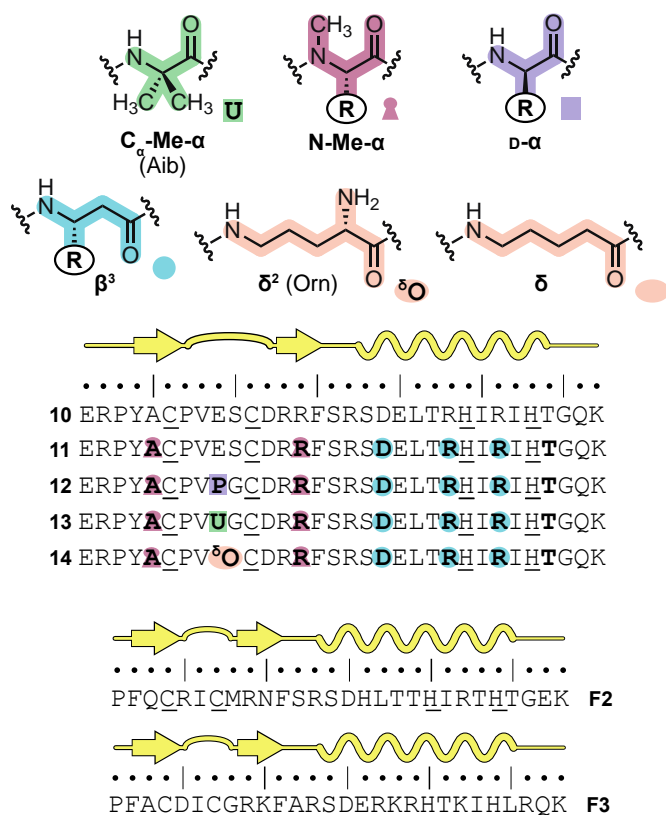
Figure 30. Sequence of multi-domain Zif268 along with backbone variants of finger 1.

While one of the overarching goals of designing backbone alterations for a multi-domain zinc finger remained focused on measuring the differences between backbone altered variant-DNA interactions and native sequence-DNA interactions, the first step along this path was to ensure that the backbone alterations placed in the first finger (Zif268-1) continued to support the native tetrahedral metal coordination environment as well as the $\beta\beta\alpha$ tertiary fold seen in natural Zif268-

60

1. Along with this, examining the backbone design principles in the context of a second zinc finger further demonstrates the generality of these principles.

Native and backbone modified E^1-K^{32} residues of finger one were synthesized and a primary UV-vis spectroscopy assay was carried out in the presence of excess Co^{+2} in 20 mM HEPES at pH 7.0 with ~ 100 µM peptide. The peptide samples used for each of these assays were not entirely pure according to MS analysis. Each sample contained varying degrees of single residue deletion sequences as well as other SPPS byproducts (*vida infra*). Never the less, the assay results are informative. The shape, intensity and position of diagnostic bands in the 500-700 nm region suggest the tetrahedral coordination of Co^{+2} for variants **10**, **11**, **12** and **13** (Figure 31) and provide compelling evidence for a native-like ββα tertiary fold in the context of the first finger.
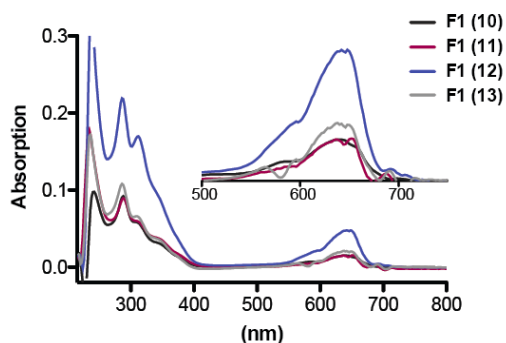
Figure 31. Absorption spectra with saturating concentrations of CoCl₂.

While confirmation with pure samples would be needed for any definitive conclusions, data supporting the fold of variant F1 (**11**) in which lessons from backbone alterations in the Sp1-3 context were applied to Zif268-1, suggests significant generality of design principals for backbone alterations developed in the Sp1-3 system. In addition, UV-vis spectroscopy data indicating the tetrahedral coordination of cobalt by variants F1 (**12**) and F1 (**13**) supports that both D-ProGly and AibGly turn modifications are tolerated in the ββα fold of Zif268-1. This is in direct

61

contrast to what was seen in turn-modified Sp1-3 variants, which lacked the propensity to bind cobalt. It is possible that the additional two residues spanning the CX_4C turn of Zif268-1 create enough distance to ensure the turn modifications do not perturb cysteine coordination. Likewise, it also possible that the mirror image type-I′ and II′ turn conformations promoted with the turn mimics are better tolerated in the irregular type I turn of Zif268-1 than in the double turn of Sp1-3.

3.3   BARRIERS MET IN ZIF268-1 SYNTHESIS

Upon attempting to produce larger quantitates of the E^1-K^{32} domain of Zif268-1 and the variants of Zif268-1, a number of barriers were encountered which prohibited synthesis of the necessary pure material to conduct UV-vis spectroscopy as well as additional analysis. While solid phase peptide synthesis (SPPS) is a well-studied procedure, it remains a time consuming and non-trivial aspect of studying heterogeneous backbone designs. Below, I survey key technical obstacles encountered in the synthesis of Zif268-1 along with efforts to overcome these issues.

One barrier met during the pursuit of additional Zif268-1 material was the co-elution of single-residue deletion products with desired material upon HPLC purification. In Figure 32a overlapping peaks in the HPLC trace can be seen; Figure 32b serves as a comparison in which one main peak can easily be isolated. The cause of single-residue deletions that coelute with the desired product appears to originate from difficult couplings to Cys and to Pro. The former is due to the steric hindrance of the trityl (Trt) side-chain protecting group and the latter, the decreased nucleophilicity of proline's secondary amine. Both these factors decrease the efficiency of coupling, which in turn increases the possibility that a percentage of the growing peptide chain

will not complete a coupling reaction. As the chain continues to grow with additional amino acid acylations the resulting sequence, which lacks complete residue coverage, can at times co-elute along with the desired sequence during HPLC purification. When coelution occurs, the inability to isolate the desired sequence in high purity prevents further biophysical analysis. Once an incomplete coupling has been identified either through analytical techniques such as a MALDI of a small batch cleavage of the growing peptide chain or upon final purification steps, future syntheses can incorporate a second coupling reaction at the challenging site as well as a capping reaction to intentionally terminate free amines for which two couplings reactions were insufficient for acylation. Both of these approaches, when applied to each difficult coupling, were successful in ensuring challenging couplings went to completion and chains that did not complete the coupling reaction did not coelute upon HPLC purification.
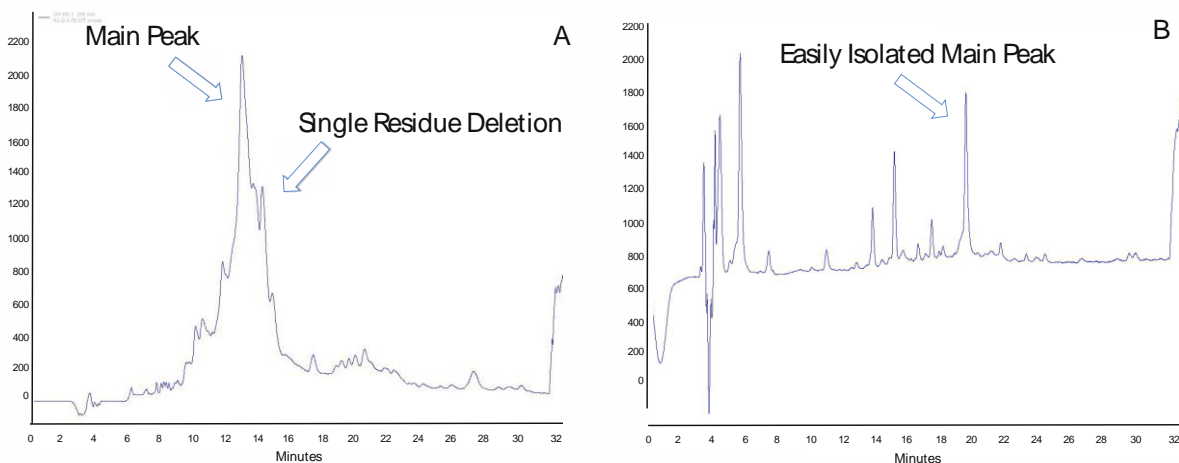
Figure 32. HPLC traces comparing the overlapping elution of peaks (A) to a single peaks able to be isolated with high purity (B).

An additional barrier encountered was the production of a significant byproduct attributed N to O acyl migration during the TFA cleavage of peptide from resin. HPLC traces of a synthesis

in which an N to O acyl migration occurred can be seen in Figure 33 as a shoulder along with a second shoulder due to a single residue deletion. The proposed mechanism of this side reaction is shown in Figure 35. Acyl shifts can occur at serines or threonines in Fmoc based SPPS; the likelihood of occurrence is dependent on the specific peptide sequence.[174] In the presence of typical α-residues, the hydroxyl side chain of threonine may attack the carbonyl of the previous N-terminal residue cleaving the peptide chain at the amide bond between Thr and the residue N-terminal to Thr. This cyclization step is reversible and the conversion back to the native backbone is far faster than the cleavage pathway. In the context of heterogeneous backbones, the extended backbone of a β^3-threonine introduces an additional irreversible path leading towards N to O acyl shifts in which the hydroxyl side chain of threonine may be prone to attack the intraresidue carbonyl. Due to the additional carbon along the backbone, this forms a five membered lactone and results in cleaving all residues C-terminal to the β^3-threonine. The resulting section of the desired peptide can be identified via MALDI. To reduce the occurrence of N to O acyl shifts in Zif268-1, a native α-threonine was returned to the sequence in place of the β threonine. Upon re-synthesis with the native α-threonine in this position, chain fragmentations due to N to O acyl shifts were not indicated in the mass spectrometry analysis.
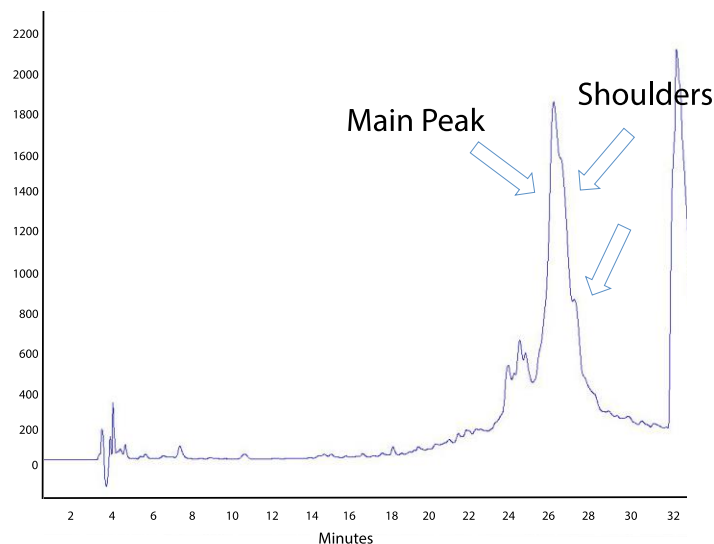
Figure 33. HPLC trace showing two shoulders adjacent to the main peak. Upon MALDI ionization, molecular weights indicated one shoulder was due to an N to O acyl migration, whereas the other shoulder is due to a single residue deletion.

Despite resolving issues with both single residue deletions and N to O acyl shifts, an additional barrier in the synthesis of Zif268-1 remained problematic: base catalyzed aspartimide formation. This side reaction is known to occur during Fmoc deprotection and can be identified by an -18 mass (present in MALDI analysis, Figure 34).[175] The reaction occurs through the nucleophilic attack of the aspartic acid side chain by the nitrogen atom attached to the carbonyl (Figure 35). This results in an apartimide ring that is then vulnerable to a nucleophilic attack of the imide ring by piperidine. In the each of the Zif268-1 sequences the initial aspartimide ring was indicated by mass spectrometry; however no additional piperidine attack was observed. A known solution to aspartimide formation is to reduce the basicity of the deprotection solution. This minimizes the likelihood that piperidine will deprotonate the backbone amide nitrogen, which decreases the likelihood for attack of the side chain ester. 0.1 M 1-hydroxybenzotriazole (HOBt) can be added to the Fmoc deprotection solution to this end.[175] In the case of the Zif268-1 variant

65

containing the D-ProGly turn (**12**), the addition of HOBt to the deprotection solution was met with success in minimizing the aspartimide formation observed during formation. However, aspartimide formation remained problematic in the wild-type Zif268-1 sequence even in the presence of HOBt.
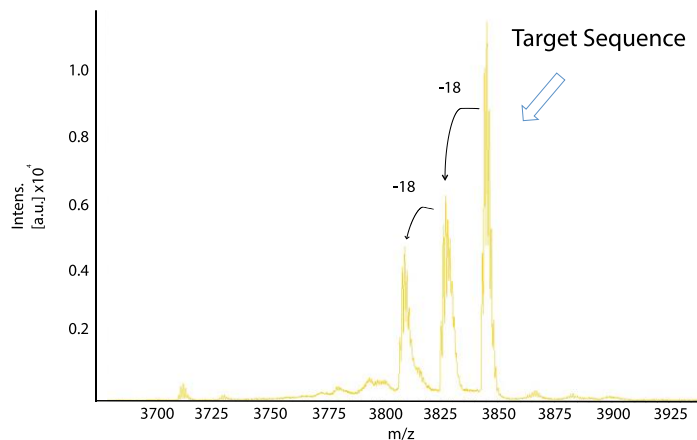
Figure 34. MALDI-MS data showing target sequence along with -18 and -36 masses.

Several alternative pathways for future mitigation of aspartimine formation in the Zif268-1 wild-type sequence exist. Purchasing fresh HOBt is one example. HOBt is light sensitive and the material used in the previous reactions was multiple years old. Additionally, ethyl 2-cyano-2-(hydroxyimino)- acetate (Oxyma) has also been shown to reduce aspartimine formation; Oxyma also has the added benefit of lacking the explosive potential of HOBt.[176,177] The addition of 5% v/v formic acid to the deprotection solution has also been shown to mitigate aspartimide formation.[177] Alternatively, aspartic acid with sterically demanding protecting groups such as 3-methylpent-3-yl (OMpe)[178] and β-2,3,4-trimethyl-pent-3-yl[179] can also be utilized. Future synthesis of the wild-type Zif268-1 sequence could incorporate one or several of these pathways to mitigate aspartimide formation.
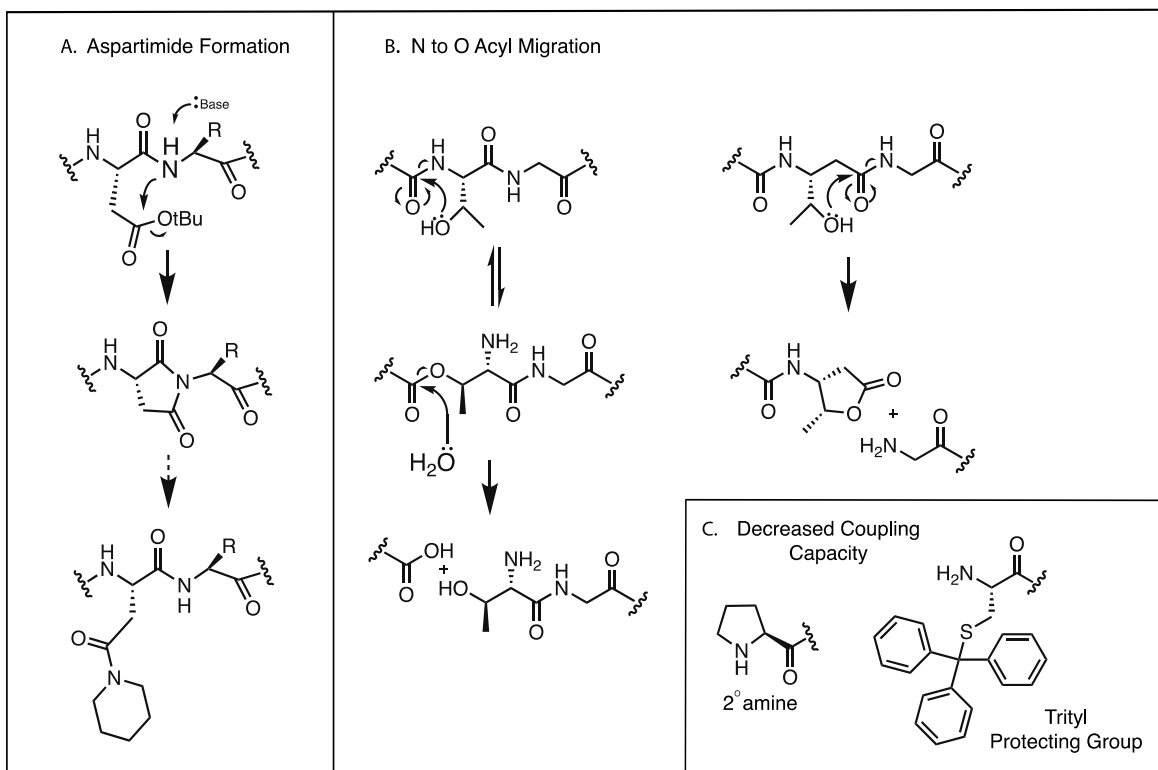
Figure 35. Mechanisms and structures for select barriers in SPPS a) Aspartimine formation b) potential pathways for N to O acyl migrations in both ☐ and ☐ backbones c) Structures underlying the steric hindrance of the trityl protecting group and the secondary amine of proline; each of which can lead to decreased coupling capacity and single residue deletions.

At this point, it was determined that an efficient synthesis of Zif268-1 necessitated 1) double coupling/capping after each cysteine and proline residue to prevent single residue deletions, 2) a native α-backbone threonine rather than β threonine at residue 29 to avoid N to O acyl shifts, 3) 0.1 M HOBt in the deprotection solution to suppress aspartimide formation in sequences with D-Pro-Gly as the turn and 4) continued efforts to mitigate aspartimide formation in the wildtype Zif268 sequence.

Steps 1 through 3 were each shown to minimize or eliminate the given barrier while several possible approaches to step four yet exist for optimization of native Zif268-1. Despite the strides

taken to determine these synthesis parameters, sufficient material was not produced in time to carry studies into the next steps, which were planned to focus directly on confirming the tertiary structure of both the backbone altered and α backbone Zif268-1 finger through 2D NMR experiments. Confirming the extent to which the structure of the first finger of Zif268 was altered or not altered would then lead to the next step: synthesizing the full 89 residue three-domain sequence with backbone alterations in the first finger with the overall aim of determining the effect of backbone alterations on DNA binding via circular dichroism (CD).

3.4 MODULAR APPROACHES TO THE MULTI-DOMAIN SEQUENCE

In parallel to efforts aimed at optimization of the synthesis of Zif268-1 and variants, we sought to establish a modular approach to the full three-finger Zif268 protein. Through this work our goal was to expand the length of potential sequences amenable to heterogenous design principles and to determine how backbone alterations impact Zif268 - DNA binding in comparison to the wild type sequence. We pursued two courses: one based on total chemical synthesis through iterative NCL reactions and the other based on expressed protein ligation involving synthesized and bacterially produced fragments of the full length protein.

3.4.1 Approach A

ERPYA**C**PVES**C**DRRFSRSDELTR**H**IRI**H**TGQKPFQ**C**RI**C**MRNFSRSDHLTT**H**IRT**H**TGEKPFA**C**DI**C**GRKFARSDERKR**H**TKI**H**LRQK

ERPYA**A**CPV**P**G**C**DR**R**FSRS**D**ELTR**R**HI**R**I**H**T**T**GQKPFQ**C**RI**C**MRNFSRSDHLTT**H**IRT**H**TGEKPFA**C**DI**C**GRKFARSDERKR**H**TKI**H**LRQK

ligation site = ⁞

NCL

E1-Q35 (or) E1-Q35

R37 - K88

E1 - K88
(and)
E1-Q35 - K88
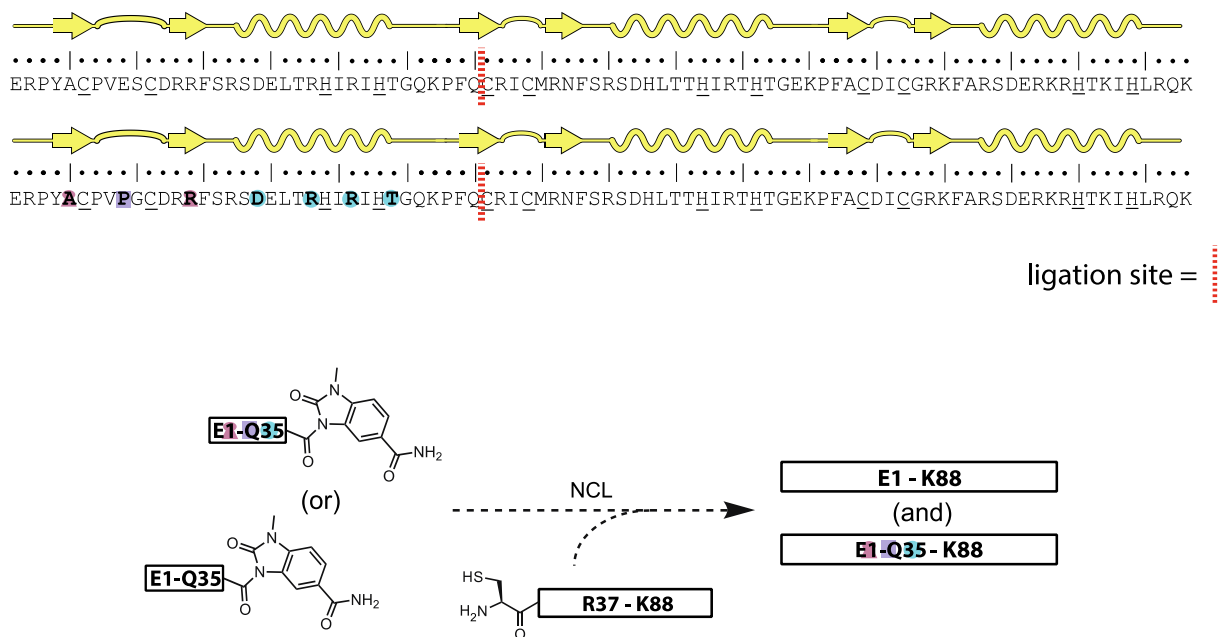
Figure 36. Schematic of "Approach A" which includes and expressed C-terminal segment combined with a chemically synthesized N-terminal segment, ligated together at one single site.

The initial design toward the 88 residue Zif268 sequence, "Approach A", aimed to utilize a single native chemical ligation site at the intersection of residues Q^{35} and C^{36}. Fragment E1-Q^{35} was to be synthesized through Fmoc-SPPS and on-resin generation of *o*-aminoaniline N-acylurea (Dbz) moiety, enabling the formation of a thioester at the C terminal Q^{35} (Figure 36). Residues C^{36}-K^{88} were to be produced via protein expression. We envisioned this modular approach would enable multiple backbone-modified variants to be synthesized and ligated in parallel to a single batch of expressed native C^{36}-K^{88} thus creating an efficient path towards production of several multi-domain variant designs. While several ligations sites were possible, position Q^{35} - C^{36} was chosen due to the ease of ligating directly with a C-terminal cysteine

(eliminating the need for post synthesis desulfurization steps) and with an eye to minimize additional steps at the thioester. The sterically unencumbered C-terminal Gln should also give a faster reaction than would occur with beta branched residues such as Ile, Thr and Val.[180]
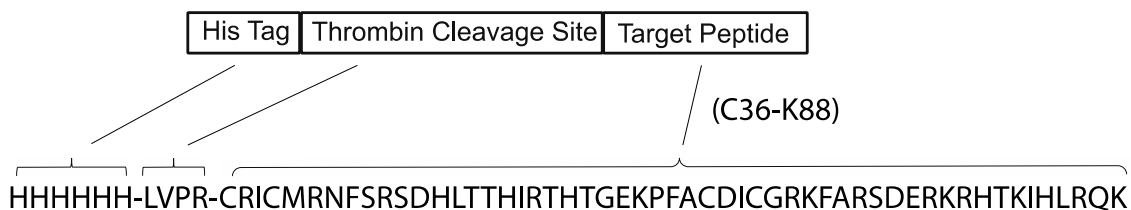
Figure 37. pET vector design with histidines tag, cleavage site and C36-K88.

A pET vector containing a histidine tag for nickel column purification along with a thrombin cleavage site and Zif268 residues C36 through K88 was designed and purchased in a stab culture of *E. coli* strain Stbl3 (Figure 37). Expression in BL21(DE3) *E. coli* and subsequent Ni column purification of this "His$_6$-ThrombinCleavage-C36-K88" peptide sequence resulted in a large batch of protein necessitating little optimization apart from increasing the concentration of imidazole to outcompete with the zinc finger during nickel column purification. SDS-PAGE after initial expression and purification with a nickel column is shown in Figure 38 along with a second SDS-PAGE showing an optimized protocol for nickel column purification consisting of an imidazole gradient in place of static imidazole concentration. With the high affinity zinc fingers have for divalent metals, additional imidazole was advantageous to maximize elution of the target sequence. As is visible in B2, even after nine elution runs with a maximum of 0.8M imidazole, a great deal of the target sequence can be seen eluting in the tenth elution with 6M imidazole. Future nickel column purifications would benefit from increased gradients of imidazole due to the additional target sequence peptide still eluting in E10 elution (Figure 38 B2).
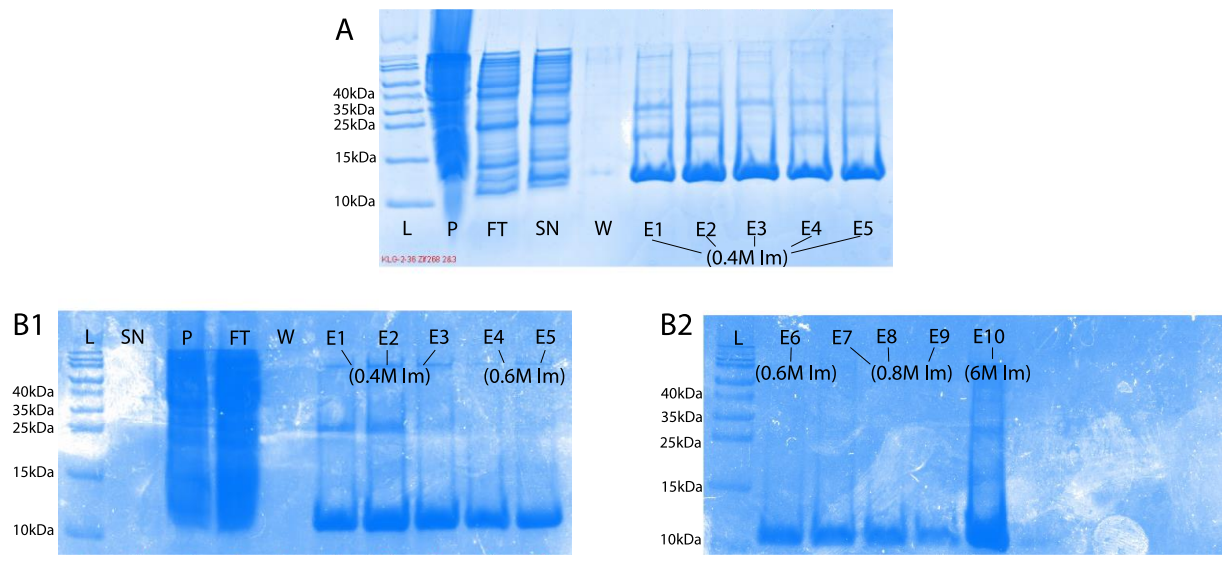
Figure 38. Two examples of SDS-PAGE run post expression of His₆-Thombincleavage-C36-K88. A) A static concentration of 0.4M imidazole is used to elute target sequence from nickel column. B1 and B2) A concentration gradient of imidazole allows increasing amounts of the target sequence to elute from the nickel column. L=ladder, SN=supernatant, P=pellet, FT=flow through, W=wash, E# = elution with Imidazole.

Elutions 1 through 9 were collected and analyzed with HPLC (Figure 39). One main peak can be seen eluting around 30% B. Attempts to confirm the molecule though MALDI-MS, were unsuccessful, as ionization proved challenging. Nonetheless, the appropriate MW observed in the gel match expectation for the construct. Thus, we proceeded to examine steps toward thrombin cleavage of the expression tag.
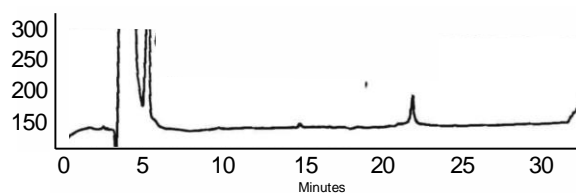
Figure 39. Analytical HPLC trace of collected elutions 1 through 9

71

In effort to determine optimal conditions for proteolytic cleavage of the expression tag, a large number of small-scale reactions were run with varied thrombin concentration, peptide concentration, temperature and reaction time. The final conditions found to be most successful were 10°C, 50 mM Tris HCl, pH 8, 100 mM NaCl, 5 mM $CaCl_2$, 20 units thrombin per 1 mg peptide. Reactions were monitored via HPLC with the mass of both the uncleaved peptide and the histag/thrombin cleavage segment able to be identified via MALDI. It is likely that the large peak growing in at elution time 28 minutes in Figure 40 is the desired C^{36}-K^{88} product yet no definitive confirmation through mass spectrometry could be made either by MALDI or by ESI.
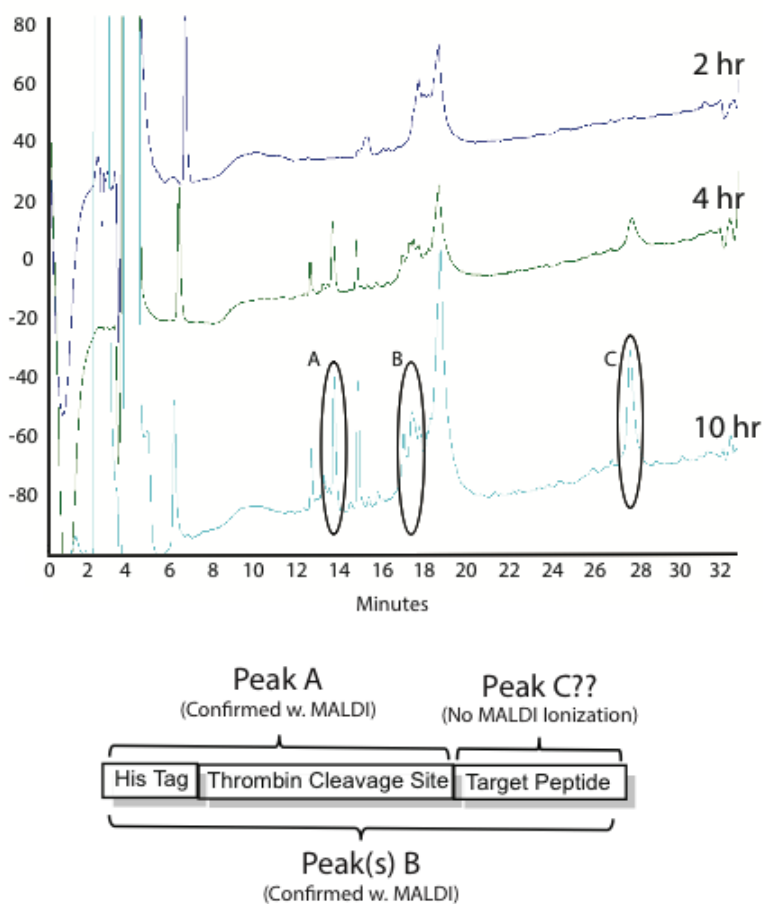
Figure 40. Analytical HPLC traces post Thrombin cleavage at designated time points.

Next steps to confirm the identity of the presumed desired product would include SDS-PAGE to confirm the relative molecular weight of the presumed peak. Though given the similarities in weight, the starting material (MW = 7818 g/mol) and the target sequence (MW = 6380 g/mol) may be difficult to differentiate through SDS-PAGE. An additional test reaction between the presumed peak and a compound containing a thioester followed by HPLC purification would also aid in determining if an N-terminal cysteine is present in the peak eluting at 28 minutes. If no change in retention time is present then it is likely the peptide eluting at 28 minutes does not contain a free cysteine and therefore is not the desired C^{36}-K^{88} product.

3.4.2 Approach B

Given the difficulties encountered with the thrombin cleavage in Approach A, an alterative strategy toward the multi domain zinc finger protein, "Approach B", was designed. While expressed protein ligation was a fundamental aspect of Approach A, Approach B aimed to chemically synthesize three separate segments of the multi-domain zinc finger via SPPS and utilize NCL at two separate ligation sites: Q^{35}/C^{36} and A^{63}/C^{64} (Figure 41).

73
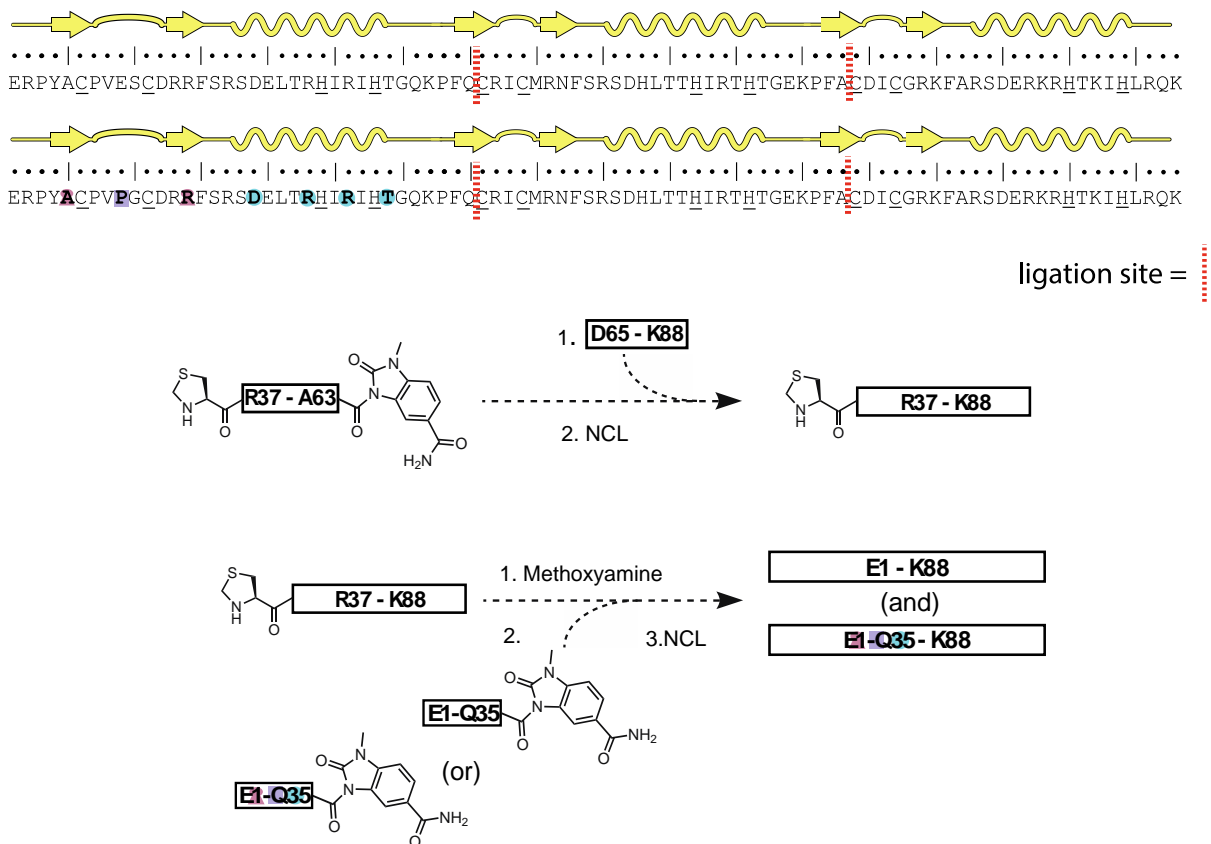
Figure 41. Schematic for "Approach B" in which three segments are chemically synthesized and ligated together at two ligation sties.

Two peptide segments with C terminal thioesters were necessary for Approach B: one C-terminal to Q^{35} and the other C-terminal to A^{63}. Both thioesters necessary for NCL were formed via a Dbz residue at the C-terminus, much like what was done in Approach A. Rather than utilizing the Dbz residue to form the thioester, a modified Dbz residue was used: N-Me Dbz. The N-methyl group on the *para*-amino group of the residue significantly decreases the sensitivity for undesired acylation while retaining enough reactivity at the remaining amine to couple the following residue with HATU in anhydrous DMF.[181] An additional optimization in synthesis toward a thioester-bearing segment included a glycine spacer coupled to the resin prior to the dbz residue. This

glycine minimizes steric hindrance in the resin loading reaction. Upon ligation, the glycine is lost with the Dbz moiety as equilibration is reached with the added thioesters.

To facilitate a one-pot sequential ligation, C^{36}-A^{63}-NMeDbz was prepared with an N-terminal 1,3-thiazolidine-4-carboxo (Thz) protecting group.[182,183] Following ligation with D^{65}-K^{88}, the free thiol should be easily unmasked with methoxyamine to enable the extension in the N-terminal direction (Figure 42).
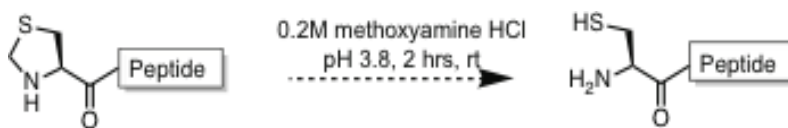
Figure 42. Transformation of Thz to a free cysteine with methoxyamine

Given the methods laid out above, Approach B was pushed forward with the synthesis and TFA cleavage of E^1-Q^{35}-Dbz-G on rink resin with Fmoc SPPS. Microcleavages at regular intervals during synthesis confirmed successful couplings up to F^{15}. Further microcleavages after complete synthesis and before Dbz cyclization showed promising yet inconclusive masses in the region of the expected weight (Figure 43a). A subsequent analytical HPLC trace post-cyclization of Dbz resin also showed a promising main peak followed by a large amount of overlapping material (Figure 43a). Neither MALDI nor ESI mass spectrometry of the isolated peak resulted in ionization (Figure 42b). Thus identification of the final E^1-Q^{35}-Dbz-G was not confirmed.
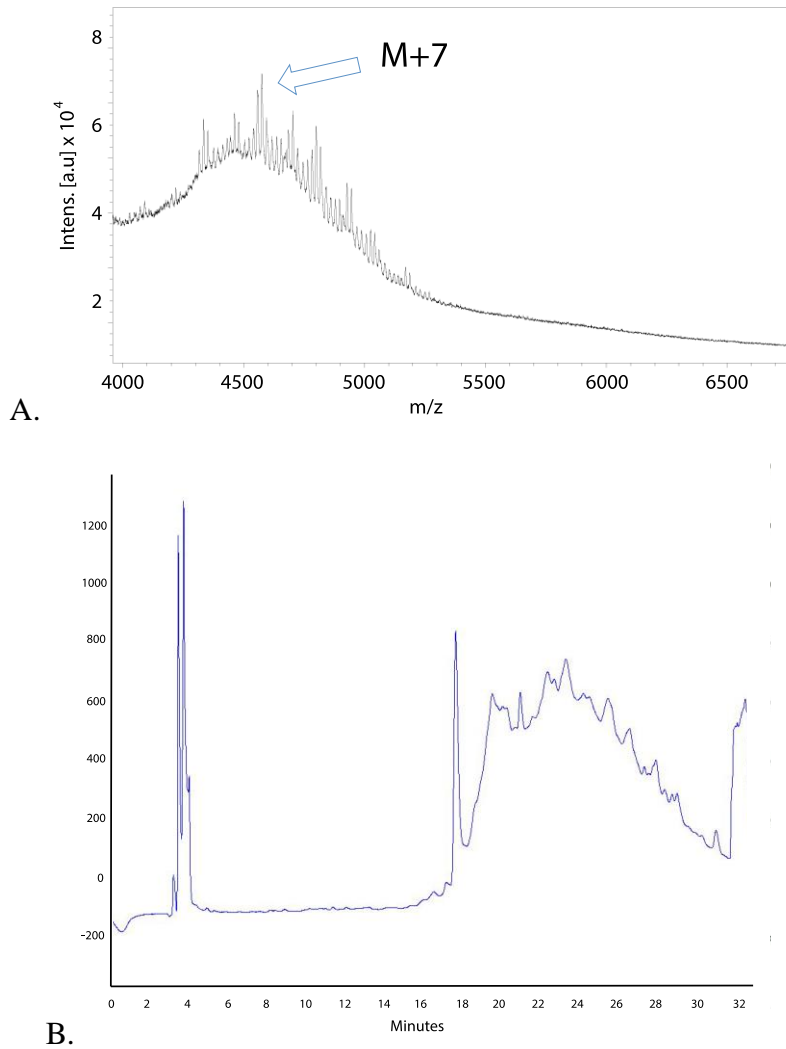
Figure 43. A) Inconclusive MALDI data after complete synthesis of E1-Q35-Dbz-G .B)Analytical HPLC of

E1-Q35-Dbz-G

Next steps for progress in the complete synthesis of the three finger Zif268 include improved confirmation of E^1-Q^{35}-Dbz.

3.5   CONCLUSIONS

Progress on several fronts suggested the synthesis of the multi-domain zinc finger foldamer was on the right track. Yet, unanticipated challenges, detailed above, limit us to a discussion of barriers remaining rather than of a decisive test of the hypothesis of multidomain zinc finger protein mimicry. The difficulty producing a multi-domain zinc finger through "Approach A" was primarily due to complications with confirming the identity of the target sequence after thrombin cleavage. The peak eluting at 28 minutes showed promise (Figure 40). Though, it would need to be confirmed through optimization of MS enabled methodology or a test reaction with a compound containing a thioester. Confirmation of the target sequence was also the limiting step for "Approach B". In this case it was E^1-Q^{35}-Dbz that was potentially synthesized but unable to be confirmed.

The preliminary spectroscopic data for Zif-268 finger 1 variants (Figure 31) do, however, suggest that our initial design principles for backbone alterations in Cys_2His_2 zinc fingers can be applied to a second native zinc finger sequence. This demonstrates the generality of the initial design principles. The tetrahedral coordination of metal by variants with backbone alterations in the turn region also further validates the design principals set out in Chapter 2, as these proved challenging in the first Sp1-3 sequence due to the tight turn. These variant's ability to mimic the structure and binding of native Zif268-1 suggests that if Approach A or B of the total synthesis with backbone alterations in the first finger were to be complete, there is high potential that the three fingers would each fold into a native like tertiary structure. The DNA binding propensity could then be determined with a simple CD scan.

3.6   METHODS

Peptide Synthesis and Purification All peptides were prepared using microwave-assisted Fmoc solid-phase peptide synthesis techniques on NovaPEG Rink Amide Resin. Microwave couplings were performed at 70□C for 4 minutes with N-□-Fmoc-protected amino acids (6 eq), HCTU (6 eq), DIEA (10 eq) in NMP for all amino acids except Fmoc-Cys(Trt)-OH, Fmoc-His(Trt)-OH and Fmoc-Arg(Pbf)-OH which were performed at room temperature. Residues coupled to N-Me were coupled with 6 eq PyAOP and 10 eq DIEA in NMP at 70□C for 4 minutes. The pseudoproline dipeptide Fmoc-Leu-Ser-OH was utilized to prevent aggregation at residues L17 and S18 and was coupled with PyAOP (6 eq), DIEA (10 eq) in NMP at room temperature for one hour (Figure 25). Fmoc deprotection was achieved with 20% 4-methylpiperidine in DMF at 80□C for 2 minutes. Resin was washed 3 x 1 minute with DMF after each coupling and deprotection step. Figure 25. Structure and placement of LS pseudoproline dipeptide. Cleavage from resin was achieved with a solution of TFA/EDT/H_2O/TIS (94:2.5:2.5:1) over a period of 3 hours followed by precipitation in cold ethyl ether. Pellets were collected via centrifugation and dissolved in solutions of 0.1% TFA in water and acetonitrile. Crude peptide was purified on C_{18} HPLC columns using 0.1% TFA with a gradient of acetonitrile. Purified peptides were identified with MALDI mass spectrometry.

Preparation and Storage of Stock Solutions. Peptides were stored as lyophilized powders under vacuum to reduce contact with oxygen. Peptide stock solutions were prepared fresh prior to each measurement. Metal stock solutions were stored at room temperature for a maximum of 3 weeks. Solvents used for sample preparation were purged by freeze-pump-thaw cycles and/or argon

bubbling. The absence of rigorous efforts to exclude oxygen in solution preparation and manipulation led to rapid peptide oxidation on a time scale that varied depending on sequence, pH and temperature in the range of minutes to hours. Concentration of peptide solutions used in biophysical experiments were quantified by DTNB assay (aka Ellman's assay) according to the procedure given by Thermo Scientific.[147] Cobalt and zinc stock solution concentrations were determined either by ICP-AES, ICP-MS, or spectrophometrically with Zincon monosodium salt according to published methods.[148]

Peptide Expression and Purification. The N-terminal 6xHis-tagged –Thrombin Cleavage-Zif268 (C36-K88) expression construct was obtained from Cyagen Biosciences. The plasmid was transformed into E. Coli BL21 codon plus (DE3) competent cells (Invitrogen) using a pET vector. A single colony was picked up and grown overnight at 37 °C in 10 mL of Luria-Bertani (LB) broth in presence of 50 μg/mL ampicillin and 35 μg/mL chloramphenicol. The culture was diluted 100-fold and allowed to grow at 37 °C to an optical density (OD_{600}) of 0.8, and protein expression was induced overnight at 18 °C with 1 mM IPTG in an Innova 44® Incubator shaker (New Brunswick Scientific). Proteins were purified as follows: harvested cells were resuspended in 15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, 25 mM imidazole, Lysozyme, DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700), and centrifuged at 13000 rpm for 40 min at 4 °C. The soluble extracts were subject to Ni-NTA agarose resin (Thermo) according to manufacturer's instructions. After passing 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 25 mM imidazole), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 400 – 800 mM imidazole. The protein concentration was determined using Bradford assay kit

(BioRad Laboratories) with BSA as a standard. The concentrated proteins were stored at -80 ºC before use.

UV-vis Spectroscopy. UV-vis spectra were acquired at 20 °C on an HP 8452A Diode Array Spectrophotometer in a 1 cm path length cuvette. Solutions contained 100 μM peptide, 20 mM HEPES (pH 7.0), 1.8 mM TCEP with 1.3 equivalents of Co^{2+}. Analytical HPLC prior to and post each experiment confirmed peptide was fully reduced.

Dbz Resin Loading. Dawson Dbz NovaSynTGR resin was swelled in DCM for 60 minutes and washed with DMF. Fmoc deprotection was achieved with 20% 4-methylpiperidine in DMF at 80□C for 2 minutes**.** A room temperature coupling was performed overnight with N-□-Fmoc-protected amino acids (6 eq), HATU (6 eq), DIEA (10 eq) in anhydrous DMF. Resin loading was measured via UV-vis with 4-methylpiperidine. An alternative procedure included coupling an N-MeDbz linker to a rink amide resin preloaded with glycine. Coupling to the N-MeDbz linker was preformed overnight with N-□-Fmoc-protected amino acids (6 eq), HATU (6 eq), DIEA (10 eq) in anhydrous DMF. SPPS proceeded as stated above. rThe final residue of C-terminal Dbz peptides was introduced as a Boc protected amino acid.

4.0   BROADER OUTLOOK

The preceding discussion focuses on and contributes to the development of design principles for increasingly diverse peptide-like entities. Potential applications of such entities fall overwhelmingly in the therapeutic sphere. Below I provide a brief overview of the chemical space of current therapeutics as well as the potential for and barriers to market entry in the case of both native and backbone-modified peptides.

Small molecules have dominated the therapeutic market for the past century.[184,185] Their ideal characteristics of membrane permeability, metabolic stability, oral bioavailability and a relatively low production cost (in comparison to other types of drugs) have allowed for remarkable advances in medicine.[184] Yet, small molecules can suffer from low target selectivity that results in a high off-target binding and manifests as side effects. Due to their small size, they also fall short when it comes to targets with large, flat and nonpolar surfaces – a class that is rapidly growing.[186]

Advances in recombinant protein expression, purification and characterization in the past 30 years have gradually increased the viability of proteins as drugs and opened the door to harnessing their remarkable target selectivity and impressive ability to bind precisely the types of surfaces that had long eluded small molecules.[187] Innovations in delivery systems have mitigated previous concerns related to delivery. Continued success in transciptomics, proteomics and genomic technologies has paved a path towards a greater emphasis on proteins as drug candidates.[186,188] As of 2013, success rates of bringing a biologic to market was about twice that of a small molecule.[187]

The two main established classes of successful drugs are then small molecules, defined by molecular weight of less than 500 g/mol, and biologics, defined by a molecular weight of over

5000 g/mol.[189] A significant gap exists between these two classes. Likewise, a significant gap exists in the types of targets each of these are able to bind. It is estimated only 10% of proteins have pockets deep enough to be bound by small molecules. Further, only 10% of possible targets reside in the extracellular space accessible by biologics.[190] This still leaves the vast majority of drug targets as potentially "undruggable" with current chemical classes.

Peptide therapeutics offer the potential to fill the gap in molecular weight between biologics and small molecules while also touting an impressive combination of attributes that align well with the increasing numbers of targets with large, shallow and nonpolar binding sites. Many protein-protein interfaces (PPI) can even now be classified as "druggable" due to the availability of increasing numbers of peptide-based drugs.[191] In addition, peptides with less than 40 amino acids have the added benefit of being approved as small molecules through the 505 pathway with Center for Drug Evaluation and Research (CDER) rather than the 351 pathway with Center for Biologics Evaluation and Research (DBER) throughout the FDA approval processes. This avoids the additional regulatory hurdles imbedded in the biological pathway and the aggressive biosimilar market landscape in the US. Many successful peptide therapeutic candidates have already seen success on the market, and the residue length of approved peptide therapeutics has been increasing each decade (Figure 44).[192]
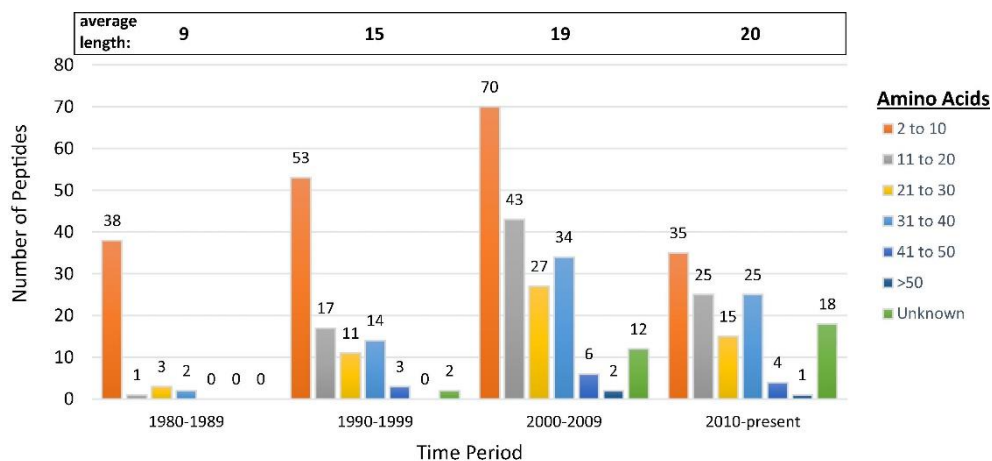
Figure 44. Length of peptides entering clinical development, by decade. Peptides with unknown length were not included in the average length calculation. Used with permission from Lau, J.L. and M.K. Dunn, *Therapeutic peptides: Historical perspectives, current development trends, and future directions.* **Bioorganic & Medicinal Chemistry, 2018.** 26(**10**): p. 2700-2707.

The potential for foldamers, as opposed to natural sequences (as was mostly referred to above), to reach the clinical market is still, at this point, a considerable challenge. Even natural sequences are limited by concerns of short plasma half-life, challenging bioavailability, immunogenicity, a potentially low yielding synthesis and the need for substantially higher financial investments (compared to small molecule candidates).[184,193] While foldamers often share many of these same challenges, they also embody the potential to address them if appropriately designed. The open questions of pharmacokinetics and ADME profiles that have yet to be examined for most foldamer designs are additional hurdles for the foldamer field but will be essential for continued progress toward any medicinal applications.[187]

Despite the above open questions, the therapeutic potential of foldamers and other unnatural peptides is a subject of much excitement,[112,184,194-196] and peptides with unnatural modifications are continuing to enter the therapeutic market (Table 5).

Table 5. Selective peptide therapeutics approved by the FDA between 2011 and 2017[196]

Generic Name	Modification
Macimorelin acetate	2-Aminoisobutyric acid, D-Trp, D- gTrp-CHO
Abaloparatide	2-Aminoisobutyric acid
Etelcalcetide	N-Terminal acetylation; C-terminal amidation
Plecanatide	Cyclization by 2 disulfide bridges
Lixisenatide	C-terminal conjugation of polylysine
Gallium Ga 68 dotatate	Cyclization, D-amino acids
Insulin degludec	Lipidation, amino acid elimination
Peginesatide	PEGylation

Progress to overcome limitations of peptide and protein therapeutics is continuing to advance. For example, the effect of multiple α to β^3 modifications within a peptide epitope was analyzed against the immune response for the parent sequence. It was found that antibody recognition raised against the parent sequence was suppressed by the periodic α to β^3 modification.[197] Work analyzing the immunogenicity of a D-protein antagonist of natural vascular endothelial growth factor A (VEGF-A) also found that while the native sequence generated strong immune response in mice, the variant was nonimmunogenetic.[198] These trends for immunogenicity of $\beta3$ and D-residue substituted peptides are promising. More data will be necessary to determine the extent to which such observations will hold in other systems. In the meantime, studies to predict which individuals will have a immune response to a particular drug are ongoing.[199] Research into expanding the range of potential residues capable of inclusion through ribosomal synthesis is also an active area of pursuit.[200,201] Still, it is estimated that of the ~70 peptide drugs (defined as less than 40 amino acids by the FDA) already on the market, about 85% are produced via synthetic rather than recombinant methods – thereby minimizing the common concern that foldamers could not be produced on a commercially viable scale.[184]

Furthermore as peptides enter the market, a trend exists towards "analog peptides" which

are based on natural sequences and contain slight alterations which improve their pharmacology profile.[192] In 2017, for example, the FDA approved 46 drugs. 12 were biologics and 34 new chemical entities (NCE). Of the 34 NCE, six were peptides which spanned the spectrum of structural complexity.[202] Five of these are produced with SPPS, and three contain an Aib residue.[202] Yet despite this, the structures of these peptides are essentially linear and nowhere near the complexity of a tertiary structure.

The number of potential drug targets has been estimated to be well over 10,000, and a recent article cites the characterization of over 14,000 PPI in the human body.[203] The reality that less than 500 targets are addressed by FDA approved drugs reveals a significant need for new therapeutic entities.[185] Several review articles detail progress towards the therapeutic space for foldamers and other unnatural peptides.[196,204,205]

In light of above, the efforts of the foldamer research community to mimic increasingly complex peptide structures embodies a wealth of knowledge that holds great potential if applied towards pressing barriers. I will continue to watch the field with excitement and high hopes for progress.

BIBLIOGRAPHY

1.   Gellman, S. H., Foldamers: A Manifesto. *Accounts of Chemical Research* **1998**, *31*, 173-180.
2.   Guichard, G.; Huc, I., Synthetic foldamers. *Chemical Communications* **2011**, *47*, 5933-5941.
3.   Martinek, T. A.; Fulop, F., Peptidic Foldamers: Ramping Up Diversity. *Chemical Society Reviews* **2012**, *41*, 687-702.
4.   Horne, W. S.; Gellman, S. H., Foldamers with Heterogeneous Backbones. *Accounts of chemical research* **2008**, *41*, 1399-1408.
5.   Johnson, L. M.; Gellman, S. H., α-Helix Mimicry with α/β-Peptides. *Methods in enzymology* **2013**, *523*, 407-429.
6.   Gellman, S. H., Foldamers: A Manifesto. *Accounts of chemical research* **1998**, *31*, 173-180.
7.   Pace, C. N.; Scholtz, J. M., A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* **1998**, *75*, 422-427.
8.   Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K., Peptoids: a modular approach to drug discovery. *Proceedings of the National Academy of Sciences* **1992**, *89*, 9367-9371.
9.   Kritzer, J. A.; Stephens, O. M.; Guarracino, D. A.; Reznik, S. K.; Schepartz, A., β-Peptides as inhibitors of protein–protein interactions. *Bioorganic & Medicinal Chemistry* **2005**, *13*, 11-16.
10.   Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H., β-Peptide Foldamers: Robust Helix Formation in a New Family of β-Amino Acid Oligomers. *J Am Chem Soc* **1996**, *118*, 13071-13072.
11.   Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D.; Oberer, L.; Hommel, U.; Amstutz, R.; Widmer, H., Probing the Helical Secondary Structure of Short-Chain β-Peptides. *Helvetica Chimica Acta* **1996**, *79*, 2043-2066.
12.   Armand, P.; Kirshenbaum, K.; Goldsmith, R. A.; Farr-Jones, S.; Barron, A. E.; Truong, K. T. V.; Dill, K. A.; Mierke, D. F.; Cohen, F. E.; Zuckermann, R. N.; Bradley, E. K., NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains. *Proceedings of the National Academy of Sciences* **1998**, *95*, 4309-4314.
13.   Kirshenbaum, K.; Barron, A. E.; Goldsmith, R. A.; Armand, P.; Bradley, E. K.; Truong, K. T. V.; Dill, K. A.; Cohen, F. E.; Zuckermann, R. N., Sequence-specific polypeptoids: A diverse family of heteropolymers with stable secondary structure. *Proceedings of the National Academy of Sciences* **1998**, *95*, 4303-4308.
14.   Checco, J. W.; Gellman, S. H., Targeting Recognition Surfaces on Natural Proteins with Peptidic Foldamers. *Current Opinion in Structural Biology* **2016**, *39*, 96-105.
15.   Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H., Helix Bundle Quaternary Structure from α/β-Peptide Foldamers. *J Am Chem Soc* **2007**, *129*, 4178-4180.
16.   Horne, W. S.; Price, J. L.; Gellman, S. H., Interplay among Side Chain Sequence, Backbone Composition, and Residue Rigidification in Polypeptide Folding and Assembly. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 9151-9156.

17.   Horne, W. S.; Boersma, M. D.; Windsor, M. A.; Gellman, S. H., Sequence-Based Design of α/β-Peptide Foldamers That Mimic BH3 Domains. *Angewandte Chemie International Edition* **2008**, *47*, 2853-2856.

18.   Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H., Structural and Biological Mimicry of Protein Surface Recognition by α/β-Peptide Foldamers. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 14751-14756.

19.   Haque, T. S.; Little, J. C.; Gellman, S. H., Stereochemical Requirements for β-Hairpin Formation: Model Studies with Four-Residue Peptides and Depsipeptides. *J Am Chem Soc* **1996**, *118*, 6975-6985.

20.   Masterson, L. R.; Etienne, M. A.; Porcelli, F.; Barany, G.; Hammer, R. P.; Veglia, G., Nonstereogenic α-aminoisobutyryl-glycyl dipeptidyl unit nucleates type I′ β-turn in linear peptides in aqueous solution. *Peptide Science* **2007**, *88*, 746-753.

21.   Nowick, J. S.; Brower, J. O., A New Turn Structure for the Formation of β-Hairpins in Peptides. *J Am Chem Soc* **2003**, *125*, 876-877.

22.   Baca, M.; Kent, S. B. H.; Alewood, P. F., Structural Engineering of the HIV-1 Protease Molecule with a β-Turn Mimic of Fixed Geometry. *Protein Science* **1993**, *2*, 1085-1091.

23.   Viles, J. H.; Patel, S. U.; Mitchell, J. B. O.; Moody, C. M.; Justice, D. E.; Uppenbrink, J.; Doyle, P. M.; Harris, C. J.; Sadler, P. J.; Thornton, J. M., Design, Synthesis and Structure of a Zinc Finger with an Artificial β-Turn. *Journal of Molecular Biology* **1998**, *279*, 973-986.

24.   Kaul, R.; Angeles, A. R.; Jäger, M.; Powers, E. T.; Kelly, J. W., Incorporating β-Turns and a Turn Mimetic out of Context in Loop 1 of the WW Domain Affords Cooperatively Folded β-Sheets. *Journal of the American Chemical Society* **2001**, *123*, 5206-5212.

25.   Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T., Protein Prosthesis: A Semisynthetic Enzyme with a β-Peptide Reverse Turn. *Journal of the American Chemical Society* **2002**, *124*, 8522-8523.

26.   Tam, A.; Arnold, U.; Soellner, M. B.; Raines, R. T., Protein Prosthesis: 1,5-Disubstituted[1,2,3]triazoles as cis-Peptide Bond Surrogates. *Journal of the American Chemical Society* **2007**, *129*, 12670-12671.

27.   Rajarathnam, K.; Sykes, B. D.; Kay, C. M.; Dewald, B.; Geiser, T.; Baggiolini, M.; Clark-Lewis, I., Neutrophil Activation by Monomeric Interleukin-8. *Science* **1994**, *264*, 90-92.

28.   Kreutzer, A. G.; Yoo, S.; Spencer, R. K.; Nowick, J. S., Stabilization, Assembly, and Toxicity of Trimers Derived from Aβ. *Journal of the American Chemical Society* **2017**, *139*, 966-975.

29.   Madine, J.; Doig, A. J.; Middleton, D. A., Design of an N-Methylated Peptide Inhibitor of α-Synuclein Aggregation Guided by Solid-State NMR. *Journal of the American Chemical Society* **2008**, *130*, 7873-7881.

30.   Cheng, P.-N.; Liu, C.; Zhao, M.; Eisenberg, D.; Nowick, J. S., Amyloid β-Sheet Mimics that Antagonize Protein Aggregation and Reduce Amyloid Toxicity. *Nature Chemistry* **2012**, *4*, 927-933.

31.   Kar, K.; Baker, M. A.; Lengyel, G. A.; Hoop, C. L.; Kodali, R.; Byeon, I.-J.; Horne, W. S.; van der Wel, P. C. A.; Wetzel, R., Backbone Engineering within a Latent β-Hairpin Structure to Design Inhibitors of Polyglutamine Amyloid Formation. *Journal of Molecular Biology* **2017**, *429*, 308-323.

32.   Etienne, M. A.; Aucoin, J. P.; Fu, Y.; McCarley, R. L.; Hammer, R. P., Stoichiometric Inhibition of Amyloid β-Protein Aggregation with Peptides Containing Alternating α,α-Disubstituted Amino Acids. *Journal of the American Chemical Society* **2006**, *128*, 3522-3523.

33.   Lengyel, G. A.; Frank, R. C.; Horne, W. S., Hairpin Folding Behavior of Mixed α/β-Peptides in Aqueous Solution. *Journal of the American Chemical Society* **2011**, *133*, 4246-4249.

34.   Lengyel, G. A.; Horne, W. S., Design Strategies for the Sequence-Based Mimicry of Side-Chain Display in Protein β-Sheets by α/β-Peptides. *Journal of the American Chemical Society* **2012**, *134*, 15906-15913.

35.   Lengyel, G. A.; Eddinger, G. A.; Horne, W. S., Introduction of Cyclically Constrained γ-Residues Stabilizes an α-Peptide Hairpin in Aqueous Solution. *Organic Letters* **2013**, *15*, 944-947.

36.   Reinert, Z. E.; Lengyel, G. A.; Horne, W. S., Protein-like Tertiary Folding Behavior from Heterogeneous Backbones. *J Am Chem Soc* **2013**, *135*, 12528-12531.

37.   Petersson, E. J.; Schepartz, A., Toward β-Amino Acid Proteins: Design, Synthesis, and Characterization of a Fifteen Kilodalton β-Peptide Tetramer. *J Am Chem Soc* **2008**, *130*, 821-823.

38.   Learte-Aymamí, S.; Curado, N.; Rodríguez, J.; Vázquez, M. E.; Mascareñas, J. L., Metal-Dependent DNA Recognition and Cell Internalization of Designed, Basic Peptides. *J Am Chem Soc* **2017**, *139*, 16188-16193.

39.   Vázquez, M. E.; Caamaño, A. M.; Mascareñas, J. L., From transcription factors to designed sequence-specific DNA-binding peptides. *Chemical Society Reviews* **2003**, *32*, 338-349.

40.   Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., Design of Artificial Transcriptional Activators with Rigid Poly-l-proline Linkers. *J Am Chem Soc* **2002**, *124*, 13067-13071.

41.   Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C., The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α Helices: Crystal structure of the protein-DNA complex. *Cell* **1992**, *71*, 1223-1237.

42.   Sonia, B.; David, B.; Diego, G. P.; Miguel, V. L.; Eugenio, V. M., Sequence-Specific DNA Recognition with Designed Peptides. *European Journal of Organic Chemistry* **2018**, *2018*, 249-261.

43.   Talanian, R.; McKnight, C.; Kim, P., Sequence-specific DNA binding by a short peptide dimer. *Science* **1990**, *249*, 769-771.

44.   Cuenoud, B.; Schepartz, A., Altered specificity of DNA-binding proteins with transition metal dimerization domains. *Science* **1993**, *259*, 510-513.

45.   Emmanuel, O.; A., P. A. F., Metal-Ion-Regulated Miniature DNA-Binding Proteins Based on GCN4 and Non-native Regulation Sites. *Chemistry – A European Journal* **2014**, *20*, 2829-2839.

46.   Chakraborty, M.; Roy, S., A peptide-based synthetic transcription factor selectively down-regulates the proto-oncogene CFOS in tumour cells and inhibits proliferation. *Chemical Communications* **2017**, *53*, 376-379.

47.   Namoto, K.; Gardiner, J.; Kimmerlin, T.; Seebach, D., Investigation of the Interactions of β-Peptides with DNA Duplexes by Circular Dichroism Spectroscopy. *Helvetica Chimica Acta* **2006**, *89*, 3087-3103.

48.   Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T., An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proceedings of the National Academy of Sciences* **1997**, *94*, 3548-3553.

49.   Gelman, M. A.; Richter, S.; Cao, H.; Umezawa, N.; Gellman, S. H.; Rana, T. M., Selective Binding of TAR RNA by a Tat-Derived β-Peptide. *Organic Letters* **2003**, *5*, 3563-3565.

50.   Akkarawongsa, R.; Potocky, T. B.; English, E. P.; Gellman, S. H.; Brandt, C. R., Inhibition of Herpes Simplex Virus Type 1 Infection by Cationic β-Peptides. *Antimicrobial Agents and Chemotherapy* **2008**, *52*, 2120-2129.

51.   Hanas, J. S.; Hazuda, D. J.; Bogenhagen, D. F.; Wu, F. Y.; Wu, C. W., Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. *Journal of Biological Chemistry* **1983**, *258*, 14120-5.

52.   Hanas, J. S.; Duke, A. L.; Gaskins, C. J., Conformational states of Xenopus transcription factor IIIA. *Biochemistry* **1989**, *28*, 4083-4088.

53.   Jay S. Hanas, J. L. L. a. J. R. H., *Zinc Finger Proteins: From Atomic Contact to Cellular Function*. Springer US: p 276.

54.   Krishna, S. S.; Majumdar, I.; Grishin, N. V., Structural classification of zinc fingers: SURVEY AND SUMMARY. *Nucleic Acids Research* **2003**, *31*, 532-550.

55.   Klug, A., The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annual Review of Biochemistry* **2010**, *79*, 213-231.

56.   Takao, K.; Manabu, T.; Yasuro, T.; Hiroko, Y.; Keiko, T.; Masao, Y., Exon 9 mutations in the WT1 gene, without influencing KTS splice isoforms, are also responsible for Frasier syndrome. *Human Mutation* **1999**, *14*, 466-470.

57.   Pelletier, J.; Bruening, W.; Kashtan, C. E.; Mauer, S. M.; Manivel, J. C.; Striegel, J. E.; Houghton, D. C.; Junien, C.; Habib, R.; Fouser, L.; Fine, R. N.; Silverman, B. L.; Haber, D. A.; Housman, D., Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* **1991**, *67*, 437-447.

58.   Nordenskjöld, A.; Friedman, E.; Anvret, M., WT1 mutations in patients with Denys-Drash syndrome: a novel mutation in exon 8 and paternal allele origin. *Human Genetics* **1994**, *93*, 115-120.

59.   Borel, F.; Barilla, K. C.; Hamilton, T. B.; Iskandar, M.; Romaniuk, P. J., Effects of Denys−Drash Syndrome Point Mutations on the DNA Binding Activity of the Wilms' Tumor Suppressor Protein WT1. *Biochemistry* **1996**, *35*, 12070-12076.

60.   Lachenmann, M. J.; Ladbury, J. E.; Phillips, N. B.; Narayana, N.; Qian, X.; Weiss, M. A., The Hidden Thermodynamics of a Zinc Finger. *Journal of Molecular Biology* **2002**, *316*, 969-989.

61.   Berg, J. M.; Godwin, H. A., LESSONS FROM ZINC-BINDING PEPTIDES. *Annual Review of Biophysics and Biomolecular Structure* **1997**, *26*, 357-371.

62.   Pavletich, N.; Pabo, C., Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **1991**, *252*, 809-817.

63.   Desjarlais, J. R.; Berg, J. M., Toward rules relating zinc finger protein sequences and DNA binding site preferences. *Proceedings of the National Academy of Sciences* **1992**, *89*, 7345-7349.

64.   Choo, Y.; Klug, A., Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proceedings of the National Academy of Sciences* **1994**, *91*, 11168-11172.

65.   Rebar, E.; Pabo, C., Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **1994**, *263*, 671-673.

66.   Desjarlais, J. R.; Berg, J. M., Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proceedings of the National Academy of Sciences* **1993**, *90*, 2256-2260.

67.   Negi, S.; Imanishi, M.; Sasaki, M.; Tatsutani, K.; Futaki, S.; Sugiura, Y., An Arginine Residue Instead of a Conserved Leucine Residue in the Recognition Helix of the Finger 3 of Zif268 Stabilizes the Domain Structure and Mediates DNA Binding. *Biochemistry* **2011**, *50*, 6266-6272.

68.   Choo, Y.; Klug, A., A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Research* **1993**, *21*, 3341-3346.

69.   Struthers, M. D.; Cheng, R. P.; Imperiali, B., Design of a Monomeric 23-Residue Polypeptide with Defined Tertiary Structure. *Science* **1996**, *271*, 342-345.

70.   Bibikova, M.; Carroll, D.; Segal, D. J.; Trautman, J. K.; Smith, J.; Kim, Y.-G.; Chandrasegaran, S., Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology* **2001**, *21*, 289-297.

71.   Havranek, J. J.; Duarte, C. M.; Baker, D., A Simple Physical Model for the Prediction and Design of Protein–DNA Interactions. *Journal of Molecular Biology* **2004**, *344*, 59-70.

72.   Sarisky, C. A.; Mayo, S. L., The ββα fold: explorations in sequence space11Edited by M. F. Summers. *Journal of Molecular Biology* **2001**, *307*, 1411-1418.

73.   Chen, K.-Y.; Knoepfler, P. S., To CRISPR and beyond: the evolution of genome editing in stem cells. *Regenerative Medicine* **2016**, *11*, 801-816.

74.   Urnov, F. D.; Rebar, E. J.; Holmes, M. C.; Zhang, H. S.; Gregory, P. D., Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **2010**, *11*, 636.

75.   Zandarashvili, L.; Esadze, A.; Vuzman, D.; Kemme, C. A.; Levy, Y.; Iwahara, J., Balancing between affinity and speed in target DNA search by zinc-finger proteins via modulation of dynamic conformational ensemble. *Proceedings of the National Academy of Sciences* **2015**, *112*, E5142-E5149.

76.   Bautista, A. D.; Craig, C. J.; Harker, E. A.; Schepartz, A., Sophistication of foldamer form and function in vitro and in vivo. *Current Opinion in Chemical Biology* **2007**, *11*, 685-692.

77.   Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F., Foldamers as versatile frameworks for the design and evolution of function. *Nature Chemical Biology* **2007**, *3*, 252-262.

78.   Seebach, D.; Gardiner, J., β-Peptidic Peptidomimetics. *Accounts of Chemical Research* **2008**, *41*, 1366-1375.

79.   Nair, R. V.; Vijayadas, K. N.; Roy, A.; Sanjayan, G. J., Heterogeneous Foldamers from Aliphatic–Aromatic Amino Acid Building Blocks: Current Trends and Future Prospects. *European Journal of Organic Chemistry* **2014**, *2014*, 7763-7780.

80.   Laursen, J. S.; Engel-Andreasen, J.; Olsen, C. A., β-Peptoid Foldamers at Last. *Accounts of Chemical Research* **2015**, *48*, 2696-2704.

81.   Robertson, E. J.; Battigelli, A.; Proulx, C.; Mannige, R. V.; Haxton, T. K.; Yun, L.; Whitelam, S.; Zuckermann, R. N., Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets. *Accounts of Chemical Research* **2016**, *49*, 379-389.

82.   Qiu, J. X.; Petersson, E. J.; Matthews, E. E.; Schepartz, A., Toward β-Amino Acid Proteins: A Cooperatively Folded β-Peptide Quaternary Structure. *J Am Chem Soc* **2006**, *128*, 11338-11339.

83.   Delsuc, N.; Léger, J.-M.; Massip, S.; Huc, I., Proteomorphous Objects from Abiotic Backbones. *Angewandte Chemie International Edition* **2007**, *46*, 214-217.

84.   Collie, G. W.; Pulka-Ziach, K.; Lombardo, C. M.; Fremaux, J.; Rosu, F.; Decossas, M.; Mauran, L.; Lambert, O.; Gabelica, V.; Mackereth, C. D.; Guichard, G., Shaping quaternary assemblies of water-soluble non-peptide helical foldamers by sequence manipulation. *Nature Chemistry* **2015**, *7*, 871-878.

85.   Cheng, P.-N.; Pham, J. D.; Nowick, J. S., The Supramolecular Chemistry of β-Sheets. *Journal of the American Chemical Society* **2013**, *135*, 5477-5492.

86.	Hegedüs, Z.; Wéber, E.; Kriston-Pál, É.; Makra, I.; Czibula, Á.; Monostori, É.; Martinek, T. A., Foldameric α/β-Peptide Analogs of the β-Sheet-Forming Antiangiogenic Anginex: Structure and Bioactivity. *Journal of the American Chemical Society* **2013**, *135*, 16578-16584.

87.	Lengyel, G. A.; Reinert, Z. E.; Griffith, B. D.; Horne, W. S., Comparison of Backbone Modification in Protein β-Sheets by α→γ Residue Replacement and α-Residue Methylation. *Organic & Biomolecular Chemistry* **2014**, *12*, 5375-5381.

88.	Cheng, R. P.; DeGrado, W. F., Long-Range Interactions Stabilize the Fold of a Non-natural Oligomer. *Journal of the American Chemical Society* **2002**, *124*, 11564-11565.

89.	Price, J. L.; Hadley, E. B.; Steinkruger, J. D.; Gellman, S. H., Detection and Analysis of Chimeric Tertiary Structures by Backbone Thioester Exchange: Packing of an α Helix against an α/β-Peptide Helix. *Angewandte Chemie International Edition* **2010**, *49*, 368-371.

90.	Delsuc, N.; Massip, S.; Léger, J.-M.; Kauffmann, B.; Huc, I., Relative Helix−Helix Conformations in Branched Aromatic Oligoamide Foldamers. *Journal of the American Chemical Society* **2011**, *133*, 3165-3172.

91.	Bou-Abdallah, F.; Giffune, T. R., The thermodynamics of protein interactions with essential first row transition metals. *Biochimica et Biophysica Acta* **2016**, *1860*, 879-891.

92.	Maayan, G., Conformational Control in Metallofoldamers: Design, Synthesis and Structural Properties. *European Journal of Organic Chemistry* **2009**, *2009*, 5699-5710.

93.	Prince, R. B.; Okada, T.; Moore, J. S., Controlling the Secondary Structure of Nonbiological Oligomers with Solvophobic and Coordination Interactions. *Angewandte Chemie International Edition* **1999**, *38*, 233-236.

94.	Zhang, F.; Bai, S.; Yap, G. P. A.; Tarwade, V.; Fox, J. M., Abiotic Metallofoldamers as Electrochemically Responsive Molecules. *Journal of the American Chemical Society* **2005**, *127*, 10590-10599.

95.	Zhao, Y.; Zhong, Z., Tuning the Sensitivity of a Foldamer-Based Mercury Sensor by Its Folding Energy. *Journal of the American Chemical Society* **2006**, *128*, 9988-9989.

96.	Tashiro, S.; Matsuoka, K.; Minoda, A.; Shionoya, M., Metallo-Foldamers with Backbone-Coordinative Oxime Peptides: Control of Secondary Structures. *Angewandte Chemie International Edition* **2012**, *51*, 13123-13127.

97.	Viles, J. H.; Patel, S. U.; Mitchell, J. B. O.; Moody, C. M.; Justice, D. E.; Uppenbrink, J.; Doyle, P. M.; Harris, C. J.; Sadler, P. J.; Thornton, J. M., Design, synthesis and structure of a zinc finger with an artificial β-turn1. *Journal of Molecular Biology* **1998**, *279*, 973-986.

98.	Lelais, G.; Seebach, D.; Jaun, B.; Mathad, R. I.; Flögel, O.; Rossi, F.; Campo, M.; Wortmann, A., β-Peptidic Secondary Structures Fortified and Enforced by Zn2+ Complexation – On the Way to β-Peptidic Zinc Fingers? *Helvetica Chimica Acta* **2006**, *89*, 361-403.

99.	Delsuc, N.; Hutin, M.; Campbell, V. E.; Kauffmann, B.; Nitschke, J. R.; Huc, I., Metal-Directed Dynamic Formation of Tertiary Structure in Foldamer Assemblies: Orienting Helices at an Angle. *Chemistry, A European Journal* **2008**, *14*, 7140-7143.

100.	Lee, B.-C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N., Biomimetic Nanostructures: Creating a High-Affinity Zinc-Binding Site in a Folded Nonbiological Polymer. *Journal of the American Chemical Society* **2008**, *130*, 8847-8855.

101.	Schafmeister, C. E.; Belasco, L. G.; Brown, P. H., Observation of Contraction and Expansion in a Bis(peptide)-Based Mechanical Molecular Actuator. *Chemistry, A European Journal* **2008**, *14*, 6406-6412.

102.   Sénèque, O.; Bourlès, E.; Lebrun, V.; Bonnet, E.; Dumy, P.; Latour, J.-M., Cyclic Peptides Bearing a Side-Chain Tail: A Tool to Model the Structure and Reactivity of Protein Zinc Sites. *Angewandte Chemie International Edition* **2008**, *47*, 6888-6891.

103.   Sénèque, O.; Bonnet, E.; Joumas, F. L.; Latour, J.-M., Cooperative Metal Binding and Helical Folding in Model Peptides of Treble-Clef Zinc Fingers. *Chemistry, A European Journal* **2009**, *15*, 4798-4810.

104.   Ma, Z.; Olechnowicz, F.; Skorik, Y. A.; Achim, C., Metal Binding to Ligand-Containing Peptide Nucleic Acids. *Inorganic Chemistry* **2011**, *50*, 6083-6092.

105.   Miller, J. P.; Melicher, M. S.; Schepartz, A., Positive Allostery in Metal Ion Binding by a Cooperatively Folded β-Peptide Bundle. *Journal of the American Chemical Society* **2014**, *136*, 14726-14729.

106.   Baskin, M.; Maayan, G., A rationally designed metal-binding helical peptoid for selective recognition processes. *Chemical Science* **2016**, *7*, 2809-2820.

107.   Reinert, Z. E.; Horne, W. S., Protein backbone engineering as a strategy to advance foldamers toward the frontier of protein-like tertiary structure. *Organic & Biomolecular Chemistry* **2014**, *12*, 8796-8802.

108.   Werner, H. M.; Cabalteja, C. C.; Horne, W. S., Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution. *ChemBioChem* **2016**, *17*, 712-718.

109.   Cheloha, R. W.; Maeda, A.; Dean, T.; Gardella, T. J.; Gellman, S. H., Backbone modification of a polypeptide drug alters duration of action in vivo. *Nature Biotechnology* **2014**, *32*, 653-655.

110.   Checco, J. W.; Kreitler, D. F.; Thomas, N. C.; Belair, D. G.; Rettko, N. J.; Murphy, W. L.; Forest, K. T.; Gellman, S. H., Targeting diverse protein–protein interaction interfaces with α/β-peptides derived from the Z-domain scaffold. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112*, 4552-4557.

111.   Werner, H. M.; Horne, W. S., Folding and function in α/β-peptides: targets and therapeutic applications. *Current Opinion in Chemical Biology* **2015**, *28*, 75-82.

112.   Gopalakrishnan, R.; Frolov, A. I.; Knerr, L.; Drury, W. J.; Valeur, E., Therapeutic Potential of Foldamers: From Chemical Biology Tools To Drug Candidates? *Journal of Medicinal Chemistry* **2016**, *59*, 9599-9621.

113.   Checco, J. W.; Gellman, S. H., Targeting recognition surfaces on natural proteins with peptidic foldamers. *Current Opinion in Structural Biology* **2016**, *39*, 96-105.

114.   Checco, J. W.; Gellman, S. H., Iterative Nonproteinogenic Residue Incorporation Yields α/β-Peptides with a Helix–Loop–Helix Tertiary Structure and High Affinity for VEGF. *Chembiochem* **2017**, *18*, 291-299.

115.   Reinert, Z. E.; Horne, W. S., Folding thermodynamics of protein-like oligomers with heterogeneous backbones. *Chemical Science* **2014**, *5*, 3325-3330.

116.   Tavenor, N. A.; Reinert, Z. E.; Lengyel, G. A.; Griffith, B. D.; Horne, W. S., Comparison of design strategies for [small alpha]-helix backbone modification in a protein tertiary fold. *Chemical Communications* **2016**, *52*, 3789-3792.

117.   Newberry, R. W.; VanVeller, B.; Raines, R. T., Thioamides in the collagen triple helix. *Chemical Communications* **2015**, *51*, 9624-9627.

118.   Zhang, Y.; Malamakal, R. M.; Chenoweth, D. M., Aza-Glycine Induces Collagen Hyperstability. *Journal of the American Chemical Society* **2015**, *137*, 12422-12425.

119.   Kreitler, D. F.; Mortenson, D. E.; Forest, K. T.; Gellman, S. H., Effects of Single α-to-β Residue Replacements on Structure and Stability in a Small Protein: Insights from Quasiracemic Crystallization. *J Am Chem Soc* **2016**, *138*, 6498-6505.

120.   Simon, M. D.; Maki, Y.; Vinogradov, A. A.; Zhang, C.; Yu, H.; Lin, Y.-S.; Kajihara, Y.; Pentelute, B. L., d-Amino Acid Scan of Two Small Proteins. *Journal of the American Chemical Society* **2016**, *138*, 12099-12111.

121.   Walters, C. R.; Szantai-Kis, D. M.; Zhang, Y.; Reinert, Z. E.; Horne, W. S.; Chenoweth, D. M.; Petersson, E. J., The effects of thioamide backbone substitution on protein stability: a study in [small alpha]-helical, [small beta]-sheet, and polyproline II helical contexts. *Chemical Science* **2017**, 2868-2877.

122.   Hanas, J. S.; Larabee, J. L.; Hocker, J. R., *Zinc Finger Proteins: From Atomic Contact to Cellular Function*. Springer US: 2005.

123.   Gersbach, C. A.; Gaj, T.; Barbas, C. F., Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. *Accounts of chemical research* **2014**, *47*, 2309-2318.

124.   Reynolds, L.; Ullman, C.; Moore, M.; Isalan, M.; West, M. J.; Clapham, P.; Klug, A.; Choo, Y., Repression of the HIV-1 5′ LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 1615-1620.

125.   Rebar, E. J.; Huang, Y.; Hickey, R.; Nath, A. K.; Meoli, D.; Nath, S.; Chen, B.; Xu, L.; Liang, Y.; Jamieson, A. C.; Zhang, L.; Spratt, S. K.; Case, C. C.; Wolffe, A.; Giordano, F. J., Induction of angiogenesis in a mouse model using engineered transcription factors. *Nature Medicine* **2002**, *8*, 1427-1432.

126.   Gaj, T.; Liu, J.; Anderson, K. E.; Sirk, S. J.; Barbas, C. F., Protein Delivery Using Cys2–His2 Zinc-Finger Domains. *ACS Chemical Biology* **2014**, *9*, 1662-1667.

127.   Narayan, V. A.; Kriwacki, R. W.; Caradonna, J. P., Structures of Zinc Finger Domains from Transcription Factor Sp1: INSIGHTS INTO SEQUENCE-SPECIFIC PROTEIN-DNA RECOGNITION. *Journal of Biological Chemistry* **1997**, *272*, 7801-7809.

128.   Sénèque, O.; Latour, J.-M., Coordination Properties of Zinc Finger Peptides Revisited: Ligand Competition Studies Reveal Higher Affinities for Zinc and Cobalt. *J Am Chem Soc* **2010**, *132*, 17760-17774.

129.   Posewitz, M. C.; Wilcox, D. E., Properties of the Sp1 Zinc Finger 3 Peptide: Coordination Chemistry, Redox Reactions, and Metal Binding Competition with Metallothionein. *Chemical Research in Toxicology* **1995**, *8*, 1020-1028.

130.   Giedroc, D. P.; Keating, K. M.; Williams, K. R.; Konigsberg, W. H.; Coleman, J. E., Gene 32 protein, the single-stranded DNA binding protein from bacteriophage T4, is a zinc metalloprotein. *Proceedings of the National Academy of Sciences* **1986**, *83*, 8452-8456.

131.   Chen, X.; Chu, M.; Giedroc, D. P., Spectroscopic characterization of Co(II)-, Ni(II)-, and Cd(II)-substituted wild-type and non-native retroviral-type zinc finger peptides. *JBIC Journal of Biological Inorganic Chemistry* **2000**, *5*, 93-101.

132.   May, S. W.; Kuo, J.-Y., Preparation and properties of cobalt(II) rubredoxin. *Biochemistry* **1978**, *17*, 3333-3338.

133.   Vasak, M.; Kaegi, J. H. R.; Holmquist, B.; Vallee, B. L., Spectral studies of cobalt(II)- and nickel(II)-metallothionein. *Biochemistry* **1981**, *20*, 6659-6664.

134.   Hutchinson, E. G.; Thornton, J. M., A revised set of potentials for β-turn formation in proteins. *Protein Science* **1994**, *3*, 2207-2216.

135.   Makhatadze, G. I.; Privalov, P. L., Energetics of Protein Structure. In *Advances in Protein Chemistry*, Anfinsen, C. B.; Richards, F. M.; Edsall, J. T.; Eisenberg, D. S., Eds. Academic Press: 1995; Vol. 47, pp 307-425.

136.   Dunitz, J. D., Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chemistry & Biology* **1995**, *2*, 709-712.

137.   De Filippis, V.; De Antoni, F.; Frigo, M.; Polverino de Laureto, P.; Fontana, A., Enhanced Protein Thermostability by Ala → Aib Replacement. *Biochemistry* **1998**, *37*, 1686-1696.

138.   Rich, A. M.; Bombarda, E.; Schenk, A. D.; Lee, P. E.; Cox, E. H.; Spuches, A. M.; Hudson, L. D.; Kieffer, B.; Wilcox, D. E., Thermodynamics of Zn2+ binding to Cys2His2 and Cys2HisCys zinc fingers and a Cys4 transcription factor site. *J Am Chem Soc* **2012**, *134*, 10405-18.

139.   Reddi, A. R.; Guzman, T. R.; Breece, R. M.; Tierney, D. L.; Gibney, B. R., Deducing the Energetic Cost of Protein Folding in Zinc Finger Proteins Using Designed Metallopeptides. *J Am Chem Soc* **2007**, *129*, 12815-12827.

140.   Miloch, A.; Krezel, A., Metal binding properties of the zinc finger metallome - insights into variations in stability. *Metallomics* **2014**, *6*, 2015-2024.

141.   Michael, S. F.; Kilfoil, V. J.; Schmidt, M. H.; Amann, B. T.; Berg, J. M., Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89*, 4796-4800.

142.   Shi, Y.; Beger, R. D.; Berg, J. M., Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt(II) as a spectroscopic probe. *Biophysical Journal* **1993**, *64*, 749-753.

143.   Rich, A. M.; Bombarda, E.; Schenk, A. D.; Lee, P. E.; Cox, E. H.; Spuches, A. M.; Hudson, L. D.; Kieffer, B.; Wilcox, D. E., Thermodynamics of Zn2+ binding to Cys2His2 and Cys2HisCys zinc fingers and a Cys4 transcription factor site. *Journal of the American Chemical Society* **2012**, *134*, 10405-18.

144.   Miura, T.; Satoh, T.; Takeuchi, H., Role of metal–ligand coordination in the folding pathway of zinc finger peptides. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998**, *1384*, 171-179.

145.   Wu, Y.-D.; Gellman, S., Peptidomimetics. *Accounts of chemical research* **2008**, *41*, 1231-1232.

146.   Werle, M.; Bernkop-Schnürch, A., Strategies to improve plasma half life time of peptide and protein drugs. *Amino Acids* **2006**, *30*, 351-367.

147.   Thermo Scientific: Ellmans Reagent Instructions. https://tools.thermofisher.com/content/sfs/manuals/MAN0011216_Ellmans_Reag_UG.pdf.

148.   Sabel, C. E.; Neureuther, J. M.; Siemann, S., A spectrophotometric method for the determination of zinc, copper, and cobalt ions in metalloproteins using Zincon. *Analytical Biochemistry* **2010**, *397*, 218-26.

149.   Kuzmic, P., Program DYNAFIT for Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. *Analytical Biochemistry* **1996**, *237*, 260-273.

150.   Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J., Sequence-Dependent Correction of Random Coil NMR Chemical Shifts. *Journal of the American Chemical Society* **2001**, *123*, 2970-2978.

151.   Kjaergaard, M.; Brander, S.; Poulsen, F. M., Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of Biomolecular NMR* **2011**, *49*, 139-149.

152.   Kjaergaard, M.; Poulsen, F. M., Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of Biomolecular NMR* **2011**, *50*, 157-165.

153.   Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallographica Section D: Biological Crystallography* **1998**, *54*, 905-921.

154.   Brunger, A. T., Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2007**, *2*, 2728-2733.

155.   Liu, Y.; Liu, Z.; Androphy, E.; Chen, J.; Baleja, J. D., Design and Characterization of Helical Peptides that Inhibit the E6 Protein of Papillomavirus. *Biochemistry* **2004**, *43*, 7421-7431.

156.   Oka, S.; Shiraishi, Y.; Yoshida, T.; Ohkubo, T.; Sugiura, Y.; Kobayashi, Y., NMR Structure of Transcription Factor Sp1 DNA Binding Domain. *Biochemistry* **2004**, *43*, 16027-16035.

157.   Wang, C.; Vernon, R.; Lange, O.; Tyka, M.; Baker, D., Prediction of structures of zinc-binding proteins through explicit modeling of metal coordination geometry. *Protein Science* **2010**, *19*, 494-506.

158.   Modell, A. E.; Blosser, S. L.; Arora, P. S., Systematic Targeting of Protein-Protein Interactions. *Trends in pharmacological sciences* **2016**, *37*, 702-713.

159.   Teyssières, E.; Corre, J.-P.; Antunes, S.; Rougeot, C.; Dugave, C.; Jouvion, G.; Claudon, P.; Mikaty, G.; Douat, C.; Goossens, P. L.; Guichard, G., Proteolytically Stable Foldamer Mimics of Host-Defense Peptides with Protective Activities in a Murine Model of Bacterial Infection. *Journal of Medicinal Chemistry* **2016**, *59*, 8221-8232.

160.   Rudi, F.; A., D. R. L.; Kerstin, M.; Oliver, Z.; W., V. J.; Daniel, O.; A., R. J., Using a β-Hairpin To Mimic an α-Helix: Cyclic Peptidomimetic Inhibitors of the p53–HDM2 Protein–Protein Interaction. *Angewandte Chemie International Edition* **2004**, *43*, 2109-2112.

161.   Cheloha, R. W.; Chen, B.; Kumar, N. N.; Watanabe, T.; Thorne, R. G.; Li, L.; Gardella, T. J.; Gellman, S. H., Development of Potent, Protease-Resistant Agonists of the Parathyroid Hormone Receptor with Broad β Residue Distribution. *Journal of Medicinal Chemistry* **2017**, *60*, 8816-8833.

162.   Johnson, L. M.; Barrick, S.; Hager, M. V.; McFedries, A.; Homan, E. A.; Rabaglia, M. E.; Keller, M. P.; Attie, A. D.; Saghatelian, A.; Bisello, A.; Gellman, S. H., A Potent α/β-Peptide Analogue of GLP-1 with Prolonged Action in Vivo. *J Am Chem Soc* **2014**, *136*, 12848-12851.

163.   Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J Am Chem Soc* **1963**, *85*, 2149-2154.

164.   Donald, Y.; James, B., USE OF THIOL ACIDS IN PEPTIDE SEGMENT COUPLING IN NON-AQUEOUS SOLVENTS. *International Journal of Peptide and Protein Research* **1981**, *18*, 383-392.

165.   Saburo, A., Polypeptide synthesis by the thioester method. *Peptide Science* **1999**, *51*, 247-265.

166.   Schnolzer, M.; Kent, S., Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **1992**, *256*, 221-225.

167.   Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S., Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776-779.

168.   Muir, T. W., Semisynthesis of Proteins by Expressed Protein Ligation. *Annual Review of Biochemistry* **2003**, *72*, 249-289.

169.   Ge-Min, F.; Yi-Ming, L.; Fei, S.; Yi-Chao, H.; Jia-Bin, L.; Yun, L.; Hong-Kui, C.; Lei, L., Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angewandte Chemie International Edition* **2011**, *50*, 7645-7649.

170.   Agouridas, V.; El Mahdi, O.; Cargoët, M.; Melnyk, O., A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorganic & Medicinal Chemistry* **2017**, *25*, 4938-4945.

171.   Viles, J. H.; Patel, S. U.; Mitchell, J. B. O.; Moody, C. M.; Justice, D. E.; Uppenbrink, J.; Doyle, P. M.; Harris, C. J.; Sadler, P. J.; Thornton, J. M., Design, synthesis and structure of a zinc finger with an artificial β-turn11Edited by M. F. Moody. *Journal of Molecular Biology* **1998**, *279*, 973-986.

172.   S., B. G.; E., D. P., Synthesis of a three zinc finger protein, Zif268, by native chemical ligation. *Peptide Science* **1999**, *51*, 363-369.

173.   Friederike, F.; André, N.; Florian, B.; Ivo, F.; Ulf, D., Semi-Synthesis and Analysis of Chemically Modified Zif268 Zinc-Finger Domains. *ChemistryOpen* **2012**, *1*, 26-32.

174.   Carpino, L. A.; Krause, E.; Sferdean, C. D.; Bienert, M.; Beyermann, M., Dramatically enhanced N→O acyl migration during the trifluoroacetic acid-based deprotection step in solid phase peptide synthesis. *Tetrahedron Letters* **2005**, *46*, 1361-1364.

175.   Dolling, R.; Beyermann, M.; Haenel, J.; Kernchen, F.; Krause, E.; Franke, P.; Brudel, M.; Bienert, M., Piperidine-mediated side product formation for Asp(OBut)-containing peptides. *Journal of the Chemical Society, Chemical Communications* **1994**, 853-854.

176.   Ramon, S. F.; Ayman, E. F.; Fernando, A., Use of Oxyma as pH modulatory agent to be used in the prevention of base-driven side reactions and its effect on 2-chlorotrityl chloride resin. *Peptide Science* **2012**, *98*, 89-97.

177.   Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W., Acid-Mediated Prevention of Aspartimide Formation in Solid Phase Peptide Synthesis. *Organic Letters* **2012**, *14*, 5218-5221.

178.   M., M.; F., D.; B., S.; P., W.; T., V., The aspartimide problem in Fmoc-based SPPS. Part I. *Journal of Peptide Science* **2003**, *9*, 36-46.

179.   M., M.; F., D., The aspartimide problem in Fmoc-based SPPS. Part III. *Journal of Peptide Science* **2005**, *11*, 650-657.

180.   Hackeng, T. M.; Griffin, J. H.; Dawson, P. E., Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10068-10073.

181.   Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E., Chemical Protein Synthesis Using a Second-Generation N-Acylurea Linker for the Preparation of Peptide-Thioester Precursors. *J Am Chem Soc* **2015**, *137*, 7197-7209.

182.   Villain, M.; Vizzavona, J.; Rose, K., Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chemistry & Biology* **2001**, *8*, 673-679.

183.   Duhee, B.; H., K. S. B., A One-Pot Total Synthesis of Crambin. *Angewandte Chemie International Edition* **2004**, *43*, 2534-2538.

184.   Di, L., Strategic Approaches to Optimizing Peptide ADME Properties. *The AAPS Journal* **2015**, *17*, 134-143.

185.   Bunnage, M. E., Getting pharmaceutical R&D back on target. *Nature chemical biology* **2011**, *7*, 335.

186. Ahn, N. G.; Wang, A. H. J., Proteomics and genomics: perspectives on drug and target discovery. *Current Opinion in Chemical Biology* **2008**, *12*, 1-3.

187. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The Future of Peptide-based Drugs. *Chemical Biology & Drug Design* **2013**, *81*, 136-147.

188. Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F., Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chemical Reviews* **2014**, *114*, 901-926.

189. Wójcik, P.; Berlicki, Ł., Peptide-based inhibitors of protein–protein interactions. *Bioorganic & Medicinal Chemistry Letters* **2016**, *26*, 707-713.

190. Verdine, G. L.; Walensky, L. D., The Challenge of Drugging Undruggable Targets in Cancer: Lessons Learned from Targeting BCL-2 Family Members. *Clinical Cancer Research* **2007**, *13*, 7264-7270.

191. Ran, X.; Gestwicki, J. E., Inhibitors of protein–protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Current Opinion in Chemical Biology* **2018**, *44*, 75-86.

192. Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* **2018**, *26*, 2700-2707.

193. Recio, C.; Maione, F.; Iqbal, A. J.; Mascolo, N.; De Feo, V., The Potential Therapeutic Application of Peptides and Peptidomimetics in Cardiovascular Disease. *Frontiers in Pharmacology* **2016**, *7*, 526.

194. Paradís-Bas, M.; Tulla-Puche, J.; Albericio, F., The road to the synthesis of "difficult peptides". *Chemical Society Reviews* **2016**, *45*, 631-654.

195. Cunningham, A. D.; Qvit, N.; Mochly-Rosen, D., Peptides and peptidomimetics as regulators of protein–protein interactions. *Current Opinion in Structural Biology* **2017**, *44*, 59-66.

196. Erak, M.; Bellmann-Sickert, K.; Els-Heindl, S.; Beck-Sickinger, A. G., Peptide chemistry toolbox – Transforming natural peptides into peptide therapeutics. *Bioorganic & Medicinal Chemistry* **2018**, *26*, 2759-2765.

197. Cheloha, R. W.; Sullivan, J. A.; Wang, T.; Sand, J. M.; Sidney, J.; Sette, A.; Cook, M. E.; Suresh, M.; Gellman, S. H., Consequences of Periodic α-to-β(3) Residue Replacement for Immunological Recognition of Peptide Epitopes. *ACS Chemical Biology* **2015**, *10*, 844-854.

198. Uppalapati, M.; Lee, D. J.; Mandal, K.; Li, H.; Miranda, L. P.; Lowitz, J.; Kenney, J.; Adams, J. J.; Ault-Riché, D.; Kent, S. B. H.; Sidhu, S. S., A Potent d-Protein Antagonist of VEGF-A is Nonimmunogenic, Metabolically Stable, and Longer-Circulating in Vivo. *ACS Chemical Biology* **2016**, *11*, 1058-1065.

199. Pratt, K., Anti-Drug Antibodies: Emerging Approaches to Predict, Reduce or Reverse Biotherapeutic Immunogenicity. *Antibodies* **2018**, *7*, 19.

200. Rogers, J. M.; Kwon, S.; Dawson, S. J.; Mandal, P. K.; Suga, H.; Huc, I., Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids. *Nature Chemistry* **2018**, *10*, 405-412.

201. Schepartz, A., Foldamers wave to the ribosome. *Nature Chemistry* **2018**, *10*, 377-379.

202. de la Torre, B.; Albericio, F., The Pharmaceutical Industry in 2017. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2018**, *23*, 533.

203. Rolland, T.; Taşan, M.; Charloteaux, B.; Pevzner, Samuel J.; Zhong, Q.; Sahni, N.; Yi, S.; Lemmens, I.; Fontanillo, C.; Mosca, R.; Kamburov, A.; Ghiassian, Susan D.; Yang, X.; Ghamsari, L.; Balcha, D.; Begg, Bridget E.; Braun, P.; Brehme, M.; Broly, Martin P.; Carvunis, A.-R.; et al., A Proteome-Scale Map of the Human Interactome Network. *Cell* **2014**, *159*, 1212-1226.

204. Horne, W. S., Peptide and peptoid foldamers in medicinal chemistry. *Expert Opinion on Drug Discovery* **2011**, *6*, 1247-1262.

205.   Cabrele, C.; Martinek, T. A.; Reiser, O.; Berlicki, Ł., Peptides Containing β-Amino Acid Patterns: Challenges and Successes in Medicinal Chemistry. *Journal of Medicinal Chemistry* **2014**, *57*, 9718-9739.