Fluorescence-based characterization of genetically encoded peptides that fold in live cells: progress towards a generic hairpin scaffold

Zihao Cheng and Robert E. Campbell*[*]
Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

ABSTRACT

Binding proteins suitable for expression and high affinity molecular recognition in the cytoplasm or nucleus of live cells have numerous applications in the biological sciences. In an effort to add a new minimal motif to the growing repertoire of validated non-immunoglobulin binding proteins, we have undertaken the development of a generic protein scaffold based on a single β-hairpin that can fold efficiently in the cytoplasm. We have developed a method, based on the measurement of fluorescence resonance energy transfer (FRET) between a genetically fused cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP), that allows the structural stability of recombinant β-hairpin peptides to be rapidly assessed both *in vitro* and *in vivo*. We have previously reported the validation of this method when applied to a 16mer tryptophan zipper β-hairpin. We now describe the use of this method to evaluate the potential of a designed 20mer β-hairpin peptide with a 3rd Trp/Trp cross-strand pair to function as a generic protein scaffold. Quantitative analysis of the FRET efficiency, resistance to proteolysis (assayed by loss of FRET), and circular dichroism spectra revealed that the 20mer peptide is significantly more tolerant of destabilizing mutations than the 16mer peptide. Furthermore, we experimentally demonstrate that the *in vitro* determined β-hairpin stabilities are well correlated with *in vivo* β-hairpin stabilities as determined by FRET measurements of colonies of live bacteria expressing the recombinant peptides flanked by CFP and YFP. Finally, we report on our progress to develop highly folded 24mer and 28mer β-hairpin peptides through the use of fluorescence-based library screening.

Keywords: Fluorescent proteins, fluorescence resonance energy transfer, peptide engineering, protein scaffolds.

1. INTRODUCTION

Ribosomally synthesized peptides and proteins that are capable of specific molecular recognition in the cytoplasm or nucleus of the living host cell are valuable research reagents that provide an effective means to inhibit or modulate specific pathways,[1, 2] or localize specific cellular components.[3, 4] The vast majority of published research directed towards the development of such reagents has focused on the use of recombinant antibody fragments[5], so called intrabodies, that can be expressed in living cells and retain the binding specificity of the intact antibody from which they were derived. Two serious limitations of intrabodies are their rather large size (~260 residues) and their notoriously poor folding efficiency in the reducing environment of the cytoplasm due to an inability to form the intra-domain disulfide bond that is critical for their stability.[6] One approach towards addressing this limitation has been to engineer highly stable antibody fragments that reliably fold even in the absence of a disulfide bond.[7, 8] Another approach is to turn to alternative protein structures that lack disulfide bonds and engineer them to have 'unnatural' molecular recognition functions.[9, 10] This second approach sacrifices the convenience of being able to generate high affinity monoclonal binding domains by traditional immunization and hybridoma techniques.[11] However, this drawback has been more than compensated for by continued development of phage display[12] and the advent of powerful *in vitro* display and selection technologies[13] that have allowed a large number of non-immunoglobulin protein architectures to be employed as molecular recognition domains.[9, 10]

The non-immunoglobulin domain proteins best suited for conversion into molecular recognition domains have been referred to as 'generic protein scaffolds': proteins that have effectively separated the parts of their tertiary structures that confer structural stability and the parts that confer the molecular recognition function.[9, 10] Desirable features of a generic protein scaffold intended for intracellular applications include a relatively small size (to reduce non-specific binding and allow access to epitopes that are structurally occluded from larger domains), and the absence of a disulfide bond. To date, the smallest validated generic protein scaffold that meets these criteria are zinc-fingers at just 26 amino acids in length.[14] Despite their small size, these mini-proteins meet the basic criteria of a generic proteins scaffold because the

[*] robert.e.campbell@ualberta.ca; phone 1 780 492-1849; fax 1 780 492-8231; www.chem.ualberta.ca/~campbell/

Genetically Engineered and Optical Probes for Biomedical Applications IV, Samuel Achilefu, Darryl J. Bornhop, Ramesh Raghavachari, Alexander P. Savitsky, Rebekka M. Wachter, Eds., Proc. of SPIE Vol. 6449, 64490S, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.717274

part of the structure that confers structural stability (the Cys_2His_2 coordination of a zinc ion) is completely independent of the part the structure used for the molecular recognition function (5 contiguous residues on the face of an α-helix). Zinc-fingers are known to be expressed and fold intracellularly so variants derived from such libraries should be suitable for molecular recognition in the cytoplasm or nucleus.[15] An alternative approach that enables the intracellular display of variable loops in a 31 residue 'stem-loop' scaffold[16] has recently been described.[17] These zinc-finger and 'stem-loop' examples notwithstanding, non-cyclized peptides generally do not fall within the definition of a 'generic protein scaffold' due to their limited ability to form stable folded structures independent of broad amino acid diversity at least 5 or 6 contiguous residues. Note that complete amino acid diversity at 6 contiguous residues would provide a theoretical library diversity of ~10^8 amino acid sequences: a large but manageable library size for either phage or *in vitro* display approaches.

In an effort to expand the repertoire of generic protein scaffolds that can be expressed and function in the cytoplasm or nucleus of living cells, we have undertaken the development of a non-disulfide-containing β-hairpin peptide that folds into a well-defined structure and tolerates broad amino acid diversity on the contiguous face of the potential binding site. Stable non-cyclized β-hairpin peptides, particularly those of the so-called 'tryptophan zipper' (trpzip) type,[18] are extremely promising candidates for this particular application. Notably, trpzip peptides are only 16 to 20 amino acids in length, soluble, monomeric, and form highly stable β-hairpins in aqueous solution. The stability of these peptides arises from two interdigitating cross-strand tryptophan-tryptophan (Trp/Trp) pairs on one face (the Trp-face) of the hairpin (Fig. 1A).[18] Related β-hairpin structures that contain trptophan-lysine cross strand pairs have been shown to selectively bind ATP,[19] flavin,[20] and single stranded DNA.[21] We reasoned that trpzip-type peptides could be a minimal and highly effective generic scaffold if the face of the hairpin that opposes the Trp-face (the non-Trp-face) is able to tolerate a diverse array of substitutions without substantial destabilization of the β-hairpin. In most trpzip peptides reported to date, the residues with side chains pointing towards the non-Trp-face have high β-sheet propensities.[18, 22-24]

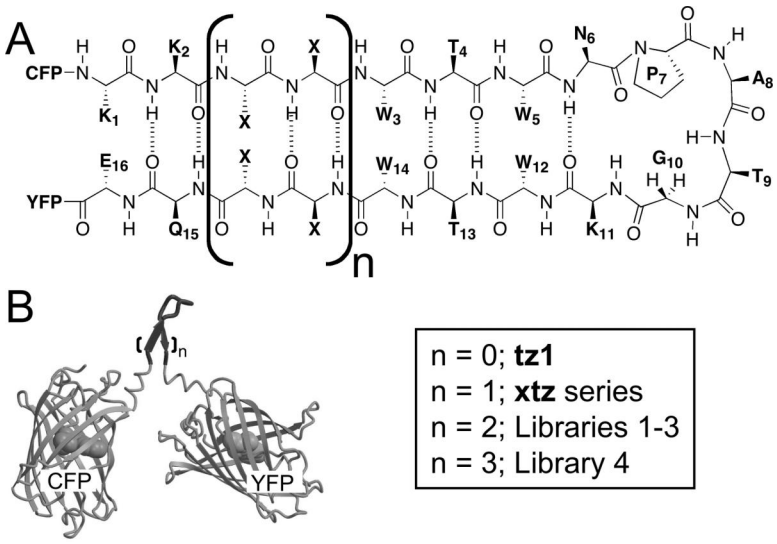

Fig. 1. Representations of the β-hairpin peptide in the context of the CFP/YFP fusion protein. A. Schematic representation emphasizing the location of inserted residue pairs (n = 0, 1, 2, or 3). Residues are denoted by standard single letter amino abbreviations. Residues labeled as 'X' could be any one of the 20 natural amino acids. B. 'To-scale' cartoon of a complete FRET construct with the location of inserted residues pairs indicated.

In order to pursue the development of trpzip-type peptides as a generic protein scaffold we required an effective means of evaluating the stability of folded peptides in the cytoplasm of living cells. The conventional approach for development of β-hairpin peptides generally involves chemical synthesis of a number of rationally and/or systematically modified peptide variants that are individually characterized by circular dichroism (CD) and/or NMR spectroscopy.[18, 22-26] A combinatorial approach based on the CD spectroscopy-guided deconvolution of peptide mixtures has also proven an effective means of developing peptides sequences that fold into β-hairpin peptides.[27, 28] A recently reported approach for characterization of β-hairpin peptide structure is to synthetically tether both a fluorophore and a quencher moiety onto

the peptide and use the rate of end-to-end collisions to study the dynamic flexibility.[29] Although these previously reported approaches have proven highly successful at developing and characterizing highly stable β-hairpin peptides, they cannot be readily adapted to the study of β-hairpin peptides in the cytoplasm of living cells. We have therefore developed a versatile and robust platform technology that inherently combines the convenience of recombinant protein expression with the versatility of fluorescence-based assays for rapidly evaluating β-hairpin structure. Our strategy for characterizing the structural stability of β-hairpins peptides is based on the presumption that attaching CFP and YFP to the ends of an intervening peptide sequence would provide a reliable and convenient fluorescence-based probe of the end-to-end distance (Fig. 1B). The CFP to YFP FRET efficiency is strongly distance dependent and we anticipated that the strength of the FRET would be higher for peptides that existed primarily in a folded β-hairpin conformation than for unstructured peptides of the same length. We have recently reported the use of this method to evaluate the *in vivo* fold-stability of a series of 16mer β-hairpin peptides (the trpzip (**tz**)-series)[30] (Fig. 1A, n = 0) and as the basis for screening of large libraries to identify a highly folded 20mer β-hairpin peptide as present in the protein known as extended trpzip 1 (**xtz1)** (Fig. 1A, n = 1; Table 1).[31]

In this manuscript we now describe the use of this method to evaluate the potential of a designed 20mer β-hairpin peptide with a 3rd Trp/Trp cross-strand pair to function as a generic protein scaffold. Quantitative analysis of the FRET efficiency, resistance to proteolysis (assayed by loss of FRET), and circular dichroism (CD) spectra revealed that the 20mer peptide is significantly more tolerant of destabilizing mutations than the 16mer peptide previously investigated.[30] Furthermore, we experimentally demonstrate that the *in vitro* determined β-hairpin stabilities are well correlated with *in vivo* β-hairpin stabilities as determined by FRET measurements of colonies of live bacteria expressing the recombinant peptides flanked by CFP and YFP. Finally, we report on our progress towards developing highly folded 24mer (Fig. 1A, n = 2) and 28mer (Fig. 1A, n = 3) β-hairpin peptides through the use of fluorescence-based library screening.

2. METHODOLOGY

2.1. General methods.

All procedures were performed essentially as previously described.[30, 31] Synthetic DNA oligonucleotides were purchased from Sigma-Genosys Canada or Integrated DNA Technologies. Pfu polymerase (Fermentas) was used for all PCR amplifications in the buffer supplied by the manufacturer. PCR products and products of restriction digest were routinely purified using the QIAquick PCR purification kit (Qiagen). Plasmids were purified using GeneJet plasmid miniprep kit (Fermentas) according to the manufacturers' protocols. Restriction enzymes were purchased from either Invitrogen or New England Biolabs. The identity of all cDNA constructs was confirmed by dye terminator cycle sequencing using the DYEnamic ET kit (Amersham Biosciences). Protein samples for all *in vitro* spectral characterization experiments were in 50mM Tris, pH 7.5.

2.2. Construction of CFP-peptide-YFP library expression vectors.

The expression vector pZC1 (derived from pBAD/His B (Invitrogen)) and methods used for cloning and expression of all peptides and peptide libraries flanked by CFP and YFP have been previously described.[30, 31] Briefly, double stranded DNA encoding the peptide libraries with appropriate sticky ends were created by slow cooling pairs of complementary single stranded oligonucleotides from 95 °C to room temperature. This synthetic annealed dsDNA was ligated into pZC1 that had been digested with Sac1 and Kpn1. Ligated product was used to transform *E. coli* DH10B (Invitrogen) by electroporation. After overnight incubation (37 °C, shaking at 225 rpm) in LB media (5 mL) containing ampicillin (0.1 mg/ml), plasmid DNA was isolated from the culture (2 mL). The resulting plasmids encode proteins of the general structure N_{term}-His_6-EK-CFP(1-230)-TSGAQ-peptide-GTSAE-YFP(5-238)-C_{term}, where 'His_6' represents six consecutive histidine residues that facilitate metal affinity purification,[32] 'EK' is the recognition sequence for the protease enterokinase, and 'peptide' is a 16mer , 20mer, 24mer or 28mer sequence as provided in Table 1.

2.3. Library screening.

The system for imaging the fluorescence of bacterial colonies grown on 10 cm Petri dishes has previously been described in detail.[30] Briefly, the light from a 175W xenon-arc lamp (Sutter Instrument Company, Novato, CA) is passed through a filter wheel holding either a 426 nm to 446 nm bandpass filter (Chroma Technology Corp., Rockingham, VT) or a 490 nm to 510 nm bandpass filter for excitation of CFP or YFP respectively. The filtered light then enters a

bifurcated fiber optic bundle (Newport Coporation, Stratford, CT) that is positioned to illuminate a 10 cm dish placed in a recessed holder on the benchtop. The fluorescence emission is filtered through a second filter wheel that is directly above the plate and that is holding either a 460 nm to 500 nm bandpass filter for CFP or a 520 nm to 550 nm bandpass filter for YFP. Images are acquired with a Retiga 1300i 12-bit CCD camera (QImaging, Burnaby, BC) positioned immediately behind the emission filter. With the exception of the lamp, the whole system is shrouded in black fabric to eliminate stray room light. Image Pro Plus (Media Cybernetics Inc., Silver Spring, MD) is used for camera control and digital image processing. Typical exposure times are 1 second and all images are background corrected with identically acquired images of a Petri dish containing LB/agar but lacking bacterial colonies. Individual colonies are automatically identified within the digital images and the pixel intensities are summed in each of 3 images corresponding to the CFP channel, the FRET channel, and direct YFP channel. For each colony the YFP/CFP ratio is calculated by dividing the average intensity in the FRET channel by the average intensity in the CFP channel. The YFP/CFP ratio is directly proportional to the FRET efficiency. The intensity of the YFP channel is directly proportional to the expression level of the fusion construct.

For library screening, colonies (typically 10-20) that exhibit a high YFP/CFP ratio are picked using a sterile tip and used to inoculate LB/ampicillin (5 ml). All cultures are grown overnight at 37 °C with shaking at 225 rpm. Protein is extracted from 1 ml of each culture using B-per II (Pierce) and emission spectra acquired using a Saphire2 fluorescence 96-well platereader equipped with monochromators (Tecan). Plasmid DNA is isolated from the culture of clones that exhibit the highest YFP/CFP emission ratio. The sequence of the peptide portion of the selected proteins is determined by DNA sequencing with a forward primer (5'-CCCTCGTGACCACCCTGACCTGG-3') that anneals to the chromophore region of the gene for CFP.

2.4. Protein expression and purification.

The pZC1 vector used for all cloning and expression allows proteins to be expressed at high levels under control of the arabinose operon. A typical protein purification procedure started with the inoculation of 1 L LB media containing ampicillin (0.1 mg/mL) and arabinose (0.2%) with a single colony of *E. coli* DH10B expressing the protein of interest. Cultures were grown overnight at 37 °C with shaking at 225 rpm. Cultures were cooled to 4 °C on ice and harvested by centrifugation (10 min, 5000 rpm). The cell pellet was resuspended in PBS buffer (130 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.6) and the cells lysed by a single passage through a French Press (Thermo Electron). Insoluble cell debris was pelleted by centrifugation at 4 °C (10 min, 10 000 rpm) and 1 ml of Ni-NTA resin (Qiagen) was added to the supernatant. Following one hour of gentle mixing at 4 °C, the supernatant was loaded onto a 6 ml polypropylene column, washed, and gravity eluted in 100 mM imidazole (pH 7.5) according to the manufacturers protocol. FRET constructs were further purified on an AKTAdesign chromatography system (Amersham Biosciences) equipped with a Hiload 16/60 Superdex 75 prepgrade gel filtration column (Amersham Biosciences) that was equilibrated with 50 mM Tris pH 7.5. Isolated proteins were concentrated with a Centricon centrifugal filter YM-30 (Millipore) and stored at 4 °C.

2.5. Spectroscopy.

Steady-state fluorescence spectra for all **xtz** series proteins (0.04 μM) were recorded on a QuantaMaster spectrofluorometer (Photon Technology International) equipped with a Xenon arc lamp. Relative rates of trypsinolysis were determined with a Safire2 monochromator-based 96-well platereader (Tecan). To a solution of each purified FRET construct (0.08 μM) was added a buffered solution of trypsin to a final concentration of 0.5 μg/ml, and the fluorescence intensity at 530 and 480 nm (excitation at 430 nm) recorded every 5 seconds. All CD spectra were obtained with 8 μM protein in a 1 mm path length cuvette on an Olis DSM 17 CD spectrometer (Olis).

3. RESULTS AND DISCUSSION

3.1. Validation of the FRET-based method with a series of 16mer β-hairpins.

In recently published work,[30] we asked whether the structural integrity of trpzip-type β-hairpins could be reliably assessed through the use of FRET measurements between a covalently linked CFP donor and YFP acceptor. To answer this question, we designed, genetically constructed, and purified a series of 6 proteins (the **tz** series) consisting of an N-terminal CFP and a C-terminal YFP flanking 16mer peptide sequences expected to fold into β-hairpin structures of various stabilities. The parent protein (that is, the one harboring the most stably folded of the β-hairpin peptides) of this

series is known as **tz1** (Fig. 1A, n = 0; Table 1) and has a peptide portion that is identical in sequence to the previously reported β-hairpin peptide HP5W4.[24] HP5W4 has been reported to be greater than 96% folded at 298 K and have a T_m of 85 °C.[24] This work led us to conclude that our FRET-based approach was generally effective at assessing the fold stability of β-hairpin peptides in live cells. An important caveat is that, in certain cases, the experimental FRET efficiencies can be perturbed by sequence-dependent factors (presumably due to effects on fluorescent protein orientation) that are irrelevant to fold stability and difficult to account for.

Table 1. Sequences of peptide and peptide libraries investigated in this work.

Name	Peptide sequence fused to N-terminal CFP and C-terminal YFP[a]			YFP/ CFP ratio	Comments
tz1	CFP	KK------WTWNPATGKWTW------QE	YFP	3.8[b]	Peptide portion identical to **HP5W4**.[24, 30]
xtz1	CFP	K*A*----*WT*WTWNPATGKWTW*RK*----*N*E	YFP	11.4[b]	'Optimized' peptide sequence identified by extensive library screening.[31]
xtz2	CFP	KK----WTWTWNPATGKWTWTW----QE	YFP	3.7[b]	**tz1** with a potential 3rd cross-strand Trp/Trp pair**.**
xtz3	CFP	KK----GTGTGNPATGKGTGTG----QE	YFP	2.2[b]	'Unstructured' version of **xtz2**
xtz4	CFP	KK----WAWAWAPATGAWAWAW----QE	YFP	3.5[b]	'Destabilized' version of **xtz2**
xtz5	CFP	KK----WTWAWAPAAGAWAWTW----QE	YFP	3.1[b]	'Destabilized' version of **xtz2**
Library 1	CFP	KA--**UX**WTWTWNPATGKWTWRK**XU**--NE	YFP	≤ 3[c]	Further extension of **xtz1**; no clones picked.
Library 2	CFP	KA--**Z**TWTWTWNPATGKWTWRKT**Z**--NE	YFP	≤ 3[c]	Further extension of **xtz1**; the two highest ratio clones (**L2a)** were identical.
Library 3	CFP	KA--**ZZ**WTWTWNPATGKWTWRK**ZZ**--NE	YFP	≤ 5[c]	Further extension of of **xtz1**; clones **L3a-e.**
Library 4	CFP	KKWTWT**Z**TWTWNPATGKWTWT**Z**TWTWQE	YFP	≤ 2[c]	Clones picked but no soluble proteins identified.
L2a	CFP	KA--*W*TWTWTWNPATGKWTWRKT*L*--NE	YFP	2.6[b]	Soluble.
L3a	CFP	KA--*NP*WTWTWNPATGKWTWRK*NF*--NE	YFP	4.0[b]	Soluble.
L3b	CFP	KA--*SP*WTWTWNPATGKWTWRK*NS*--NE	YFP	4.5[b]	Soluble.
L3c	CFP	KA--*NP*WTWTWNPATGKWTWRK*TH*--NE	YFP	4.8[b]	Soluble.
L3d	CFP	KA--*RD*WTWTWNPATGKWTWRK*ES*--NE	YFP	5.2[b]	Soluble.
L3e	CFP	KA--*NP*WTWTWNPATGKWTWRK*EG*--NE	YFP	5.5[b]	Soluble.

[a] **U** = amino acids Trp/Lys/Arg/stop, (codon WRG); **X** = amino acids Ala/Val/Gly/Thr/Met/Arg (codon RBG); **Z** = all 20 amino acids (codon NNK). Underlined residues are mutations or insertions relative to **tz1**. Italicized amino acids were selected from library screening. [b] Ratios are determined with soluble protein extracted from overnight cultures of single clones of bacteria. [c] Ratios are 'in colony' measurements and are a reliable measure of relative FRET efficiency when compared to other 'in colony' measurements.

3.2. An 'irrational' approach to developing an extended and highly folded 20mer β-hairpin.

In other recent work we have built upon our promising results with the **tz** series of proteins and used FRET to exhaustively screen an iterative series of 3 large libraries of peptide variants expressed in colonies of bacteria as fusions to both CFP and YFP.[31] The protein progenitor of the variants in these libraries was conceptually identical to **tz1** (Fig. 1A, n = 0) with each strand of the β-hairpin extended by 2 residues inserted near the bottom of the hairpin's stem (Fig. 1A, n = 1). This work resulted in the development of the protein known as **xtz1** (Table 1) that, in turn, serves as the protein progenitor for the further extended variants (Fig. 1A, n = 2 or 3) that will be described below. The development and characterization of **xtz1** will be published elsewhere[31] and will not be discussed further.

3.3. Rational design and *in vitro* characterization of a series of 20mer β-hairpins.

In parallel with our 'irrational' library-screening effort that ultimately produced **xtz1**,[31] we also attempted the rational design of what we expected might be a highly folded 20mer β-hairpin. Given the generally favorable properties and high fold stability of trpzip-type peptides, we reasoned that an elongated β-hairpin with a 3rd cross-strand Trp/Trp pair could potentially be even more stable than a comparable β-hairpin with only 2 such cross-strand pairs. We therefore designed a new protein, designated **xtz2** (Table 1), which contained a 20mer β-hairpin that was 'extended' by 4 residues relative to its **tz1** counterpart. The four additional residues in **xtz2** are a Trp/Thr pair inserted between the 2nd and 3rd residues and a Thr/Trp pair inserted between the 14th and 15th residues of the **tz1** β-hairpin (refer to Fig. 1A). In analogy to the **tz** series proteins[30] we constructed 3 additional variants of **xtz2**: one unstructured control peptide lacking tryptophan (**xtz3**); one with 6 destabilizing mutations (**xtz4**); and one with 5 destabilizing mutations (**xtz5**). The sequence of the peptide portions of the **xtz** series proteins is provided in Table 1.

Proteins **xtz2**, **xtz3**, **xtz4**, and **xtz5** were expressed in *E. coli* and purified by a procedure identical to that previously reported for the **tz** series proteins.[30] The yields of purified **xtz** series proteins were typically ~50% of their **tz** series counterparts. As shown in Table 2, the experimentally determined FRET efficiency was highest for the **xtz2** protein containing the non-destabilized parent peptide and decreased in the order (highest FRET) **xtz2** > **xtz5** > **xtz4** > **xtz3** (lowest FRET). The rates of proteolysis relative to **tz1** were inversely correlated with FRET efficiency (Table 2 and Fig. 2B), increasing in the order (slowest proteolysis) **xtz2** < **xtz5** < **xtz4** < **xtz3** (fastest proteolysis). By analogy with the **tz** series which contained variants with similar destabilizing mutations,[30] we had expected a stability ranking of (most stable) **xtz2** > **xtz5** > **xtz4** > **xtz3** (unstructured). Indeed, this ordering is consistent with both the FRET and resistance to proteolysis data.

Table 2. FRET efficiencies and rates of proteolysis for proteins **tz1** and **xtz2** through **xtz5**.

Protein	FRET efficiency (%)	Relative rate of trypsinolysis per K/R[a]
tz1	59.0 (± 0.7)[b]%	1.00 (± 0.08)[b]
xtz2	61.7 (± 1.2)%	0.14 (± 0.01)
xtz3	44.0 (± 0.8)%	12.8 (± 0.6)
xtz4	53.1 (± 0.4)%	8.76 (± 0.2)
xtz5	58.3 (± 0.7)%	3.26 (± 0.9)
[a] Rates are the initial velocities of trypsinolysis divided by the total number of lysine and arginine residues in the peptide sequence, relative to **tz1** which was assigned a value of 1.00. [b] Errors for three or more independent measurements are reported in parenthesis as ± standard deviation.		

Trpzip-type peptides have strong bands in their CD spectra at 215 and 229 nm due to exciton coupling of the two pairs of edge-to-face packed indole sidechains of tryptophan.[18, 33] Fluorescent proteins have far UV CD spectra that are typical[34] for a protein with high β-sheet content.[35, 36] Accordingly, the CD spectrum of the ~520 residue **xtz2** protein is dominated by the primarily β-sheet structure[34] of CFP and YFP which together account for ~88% of the residues of the full length protein.[30] However, close inspection of the **xtz2** CD spectrum reveals an additional positive component that is coincident with the positive component of the exciton couplet at 229 nm.[18] Subtraction of the spectrum of **xtz3**, which lacks the 4 tryptophan residues and is expected to be completely unstructured, from the spectrum of **xtz2** unmasks the expected exciton couplet in the difference CD spectrum (Fig. 2A). Likewise, the CD spectrum of **xtz3** was subtracted from those of proteins **xtz4** and **xtz5** (Fig. 2A). The strength of the exciton couplet decreased in the order (strongest couplet) **xtz2** > **xtz5** ~ **xtz4** (weakest couplet). The relative strengths of the exciton couplets reflects the relatively fold stability of the β-hairpin peptides and provides further support for the stability ranking provided above.

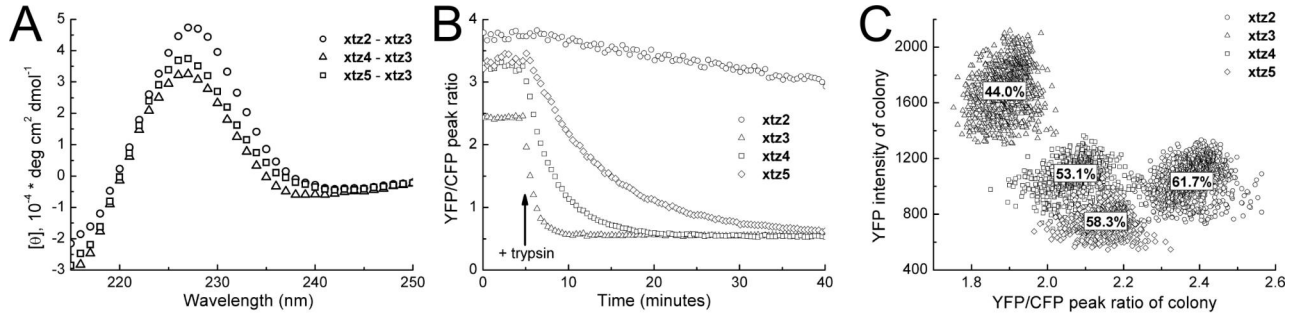

Fig. 2. Characterization of the **xtz** series proteins. A. Difference CD spectra obtained by subtracting the CD spectrum of **xtz3** from each of **xtz2**, **xtz4**, and **xtz5**. B. Proteolysis of proteins **xtz2** through **xtz5** with identical concentrations of trypsin. Trypsinolysis cleaves the acceptor YFP from the donor CFP and results in a loss of FRET. The final ratio (~0.5 in all cases) corresponds to the donor CFP only and is stable over many hours indicating that the CFP portion of the protein is not susceptible to trypsinolysis under the conditions of the experiment. The relative initial rates of proteolysis are provided in Table 2. C. The YFP/CFP ratio of proteins **xtz2** through **xtz5** in live cells. The YFP/CFP peak ratio for each individual colony was determined by ratioing the average intensity in the YFP channel by the average intensity in the CFP channel. The YFP intensity for each individual colony is the average intensity for YFP emission when YFP is directly excited at 500 nm. Superimposed on each set of data points is the *in vitro* determined FRET efficiency for the purified protein.

3.4. Assessing the fold stability of the xtz series proteins in live cells.

The *in vitro* characterization of the **xtz** series proteins provided strong support for the β-hairpin portion of **xtz2** being highly structured, the β-hairpin portions of **xtz4** and **xtz5** somewhat less so, and the peptide portion of **xtz3** completely unstructured. We next asked whether this trend would also be true of these peptides when presented in the intracellular milieu. To answer this question we expressed each protein in the **xtz** series in *E. coli* and, using a custom-built digital imaging system, we determined the ratio of YFP to CFP fluorescence (excitation of CFP at 430 nm) in individual colonies. This approach does not provide the actual FRET efficiency but it does provide a YFP/CFP intensity ratio that is proportional to the FRET efficiency. In practice, we find that it is difficult to obtain a perfect correlation between the 'in colony' ratios and the *in vitro* FRET efficiencies, apparently because of an age-dependent colony autofluorescence and inner filter effects that we have been unable to properly correct for.[30] However, if the bacteria are plated at a consistent density (100-250 per plate), and all plates are treated in an identical fashion, ratios measured on different plates display excellent qualitative agreement with the *in vitro* FRET efficiencies. Shown in Fig. 2C is representative ratio data obtained for all **xtz** series proteins on separate plates and imaged 'in colony'. We have plotted this data as colony ratio vs. colony intensity to better represent the intrinsic variability observed for these homogeneous populations of bacteria. We attribute the pronounced and reproducible differences in colony intensity to protein-specific differences in expression levels. It is interesting to note that the proteins containing less structured peptide sequences tend to be expressed at higher levels. From these results we conclude that the relative stability of the **xtz** series β-hairpins in the intracellular environment is similar to their *in vitro* stability.

3.5. Characterization of the isolated peptide portion of xtz2.

Encouraged by our apparent success at designing a highly stable extended β-hairpin structure, we recombinantly expressed and purified the peptide portion of **xtz2** (designated **xtz2-peptide**) with no fluorescent proteins attached to either end. This work will be described in detail elsewhere[31] and will only be summarized here. Trpzip-type β-hairpin peptides are typically soluble up to at least millimolar concentrations in buffered aqueous solution and show the characteristic exciton couplet at 215 nm and 227 nm in their CD spectrum.[18] Since the **xtz2-peptide** has significantly more exposed hydrophobic surface than other previously reported trpzip-type peptides, the fact that it was not soluble at concentrations greater than ~70 μM did not come as a great surprise. However, we were quite surprised to find that the CD spectrum of this relatively dilute peptide showed no peak attributable to the exciton couplet. Addition of 50% trifluoroethanol (TFE) slightly improved the solubility and induced the formation of the trpzip-type β-hairpin structure as revealed by the appearance of the characteristic exciton couplet in the CD spectrum. It is apparent that the **xtz2-peptide** in aqueous solution only folds into a stable β-hairpin structure when in the context of the full-length **xtz2** fusion protein with CFP fused on one end and YFP on the other. So, despite its excellent tolerance of multiple destabilizing mutations, the poor solubility and context-dependent folding of the **xtz2-peptide** eliminate it as a candidate generic protein scaffold.

3.6. Beyond the 20mer: efforts directed towards the development of 24mer and 28mer β-hairpins.

The poor solubility of **xtz2-peptide** rendered it a 'dead-end' for development as a generic peptide scaffold or as the basis for further extension of the hairpin. In contrast, the peptide portion of **xtz1** does appear to hold a lot of promise for further work in these directions. In recent work[31] we have shown that the 20mer β-hairpin of **xtz1** is highly soluble in water and highly structured. For these reasons, the peptide portion of **xtz1** has served as the progenitor for our continuing efforts to create highly folded 24mer and 28mer β-hairpin peptides. Since the 'rational' design-based approach had not proven particularly successful, we decided to return to the 'irrational' screening-based approach for the identification of the further extended peptides. In our first such effort, we constructed a library of protein variants (**Library 1**, Table 1) in which 2 partially randomized residues were inserted before the 3rd and after the 18th residues of the peptide portion of **xtz1**. The particular subsets of amino acids represented at each position were chosen because these amino acids had frequently been identified at cross-strand positions during screens directed towards the development of **xtz1**. Several plates of colonies expressing members of **Library 1** were imaged and the YFP/CFP ratios determined for each individual colony. Disappointingly, the highest 'in colony' ratios observed were ~3; only slightly greater than the 'in colony' ratio of ~2 observed for similar constructs (such as **xtz3**) that have unstructured peptides linking CFP and YFP.

For our second attempt, we constructed a similar library of protein variants (**Library 2**, Table 1) in which the inserted residues directed towards the non-Trp-face were threonine and the inserted residues directed towards the Trp-face were completely randomized. As with **Library 1** the highest 'in colony' ratios were only ~3, however we did pick and sequence two of the highest ratio clones. Interestingly, both clones (**L2a**, Table 1) were found to have identical residues at the randomized positions; a tryptophan at position 3 and a leucine at position 22. The hydrophobic character of both residues, and their expected close proximity in cross-strand positions, suggests that these two residues could be stabilizing the hairpin structure through Van der Waals' contact.

For our third attempt, all four inserted residues were randomized to all possible amino acids (**Library 3**, Table 1). We were encouraged to observe a range of 'in colony' YFP/CFP ratios up to values of ~5 for this larger library. Five clones (**L3a-e**, Table 1) with high YFP/CFP ratio were picked and the DNA sequenced to reveal the composition of the peptide sequence. The peptide sequences did not reveal a single clear consensus, though there was a strong preference for asparagine and proline at positions 3 and 4, respectively. While none of these 24mer clones exhibit a YFP/CFP ratio as high as that observed for **xtz1**, we find these preliminary results very encouraging. Further work to optimize the sequences of the 24mer peptides as well as to structurally characterize the isolated peptide portions by CD and NMR spectroscopy are underway.

In our sole attempt at creating a highly folded 28mer hairpin sequence, we semi-rationally designed a library of trpzip-type variants with 4 Trp/Trp cross-strand pairs interrupted by a central cross-strand Trp-face pair of randomized residues (**Library 4**, Table 1). The 'in colony' ratios for this library were very low (~2). Several clones were picked but it was not possible to extract any soluble protein from overnight cultures.

With the benefit of hindsight and experience, we now expect that the most effective approach for development of 28mer or longer hairpin peptides will be an incremental approach with extensive optimization of the intermediates. We also expect that these efforts will be greatly facilitated by a complete structural characterization of the intermediates by 2-dimensional NMR spectroscopy. In the absence of structural information, it is possible (or likely) that subsequent extension by 'blind' insertion of new residues could disrupt or completely abolish the stabilizing interactions identified in earlier rounds of screening.

CONCLUSION

The β-hairpin peptide portion of **xtz2** exhibits some, but not all, of the properties desirable in a generic protein scaffold. This 20mer β-hairpin could tolerate as many as 6 destabilizing mutations without significant loss of secondary structure, but once removed from the context of the CFP and YFP fusion protein, it exhibited very poor solubility and did form the expected β-hairpin structure in the absence of TFE. Our ideal β-hairpin scaffold would be highly solubility, exhibit efficient and context-independent folding, and have a high tolerance for destabilizing mutations. Of the 3 candidate peptides we have now investigated in detail (that is, the peptide portions of **tz1**, **xtz1**, and **xtz2**)[30, 31] we currently favor the **xtz1-peptide** as the most promising potential generic peptide scaffold. Accordingly, we are now proceeding with the construction of phage-displayed[12] libraries of **xtz1-peptide** in which 6 positions have been

randomized to all possible amino acids. These libraries will be panned against a variety of intracellular protein targets in order to select for high-affinity binding peptides.

In summary, we have developed and validated a new method for rapidly assessing the structural stability of trpzip-type β-hairpin peptides both *in vitro* and in live bacteria cells. This system has been used to characterize a previously unreported series of 20mer tryptophan zipper β-hairpins that contain a third cross-strand Trp/Trp pair. Furthermore, we have used this system to screen large number of variants and identify 24mer peptides that show promise as extended and folded β-hairpins. Although our current efforts are focused on tryptophan zipper β-hairpin peptides, we expect that this platform technology could be used to screen libraries of practically any small genetically encoded peptide motif for variants that exhibit higher folding efficiency or stability in the cytoplasm of live cells.

ACKNOWLEDGMENTS

This research was made possible with support from the University of Alberta, the Canadian Foundation for Innovation, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and an Alberta Ingenuity New Faculty Grant. The authors thank Wayne Moffat for assistance with CD spectroscopy, and the University of Alberta Molecular Biology Service Unit for analysis of DNA sequencing reactions. REC holds a Canada Research Chair in Bioanalytical Chemistry.

REFERENCES

1. W. A. Marasco, "Intracellular antibodies (intrabodies) as research reagents and therapeutic molecules for gene therapy," *Immunotechnology* 1(1), 1-19 (1995).
2. R. E. Kontermann, "Intrabodies as therapeutic agents," *Methods* 34(2), 163-170 (2004).
3. C. Nizak, S. Martin-Lluesma, S. Moutel, A. Roux, T. E. Kreis, B. Goud, and F. Perez, "Recombinant antibodies against subcellular fractions used to track endogenous Golgi protein dynamics in vivo," *Traffic* 4(11), 739-753 (2003).
4. C. Nizak, S. Monier, E. del Nery, S. Moutel, B. Goud, and F. Perez, "Recombinant antibodies to the small GTPase Rab6 as conformation sensors," *Science* 300(5621), 984-987 (2003).
5. R. E. Bird, K. D. Hardman, J. W. Jacobson, S. Johnson, B. M. Kaufman, S. M. Lee, T. Lee, S. H. Pope, G. S. Riordan, and M. Whitlow, "Single-chain antigen-binding proteins," *Science* 242(4877), 423-426 (1988).
6. R. Glockshuber, T. Schmidt, and A. Pluckthun, "The disulfide bonds in antibody variable domains: effects on stability, folding in vitro, and functional expression in Escherichia coli," *Biochemistry* 31(5), 1270-1279 (1992).
7. K. Proba, A. Worn, A. Honegger, and A. Pluckthun, "Antibody scFv fragments without disulfide bonds made by molecular evolution," *J. Mol. Biol.* 275(2), 245-253 (1998).
8. E. Ohage and B. Steipe, "Intrabody construction and expression. I. The critical role of VL domain stability," *J. Mol. Biol.* 291(5), 1119-1128 (1999).
9. A. Skerra, "Engineered protein scaffolds for molecular recognition," *J. Mol. Recognit.* 13(4), 167-187 (2000).
10. P. A. Nygren and A. Skerra, "Binding proteins from alternative scaffolds," *J. Immunol. Methods* 290(1-2), 3-28 (2004).
11. G. Winter and C. Milstein, "Man-made antibodies," *Nature* 349(6307), 293-299 (1991).
12. G. P. Smith and V. A. Petrenko, "Phage Display," *Chem. Rev.* 97(2), 391-410 (1997).
13. D. Lipovsek and A. Pluckthun, "In-vitro protein evolution by ribosome display and mRNA display," *J Immunol Methods* 290(1-2), 51-67 (2004).
14. E. Bianchi, A. Folgori, A. Wallace, M. Nicotra, S. Acali, A. Phalipon, G. Barbato, R. Bazzo, R. Cortese, F. Felici, and et al., "A conformationally homogeneous combinatorial peptide library," *J. Mol. Biol.* 247(2), 154-160 (1995).
15. A. Klug, "Zinc finger peptides for the regulation of gene expression," *J. Mol. Biol.* 293(2), 215-218 (1999).
16. T. L. Gururaja, S. Narasimhamurthy, D. G. Payan, and D. C. Anderson, "A novel artificial loop scaffold for the noncovalent constraint of peptides," *Chem. Biol.* 7(7), 515-527 (2000).
17. K. M. Marks, M. Rosinov, and G. P. Nolan, "In vivo targeting of organic calcium sensors via genetically selected peptides," *Chem. Biol.* 11(3), 347-356 (2004).
18. A. G. Cochran, N. J. Skelton, and M. A. Starovasnik, "Tryptophan zippers: stable, monomeric beta -hairpins," *Proc. Natl. Acad. Sci. U.S.A.* 98(10), 5578-5583 (2001).

19. S. M. Butterfield and M. L. Waters, "A designed beta-hairpin peptide for molecular recognition of ATP in water," *J. Am. Chem. Soc.* 125(32), 9580-9581 (2003).
20. S. M. Butterfield, C. M. Goodman, V. M. Rotello, and M. L. Waters, "A peptide flavoprotein mimic: flavin recognition and redox potential modulation in water by a designed beta hairpin," *Angew. Chem., Int. Ed.* 43(6), 724-727 (2004).
21. S. M. Butterfield, W. J. Cooper, and M. L. Waters, "Minimalist protein design: a beta-hairpin peptide that binds ssDNA," *J. Am. Chem. Soc.* 127(1), 24-25 (2005).
22. S. J. Russell, T. Blandl, N. J. Skelton, and A. G. Cochran, "Stability of cyclic beta-hairpins: asymmetric contributions from side chains of a hydrogen-bonded cross-strand residue pair," *J. Am. Chem. Soc.* 125(2), 388-395 (2003).
23. N. J. Skelton, T. Blandl, S. J. Russell, M. A. Starovasnik, and A. G. Cochran, "beta-hairpin polypeptides by design and selection," *Spectroscopy* 17(2-3), 213-230 (2003).
24. R. M. Fesinmeyer, F. M. Hudson, and N. H. Andersen, "Enhanced hairpin stability through loop design: the case of the protein G B1 domain hairpin," *J. Am. Chem. Soc.* 126(23), 7238-7243 (2004).
25. A. G. Cochran, R. T. Tong, M. A. Starovasnik, E. J. Park, R. S. McDowell, J. E. Theaker, and N. J. Skelton, "A minimal peptide scaffold for beta-turn display: optimizing a strand position in disulfide-cyclized beta-hairpins," *J. Am. Chem. Soc.* 123(4), 625-632 (2001).
26. B. Ciani, M. Jourdan, and M. S. Searle, "Stabilization of beta-hairpin peptides by salt bridges: role of preorganization in the energetic contribution of weak interactions," *J. Am. Chem. Soc.* 125(30), 9038-9047 (2003).
27. M. T. Pastor, M. Lopez de la Paz, E. Lacroix, L. Serrano, and E. Perez-Paya, "Combinatorial approaches: a new tool to search for highly structured beta-hairpin peptides," *Proc. Natl. Acad. Sci. U.S.A.* 99(2), 614-619 (2002).
28. M. T. Pastor, P. Mora, A. Ferrer-Montiel, and E. Perez-Paya, "Design of bioactive and structurally well-defined peptides from conformationally restricted libraries," *Biopolymers* 76(4), 357-365 (2004).
29. F. Huang, R. R. Hudgins, and W. M. Nau, "Primary and secondary structure dependence of peptide flexibility assessed by fluorescence-based measurement of end-to-end collision rates," *J. Am. Chem. Soc.* 126(50), 16665-16675 (2004).
30. Z. Cheng and R. E. Campbell, "Assessing the structural stability of designed beta-hairpin peptides in the cytoplasm of live cells," *ChemBioChem* 7(8), 1147-1150 (2006).
31. Z. Cheng, M. Miskolzie, and R. E. Campbell, "In vivo screening for elongated and highly folded β-hairpin peptides," *ChemBioChem* submitted (2006).
32. E. Hochuli, W. Bannwarth, H. Dobeli, R. Gentz, and D. Stuber, "Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent," *Bio-Technology* 6(11), 1321-1325 (1988).
33. I. B. Grishina and R. W. Woody, "Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators," *Faraday Discuss.* (99), 245-262 (1994).
34. N. V. Visser, M. A. Hink, J. W. Borst, G. N. van der Krogt, and A. J. Visser, "Circular dichroism spectroscopy of fluorescent proteins," *FEBS Lett.* 521(1-3), 31-35 (2002).
35. M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington, "Crystal structure of the Aequorea victoria green fluorescent protein," *Science* 273(5280), 1392-1395 (1996).
36. F. Yang, L. G. Moss, and G. N. Phillips, Jr., "The molecular structure of green fluorescent protein," *Nat. Biotechnol.* 14(10), 1246-1251 (1996).