DDR

DRUG DEVELOPMENT RESEARCH 78 : 268–282 (2017)

Research Overview

Small Scaffolds, Big Potential: Developing Miniature Proteins as Therapeutic Agents

Justin M. Holub*

Department of Chemistry and Biochemistry, Ohio University, Athens, OH, 45701, USA

Strategy, Management and Health Policy				
Enabling Technology, Genomics, Proteomics	Preclinical Research	Preclinical Development Toxicology, Formulation Drug Delivery, Pharmacokinetics	Clinical Development Phases I-III Regulatory, Quality, Manufacturing	Postmarketing Phase IV

ABSTRACT Miniature proteins are a class of oligopeptide characterized by their short sequence lengths and ability to adopt well-folded, three-dimensional structures. Because of their biomimetic nature and synthetic tractability, miniature proteins have been used to study a range of biochemical processes including fast protein folding, signal transduction, catalysis and molecular transport. Recently, miniature proteins have been gaining traction as potential therapeutic agents because their small size and ability to fold into defined tertiary structures facilitates their development as protein-based drugs. This research overview discusses emerging developments involving the use of miniature proteins as scaffolds to design novel therapeutics for the treatment and study of human disease. Specifically, this review will explore strategies to: (i) stabilize miniature protein tertiary structure; (ii) optimize biomolecular recognition by grafting functional epitopes onto miniature protein scaffolds; and (iii) enhance cytosolic delivery of miniature proteins through the use of cationic motifs that facilitate endosomal escape. These objectives are discussed not only to address challenges in developing effective miniature protein-based drugs, but also to highlight the tremendous potential miniature proteins hold for combating and understanding human disease. Drug Dev Res 78 : 268-282, 2017.

Key words: miniature protein; protein scaffold; protein-based therapeutics; tertiary structure; protein grafting

INTRODUCTION

Protein-based therapeutics represent one of the fastest growing sectors of the pharmaceutical industry. In 2016, 9 of the top 15 best-selling drugs were protein-based, with combined global sales exceeding $70 billion (Philippidis, 2017). There are currently over 100 protein-based therapeutics on the market that have been approved to treat a wide range of diseases including cancer, metabolic dysfunction, infectious disease, and autoimmune disorders (Dimitrov, 2012). Protein drugs are classified into nine groups according to their primary mode of activity: replacing deficient or abnormal proteins (Group Ia); augmenting an existing pathway (Ib); providing a novel function (Ic); interfering with a molecule or organism (IIa); delivering compounds or other proteins (IIb); protecting against pathogenic foreign agents (IIIa); treating autoimmune disease (IIIb); treating cancer (IIIc); and protein diagnostics (IV) (Leader et al., 2008). Of these specific groups, protein therapeutics in Group I and II make up the majority of approved protein-based drugs, constituting over 80% of the total market share (Raghava, 2017).

Despite high success rates in the clinic and a growing interest in developing novel protein-based therapies, the pharmaceutical industry has struggled

*Correspondence to: Justin M. Holub, Department of Chemistry and Biochemistry, Ohio University, Athens, OH 45701, USA. E-mail: holub@ohio.edu

Received 20 June 2017; Accepted 26 July 2017

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/ddr.21408

to develop highly efficacious protein drugs that extend beyond Group I and II classifications. Two primary difficulties hampering large-scale development of protein-based therapeutics are: (i) end-product heterogeneity, which can lead to adverse side-effects upon treatment with the desired therapeutic; and (ii) inefficient cellular delivery due to large proteins being unable to traverse the cell membrane. Indeed, more than 60% of all protein drugs currently in use target extracellular receptors or cofactors in the serum (Raghava, 2017). It should also be noted that over 85% of therapeutic proteins are greater than 60 amino acids in length (Raghava, 2017), effectively limiting their production to recombinant expression procedures. Peptide-based molecules have also enjoyed a recent surge in popularity as potential therapeutics because of their relatively small size and exquisite biocomplementarity (Fosgerau and Hoffmann, 2015). However, isolated peptide sequences are highly susceptible to hydrolysis or oxidation and their inability to fold often compromises their ability to effectively target therapeutically relevant biomolecular interactions (Bak et al., 2015). These drawbacks have presented modern bioorganic and medicinal chemists with a striking design challenge to develop therapeutic proteins that are at once structured, homogeneous and able to be delivered to intracellular compartments.

Miniature proteins are small, well-folded peptide oligomers that display protein-like behavior, including cooperative melting transitions and ability to interact with other biomolecules (Horng et al., 2003; Kritzer et al., 2006). In general, proteins classified as "miniature" are less than 50 amino acids in length and under 10 kDa in molecular weight. Compared to linear peptides and large recombinant proteins, miniature proteins possess desirable attributes that potentiate their development as protein-based therapeutics. For example, the natural folding of miniature proteins allows for discrete recognition elements to be displayed in patterns that are able to interact with native biopolymer surfaces (Zondlo and Schepartz, 1999). Miniature proteins are also able to be chemically synthesized by solid-phase peptide synthesis procedures (Nicoll et al., 2005), thus affording researchers a high level of control over the sequence and homogeneity of the final product. In addition to being synthetically tractable, miniature proteins are genetically encodable, further expanding their utility as tools to study intricate biochemical pathways (Bottger et al., 1997). Finally, miniature proteins are more resistant to proteolysis compared to extended peptides (Martin and Vita, 2000; Tyndall et al., 2005) and can be engineered to gain cell entry (Daniels and Schepartz, 2007). Miniature proteins therefore occupy a unique "middle space" between peptides and large proteins, taking advantage of the beneficial physicochemical properties of these molecules, including sequence specificity and biomolecular recognition, while circumventing the unwanted attributes, such as proteolytic susceptibility and end product heterogeneity. This research overview will focus on recent strategies to develop miniature proteins as therapeutics and as tools to study the molecular nature of disease. Emphasis will be placed on techniques used to: (i) stabilize miniature protein tertiary structure; (ii) engineer miniature proteins for optimized biomolecular recognition; and (iii) enhance delivery of miniature proteins to the cytoplasm of live cells. Numerous reviews have been published that explore the nature and efficacy of peptide- and protein-based therapeutics (Leader et al., 2008; Chalker, 2013; Craik et al., 2013). However, these reviews have focused primarily on constructs that have been developed into therapeutics in their natural state; that is, without extensive modification of the primary sequence. This research overview will discuss miniature proteins that are used as scaffolds to develop novel constructs designed to target therapeutically relevant biomolecular interactions. Furthermore, only miniature proteins that are composed entirely of natural amino acids are included; therefore, stapled peptides (Walensky et al., 2004), peptides containing surrogate hydrogen bonds (Chapman et al., 2004) and miniature proteins containing D- or β-amino acids (Pouny and Shai, 1992; Appella et al., 1997) will not be discussed.

STABILIZING MINIATURE PROTEIN FOLDS

An intriguing hallmark of miniature proteins is their ability to fold into stable, protein-like architectures despite their relatively short sequence lengths. Indeed, defined elements of secondary and tertiary structures are precisely what endow miniature proteins with biological function. For many years, it was thought that only proteins longer than 40 amino acids could adopt well-folded tertiary structures under physiological conditions. Now, it is widely accepted that much shorter peptide oligomers, some as small as 12 amino acids (Cochran et al., 2001), can adopt protein-like architectures in aqueous environments. Several factors have played a role in this paradigm shift, including the accelerated discovery of miniature proteins from natural sources, the ability to chemically synthesize short peptide oligomers that fold into stable tertiary structures, and the refinement of computational simulations that predict the nature of protein folds.

Despite their small size, miniature proteins rely on the same intramolecular forces that drive the folding of larger proteins, such as hydrophobic interactions, metal ion complexation and covalent linkages. Although there is no universally accepted definition

for a self-folding domain, there are several features of the protein primary sequence that can be used as diagnostics to predict which sequences will adopt a stable protein-like fold (Neidigh et al., 2002). These features include: (i) multiple secondary structure folds in close proximity; (ii) pronounced dispersion of chemical shift that reflects tertiary interactions; (iii) side chain packing that produces well-defined χ^1 and χ^2 values; (iv) protection of backbone amide exchange; and (v) cooperative folding and resistance to thermal unfolding. As strategies for predicting and achieving stable folds in miniature proteins have become more sophisticated, researchers are discovering very short peptide sequences that can adopt protein-like architectures. Moreover, the discovery of endogenous miniature proteins from natural sources has enhanced understanding of how Nature has evolved to fold very short peptide sequences with high efficiency. The following section discusses the intramolecular interactions used to stabilize miniature protein tertiary structure. It should be noted that several insightful reviews have been published on this subject previously (Polticelli et al., 2001; Yin, 2012). However, since these reports were published, additional developments have followed at a rapid pace and a brief update on this subject is warranted.

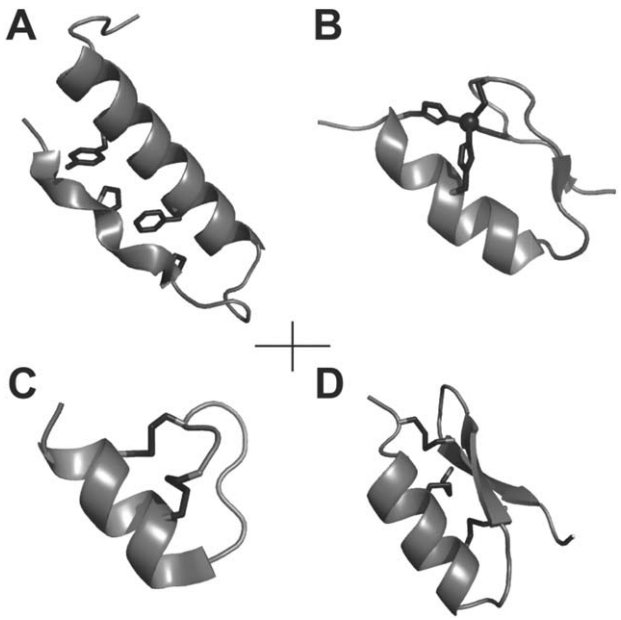

Fig. 1. Ribbon diagrams of naturally occurring miniature proteins; side chains contributing to the protein folds are shown as dark grey sticks. (**A**) Avian pancreatic peptide (PDB ID: 1PPT); Tyr, Pro and Phe residues form the hydrophobic core of the globular protein. (**B**) Zinc-finger domain (PDB ID: 1A1L); Zn ion is shown as a dark grey sphere complexed by two His and two Cys residues. (**C**) Apamin (PDB ID: 3IUX); two disulfide linkages contributing to the α-helical fold are shown as dark grey sticks. (**D**) Scyllatoxin (PDB ID: 1SCY); three disulfide linkages contributing to the *α/β* fold are shown as dark grey sticks.

Hydrophobic Interactions

The hydrophobic effect is a major driving force in the formation of protein tertiary structure. Water-soluble proteins typically contain a core of hydrophobic amino acids that are sequestered from the aqueous environment, stabilizing the overall fold of the protein. Although other intramolecular interactions, such as hydrogen bonds and salt bridges, contribute significantly to protein stability, minimizing the exposure of hydrophobic side chains to water is thought to be the primary force that drives the protein folding process. The hydrophobic effect is often utilized to drive structure formation and impart protein-like folding behavior to small proteins; in fact, hydrophobic interactions have been exploited in the design of entirely synthetic miniature proteins that adopt novel folds not found in natural domains (Neidigh et al., 2002).

Hydrophobic interactions are ubiquitous among miniature proteins and many miniature protein scaffolds require hydrophobic cores to adopt three-dimensional structures. One such class of miniature protein is the avian pancreatic peptide (aPP), a 36-residue miniature protein that adopts a fold dominated by an α-helix and a poly-proline type II (PPII) helix (Fig. 1A) (Blundell et al., 1981). The globular fold of aPP is stabilized by the presence of a hydrophobic core of stacked aromatic groups, branched aliphatic amino acids and prolines. Another miniature protein that is stabilized primarily by hydrophobic interactions includes the villin headpiece. This 35-residue subdomain is contained at the extreme C-terminal end of the chicken F-actin bundling protein villin. When isolated, the villin headpiece was found to fold into a stable three-dimensional structure containing three short α-helices (Chiu et al., 2005). The hydrophobic core of the villin headpiece is made up primarily of three stacked Phe residues and several other branched aliphatic side chains. Trp-cage proteins were discovered as minimized versions of exendin-4 (EX4), a 39-residue protein isolated from Gila monster saliva. EX4 is poorly folded in aqueous solution, however, the 20-residue C-terminal end was found to organize into a compact fold containing two short helices stabilized through close association of one Gly, one Trp, and three Pro residues (Neidigh et al., 2002). Another small protein that is stabilized through hydrophobic interactions is known as the Trp-zipper. This *β*-hairpin fold is characterized by cross-strand Trp residues that interdigitate, forming a

large hydrophobic cluster on one side of the miniature protein (Cochran et al., 2001). It was noted that replacing any of the Trp residues with other aromatic amino acids significantly reduces the stability of the structure, suggesting that the hydrophobic stacking interactions are what drive the folding of this particular sequence. Remarkably, Trp-zipper motifs are able to induce tertiary structures in peptides as short as 12 amino acids when the Trp residues are placed at positions 2, 4, 9, and 11 (Cochran et al., 2001), making these among the smallest miniature proteins ever reported.

Metal Ion Complexation

Metal ions play crucial roles in myriad biological processes, including enzyme catalysis, propagation of action potentials, and gas transport. It has been estimated that nearly one-half of all known proteins contain at least one metal ion (Thomson and Gray, 1998) and that nearly one-third of all proteins require a metal to carry out their function (Waldron and Robinson, 2009). In proteins, cationic metal ions including Mg^{+2}, Ca^{+2}, Cu^{+2}, and Zn^{+2} can be used to stabilize protein folds. Such metals are often coordinated by ionized side chains or imidazole rings that are judiciously placed within the protein sequence to ensure proper folding of the peptide backbone upon complexation. Metal ions are often used to stabilize the folding of small, isolated sub-domains of large proteins that are not integrated into the core protein architecture. Such metal-chelating subdomains, which include zinc-fingers and EF hand motifs, have been found to fold into stable tertiary structures in the presence of metals when isolated from the parent protein (Ye et al., 2001; Reddi et al., 2007). These constructs represent a unique class of miniature protein in which the three-dimensional fold is stabilized by metal ion complexation.

Zinc-finger domains are arguably the most well-characterized metal binding miniature proteins. Zinc-finger motifs are generally between 25 and 30 amino acids in length and often occur as tandem repeats in large DNA-binding proteins including transcription factors, nuclear hormone receptors and nucleases. These small domains fold into highly stable structures consisting of an α-helix and a short anti-parallel β-sheet (Fig. 1B) (Elrod-Erickson et al., 1998). Zinc-fingers were originally classified by the number and order of the residues used to chelate zinc (such as Cys_2-His_2, Cys_4, and Cys_6), however, this nomenclature has been replaced by a more systematic method that classifies zinc-fingers into "fold groups" based on the overall shape of their fold. Common zinc-finger folds include Cys_2-His_2 ("classic zinc-finger"), gag knuckle, treble clef, zinc ribbon, and Zn_2-Cys_6 (Krishna et al., 2003). Notably, the Cys_2-His_2 zinc-finger motif has been used as a scaffold to develop small protein-based therapeutics that inhibit gene transcription (Mino et al., 2006; Kang et al., 2008) and for use as tools to study mechanisms of cell-uptake (Appelbaum et al., 2012; Holub et al., 2013). Among other miniature proteins that bind metal ions include the EF hand motif. The EF hand is a helix–loop–helix structure that is found in many calcium binding proteins including calmodulin and calerythrin (Tossavainen et al., 2003). Isolated EF hand motifs are typically between 12 and 20 amino acid residues in length and bind Ca^{+2} primarily through side-chain carboxylates at positions 1, 3, 5, 12, or in pseudo EF hands at positions 1, 4, 6, 9 (Zhou et al., 2006). Examples of isolated EF-hand motifs are still relatively rare in the literature. However, there are groups currently working to develop EF hand motifs as protein-based therapeutics that serve to chelate calcium (Aggarwal et al., 2012; Davis et al., 2016).

Covalent Linkages

Covalent bonds are among the strongest forces contributing to protein tertiary structure. Apart from the covalent bonds linking atoms within the peptide backbone and side-chain functional groups, globular protein folds can be stabilized through covalent linkages between side chains. Cysteine is the only naturally occurring amino acid that can form disulfide bonds with other cysteines. Other covalent bonds, such as desmosine and isodesmosine, are formed from the crosslinking of allysine and lysine in the production of elastin (Umeda et al., 2001). However, desmosine-type of crosslinks have not been found in other proteins and disulfide linkages remain the most common covalent bonds used to stabilize protein tertiary structure.

Among the smallest proteins that utilize disulfide linkages to stabilize tertiary folds are the α-conotoxins. These are small (10–30 amino acid) proteins isolated from marine cone snails that include one or more disulfide bonds. Notably, these proteins display remarkable structural and sequence heterogeneity; the smallest conotoxins may contain only one turn of an α-helix as the singular element of secondary structure, while larger conotoxins can include helices, sheets, and turns (Santos et al., 2004). Over the past decades, conotoxins have been discovered to exert remarkable pharmacological utility, including the ability to target nicotinic acetylcholine receptors, voltage gated Na^+ channels and NMDA receptors (Lewis et al., 2012). Other prominent examples of miniature proteins that contain disulfide linkages include toxins found in venoms of arthropods. This diverse array of miniature proteins

includes apamin, an 18-amino acid peptide found in bee venom that contains a short α-helix and a β-turn stabilized by two disulfide linkages (Fig. 1C) (Li et al., 2009); scyllatoxin (ScTx), a 31-amino acid peptide isolated from the venom of scorpions that folds into an α/β structural motif stabilized by three disulfide linkages (Fig. 1D) (Martins et al., 1995); and hanatoxin a 35-amino acid protein found in tarantula venom that adopts a folded structure designated an "inhibitor cysteine knot" (ICK) (Takahashi et al., 2000). The ICK fold is a motif that is commonly found in invertebrate toxins containing two to four disulfide linkages and up to three β-strands. The predominant biological function of these toxins and their derivatives is to inhibit voltage-gated ion channels on the surface of neurons. This inherent ability that has led to their use to study nerve cell conductance (Gilchrist et al., 2014; Morales-Lazaro et al., 2015). Furthermore, this mechanism of action has led to the development of such toxins as potential therapeutics in treating cancer and cardiovascular disease (Ding et al., 2014; Bhavya et al., 2016).

Disulfide linkages are also used to stabilize the folds of endogenous miniature proteins found in humans, including peptide hormones and growth factors. Some notable examples of disulfide-containing miniature proteins include: somatostatin, vasopressin, oxytocin, insulin, and epidermal growth factor (EGF). Somatostatin is a macrocyclic, 14-amino acid peptide hormone containing a single intramolecular disulfide bond that acts to inhibit glucagon and insulin secretion (Harris, 1994). Vasopressins are small (9-amino acid) proteins that act to regulate water retention by causing the kidneys to conserve water (Oosthuyzen et al., 2016). Oxytocin is 9-amino acid peptide hormone that is similar in structure and sequence to vasopressin, however, it primarily plays a role in sexual reproduction and social bonding (Neumann, 2008). Vasopressin and oxytocin folds are stabilized by a single disulfide linkage, and are so similar in sequence and structure that each can cause cross-reactivity when administered exogenously (Li et al., 2008). Insulin is a small protein composed of 51-amino acids containing two peptide chains connected by two disulfide bonds (Vinther et al., 2013). This well-characterized biomolecule functions by regulating the level of glucose in the bloodstream (De Meyts, 2004). EGF is a 53-amino acid protein that folds into a three-dimensional motif stabilized by three intramolecular disulfide linkages (Ogiso et al., 2002). The primary biological function of EGF is to stimulate cell growth, proliferation, and differentiation (Carpenter and Cohen, 1990). The therapeutic relevance of endogenous miniature proteins cannot be overstated. Indeed, many such miniature proteins have been developed as therapeutics to treat cancers (Rai et al., 2015), diabetes (Meah and Juneja, 2015) and, in the case of oxytocin, as a medication to facilitate childbirth (Shyken and Petrie, 1995).

RE-ENGINEERING MINIATURE PROTEINS FOR BIOMOLECULAR RECOGNITION

Proteins rely on their molecular architecture and sequence specificity to perform biological functions. To this end, the structural components of proteins such as helices, sheets, and turns are often used in concert to maximize the efficiency of biomolecular interactions. Natural proteins have evolved the ability interact with a wide variety of biomolecules including carbohydrates, nucleic acids, lipids, small molecule metabolites, and other proteins. Miniature protein folds that include elements of secondary and tertiary structure can, in theory, be developed as modulators of such interactions. Importantly, miniature protein folds allow for discrete binding epitopes to be displayed in patterns that are complementary to native biopolymer surfaces. The *de novo* synthesis of miniature proteins has allowed researchers to target many therapeutically relevant biomolecular interactions with re-engineered miniature protein scaffolds that now serve as attractive leads for potential therapeutics. This section discusses current techniques to develop and engineer small protein scaffolds that target discrete biomolecular interactions. It should be noted that not all miniature proteins discussed herein have seen use in the clinic. However, emphasis is placed on those miniature proteins that target therapeutically relevant biomolecular interactions and are therefore considered to have significant potential to be developed as drugs.

Endogenous Miniature Protein Therapeutics

Arguably, miniature protein-based therapeutics have existed for nearly 100 years. The first therapeutic miniature protein to achieve widespread clinical use was insulin (De Meyts, 2004). Insulin isolated from animal pancreatic cells has been used to treat type 1 diabetes in humans since 1922 (Quianzon and Cheikh, 2012). However, pure insulin was difficult to isolate from whole organisms and often caused adverse immunogenic responses in patients. Subsequent advancements in the chemical synthesis of proteins in the 1960s afforded the development of synthetic "human" insulin that significantly increased potency and reduced immunogenicity. Biosynthetic insulin is currently produced through recombinant DNA techniques. Despite these advances, however, synthetic yields of insulin have remained modest because of the challenging structure of the hormone and many biologically relevant analogs have been

difficult to produce using recombinant or synthetic methods. In 2014, Zaykov and co-workers used a structure-based approach to identify an optimized linear single chain insulin precursor that proved of high utility in the preparation of previously unattainable analogs (Zaykov et al., 2014). In this study, the authors used the insulin protein as a scaffold to systematically design an insulin precursor dubbed DesDi. The DesDi miniature protein was developed by directly linking the C-terminus of the B chain with the N-terminus of the A chain. This 49-amino acid peptide folds with high efficiency into a stable, insulin-like structure that was highly tolerant to amino acid substitution at various corresponding insulin residues. Furthermore, DesDi could be treated with endoproteinases to be processed into a fully functional, insulin-like miniature protein. Several sequence analogs were developed that optimized folding and allowed the researchers to study the influence of various amino acid side chains on biological activity. The DesDi peptide is now being used as a scaffold to develop analogs that have improved glycemic control and bioavailability over natural insulin (Zaykov et al., 2016).

Another naturally occurring miniature protein that has seen use in the clinic is EGF. Since first being isolated over 50 years ago from the submaxillary gland of mice, studies on the biological effects of EGF have followed at a rapid pace (Cohen, 1962; Hardwicke et al., 2008). Endogenous EGF is a 6 kDa protein that stimulates cellular growth, proliferation, and differentiation in normal tissue through interaction with the EGF receptor. As early as 1973, EGF was being investigated as a potential therapeutic to facilitate wound healing because of its ability to stimulate proliferation and migration of keratinocytes (Savage and Cohen, 1973). Treatment results in this area have generally been positive and recombinant EGF is now produced by several pharmaceutical companies as a therapeutic to enhance wound healing (Mola et al., 2003; Frew et al., 2007). Despite these successes however, the mitogenic effects associated with EGF have been implicated in the spread of epithelial malignancies and the majority of research devoted to EGF is focused on its detrimental effects (Velu et al., 1987; Lemoine et al., 1991; Navas et al., 2012). Nevertheless, recent advancements in polymer drug delivery systems designed to stabilize and enhance delivery of therapeutic agents to affected areas has led to renewed interest in the use of EGF as a potential therapy to enhance wound healing. In theory, such polymers could be used to encapsulate EGF and deliver the drug only to areas of increased vascular permeability, thus decreasing the deleterious effects of exposing healthy tissue to EGF. Research in this area is currently ongoing and several groups have reported on the beneficial effects of using polymer-based drug delivery systems to enhance the efficacy of therapeutic EGF proteins (Lai et al., 2014; Chereddy et al., 2015).

Protein Grafting onto Miniature Protein Scaffolds

In addition to using natural miniature proteins as therapeutic agents, researchers have utilized a technique known as protein grafting to re-engineer existing miniature protein scaffolds to target therapeutically relevant biomolecular interactions. Protein grafting is defined as the transfer of a sequence-specific binding epitope of one protein to the surface of another protein (Fig. 2). Typically, the epitope is grafted onto a surface that shares a high degree of structural homology with the original domain. Protein grafting has been used extensively in the development of non-natural miniature proteins that can target a wide variety of biomolecules, including DNA, RNA, carbohydrates, small molecule metabolites and other proteins (Cooper and Waters, 2005). One of the earliest reports of using protein grafting to develop inhibitors of protein–protein interactions was published over 20 years ago by Smith et al., (1995). This group used phage display to replace a surface loop on the human tissue-type plasminogen activator (t-PA) with the HCDR3 sequence of the monoclonal antibody Fab-9. Fab-9 antibodies are known to target the platelet integrin $\alpha_{IIb}\beta_3$ and have been used in the development therapeutic antibodies (Smith et al., 1994). Notably, the resulting variant of t-PA protein bound $\alpha_{IIb}\beta_3$ integrin with nanomolar affinity and retained full enzymatic activity. More recently, loop grafting has been used to develop miniature proteins that bind $\alpha_v\beta_3$ integrins (Silverman et al., 2009). In this report, the authors used agouti-related protein (AgRP), a 4 kDa cysteine-knot protein that contains four solvent exposed loops, as a scaffold to develop a miniature protein that was able to target $\alpha_v\beta_3$ integrins. Yeast display was used to generate a library of variants that contained an RGD-integrin recognition motif that was flanked by randomized residues. The authors were successful at designing highly specific miniature proteins that bound $\alpha_v\beta_3$ with sub-nanomolar affinity. Efforts are currently underway to develop these miniature proteins into therapeutics that inhibit the migration of cancer cells in humans.

Developing Miniature Proteins to Target DNA

Apart from loop sequences, protein grafting has also been applied to other elements of secondary structure, including α-helices, poly-proline helices, and β-sheets. In 1999, Zondlo *et al*. used aPP as a scaffold

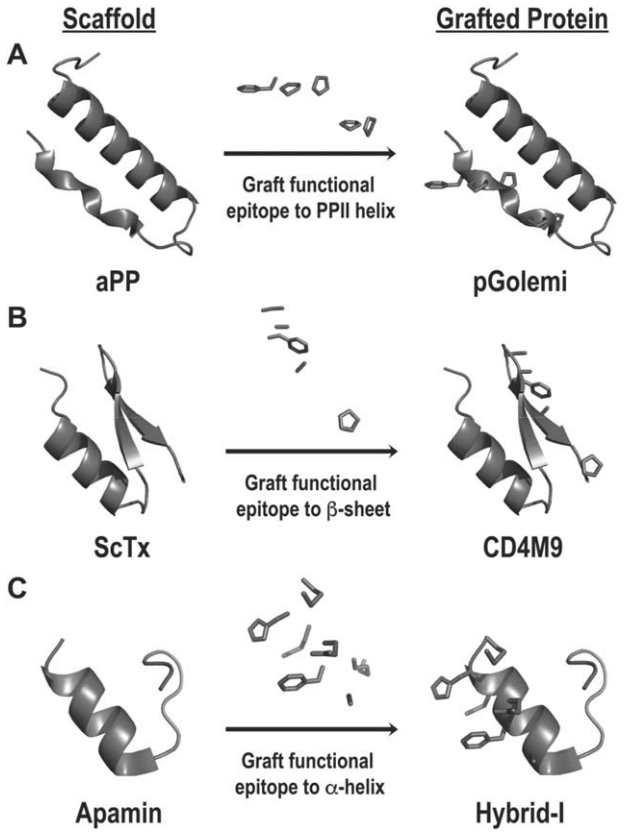

Fig. 2. Generalized strategies for protein grafting onto different structural folds. (**A**) The FPPPP epitope of ActA is grafted onto the PPII helix of aPP to generate the functional miniature protein pGolemi, which inhibits bacterial motility in biological extracts (Golemi-Kotra et al., 2004). (**B**) Residues from CD4 are grafted onto the β-sheet of ScTx to generate CD4M9, a miniature protein that targets the HIV-1 envelope glycoprotein gp120 (Vita et al., 1999). (**C**) Residues from the S-peptide of RNase A are grafted onto the α-helix of apamin to generate the versatile miniature protein Hybrid-I, which can reactivate nuclease activity and induce antibodies that cross-react with RNase (Pease et al., 1990).

to develop a miniature protein capable of targeting DNA with high specificity (Zondlo and Schepartz, 1999). In this study, amino acid residues of the helical transcription factor GCN4 required for DNA recognition were grafted onto the α-helix of aPP. The authors used phage display to randomize residues along the surface of the aPP α-helix with amino acids used by GCN4 to recognize the CRE half-site (hsCRE) of DNA. Importantly, residues required for aPP folding were not modified. Following multiple rounds of selection, the aPP-based miniature protein PPBR4^{SR} was found to bind the hsCRE DNA sequence with 1.5 nM affinity. Interestingly, the α-helix of PPBR4^{SR} had more than 60% of its residues modified from wild-type aPP and only attained a fully α-helical conformation in the presence of hsCRE DNA sequences. Nevertheless, this report showed that it was possible to use protein grafting in the development of a miniature protein that can target specific DNA sequences. In addition to aPP, the villin headpiece domain has been re-engineered to target DNA. Morii and co-workers used a structure-based design strategy to construct a miniature protein based on the villin headpiece domain that binds to GCN4-binding sequences of DNA (Morii et al., 2002). The authors noted that the small, well-folded α3 helix of the villin headpiece would serve as an ideal scaffold on which to graft the DNA-binding epitope of GCN4. Sequence alignments between villin and GCN4 showed that helix α3 needed to be modified at seven positions to display the GCN4 binding epitope. Furthermore, the authors extended the villin headpiece by six residues in order to include the entire DNA-binding region of GCN4. Once designed, plasmids coding for the DNA-binding villin headpiece (bVIL) were generated and expressed from bacterial cells. Following expression and purification, the ability for bVIL to bind DNA using gel-shift assays was evaluated. The authors found that bVIL bound to DNA in a ridged-body mechanism and did not show affinity or selectivity for DNA higher than that of the GCN4 basic region peptide. However, the addition of a leucine zipper region to the folded villin headpiece domain (bVILZIP) restored sequence-specific DNA binding comparable to that of GCN4. It was also demonstrated that the dimeric bVILZIP construct undergoes a significant conformational change upon DNA binding, indicating that structural transitions play a role in the induced-fit binding mechanism of GCN4•DNA interactions. From a therapeutic standpoint, such DNA-binding miniature proteins have enormous potential to be developed as drugs that can influence transcription patterns in cells.

Targeting Protein–Protein Interactions Using Re-Engineered aPP Scaffolds

The aPP miniature protein has been used as a scaffold to develop ligands that target therapeutically relevant protein–protein interactions. Bcl-2 is an anti-apoptotic mitochondrial membrane protein that functions along the intrinsic apoptosis pathway (Cory et al., 2003). Helical pro-apoptotic proteins, such as Bak, bind and neutralize anti-apoptotic Bcl-2 proteins on the mitochondrial membrane, facilitating the release of cytochrome c. In 2001, Chin *et al*. used phage display to graft residues from the helical Bcl-2 interaction domain of Bak required for Bcl-2 recognition to the α-helix of aPP (Chin and Schepartz, 2001). One construct, designated PPBH3–1, adopted a stable aPP-like fold and bound Bcl-2 with a K_d of 52 nM. By contrast,

native Bak peptides were intrinsically disordered in solution and bound Bcl-2 with a K_d of 5 μM. The authors concluded that the enhanced affinity for the aPP-like molecule to Bcl-2 over its natural ligand was due to pre-organization of the otherwise unstructured Bak functional epitope. Rutledge, *et al*. used a similar strategy to graft residues from the kinase-inducible activation domain (KID) of the transcription factor CREB to the α-helix of aPP (Rutledge et al., 2003). CREB KID binds the KIX domain of the transcriptional coactivator protein CBP through a relatively shallow binding groove on the CBP surface, representing a challenging target for molecules designed to inhibit this interaction. Nevertheless, the authors were successful in designing several aPP-based miniature proteins (PPKID1–5) that bound the CBP KIX domain with affinities ranging from 515 nM to 1.2 μM. Finally, the poly-proline type II (PPII) helix of aPP has been used as a scaffold for protein grafting. In 2003, Golemi *et al*. reported on the development of an aPP-like protein (pGolemi) that bound the EVH1 domain of the protein Mena (Golemi-Kotra et al., 2004). Mena proteins are involved in regulating actin cytoskeletal dynamics and bind the PPII helices of proline-rich proteins, including the FPPPP (FP_4) motif of ActA. ActA is a protein used by bacteria to propel themselves through infected host cells (Machner et al., 2001). In this study, the authors grafted the FP4 motif of ActA to the PPII helix of aPP to generate pGolemi (see Fig. 2A). The pGolemi miniature protein was found to bind Mena with a K_d of 700 nM and was observed to significantly inhibit the motility of *L. monocytogenes* bacteria in *Xenopus* egg cytoplasmic extract. Taken together, these studies demonstrate the utility of grafting functional epitopes onto a pre-organized miniature protein scaffold to develop ligands that can target and disrupt therapeutically relevant protein–protein interactions.

Miniature Protein-Based Antiviral Agents

Protein grafting has also been utilized to develop miniature proteins that function as antiviral agents. One example of this strategy includes the work of Sia *et al*. in which the complete binding epitope of an HIV-1 C-peptide was grafted onto the surface of a GCN4 leucine zipper protein (Sia and Kim, 2003). The C-peptide is derived from the C-terminal region of the HIV-1 gp14 protein and has been used as a potent inhibitor of viral entry into cells (D'Souza et al., 2000). However, the native C-peptide is inherently unstructured in solution and is susceptible to proteolytic degradation. By grafting the C-peptide binding epitope to a structured protein such as GCN4, the authors were able to develop a well-folded miniature protein that exhibited antiviral activity approaching that of the natural ligand. The miniature protein, designated C34-GCN4, was 34 residues in length and inhibited HIV-1 viral infection in cells with an IC_{50} of 4.6 nM. Notably, this inhibitory concentration compared well with the IC_{50} of 3.1 nM for the natural ligand. While there is enormous therapeutic potential of such miniature proteins to inhibit viral fusion through direct interaction with the gp14 viral protein, the authors noted that C34-GCN4 miniature proteins could also be developed as immunogens to elicit antibodies that neutralize HIV-1 infection. Another approach to develop miniature proteins as antiviral agents involved the rational design of a small protein that inhibits the interaction between the gp120 envelope glycoprotein of HIV-1 and CD4 receptors on the surface of T-helper cells (Vita et al., 1999). Analysis of the crystal structure of CD4 bound to gp120 has indicated that the CD4 binding surface is centered on a protruding β-hairpin motif that binds to a large (∼800 $Å^2$) depression denoted the "Phe43 cavity" (Kwong et al., 1998). For this study, researchers grafted the CD4 binding epitope to a structurally equivalent region of ScTx (see Fig. 2B). This protein, dubbed CD4M9, was shortened to 28 amino acids and was shown to inhibit gp120 binding to soluble CD4 with an IC_{50} that was 100-fold higher than that of native CD4, indicating suboptimal interaction with the target protein. Further refinement of the miniature protein sequence revealed that replacing Phe23 of CD4M9 with a biphenylalanine enhanced the binding affinity for gp120 by more than 200-fold to ∼7.5 nM.

Zinc-finger domains have also been utilized as scaffolds to develop antiviral agents. In 1999, McColl and co-workers used a structure-based approach to design RNA-binding zinc-finger proteins that target the HIV-1Rev response element (RRE) (McColl et al., 1999). The Rev protein is a transactivating protein that binds to RREs on HIV-1 RNA. In isolation, however, the RRE-binding peptide of Rev binds modestly to RNA with affinity that is proportional to its helical content (Tan et al., 1993). McColl *et al*. hypothesized that grafting the RRE-binding peptide sequence onto a well-folded miniature protein scaffold would greatly enhance its stability and lead to high-affinity binders of RRE. To accomplish this, the authors grafted the arginine-rich helix from HIV-1 Rev protein into the α-helix of a 27-amino acid zinc-finger protein. The resultant miniature protein, ZF2-REV, was shown to bind specifically to the RRE IIB site with an apparent K_d of 330 nM. Notably, this effect was zinc dependent, as the binding affinity of ZF2-REV to RRE IIb was three-fold lower in the absence of zinc. Liu *et al*. reported on the development of artificial zinc finger and Trp cage proteins to target the E6 oncoprotein of human papillomavirus (HPV) (Liu et al., 2004). HPV is

an oncovirus that infects epithelial cells and has been implicated in the development of cervical cancer (Petry, 2014). The E6 protein of HPV is believed to promote tumorigenesis by facilitating intracellular degradation of the tumor suppressor, p53. In this study, researchers developed E6-binding miniature proteins by grafting minimal consensus sequences onto the α-helix of a zinc-finger domain and a Trp-cage protein. The critical residues (LQELLGE) of the E6-binding motif from E6AP were incorporated into the helix of the third zinc finger of the Sp1 protein and a 26-residue Trp-cage scaffold derived from EX4. As with the McColl study, binding of zinc stabilized the three-dimensional structure of the zinc-finger protein. The E6-binding epitope was placed near the N-terminus of the Trp-cage miniature protein entirely within the α-helix. Interestingly, it was determined that placing the E6 binding epitope in the rigidified region of the zinc-finger α-helix abolished binding to E6. However, moving the E6-binding epitope towards the C-terminus of the protein, where a small portion of the epitope was extended beyond the helix into an unstructured region, significantly enhanced binding. This indicated that an induced-fit binding mechanism may be required for favorable E6•E6AP interactions. Of their most active miniature proteins, the authors reported that the zinc-finger protein E6apc2 inhibited the E6•E6AP interaction with an IC_{50} of 19.3 μM and the Trp-cage protein E6apn2 displayed an IC_{50} of 26.2 μM. Each of these activities were close to that of the parent peptide that this group characterized previously (Be et al., 2001).

Miniature Protein Toxins as Potential Therapeutics

Invertebrate venoms contain small proteins that function to inhibit voltage-gated ion channels (Gilchrist et al., 2014). Many of these toxic substances have been used as probes to study cell physiology and have revealed a host of information about ion channel structure/function relationships (Morales-Lazaro et al., 2015). As scaffolds, such proteins offer a unique platform from which to design potential therapeutics because their folds are often stabilized by disulfide linkages, allowing the proteins to retain their native folds even with major changes to the amino acid sequence (Vita et al., 1995, 1998). Interestingly, many protein toxins found in nature seem to occupy similar structural folds despite a high degree of sequence heterogeneity. For example, the α-KTx family of potassium channel blockers found in scorpion venom includes more than distinct 50 members (Tytgat et al., 1999). However, the majority of these toxins fold into α/β structural motifs containing a single α-helix and a two- or three-strand β-sheet linked by two to four disulfide linkages.

An early study using a toxin as a scaffold to display other protein epitopes reported on using the bee venom protein, apamin to display the S peptide from ribonuclease A (Pease et al., 1990). In this report, the authors created a hybrid apamin protein that was grafted with the peptide S sequence of ribonuclease A (see Fig. 2C). This 25-residue miniature protein contained 10 residues from the S-peptide sequence grafted onto the C-terminus of apamin; native cysteine residues that contributed to structural stability were not modified. The authors found that this hybrid protein retained its native fold and elicited an immune response in rabbits, which produced antibodies that cross-reacted with native ribonuclease A. Importantly, this finding demonstrated that protein hybrids could be used to generate antibodies against an entirely different protein with a similar sequence. More recently, apamin has been used as a template for the rational design of miniature protein-based activators of the tumor suppressor, p53 (Li et al., 2009). In this report, Li *et al*. used a structure-based approach to graft on the FYWL epitope from the N-terminal transactivation domain of p53 to the C-terminal α-helix of apamin. These apamin variants, dubbed "stingins", were between 16 and 18 amino acids in length and were found to adopt a similar fold to that of native apamin. Several stingin variants were able to bind MDM2 proteins with K_ds values ranging from 17.7 to 83.2 nM. The authors note that stingins represent a novel class of p53 activators that may have the potential to be used as antitumor agents, provided they are coupled with a therapeutically viable delivery method.

Scorpion toxins have also been gaining influence as scaffolds to develop miniature protein-based therapeutics. Early reports using scorpion toxin proteins as scaffolds to develop biologically active compounds found that the folded structure of these miniature proteins are able to tolerate drastic variations in amino acid sequence, provided the position of the structural cysteines within the primary sequence remains unchanged (Vita et al., 1995). Since this study was published, scorpion toxins have been developed into therapeutics in the form of ion channel blockers, antitumor agents and antimicrobial agents (Bhavya et al., 2016). One group reported on the development of a miniature protein derived from the potassium channel blocker BmBKTx1 that binds MDM2 and efficiently kills tumor cells (Li et al., 2008). BmBKTx1 adopts a typical α/β toxin fold that contains an α-helix and a three-strand β-sheet stabilized by three disulfide linkages. In this study, the authors designed a novel MDM2-binding protein by grafting the FLWL MDM2 binding epitope from p53

onto the α-helix of BmBKTx1. These molecules were termed "stoppin" for scorpion toxin-derived potent p53-MDM2/MDMx inhibitor. Stoppin-1 bound MDM2 *in vitro* with a K_d of 790 nM, thus validating this approach using protein grafting to turn BmBKTx1 into an effective inhibitor of the p53•MDM2/MDMx interaction. Despite potent *in vitro* activity, however, stoppin-1 was ineffective at killing tumor cells in culture and the authors attributed this to the inability of stoppin-1 to traverse the cell membrane. To circumvent this issue, five native residues within the β-sheet portion of the scaffold were replaced with arginine in order to facilitate the delivery of these molecules to their cytosolic target. Notably, this cationic miniature protein, termed "stoppin-2", was effective at killing HTC116-$p53^{+/+}$ cells at concentrations of 100 μg/mL. Scorpion toxins have also been used as scaffolds to target proteins along the Bcl-2 pathway. Work in the Holub laboratory has focused on the development of ScTx-based BH3 domain mimetics that target anti-apoptotic Bcl-2 proteins with submicromolar affinity (Harris et al., 2016). In this work, ScTx-based BH3 domain mimetics were re-engineered by grafting residues from the helical BH3 interaction domains of the pro-apoptotic Bcl-2 proteins Bax and Bak to the α-helix of ScTx. This study revealed that ScTx–Bax and ScTx–Bak mimetics containing three disulfide linkages folded into structures similar to that of wild-type scyllatoxin, but did not bind Bcl-2 proteins *in vitro*. Structural variants containing no disulfides, termed ScTx–$Bax^{\Delta\Delta\Delta}$ and ScTx–$Bak^{\Delta\Delta\Delta}$, were found exist as random coil structures in solution but were able to bind Bcl-2 with respective K_d values of 287 and 827 nM. Notably, the unstructured ScTx-BH3 domain mimetics were able to adopt an α/β motif similar to that of wild-type ScTx in the presence of structure-inducing co-solvents. Collectively, these results indicated that an induced-fit binding mechanism is required for favorable BH3•Bcl-2 interactions. The Holub group recently expanded on these findings by determining the role of single disulfide linkages on the folding and activity of ScTx-based BH3 domain mimetics (Arachchige et al., 2017). In this study, five ScTx-Bax variants were synthesized that varied in the number and position of disulfide linkages within their primary sequence. Two analogs: ScTx–Bax and ScTx–$Bax^{\Delta\Delta,8-26}$ were found to fold into structures reminiscent of wild-type ScTx. Among the folded ScTx–Bax variants, only ScTx–$Bax^{\Delta\Delta,8-26}$ bound Bcl-2 *in vitro*, displaying a K_d of 277.5 nM. Two other analogs: ScTx–$Bax^{\Delta\Delta\Delta}$ and ScTx–$Bax^{\Delta\Delta,12-28}$ were found to bind Bcl-2 with respective K_ds of 105.7 and 377.6 nM despite showing no evidence of folded structure in solution. The final analog: ScTx–$Bax^{\Delta\Delta,3-21}$ did not bind Bcl-2 *in vitro* and did not fold into a structure resembling wild-type ScTx. These results indicated that ScTx–Bax constructs containing one disulfide target Bcl-2 with higher affinity when the disulfide is placed near the middle or C-terminus of the helical BH3 domain. Thus, in the context of ScTx-based BH3-domain mimetics, the N-terminus of the helical BH3 domain benefits from more flexibility when targeting Bcl-2 proteins *in vitro*. Furthermore, these results stress the importance of judicious placement of covalent linkages within BH3 domain mimetics designed to target anti-apoptotic Bcl-2 proteins. Taken together, these studies indicate that proteins found in arthropod venoms offer unique and powerful platforms for studying the molecular nature of therapeutically relevant protein–protein interactions. It should also be noted that many protein toxins isolated from arthropod venom contain modular disulfide networks, allowing researchers to add or remove covalent linkages within the miniature protein to optimize binding to native protein surfaces.

ENHANCING CYTOSOLIC DELIVERY OF MINIATURE PROTEINS

As discussed in this research overview, miniature proteins possess many desirable attributes that potentiate their development as therapeutic agents, including synthetic tractability, proteolytic resistance, and the ability to fold into structures capable of targeting native biomolecular interactions. Despite these hallmarks, however, miniature proteins often suffer from poor bioavailability because of relatively high molecular weights, immunogenicity, and short half-lives (Bruno et al., 2013). Furthermore, miniature proteins administered to cells often become trapped in endosomes and are unable to interact with cytoplasmic targets (Fu et al., 2014). Unfortunately, this issue of endosomal entrapment severely limits the number of "druggable" biomolecular interactions that can be targeted by protein-based therapeutics. Indeed, an effective, universal strategy to deliver protein-based drugs to the cytoplasm of cells would likely revolutionize the protein therapeutic industry. Several strategies for facilitating the cytosolic access of therapeutic proteins have been proposed, including pore formation in the endosomal membrane, pH-buffering effects of protonatable groups, fusion into the lipid bilayer, and photochemical disruption of the endosome (Varkouhi et al., 2011). However, many of these approaches are toxic to cells (pore formation) or unable to be applied to deep tissue (photochemical disruption). Strategies are now emerging to incorporate functional epitopes into the primary sequence of miniature proteins that facilitate endosomal escape. This section will briefly discuss efforts to enhance the cytosolic access of miniature proteins through changes in primary sequence and will include

studies that have used miniature proteins as scaffolds to elucidate the molecular mechanisms of cytosolic delivery. Strategies to enhance bioavailability of protein-based drugs that significantly modify the protein's size and structure, such as PEGylation (Ginn et al., 2014) and palmitoylation (Oeste et al., 2014), will not be discussed here.

Protein Transduction Domains

Protein transduction domains are small peptides that are capable of traversing cell membranes. Protein transduction domains are often classified into three general categories: (i) short, cationic peptides that are composed primarily of Arg and Lys residues; (ii) hydrophobic leader sequences from cytokines and growth factors; and (iii) cell-specific peptides that have been identified through phage display libraries (Zahid and Robbins, 2012). Since the discovery of the cationic protein transduction domain Tat in the late 1980s (Ruben et al., 1989), researchers have utilized polycationic peptide sequences as a means to deliver a wide variety of biomolecules to cells, including peptides, proteins, and nucleic acids (Zahid and Robbins, 2012). This delivery method has shown that cationic peptides are able to facilitate cytosolic delivery of therapeutic agents by destabilizing the lipid bilayer (Sugita et al., 2008). While determining the exact mechanism of protein transduction remains an area of active research, it is thought that cationic residues are able to interact with proteoglycans on the negatively charged cell surface to facilitate molecular transport (Poon and Gariepy, 2007). Efficient protein transduction using such peptides often requires micromolar concentrations of the molecules being transduced and this relatively high concentration of positive charge can lead to cytotoxicity. Protein transduction domains have been used as delivery vehicles either as separate entities (Kichler et al., 2006) or as direct conjugates (Becker-Hapak et al., 2001) to the therapeutic agents being administered to the cell. However, transduction efficiency of the therapeutic agent is often reduced when the transduction domain and drug are not direct conjugates. Furthermore, covalently linking the protein transduction domain to the therapeutic agent can adversely affect trafficking and biological activity. To circumvent these issues, researchers have now turned to grafting protein transduction domains onto miniature protein scaffolds. Provided the transduction domain does not interfere with the fold of the miniature protein, this approach could, in theory, be used to transduce miniature protein-based therapeutics to cell interiors without the need for additional delivery agents. Such "scaffolded" protein transduction domains may provide researchers unique tools that can be used to study the nature of protein transduction as it relates to elements of protein structure.

Cytosolic Delivery of Miniature Proteins

The Schepartz group has reported extensively on the development of minimally cationic miniature proteins that display enhanced cell uptake and cytosolic delivery. The majority of these studies focus on re-engineering aPP miniature proteins for internalization by installing cationic residues within the PPII- or α-helix (Daniels and Schepartz, 2007, Appelbaum et al., 2012). Interestingly, their results showed that aPP scaffolds containing cationic motifs significantly surpassed the uptake efficiency of well-known protein transduction domains such as Tat and poly-arginine. Furthermore, it was demonstrated that cationic aPP-based miniature proteins were significantly more permeable to cells than their wild-type counterparts. Other studies from this group have involved varying the number and position of arginine residues within the α-helix of aPP or zinc-finger domains (Smith et al., 2008; Appelbaum et al., 2012). These studies were designed to evaluate how the stoichiometry of cationic residues within an α-helix of a miniature protein affected cell uptake and cytosolic delivery. Results from these experiments showed that up to five arginine residues could be tolerated within the α-helix of aPP without affecting the canonical aPP fold. Using innovative fluorescence-based translocation and transcriptional assays, the authors found that positioning five arginine residues on three helical faces of the α-helix significantly enhanced the cytosolic delivery of aPP and zinc-finger domains over wild-type (Appelbaum et al., 2012; Holub et al., 2013). Furthermore, it was determined that miniature proteins containing this "5.3 motif" were rapidly internalized by the cell and escaped into the cytoplasm from $Rab5^+$ endosomes. Interestingly, this trafficking pathway was different than that followed by Tat and poly-arginine, which each required transport beyond $Rab5^+$ endosomes before being released into the cytosol. The authors noted that a primary advantage of this strategy is that the cell permeability of a miniature protein could be significantly enhanced without increasing its molecular size or affecting its overall fold. In addition, these studies underscore the utility of using miniature proteins as tools to study the molecular nature of cell internalization and cytosolic delivery.

CONCLUSIONS

Over the past century, protein-based therapeutics have made a significant impact on the state of

human health. The growing influence of this unique class of drugs is evident in the increasing number of protein-based therapies that have been approved for clinical use. Proteins are considered attractive candidates for therapeutics because their natural folding and sequence specificity makes them ideal for targeting a wide range of biomolecular interactions. Miniature proteins are short, well-folded peptide oligomers that have received considerable attention because of their synthetic tractability and ability to fold into three-dimensional architectures. Despite being relatively short in sequence, miniature proteins rely on the same intramolecular interactions as larger proteins to adopt stable three-dimensional folds. Hydrophobic interactions, metal ion complexation and covalent linkages all play roles in the folding and stability of miniature proteins. It is precisely these folds that endow miniature proteins with biological function and make them attractive leads as therapeutic agents.

While the therapeutic relevance of endogenous miniature proteins such as insulin and EGF cannot be overstated, it may be argued that the true advantage of developing miniature proteins as drugs lies in their ability to be used as scaffolds that can be engineered to include functional epitopes from larger proteins. This protein grafting strategy has seen widespread use in the development of structurally diverse miniature protein scaffolds that have been re-engineered to target biomolecules including DNA, RNA, and other proteins implicated in disease. Furthermore, sequence-specific epitopes have now been grafted onto such varied folds as helices, loops, sheets, and turns, thus demonstrating the broad applicability of this approach. Given the incredible versatility of the protein grafting strategy, it may be possible, in theory, to design a complementary miniature protein for any biomolecular interaction.

Finally, a major disadvantage of protein-based drugs is their inability to traverse cell membranes. Indeed, the vast majority of protein-based therapeutics currently on the market have been designed to target extracellular receptors or serum cofactors. Unfortunately, this inherent cell-impermeability severely limits the "druggable" interactions that protein-based therapeutics are able to target. Work is now being conducted that aims to enhance the cytosolic delivery of miniature proteins through the use of unique protein transduction domains that facilitate endosomal escape. While applications of this approach to biologically functional miniature proteins remain limited, it will be exciting to see how this strategy is utilized in next-generation miniature proteins as the quest for developing ideal protein-based therapeutics continues.

AKNOWLEDGMENTS

This work was supported in part by grants from the Ohio University Research Committee, the Ohio University Baker Fund and the Office of Research and Sponsored Programs at Ohio University.

REFERENCES

Aggarwal A, Tandon S, Singla SK, Tandon C. 2012. A novel antilithiatic protein from Tribulus terrestris having cytoprotective potency. Protein Pept Lett 19:812–819.

Appelbaum JS, LaRochelle JR, Smith BA, Balkin DM, Holub JM, Schepartz A. 2012. Arginine topology controls escape of minimally cationic proteins from early endosomes to the cytoplasm. Chem Biol 19:819–830.

Appella DH, Christianson LA, Klein DA, Powell DR, Huang X, Barchi JJ Jr., Gellman SH. 1997. Residue-based control of helix shape in beta-peptide oligomers. Nature 387:381–384.

Arachchige D, Margaret Harris M, Coon Z, Carlsen J, Holub JM. 2017. Role of single disulfide linkages in the folding and activity of scyllatoxin-based BH3 domain mimetics. J Pept Sci 23:367–373.

Bak A, Leung D, Barrett SE, Forster S, Minnihan EC, Leithead AW, Cunningham J, Toussaint N, Crocker LS. 2015. Physicochemical and formulation developability assessment for therapeutic peptide delivery—a primer. Aaps j 17:144–155.

Be X, Hong Y, Wei J, Androphy EJ, Chen JJ, Baleja JD. 2001. Solution structure determination and mutational analysis of the papillomavirus E6 interacting peptide of E6AP. Biochemistry 40:1293–1299.

Becker-Hapak M, McAllister SS, Dowdy SF. 2001. TAT-mediated protein transduction into mammalian cells. Methods 24:247–256.

Bhavya J, Francois NN, More VS, More SS. 2016. Scorpion toxin polyptides as therapeutic agents: an overview. Protein Pept Lett 23:848–859.

Blundell TL, Pitts JE, Tickle IJ, Wood SP, Wu CW. 1981. X-ray analysis (1. 4-A resolution) of avian pancreatic polypeptide: small globular protein hormone. Proc Natl Acad Sci U S A 78: 4175–4179.

Bottger A, Bottger V, Sparks A, Liu WL, Howard SF, Lane DP. 1997. Design of a synthetic Mdm2-binding mini protein that activates the p53 response in vivo. Curr Biol 7:860–869.

Bruno BJ, Miller GD, Lim CS. 2013. Basics and recent advances in peptide and protein drug delivery. Ther Deliv 4:1443–1467.

Carpenter G, Cohen S. 1990. Epidermal growth factor. J Biol Chem 265:7709–7712.

Chalker JM. 2013. Prospects in the total synthesis of protein therapeutics. Chem Biol Drug Des 81:122–135.

Chapman RN, Dimartino G, Arora PS. 2004. A highly stable short alpha-helix constrained by a main-chain hydrogen-bond surrogate. J Am Chem Soc 126:12252–12253.

Chereddy KK, Lopes A, Koussoroplis S, Payen V, Moia C, Zhu H, Sonveaux P, Carmeliet P, des Rieux A, Vandermeulen G, et al. 2015. Combined effects of PLGA and vascular endothelial growth factor promote the healing of non-diabetic and diabetic wounds. Nanomedicine 11:1975–1984.

Chin JW, Schepartz A. 2001. Design and evolution of a miniature Bcl-2 binding protein. Angew Chem Int Ed Engl 40:3806–3809.

Chiu TK, Kubelka J, Herbst-Irmer R, Eaton WA, Hofrichter J, Davies DR. 2005. High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. Proc Natl Acad Sci U S A 102:7517–7522.

Cochran AG, Skelton NJ, Starovasnik MA. 2001. Tryptophan zippers: stable, monomeric beta -hairpins. Proc Natl Acad Sci U S A 98:5578–5583.

Cohen S. 1962. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the newborn animal. J Biol Chem 237:1555–1562.

Cooper WJ, Waters ML. 2005. Molecular recognition with designed peptides and proteins. Curr Opin Chem Biol 9:627–631.

Cory S, Huang DC, Adams JM. 2003. The Bcl-2 family: roles in cell survival and oncogenesis. Oncogene 22:8590–8607.

Craik DJ, Fairlie DP, Liras S, Price D. 2013. The future of peptide-based drugs. Chem Biol Drug Des 81:136–147.

D'Souza MP, Cairns JS, Plaeger SF. 2000. Current evidence and future directions for targeting HIV entry: therapeutic and prophylactic strategies. Jama 284:215–222.

Daniels DS, Schepartz A. 2007. Intrinsically cell-permeable miniature proteins based on a minimal cationic PPII motif. J Am Chem Soc 129:14578–14579.

Davis JP, Shettigar V, Tikunova SB, Little SC, Liu B, Siddiqui JK, Janssen PM, Ziolo MT, Walton SD. 2016. Designing proteins to combat disease: cardiac troponin C as an example. Arch Biochem Biophys 601:4–10.

De Meyts P. 2004. Insulin and its receptor: structure, function and evolution. Bioessays 26:1351–1362.

Dimitrov DS. 2012. Therapeutic proteins. Methods Mol Biol 899:1–26.

Ding J, Chua PJ, Bay BH, Gopalakrishnakone P. 2014. Scorpion venoms as a potential source of novel cancer therapeutic compounds. Exp Biol Med (Maywood) 239:387–393.

Elrod-Erickson M, Benson TE, Pabo CO. 1998. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Structure 6:451–464.

Fosgerau K, Hoffmann T. 2015. Peptide therapeutics: current status and future directions. Drug Discov Today 20:122–128.

Frew SE, Rezaie R, Sammut SM, Ray M, Daar AS, Singer PA. 2007. India's health biotech sector at a crossroads. Nat Biotechnol 25:403–417.

Fu A, Tang R, Hardie J, Farkas ME, Rotello VM. 2014. Promises and pitfalls of intracellular delivery of proteins. Bioconjug Chem 25:1602–1608.

Gilchrist J, Olivera BM, Bosmans F. 2014. Animal toxins influence voltage-gated sodium channel function. Handb Exp Pharmacol 221:203–229.

Ginn C, Khalili H, Lever R, Brocchini S. 2014. PEGylation and its impact on the design of new protein-based medicines. Future Med Chem 6:1829–1846.

Golemi-Kotra D, Mahaffy R, Footer MJ, Holtzman JH, Pollard TD, Theriot JA, Schepartz A. 2004. High affinity, paralog-specific recognition of the Mena EVH1 domain by a miniature protein. J Am Chem Soc 126:4–5.

Hardwicke J, Schmaljohann D, Boyce D, Thomas D. 2008. Epidermal growth factor therapy and wound healing—past, present and future perspectives. Surgeon 6:172–177.

Harris AG. 1994. Somatostatin and somatostatin analogues: pharmacokinetics and pharmacodynamic effects. Gut 35:S1–S4.

Harris MM, Coon Z, Alqaeisoom N, Swords B, Holub JM. 2016. Targeting anti-apoptotic Bcl2 proteins with scyllatoxin-based BH3 domain mimetics. Org Biomol Chem 14:440–446.

Holub JM, Larochelle JR, Appelbaum JS, Schepartz A. 2013. Improved assays for determining the cytosolic access of peptides, proteins, and their mimetics. Biochemistry 52:9036–9046.

Horng JC, Moroz V, Raleigh DP. 2003. Rapid cooperative two-state folding of a miniature alpha-beta protein and design of a thermostable variant. J Mol Biol 326:1261–1270.

Kang YA, Shin HC, Yoo JY, Kim JH, Kim JS, Yun CO. 2008. Novel cancer antiangiotherapy using the VEGF promoter-targeted artificial zinc-finger protein and oncolytic adenovirus. Mol Ther 16:1033–1040.

Kichler A, Mason AJ, Bechinger B. 2006. Cationic amphipathic histidine-rich peptides for gene delivery. Biochim Biophys Acta 1758:301–307.

Krishna SS, Majumdar I, Grishin NV. 2003. Structural classification of zinc fingers: survey and summary. Nucleic Acids Res 31:532–550.

Kritzer JA, Zutshi R, Cheah M, Ran FA, Webman R, Wongjirad TM, Schepartz A. 2006. Miniature protein inhibitors of the p53-hDM2 interaction. Chembiochem 7:29–31.

Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 393:648–659.

Lai HJ, Kuan CH, Wu HC, Tsai JC, Chen TM, Hsieh DJ, Wang TW. 2014. Tailored design of electrospun composite nanofibers with staged release of multiple angiogenic growth factors for chronic wound healing. Acta Biomater 10:4156–4166.

Leader B, Baca QJ, Golan DE. 2008. Protein therapeutics: a summary and pharmacological classification. Nat Rev Drug Discov 7:21–39.

Lemoine NR, Jain S, Silvestre F, Lopes C, Hughes CM, McLelland E, Gullick WJ, Filipe MI. 1991. Amplification and overexpression of the EGF receptor and c-erbB-2 proto-oncogenes in human stomach cancer. Br J Cancer 64:79–83.

Lewis RJ, Dutertre S, Vetter I, Christie MJ. 2012. Conus venom peptide pharmacology. Pharmacol Rev 64:259–298.

Li C, Liu M, Monbo J, Zou G, Li C, Yuan W, Zella D, Lu WY, Lu W. 2008. Turning a scorpion toxin into an antitumor miniprotein. J Am Chem Soc 130:13546–13548.

Li C, Pazgier M, Liu M, Lu WY, Lu W. 2009. Apamin as a template for structure-based rational design of potent peptide activators of p53. Angew Chem Int Ed Engl 48:8712–8715.

Li C, Wang W, Summer SN, Westfall TD, Brooks DP, Falk S, Schrier RW. 2008. Molecular mechanisms of antidiuretic effect of oxytocin. J Am Soc Nephrol 19:225–232.

Liu Y, Liu Z, Androphy E, Chen J, Baleja JD. 2004. Design and characterization of helical peptides that inhibit the E6 protein of papillomavirus. Biochemistry 43:7421–7431.

Machner MP, Urbanke C, Barzik M, Otten S, Sechi AS, Wehland J, Heinz DW. 2001. ActA from Listeria monocytogenes can interact with up to four Ena/VASP homology 1 domains simultaneously. J Biol Chem 276:40096–40103.

Martin L, Vita C. 2000. Engineering novel bioactive mini-proteins from small size natural and de novo designed scaffolds. Curr Protein Pept Sci 1:403–430.

Martins JC, Van de Ven FJ, Borremans FA. 1995. Determination of the three-dimensional solution structure of scyllatoxin by 1H nuclear magnetic resonance. J Mol Biol 253:590–603.

McColl DJ, Honchell CD, Frankel AD. 1999. Structure-based design of an RNA-binding zinc finger. Proc Natl Acad Sci U S A 96:9521–9526.

Meah F, Juneja R. 2015. Insulin tactics in type 2 diabetes. Med Clin North Am 99:157–186.

Mino T, Hatono T, Matsumoto N, Mori T, Mineta Y, Aoyama Y, Sera T. 2006. Inhibition of DNA replication of human papillomavirus by artificial zinc finger proteins. J Virol 80:5405–5412.

Mola E, Acevedo B, Silva R, Blanca T, Montero R, Herrera L. 2003. Development of cuban biotechnology. J Comm Biotechnol 9:147–152.

Morales-Lazaro SL, Hernandez-Garcia E, Serrano-Flores B, Rosenbaum T. 2015. Organic toxins as tools to understand ion channel mechanisms and structure. Curr Top Med Chem 15: 581–603.

Morii T, Sato S, Hagihara M, Mori Y, Imoto K, Makino K. 2002. Structure-based design of a leucine zipper protein with new DNA contacting region. Biochemistry 41:2177–2183.

Navas C, Hernandez-Porras I, Schuhmacher AJ, Sibilia M, Guerra C, Barbacid M. 2012. EGF receptor signaling is essential for k-ras oncogene-driven pancreatic ductal adenocarcinoma. Cancer Cell 22:318–330.

Neidigh JW, Fesinmeyer RM, Andersen NH. 2002. Designing a 20-residue protein. Nat Struct Biol 9:425–430.

Neumann ID. 2008. Brain oxytocin: a key regulator of emotional and social behaviours in both females and males. J Neuroendocrinol 20:858–865.

Nicoll AJ, Weston CJ, Cureton C, Ludwig C, Dancea F, Spencer N, Smart OS, Gunther UL, Allemann RK. 2005. De novo design of a stable N-terminal helical foldamer. Org Biomol Chem 3:4310–4315.

Oeste CL, Pinar M, Schink KO, Martinez-Turrion J, Stenmark H, Penalva MA, Perez-Sala D. 2014. An isoprenylation and palmitoylation motif promotes intraluminal vesicle delivery of proteins in cells from distant species. PLoS One 9:e107190.

Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim JH, Saito K, Sakamoto A, Inoue M, Shirouzu M, et al. 2002. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110:775–787.

Oosthuyzen W, Scullion KM, Ivy JR, Morrison EE, Hunter RW, Starkey Lewis PJ, O'Duibhir E, Street JM, Caporali A, Gregory CD, et al. 2016. Vasopressin regulates extracellular vesicle uptake by kidney collecting duct cells. J Am Soc Nephrol 27: 3345–3355.

Pease JH, Storrs RW, Wemmer DE. 1990. Folding and activity of hybrid sequence, disulfide-stabilized peptides. Proc Natl Acad Sci U S A 87:5643–5647.

Petry KU. 2014. HPV and cervical cancer. Scand J Clin Lab Invest Suppl 244:59–62; discussion 62.

Philippidis A. 2017. The top 15 best-selling drugs of 2016. Genetic Engineering & Biotechnology News http://www.genengnews.com/the-lists/the-top-15-best-selling-drugs-of-2016/77900868

Polticelli F, Raybaudi-Massilia G, Ascenzi P. 2001. Structural determinants of mini-protein stability. Biochem Mol Biol Educ 29:16–20.

Poon GM, Gariepy J. 2007. Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells. Biochem Soc Trans 35:788–793.

Pouny Y, Shai Y. 1992. Interaction of D-amino acid incorporated analogues of pardaxin with membranes. Biochemistry 31:9482–9490.

Quianzon CC, Cheikh I. 2012. History of insulin. J Community Hosp Intern Med Perspect 2: doi: 10.3402/jchimp.v2i2.18701

Raghava GPS. 2017. THPdb: A database of FDA approved therapeutic peptides and proteins. http://crdd.osdd.net/raghava/thpdb/stats.php

Rai U, Thrimawithana TR, Valery C, Young SA. 2015. Therapeutic uses of somatostatin and its analogues: current view and potential applications. Pharmacol Ther 152:98–110.

Reddi AR, Guzman TR, Breece RM, Tierney DL, Gibney BR. 2007. Deducing the energetic cost of protein folding in zinc finger proteins using designed metallopeptides. J Am Chem Soc 129:12815–12827.

Ruben S, Perkins A, Purcell R, Joung K, Sia R, Burghoff R, Haseltine WA, Rosen CA. 1989. Structural and functional characterization of human immunodeficiency virus tat protein. J Virol 63:1–8.

Rutledge SE, Volkman HM, Schepartz A. 2003. Molecular recognition of protein surfaces: high affinity ligands for the CBP KIX domain. J Am Chem Soc 125:14336–14347.

Santos AD, McIntosh JM, Hillyard DR, Cruz LJ, Olivera BM. 2004. The A-superfamily of conotoxins: structural and functional divergence. J Biol Chem 279:17596–17606.

Savage CR Jr., Cohen S. 1973. Proliferation of corneal epithelium induced by epidermal growth factor. Exp Eye Res 15:361–366.

Shyken JM, Petrie RH. 1995. Oxytocin to induce labor. Clin Obstet Gynecol 38:232–245.

Sia SK, Kim PS. 2003. Protein grafting of an HIV-1-inhibiting epitope. Proc Natl Acad Sci U S A 100:9756–9761.

Silverman AP, Levin AM, Lahti JL, Cochran JR. 2009. Engineered cystine-knot peptides that bind alpha(v)beta(3) integrin with antibody-like affinities. J Mol Biol 385:1064–1075.

Smith BA, Daniels DS, Coplin AE, Jordan GE, McGregor LM, Schepartz A. 2008. Minimally cationic cell-permeable miniature proteins via alpha-helical arginine display. J Am Chem Soc 130: 2948–2949.

Smith JW, Hu D, Satterthwait A, Pinz-Sweeney S, Barbas CF 3rd. 1994. Building synthetic antibodies as adhesive ligands for integrins. J Biol Chem 269:32788–32795.

Smith JW, Tachias K, Madison EL. 1995. Protein loop grafting to construct a variant of tissue-type plasminogen activator that binds platelet integrin alpha IIb beta 3. J Biol Chem 270: 30486–30490.

Sugita T, Yoshikawa T, Mukai Y, Yamanada N, Imai S, Nagano K, Yoshida Y, Shibata H, Yoshioka Y, Nakagawa S, et al. 2008. Comparative study on transduction and toxicity of protein transduction domains. Br J Pharmacol 153:1143–1152.

Takahashi H, Kim JI, Min HJ, Sato K, Swartz KJ, Shimada I. 2000. Solution structure of hanatoxin1, a gating modifier of voltage-dependent K(+) channels: common surface features of gating modifier toxins. J Mol Biol 297:771–780.

Tan R, Chen L, Buettner JA, Hudson D, Frankel AD. 1993. RNA recognition by an isolated alpha helix. Cell 73:1031–1040.

Thomson AJ, Gray HB. 1998. Bio-inorganic chemistry. Curr Opin Chem Biol 2:155–158.

Tossavainen H, Permi P, Annila A, Kilpelainen I, Drakenberg T. 2003. NMR solution structure of calerythrin, an EF-hand calcium-binding protein from Saccharopolyspora erythraea. Eur J Biochem 270:2505–2512.

Tyndall JD, Nall T, Fairlie DP. 2005. Proteases universally recognize beta strands in their active sites. Chem Rev 105:973–999.

Tytgat J, Chandy KG, Garcia ML, Gutman GA, Martin-Eauclaire MF, van der Walt JJ, Possani LD. 1999. A unified nomenclature for short-chain peptides isolated from scorpion venoms: alpha-KTx molecular subfamilies. Trends Pharmacol Sci 20: 444–447.

Umeda H, Takeuchi M, Suyama K. 2001. Two new elastin cross-links having pyridine skeleton. Implication of ammonia in elastin cross-linking in vivo. J Biol Chem 276:12579–12587.

Varkouhi AK, Scholte M, Storm G, Haisma HJ. 2011. Endosomal escape pathways for delivery of biologicals. J Control Release 151:220–228.

Velu TJ, Beguinot L, Vass WC, Willingham MC, Merlino GT, Pastan I, Lowy DR. 1987. Epidermal-growth-factor-dependent transformation by a human EGF receptor proto-oncogene. Science 238:1408–1410.

Vinther TN, Norrman M, Ribel U, Huus K, Schlein M, Steensgaard DB, Pedersen TA, Pettersson I, Ludvigsen S, Kjeldsen T, et al. 2013. Insulin analog with additional disulfide bond has increased stability and preserved activity. Protein Sci 22:296–305.

Vita C, Drakopoulou E, Vizzavona J, Rochette S, Martin L, Menez A, Roumestand C, Yang YS, Ylisastigui L, Benjouad A, et al. 1999. Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. Proc Natl Acad Sci U S A 96:13091–13096.

Vita C, Roumestand C, Toma F, Menez A. 1995. Scorpion toxins as natural scaffolds for protein engineering. Proc Natl Acad Sci U S A 92:6404–6408.

Vita C, Vizzavona J, Drakopoulou E, Zinn-Justin S, Gilquin B, Menez A. 1998. Novel miniproteins engineered by the transfer of active sites to small natural scaffolds. Biopolymers 47:93–100.

Waldron KJ, Robinson NJ. 2009. How do bacterial cells ensure that metalloproteins get the correct metal?. Nat Rev Microbiol 7:25–35.

Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, Wagner G, Verdine GL, Korsmeyer SJ. 2004. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. Science 305:1466–1470.

Ye Y, Lee HW, Yang W, Shealy SJ, Wilkins AL, Liu ZR, Torshin I, Harrison R, Wohlhueter R, Yang JJ. 2001. Metal binding affinity and structural properties of an isolated EF-loop in a scaffold protein. Protein Eng 14:1001–1013.

Yin H. 2012. Constrained peptides as miniature protein structures. ISRN Biochem 2012:692190.

Zahid M, Robbins PD. 2012. Protein transduction domains: applications for molecular medicine. Curr Gene Ther 12:374–380.

Zaykov AN, Mayer JP, DiMarchi RD. 2016. Pursuit of a perfect insulin. Nat Rev Drug Discov 15:425–439.

Zaykov AN, Mayer JP, Gelfanov VM, DiMarchi RD. 2014. Chemical synthesis of insulin analogs through a novel precursor. ACS Chem Biol 9:683–691.

Zhou Y, Yang W, Kirberger M, Lee HW, Ayalasomayajula G, Yang JJ. 2006. Prediction of EF-hand calcium-binding proteins and analysis of bacterial EF-hand proteins. Proteins 65:643–655.

Zondlo NJ, Schepartz A. 1999. Highly specific DNA recognition by a designed miniature protein. J Am Chem Soc 121:6938–6939.