A novel artificial loop scaffold for the noncovalent constraint of peptides

Tarikere L Gururaja[1], Shanaiah Narasimhamurthy[2], Donald G Payan[1] and DC Anderson[1]

Background: Few examples exist of peptides of < 35 residues that form a stable tertiary structure without disulfide bonds. A method for stabilization and noncovalent constraint of relatively short peptides may allow the construction and use of intracellular peptide libraries containing protein minidomains.

Results: We have examined a novel method for the noncovalent constraint of peptides by attaching the peptide EFLIVKS (single-letter amino acid code), which forms dimers, to the amino and carboxyl termini of different peptide inserts. An 18 residue random coil taken from the inhibitor loop of barley chymotrypsin inhibitor 2 was inserted between the peptides to produce a 32-mer minidomain that is attacked only slowly by elastase, has numerous slowly exchanging protons, contains a high β-structure content and has a T_m above 37°C. A point mutation disrupting the hydrophobic interior in both dimerizing peptides causes a loss of all slowly exchanging protons and of secondary structure. Adding specific charged residues to each terminus substantially increased the T_m, as did point mutants designed to add interdimerizer ion pairs. Three flexible epitope tag inserts and a nonamer insert do not appear to be folded in a stable structure by EFLIVKS. The properties of two peptides selected for expression in HeLa cells suggest they do form a stable tertiary structure.

Conclusions: Attaching short dimerizing peptides to both the amino and carboxyl termini of several 18-mer peptides appears to create stable monomeric tertiary structures. Mutations in the dimerizers can either destabilize or significantly stabilize a standard 18-mer insert. Dimerizing peptides flanking random insert sequences could be used as a strategy to generate heterogeneous peptide libraries with both extended and folded members.

Addresses: [1]Rigel, Inc., 240 E. Grand Avenue, South San Francisco, CA 94080, USA. [2]Department of Oral Biology, SUNY Buffalo, Buffalo, NY 14214, USA.

Correspondence: DC Anderson
E-mail: dcanderson@rigel.com

Key words: peptide dimerizer, protein folding, protein minidomain

Received: **10 March 2000**
Accepted: **27 April 2000**

Published: **26 June 2000**

Chemistry & Biology 2000, **7**:515–527

1074-5521/00/$ – see front matter

Introduction

Cyclic or otherwise constrained peptides have many valuable features compared with their linear analogs, including enhanced stability to proteolysis and a restricted conformation space that can result in a higher binding affinity for cognate binding proteins, because of a reduced entropic cost of binding. These constrained peptides contained in minimized proteins can be used as an intermediate step in the design of agents blocking protein–protein interactions and as the basis for the subsequent design of small molecules, which may be useful as drugs [1,2]. When the peptides are expressed in cells, they may modulate signaling pathways [3] and if the peptides are expressed in live mammalian cells, using retroviral vectors, they may be screened for defined changes in cellular phenotype. The resulting functional peptides may provide a route for the affinity isolation of their binding partners.

Unlike peptides in phage-display libraries, intracellular peptides are subject to proteolysis and thus library members should be constructed to be relatively inert to cellular proteases. Although intracellular peptide catabolism has not been well characterized, the ubiquitin–proteasome system is known to be involved in the degradation of short and long half-life proteins [4,5]. Further proteolysis, involving aminopeptidases, can result in degradation of peptides to amino acids [6]. In antigen presenting cells, short linear peptides resulting from cytoplasmic proteolysis can be removed to the endoplasmic reticulum by the peptide transporters TAP1 and TAP2 [7].

Owing to the presence of endogenous reduced glutathione in the 2–10 mM range [8,9] and the presence of thioredoxin reductase [10], standard methods of constraining peptides such as disulfide bonds are unlikely to work in the cytoplasm. Larger protein scaffolds such as thioredoxin [11] or green fluorescent protein (GFP; B. Peelle, personal communication) [12] have been used for the intracellular display of peptides, and intein-based chemistry might also be useful for such display [13]. A scaffold for the display of expressed peptides that is relatively inert to proteolysis and small enough to allow access to protein

binding sites (such as active-site crevices) may be useful for modifying the function of these proteins. If the scaffold is also flexible enough to allow the overall structure of the library member to be determined to a significant degree by the random insert sequence, a greater variety of structures could be generated.

To examine methods for the noncovalent constraint of peptides, and thus for creating a minidomain, we have added a 7-mer dimerizing peptide, EFLIVKS (single-letter amino acid code), to both the amino and carboxyl termini of a variety of test peptides. This peptide is the reversed sequence of the core dimerizing peptide of the neuropeptide head activator [14], is an equally potent dimerizer in solution with an apparent K_d at pH 7.5 of 2 μM, and appears to form dimers with β structure (T.L.G., unpublished observations). The formation of peptide dimers between the termini of an intermediate peptide sequence might be expected to constrain this sequence because of the high local concentration of the dimerizer peptides tethered by the intervening sequence. We also have examined several different insert sequences, including an 18-mer sequence taken from the protease contact loop of barley chymotrypsin inhibitor 2 (Ci2; [15]). This 18-mer analog has been reported to act as a 390 pM inhibitor of elastase [16] when constrained with a disulfide bond. As the peptide constructs examined here are not sufficiently soluble to obtain a nuclear magnetic resonance (NMR) structure, we have used proteolytic susceptibility, deuterium exchange kinetics observed by mass spectrometry (MS), and circular dichroism (CD) to examine their overall structure. These 32-mer minidomains are in the size range of the smallest reported peptides that form stable folded structures [17–21], and may have a unique structural motif relative to the simple secondary structures of other small protein domains.

Here, we show that several of these designed constructs with the Ci2 insert appear to adopt a stable tertiary structure, and two have T_m values above 50°C. We also show that two EFLIVKS-constrained peptides, present after three rounds of screening for taxol resistance in HeLa cells (D.G.P., unpublished observations), have a defined secondary structure and slowly exchanging protons, and one has a T_m of 60°C. The results suggest one mechanism for obtaining stable peptides expressed within live mammalian cells.

Results

Proteolysis of EFLIVKS-constrained chymotrypsin inhibitor 2 loop peptides and analogs

To examine the elastolytic stability of dimerizer-constrained peptides and two controls, we chose an 18-mer sequence derived from the inhibitory loop of Ci2. Both the native inhibitor and a disulfide-constrained engineered loop peptide have been reported to block elastase with a 390 pM inhibition constant [16]. To test the elastase inhibition by this disulfide-constrained loop peptide and potential inhibition by different peptide dimerizer-constrained versions of this inhibitor loop (peptides **3** and **4**–**6**, see Table 1), we incubated the peptides with elastase in the presence of a chromophoric substrate (Table 2). In each case, we observed either weak or no inhibition of the enzyme by the peptide constructs, including the disulfide-cyclized peptide **2**, which gave only 66% inhibition at a 10 μM concentration.

To see if the peptide dimerizers fused to the termini of the Ci2 insert peptide **3a** protected the peptide from elastase, we incubated peptides **3**–**6** with 100 nM elastase for 2–3 hours at pH 7.8, 25°C. Each reaction mixture was then separated and analyzed using high-performance liquid chromatography (HPLC)–MS as shown in Figure 1. As a

Table 1

Peptide sequences studied.

Peptide	Sequence	Peptide	Sequence
1	CGTIVTMEYRIDRTRSFC	**12**	MGEFLIVKSG$_4$DYKDDDDKG$_4$EFLIVKSGPPP
2	cyclic CGTIVTMEYRIDRTRSFC	**13**	MGEFLIVKSG$_4$YPYDVPDYASLG$_3$EFLIVKSGP
3	EFLIVKSVGTIVTMEYRIDRTRSFVEFLIVKS	**14**	MGEFLIVKSGDFNHFGNYLLDRRFFIAFGEFLIVKLGPP
3a	VGTIVTMEYRIDRTRSFV	**15**	MGEFLIVKSGHSSGIPVGVGWCWNSAGGGEFLIVKSGPP
4	EFLIVKSVGTIVTMEYRIDRTRSFVSKVILFE	**16**	MGEFLIVKSVGTIVTMEYRIDRTRSFVEFLIVKSGPP
5	SKVILFEVGTIVTMEYRIDRTRSFVEFLIVKS	**17**	MGEFLIVKSG$_3$ERPQEWAMEGPRDGLG$_3$EFLIVKSGPP
6	SKVILFEVGTIVTMEYRIDRTRSFVSKVILFE	**18**	KFLIVKSVGTIVTMEYRIDRTRSFVEFLIVES
7	EEFLIVKKSVGTIVTMEYRIDRTRSFVEEFLIVKKS	**19**	EFLIVESVGTIVTMEYRIDRTRSFVEFLIVES
8	K$_6$G$_4$EFLIVKSVGTIVTMEYRIDRTRSFVEFLIVKS	**20**	EKLKVKSVGTIVTMEYRIDRTRSFVEKLKVKS
9	EFLKVKSVGTIVTMEYRIDRTRSFVEFLKVKS	**21**	ESLSVKSVGTIVTMEYRIDRTRSFVESLSVKS
10	KFLIVKSVGTIVTMEYRIDRTRSFVKFLIVKS	**22**	K$_3$GSGSEFLIVKSVGTIVTMEYRIDRTRSFVEFLIVKSGSGSK$_3$
11	EFLIVKSSTKSIPPQSEFLIVKS	**23**	K$_6$GSGSEFLIVKSVGTIVTMEYRIDRTRSFVEFLIVKSGSGSK$_6$

All peptides are carboxy-terminal acids.

Table 2

Peptide inhibition of porcine pancreatic elastase.

Assay	% inhibition	n
50 nM elastase	0	
+ 500 nM peptide 2	6.2 ± 1.2	5
+ 10 μM peptide 2	66.4 ± 16.3	6
+ 10 μM peptide 3	−4.3 ± 0.63	3
+ 10 μM peptide 4	7.0 ± 1.0	3
+ 10 μM peptide 5	2.4 ± 0.29	3
+ 10 μM peptide 6	−1.4 ± 0.14	3
+ 2 μM PMSF	100	

control, the peptide CGTIVTMEYRIDRTRSFC was examined either in its unconstrained dithiothreitol-reduced form (**1**) or in its oxidized (cyclic) form (**2**). Peptide **1** (Figure 1a) was highly susceptible to elastolysis, giving around 11 different identifiable peptides (Table S1, Supplementary material). The main cleavage was after Tyr9 and additional cleavages were after Ile4, Val5, Thr6, Met7, Thr14 and Phe17. The cyclic analog **2** (Figure 1b) was cleaved more slowly than its linear analog, and after 3 hours was cleaved mainly after Tyr9 and also after Met7, Thr14 and Phe17. The proteolysis of peptide **2** by elastase provides clear evidence that it is a substrate, and the partial inhibition of elastase at the levels used in the LC/MS experiment will underestimate the proteolytic susceptibility of this peptide compared to other peptides which do not inhibit elastase. We have not done further experiments to determine the cause of the weak elastase inhibition of peptide **2**, but the observed inhibition could be due to product inhibition at higher levels of substrate.

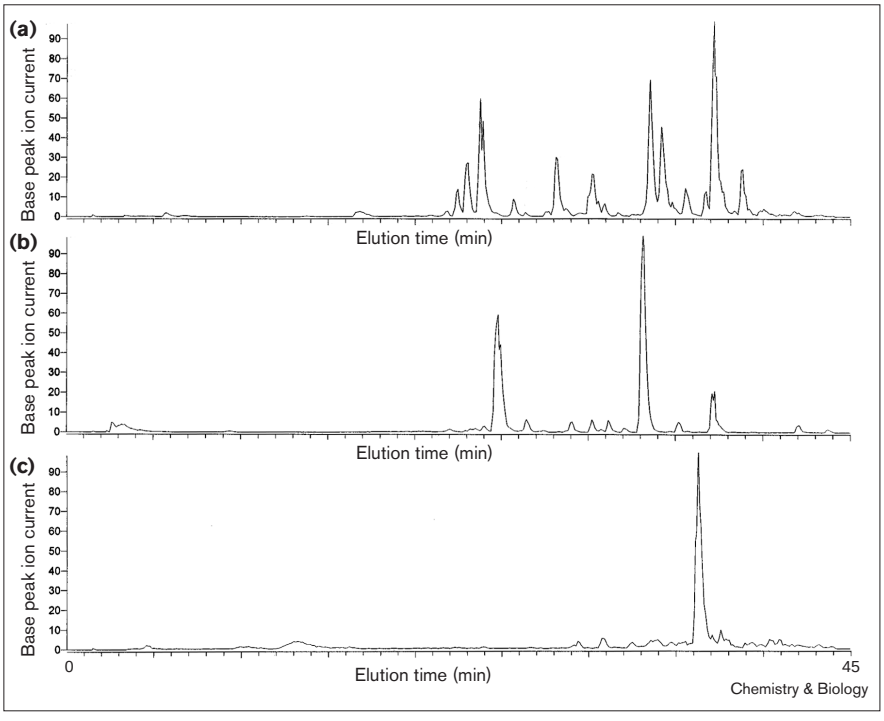
The Ci2 loop 18-mer with EFLIVKS attached to each end (**3**, Figure 1c) was not attacked to a significant degree by elastase after three hours, with only a small amount of proteolysis after Tyr16 initially observed. Most small peaks were synthetic impurities present in the absence of added elastase. After 24 hours, enough proteolysis occurred to assign additional elastolytic sites in this construct after Val5, Met14, Thr21, Phe27 and Ile29. Thus, cleavage occurred at many of the same residues as in the linear and cyclic peptides, but at a much reduced rate. Peptides **4**–**6** were also not significantly attacked by elastase after 2.5, 3.2 and 2.7 hours, respectively (data not shown).

The addition of peptide dimerizers to the ends of peptide **3a** thus appears to constrain the Ci2 peptide in a fashion that confers resistance to elastase.

Deuterium exchange and gel filtration experiments using constrained loop peptides

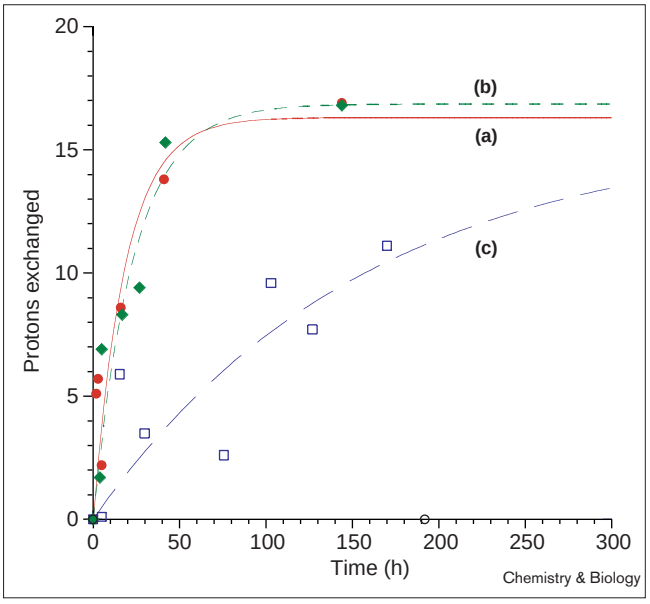
Relative to surface-exposed residues, the amide backbone protons of peptides and proteins will exchange more slowly with deuterated water when they are buried in the interior of a protein (and are inaccessible to water) or are involved in stable hydrogen bonding (see [22,23] and references

Figure 1



Elastolysis of linear, cyclic and dimerizer-constrained peptides, derived from barley Ci2, after reverse-phase high-performance liquid chromatography separation and detection using mass spectrometry. Base peak chromatograms of different test peptides (5–10 μM) were recorded after incubation with 100 nM elastase in pH 7.8 0.15 M Tris buffer at 25°C. **(a)** Peptide **1**, a linear version of the Ci2 insert, gives numerous fragments after a 2.9 h incubation with elastase. **(b)** The disulfide-cyclized peptide **2** also gives several major fragments after a 2.1 h incubation; the reaction was stopped with phenylmethanesulfonyl fluoride (PMSF) and dithiolthreitol (DTT) was added before chromatography, to reduce any disulfide bonds. **(c)** Little proteolysis is observed for peptide **3** after a 3 h incubation. Peptides **4**–**6** gave similar results to **3**.

Figure 2



Kinetics of deuterium exchange for dimerizer-constrained peptides. The fit to a first-order equation of the time-dependence of mass increase for dimerizer-constrained peptides, after a tenfold dilution in D_2O at pH 5, is plotted after subtraction of rapidly exchanging protons from the net mass gain. The mass increase was measured by mass spectrometry. **(a)** Exchange kinetics for peptide **3**, giving an exchange rate constant of 0.054 h^{-1}. A total of 16 protons exchanged in this phase, and 21 more at a rate too slow to measure. **(b)** Similar exchange kinetics were seen for peptide **6**, giving a rate constant of 0.041 h^{-1} and an amplitude of 17 protons; 18 protons exchanged at a significantly slower rate. **(c)** Exchange of protons for deuterium in peptide **15**, with a rate constant of 0.0044 h^{-1}; 21 protons exchanged in this phase, and there were no additional protons exchanging at a slower rate.

therein). Mass spectrometry has been used to examine the hydrogen exchange properties of a variety of different proteins (see [24,25] for recent examples) and the existence of slowly exchanging protons has been used to infer the existence of tertiary structure [19]. Deuterium exchange studies described here were carried out at pH 5 to allow ready observation of the peptide mass spectra, as low levels of the peptides were hard to detect at higher pH values. The pH was not reduced because some of the constrained loops do not appear to retain structure below pH 4.5, as measured by CD. To examine the compactness of the peptide dimerizer-constrained Ci2 loop peptide, the rate and stoichiometry of deuterium incorporation upon dilution into D_2O was obtained.

The kinetics of deuterium exchange for protons in several representative peptide constructs, peptides **3**, **6** and **14**, are shown in Figure 2. For these peptides, a total of 67 Da, 66 Da and 63 Da were added to the time zero mass upon complete proton exchange, respectively. In the fast phase (less than an hour), 29, 31 and 46 protons exchanged, respectively. The expected number of fast-exchanging sidechain and terminal protons in each case is 33, 33 and 27, respectively. For each peptide, a number of additional protons exchanged over a period of around 300 hours. Each curve was well described as a first-order process. The derived amplitudes of these more slowly exchanging protons were 16, 17 and 21 protons, with corresponding rate constants of 0.054, 0.041 and 0.0044 h^{-1}, respectively (Table 3). Exchangeable protons not accounted for in these two phases, inferred from the missing amplitude relative to the total number of exchangeable protons, were presumed to exchange even more slowly. These were 21, 18 and zero protons for the three peptides in Figure 2. Both classes of slowly exchanging protons exchange at a rate far slower than that measured for surface-exposed protons, taken from nearest neighbors identical to those found in EFLIVKS [26]. These control protons exchanged with deuterium with rate constants in the range of 6–60 h^{-1} at pH 5, giving an estimate of protection factors for the protons with an observable exchange rate studied here in the range of 1.2×10^2–1.2×10^3 for **3** and **6**, and 1.4×10^3–1.4×10^4 for **14**. The remaining unexchanged protons presumably have tenfold or larger protection factors.

With the Ci2 18-mer insert, similar results were obtained with a different dimerizer, EEFLIVKKS, attached to each end of the insert. For peptide **7**, 39 of the 70 protons exchanged with deuterium within an hour, eight exchanged with a rate constant of 0.15 h^{-1} and the remaining 23 protons exchanged more slowly. The total of 31 slowly exchanging protons in this analog was somewhat less than the 37 protons in **1**, suggesting some changes in structure between the two constructs. Thus, with the same Ci2 insert, several dimerizer variants produced mini-domains with slowly exchanging protons. Not all peptide dimerizers have this property, however. For the peptide **8**, with K_6–G_4 fused to the amino terminus of the parent peptide **3** to enhance solubility, all but ~5 protons exchanged within an hour. This amino-terminal fusion may thus destabilize the structure of this construct or at least increase its mobility.

Substitution of Ile4 in EFLIVKS previously has been shown to disrupt dimerization of this peptide (T.L.G., unpublished observations). We therefore tested a point mutation in each dimerizer (creating EFL*K*VKS) attached to the Ci2 18-mer insert (**9**, Table 1) and examined the effect of this mutation on the 18-mer insert structure by deuterium exchange. If this mutation disrupted the structure, the number of slowly exchanging protons would be diminished. When the exchange kinetics were examined, all but one proton exchanged within an hour. When the amino-terminal glutamate in EFLIVKS was mutated to lysine (creating *K*FLIVKS) and both mutant dimerizers were fused to the Ci2 insert in **10**, all slow-exchanging protons were also lost.

We also examined the effect of different insert sequences on the exchange kinetics of the overall peptide when EFLIVKS or its analogs were fused to both termini of these inserts. One insert, STKSIPPQS, represented an analog of the protease inhibitor cyclic CTKSIPPQC [27]. Peptide **11** (Table 3) had a total of 36 exchangeable protons, 33 of which exchanged in an hour. Thus, this peptide construct did not appear to have a stable tertiary structure. A second insert, the flag epitope tag DYKDDDDK, was flanked by four glycines on each end to increase its flexibility and to allow binding to anti-flag antibodies. This was fused to MGEFLIVKS at the amino terminus and EFLIVKSGPP at the carboxyl terminus. This construct **12** had 36 exchangeable protons, 35 of which exchanged within an hour. A similar construct **13** included the influenza hemagglutinin epitope tag YPYDVPDYASL flanked by four and three glycine residues at its amino and carboxyl termini, respectively. All protons were exchanged for deuterium within ~1 h. Thus, both a shorter insert sequence and inserts flanked by multiple glycines did not have slowly exchanging protons.

Two additional peptides, present after three rounds of peptide library screening for peptides that confer resistance to taxol (D.G.P., unpublished observations), were examined. Peptide **14**, with the insert GDFNHFGNYLLDRRFFIAFG and a slightly modified carboxy-terminal dimerizer EFLIVKLGPP, had 21 slowly exchanging protons with a rate constant tenfold slower than for the other constructs with slowly exchanging protons. Peptide **15**, with the insert GHSSGIPVGVGWCWNSAGGG, had around four measurable slowly exchanging protons and a further 17 protons exchanging at a rate too slow to measure in this experiment.

Saturated solutions of several peptides with slowly exchanging protons were examined by gel filtration at pH 7.3 for evidence of oligomerization. A standard curve was constructed from 23 different peptides and small proteins (see the Materials and methods section). Peptides **6**, **7** and **18** exhibited single main peaks with apparent molecular weights of 1600, 2540 and 2360 Da, respectively (data not shown). There were no peaks at higher integral multiples of these molecular weights. The expected molecular weights were 3778, 4293 and 3778 Da, respectively. Thus, these constructs appeared to elute as monomers as somewhat smaller peptides than their true molecular masses or than peptides used to construct the standard curve. Peptide **15**, however, exhibited a main peak with an apparent molecular mass of 2240 Da, and a significant leading shoulder with an apparent mass of 4400 Da (data not shown). It also elutes at an apparent molecular weight below its expected molecular mass of 3957 Da, but appears to form some dimer in a saturated solution at pH 7.3 as the main peak and shoulder are different in their apparent molecular masses by a factor of two.

Table 3

Deuterium exchange rates and amplitudes for different peptide constructs.

Peptide	Total exchangeable protons	Proton exchange amplitudes*	Measured exchange rate constant
3	66.5 ± 1.4	29.3 ± 1.5 fast 16 intermediate 21 slow	0.054 h^{-1}
6	66	31 fast 17 intermediate 18 slow	0.041 h^{-1}
7	70.3	39.4 fast 7.9 intermediate 23 slow	0.15 h^{-1}
8	87.7	82.8 fast	
9	71.6	70.1 fast	
10	47.1	47.7	
11	36.1 ± 1.9	32.6 ± 3.8 fast	
12	35.9	34.8 fast	
13	40.3	40.3 fast	
14	62.7	45.7 ± 0.2 fast 20.6 intermediate	0.0044 h^{-1}
15	58.6	37.2 ± 1.0 fast 4.2 intermediate 17.2 slow	

*The fast amplitude is calculated for protons exchanging within ~1 h.

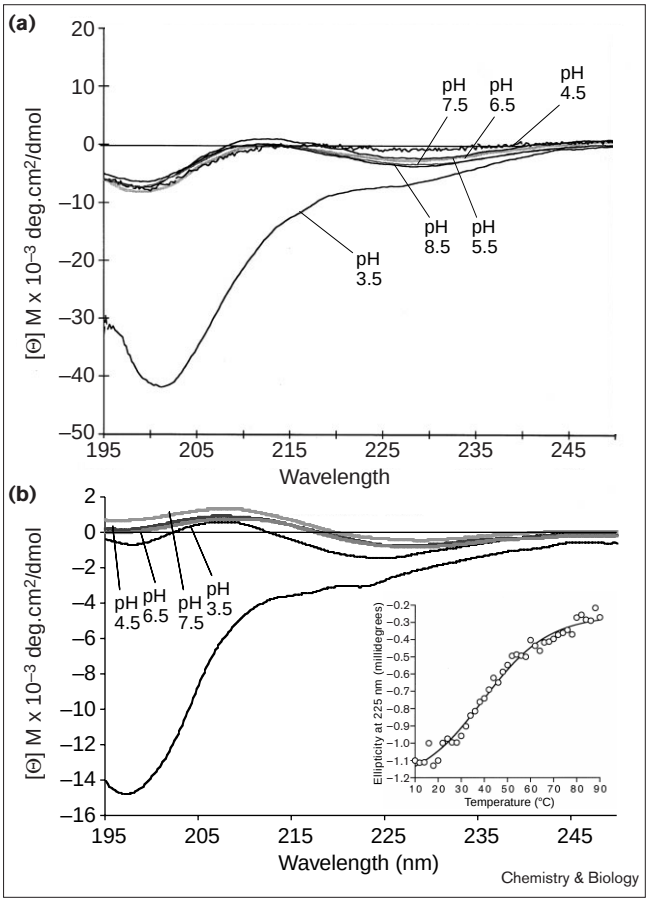
Circular dichroism studies of peptide constructs

Because the peptides studied here are not adequately soluble for NMR structure determination, we examined their solution structure using CD. These measurements are sensitive to the secondary structure of both peptides and proteins, and have been extensively used for conformational analysis [28–30]. Here, these measurements are used to examine secondary structure formation, pH dependence and stability, to compare the effects of different dimerizers on insert structure, to examine the structural effects of mutations in the dimerizers, and to look at the effects of different insert sequences on the overall structure of dimerizer-constrained loops. When these measurements are combined with measurements of proteolytic susceptibility, deuterium exchange and the results of conformational searches, they give information on the overall structure and folding of the minidomains examined here.

Peptide 11

The first insert examined was a nonamer, to test the effects of EFLIVKS constraints on a relatively short peptide. The CD spectrum was recorded between pH 3.5 and 8.5 (Figure 3a). A pH-dependent secondary structure was observed. At pH values of 4.5 and above the CD spectrum suggests the existence of some β structure, and the T_m at 225 nm is 39.6°C (data not shown). At pH 3.5, a strong minimum at 201 nm was seen, near the expected

Figure 3



CD spectra of EFLIVKS-constrained peptides. **(a)** pH-dependence of the spectrum of peptide **11**. Spectra were collected at pH 3.5–8.5, as labeled. The pH 3.5 spectrum shows a significant increase in random coil. The spectra are consistent with a low-pH transition to a structure that is stable at pH values at and above 4.5. **(b)** pH dependence of peptide **3**. The spectra were acquired at pH 3.5–7.5, as labeled. The spectrum of the 18-mer insert peptide **3a**, which is a random coil at pH 7.5, is shown at the bottom. The spectra of **3** are consistent with a low-pH transition to a structure that is more stable at pH values at and above 4.5. Inset: thermal melting curve of **3**, giving a T_m of 39.9°C.

minimum for a random coil [31] of 195–197 nm. This curve is also similar in shape to that of a type 1 β turn observed in a short peptide [32]. The protamine-like protein from Spisula also has a similar spectrum but has been shown to have tertiary structure [33]. As this CD spectrum was seen with other peptide constructs that exhibit a transition from this spectrum at lower pH values to other spectra stable at higher pH values, as its assignment to a secondary structure appears ambiguous, and as **11** was soluble at ~1 mM at pH 4, the structure of this peptide was examined using NMR (details are in the Supplementary material section).

The temperature coefficients of all amide resonances were found to be ≥0.004 ppm K^{-1} (data not shown), suggesting that the backbone NH groups are exposed to the solvent and are not involved in any intramolecular hydrogen-bonding interactions. The fast $^1H/^2H$ exchange rate observed for all backbone amide resonances provides further evidence that the amide groups are not involved in any intramolecular hydrogen bonding. The prevalence of strong $d_{\alpha N(i,\ i+1)}$ and weak $d_{\alpha N(i,\ i+1)}$ nuclear Overhauser effects (NOEs) and a continuous stretch of weak and medium $d_{\beta N(i,\ i+1)}$ and $d_{\alpha\beta(i,\ i+1)}$ NOEs in the absence of any observable d_{NN} NOE interactions indicate that the backbone dihedral angles are predominantly in the unfolded region of ϕ, φ space [34,35]. The $J_{NH-C\alpha H}$ values are in the range of 6.5–8.4 Hz for all residues except Ser7. This suggests the existence of populations of unfolded nonhydrogen bonded conformations of comparable energy with ϕ values exceeding the regular helical region [36]. Collectively, the NMR data suggest that peptide **11** is unstructured in aqueous solution at low pH.

In contrast to the results at pH 3.5, at pH 4.5–8.5 a different secondary structure was observed by CD. These CD spectra had a much diminished band at 202 nm, indicating a loss of random coil. They also had a slight positive band at around 210–215 nm, and a negative band around 228–230 nm, indicating the presence of β turns [37]. Because at pH 5.0 this construct has three or fewer slowly exchanging protons (Table 3), the peptide may be unfolded or in a molten globule state (i.e. it has no stable tertiary structure) but may have secondary structure containing some β turns and significantly less random coil than at pH 3.5. When observing the CD spectrum at 225 nm, the structure present at pH 7.5 has a T_m of 39.6 ± 1°C (data not shown). The wide transition range of ~25°C is consistent with the lack of a simple two-state transition and perhaps the presence of a molten globule [38].

Peptides 3 and 3a

The pH dependence of the CD spectrum of peptide **3** is shown in Figure 3b. At pH 3.5, the CD spectrum has a prominent trough around 200 nm indicative, in view of the above NMR results, of the presence of random coil. Also, it has a strong maximum around 210 nm and strong minimum at 225–230 nm, which is consistent with significant beta-turn content [37]. The spectra at pH 4.5–6.5 are different from that at pH 3.5 but similar to each other, with the trough around 200 nm diminished and the peak near 210 nm enhanced. At these pH values, there thus appears to be less random coil. At pH 7.5, the 200 nm minimum decreases further and the peak at ~210 nm is enhanced, suggesting a further change in structure from pH 4.5–6.5. Thus this peptide shows more structural plasticity than peptide **11**, but both have more random coil at a low pH than at higher pH values. Peptide **3** has a T_m of 39.9 ± 1.6°C when observed at 225 nm (Figure 3b, inset), and appears to melt over a significantly broader range (80°C) than peptide **11** (~40°C). Peptide **3**, unlike peptide

11, has around 37 slowly exchanging protons. The large number of slowly exchanging protons combined with the observation of significant secondary structure suggests that this construct has some stable tertiary structure. The very broad melting range and low T_m suggests that the low-temperature structure could have features of a molten globule. However, these data do not seem to fit the classic definition of a molten globule, which would not have predicted the slowly exchanging protons.

The CD spectrum of peptide **3a**, VGTIVTMEYRIDRTRSFV, was run at pH 7.5 as a control (Figure 3b). The CD spectrum shows a prominent negative band around 197 nm, consistent with a random-coil structure.

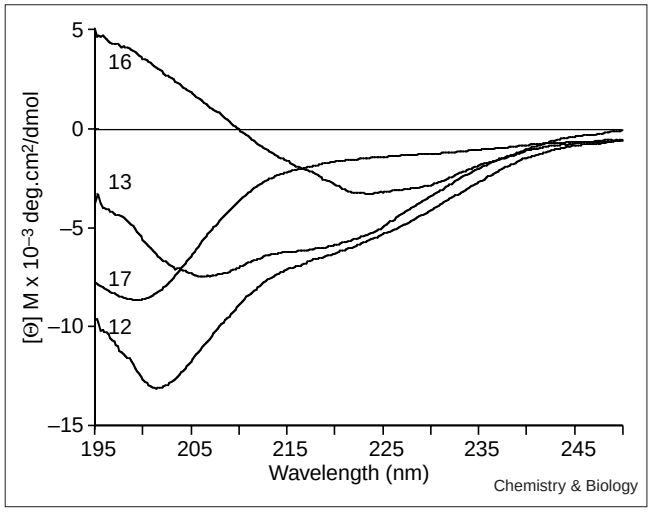
EFLIVKS constructs with an amino-terminal MG and carboxy-terminal GPP

For peptide expression in live cells, MG was added to the amino terminus and GPP to the carboxyl terminus to potentially block proteolysis by cellular carboxypeptidases [39]. The effect of these sequence additions on peptide **3** (now peptide **16**) at pH 7.5 is shown in Figure 4. The CD spectrum of **16** is similar to that of **3**, except that there is no minimum around 200 nm and only small changes are seen in the spectrum between pH 3.5 and 8.5 (data not shown). Like **3**, peptide **16** has a very broad transition range of over 80°C and a similar T_m of 38.4°C when observed at 225 nm. The amino-terminal addition of MG and addition of GPP to **3** appears to stabilize the low-pH structure and to change the higher-pH structures by eliminating random coil, without significantly changing the overall stability as measured by the T_m.

The CD spectra of a variety of other EFLIVKS-constrained peptide inserts, all with an amino-terminal MG and carboxy-terminal GPP, were also compared at pH 7.5 (Figure 4). These additional inserts were tested to compare the ability of EFLIVKS to help fold peptides different from the Ci2 insert. Peptide **13** contains an insert consisting of the influenza hemagglutinin epitope tag with glycine spacers, G_4YPYDVPDYASLG_3. The characteristic double minima in the CD spectrum at 205–207 nm and 220–225 nm indicate the presence of α helix [31]. This construct did not have slowly exchanging protons (Table 3); thus, the CD spectrum may reflect the presence of only secondary structure or mobile tertiary structure.

Peptide **12** contains the flag epitope tag with glycine spacers, G_4DYKDDDDKG_4, as an insert. The resulting CD spectrum contains a single minimum at 202 nm and another small minimum at ~220 nm (Figure 4). On the basis of the similarity of this spectrum to that of peptide **3** at low pH, this peptide appears to contain a significant percentage of random coil. The T_m at 200 nm is 27.9°C (data not shown). As peptide **12** does not have slow exchanging protons (Table 3), it does not appear to have a

Figure 4



CD spectra of dimerizer-constrained peptides with different insert sequences in the scaffold MGEFLIVKS–insert–EFLIVKSGPP. All spectra were acquired at pH 7.5. The peptides and their inserts are **16**, Ci2 18-mer, **13**, influenza hemagglutinin epitope tag, **12**, flag epitope tag, and **17**, IκB epitope tag. The spectra depend on the sequence of the insert. Peptides **12** and **17**, which have multiglycine spacers on each side of the epitope, have a significant random coil content. Peptide **13** has a significant α-helical content.

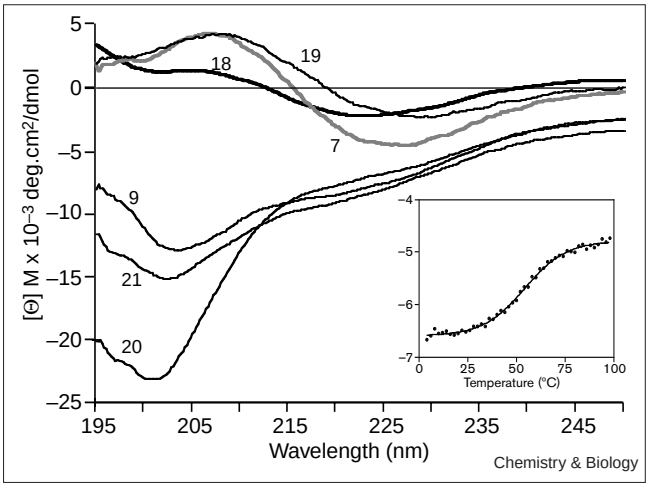
stable tertiary structure. Peptide **17** contains an insert including the IκB epitope G_3ERPQEWAMEGPRDGLG_3 with flanking glycine spacers. It gave a CD spectrum with a prominent band at 202 nm (Figure 4), which was observed at pH 3.5–8.5, and a T_m measured at 200 nm of 35.7°C (data not shown). This construct thus appears to be a random coil at all pH values examined.

Charge modifications of the peptide dimerizer

The effects of mutations in, or short additions to, the EFLIVKS sequence on the CD spectrum of the Ci2 peptide insert are shown in Figure 5. Peptide **7** has additional charges in both dimerizer regions, giving the sequence EEFLIVKKS. The 202 nm minimum, attributed to random coil, is diminished relative to peptide **3**, and the construct, with a T_m measured at 225 nm of 54.2°C (inset), is significantly more stable than peptide **3**. The spectrum shown, obtained at pH 7.5, is similar to spectra obtained from pH 3.5–8.5 (not shown), and the 220–225 nm minimum is indicative of β structure as with peptide **3**. As peptide **7** has 31 slow-exchanging protons (Table 3), it appears to have a stable tertiary structure.

In peptide **18**, single lysine and glutamate residues were switched between dimerizers, giving KFLIVKS–Ci2–insert–EFLIVES. This change was made to encourage additional interdimerizer ion pairing. The CD spectrum at pH 7.5 (Figure 5) is quite similar to that of peptide **7**, with no major negative peak around 202 nm at

Figure 5



CD spectra of peptides with mutations in, or short additions to, the dimerizer sequence at pH 7.5. All peptides have the Ci2 18-mer insert sequence between the two dimerizer sequences. Peptides **7**, **18** and **19** all lack significant random-coil content and contain β structure. Both **7** and **18** are significantly more stable than **3**, with T_m values of 54.2°C (inset) and 56.9°C, respectively. The mutations in **9**, **20** and **21**, which disrupt the hydrophobic core of each dimerizer, all appear to disrupt the structure of the overall construct, causing a significant increase in random-coil content, as shown by the significant minimum around 202 nm.

any pH (not shown), and a T_m of 56.9°C (not shown). Peptide **18** thus appears to be even more stable than peptide **7**. Both **7** and **18** have no low-pH random coil structure (not shown). In peptide **19**, the lysine residue of each dimerizer was mutated to glutamate, giving EFLIVES-Ci2 insert-EFLIVES. The CD spectrum of this peptide at pH 7.5 is similar to the CD spectra of **3**, **7** and **18**, missing the minimum around 202 nm, indicative of random coil content. The overall structure of the Ci2 construct thus appears to be maintained with several charge substitutions or additions, and the changes in **7** and **18** significantly stabilize the 32-mer construct relative to the original peptide **3**.

Modifications to the hydrophobic core of EFLIVKS

In a third set of dimerizer modifications, the hydrophobic character of the dimerizer was changed. In peptide **20** both Phe2 and Ile4 in both EFLIVKS sequences were mutated to lysine, giving a dimerizer sequence at each terminus of the Ci2 insert of E*K*L*K*VKS. The CD spectrum of this peptide at pH 7.5 (Figure 5) indicates a significant increase in random-coil content relative to **3**. In peptide **21**, both Phe2 and Ile4 were mutated to serine instead of lysine. At pH 7.5, this peptide also had a significant increase in random-coil content. In peptide **9**, only Ile4 was mutated to lysine in each dimerizer, giving EFL*K*VKS as the dimerizing sequence on each side of the Ci2 insert. The pH 7.5 CD spectrum of **9**, shown in Figure 5, also has a significant minimum at 202 nm indicating a significant loss of structure. This peptide had at most one or two slowly exchanging protons (Table 3), suggesting that this single change in the hydrophobic core of EFLIVKS is sufficient to disrupt the structure of the entire peptide construct.

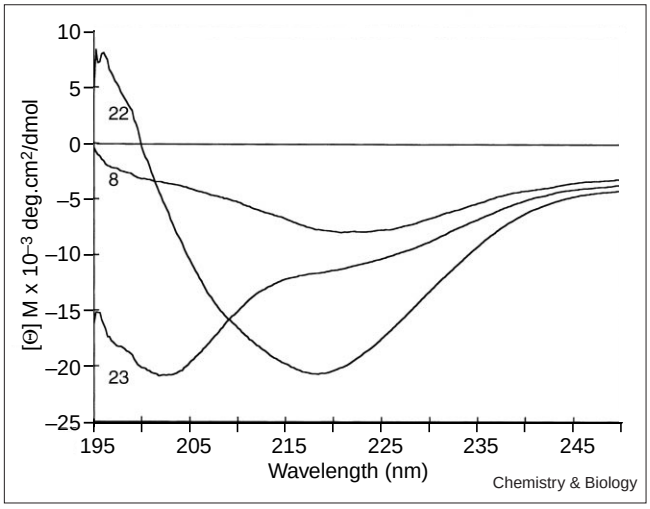
Fusion of charged sequences to one or both dimerizers of peptide 3

To examine the importance for the overall structure of an intact amino and carboxyl terminus of peptide **3**, and to potentially improve the solubility of the construct, we fused different sequences to one or both termini of this peptide. Peptide **8**, resulting from fusion of K_6G_4 to the amino terminus of **3**, had a pH 7.5 CD spectrum shown in Figure 6 that is very different than that for peptide **3**, with a broad minimum at 220–225 nm. This spectrum does not appear to be characteristic of any one dominant secondary structure, and can be deconvoluted to a mixture of β sheet and β turn (58%), α helix (14%) and 28% random coil. As this structure has at most five slowly exchanging protons (Table 3), the additional residues added to the amino terminus appear to have destabilized the tertiary structure of the control peptide, while creating a different secondary structure. In peptide **22**, fusion of K_3GSGS to the amino terminus and GSGSK_3 to the carboxyl terminus of **3** resulted in the appearance of a broad minimum in the pH 7.5 CD spectrum at 215–220 nm, which does not appear to be characteristic of a single secondary structure. In peptide **23**, K_6GSGS and GSGSK_6 were fused to the amino and carboxyl termini, respectively. The pH 7.5 CD spectrum (Figure 6) is dominated by a minimum at 202 nm which is characteristic of a random coil structure. This minimum dominates the spectrum at all pH values between pH 3.5–8.5 (data not shown). Thus fusions of 7–10mer peptides containing multiple lysines to one or both termini of **3**, intended to help solubilize this structure, have instead caused either a major shift in, or loss of, secondary structure, or a loss of tertiary structure compared with peptide **3**.

Peptides expressed in mammalian cells

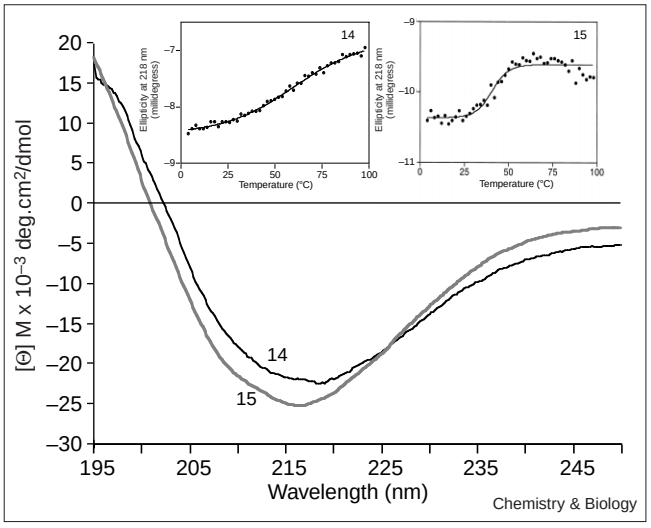
Two peptides present after three rounds of screening for resistance to taxol treatment in HeLa cells (D.G.P., unpublished observations), **14** and **15**, were also examined using CD. Both were poorly soluble, with saturated solutions in pH 7.5 phosphate-buffered saline (PBS) containing 59 and 4 μM, respectively, as measured by amino-acid analysis. The CD spectrum of peptide **15** (Figure 7) has a relatively pH-independent (not shown) broad minimum around 215 nm and a T_m of 40.6°C when measured at 218 nm (Figure 7, inset). The transition range for **15** was the narrowest observed for any EFLIVKS-constrained peptide, covering around 20°C. This peptide has a total of 21 slow exchanging protons and thus has a stable tertiary structure.

Figure 6



CD spectra of peptides resulting from the fusion of different sequences to one or both termini of peptide **3**, collected at pH 7.5. In peptide **8**, K_6G_4 was fused to the amino terminus of **3**; in peptide **22**, K_3GSGS and GSGSK_3 were fused to the amino and carboxyl termini, respectively; in peptide **23**, K_6GSGS and GSGSK_6 were fused to the amino and carboxyl termini, respectively. The fusions cause a loss of most slowly exchanging protons in **8**, the appearance of a rather different structure in **22** compared with **3**, and the appearance of random coil at all pH values tested in **23.**

Figure 7



CD spectra of peptides previously expressed in mammalian cells. The CD spectrum of peptides **14** and **15** are shown. Although these peptides have 18-mer inserts with different sequences, both spectra reflect the existence of β structure and are similar to each other. Peptide **15** has a T_m of 40.6°C when observed at 218 nm (inset). Unlike most other constructs, this peptide has a relatively narrow transition range, around 25–50°C. Peptide **14** has a T_m of 60.0°C (inset), a broad transition range, and is the most stable dimerizer-constrained peptide examined.

The spectrum of peptide **14** is nearly identical to that of **15** in spite of a very difference insert sequence (Table 1), with a pH-independent (not shown) broad minimum at 215–220 nm (Figure 7). Unlike **15**, it has a thermal transition range of around 90°C, and a T_m of 60.0°C at 218 nm (Figure 7, inset). This is the most thermostable EFLIVKS-based construct, with a T_m in the range of many proteins. This construct also had 21 protons with a slow exchange rate (Table 3) and thus a stable tertiary structure.

Discussion

Here, we have examined the secondary and tertiary structure of a variety of 32-mer peptides folded by self-binding peptides fused to their amino and carboxyl termini. As the peptides were generally soluble at concentrations less than 60 μM and at pH values where they were folded, a direct determination of structure using NMR methods would be difficult. We thus combined several different methodologies to examine the existence of tertiary structure in these peptides. An insert of roughly 18 residues may be needed for stability of the entire construct. When a shorter nonamer insert is used (peptide **11**), giving an overall length of 23 residues, a defined secondary structure is obtained, but without slowly exchanging protons. This peptide may thus be too short to form a stable tertiary structure with the scaffold used here. The smallest natural monomeric protein domains stable in the absence of disulfide bonds or the binding of stabilizing cofactors such as metal ions, are around 50–60 residues. In spite of this, smaller stable domains have been designed or reported, including a monomeric 23-mer peptide with a zinc-finger-like fold [17,40], a stable 33-mer two-helix analog of the Z-domain of protein A [18], a 35-mer monomeric stable, folded subdomain of the headpiece domain of villin [19], a stable ββα motif peptide [20], and a stable 36-mer three-helical peptide analog derived from dihydrolipoamide acetyltransferase [21]. The stable constructs reported here are thus in the size range of the smallest protein domains.

The self-dimerizing peptide EFLIVKS, when fused to the amino and carboxyl terminus of the Ci2 peptide **3a**, appears to fold the peptide into a stable structure at room temperature. It has 37 slowly exchanging protons and significant β-turn content, as measured by CD. The broad melting range of this peptide observed by CD, coupled with the observation of slow-exchanging protons, suggests that this peptide construct may have more than one slowly exchanging conformer. Although the Ci2 loop in the context of barley chymotrypsin inhibitor is a potent elastase inhibitor, the EFLIVKS-constrained Ci2 insert (peptide **3**) is not an elastase inhibitor and is quite inert to elastolysis, and the cyclic peptide **2** is not a 390 pM inhibitor as reported but is instead a substrate with weak

inhibition. Thus, the dimerizers appear to have created a unique set of structures compared with the loop structure present in barley chymotrypsin inhibitor and with the disulfide-cyclized version of the peptide.

One question is whether the Ci2 insert itself dominates the folding of its constructs with different peptide dimerizers. The CD spectrum of the insert itself, peptide **3a**, appears to contain mainly random-coil structure. Three constructs with variations in the peptide dimerizer, **3**, **6** and **7**, all have slowly exchanging protons and nonrandom-coil CD spectra. However, specific mutants in the dimerizer, Ile4→Lys and Ile29→Lys, as well as a K_6G_4–amino-terminal fusion, disrupt the structure of peptide **3**, driving it into a random coil. These data suggest that addition of the bare dimerizers to both termini plays a significant role in the folding of this insert.

Two modified dimerizers significantly stabilized the Ci2 insert relative to the EFLIVKS dimerizer. EEFLIVKKS added to both termini of **3a**, and a combination of KFLIVKS at the amino terminus and EFLIVES at the carboxyl terminus, increased the T_m of the Ci2 insert by 14.3 and 17.0°C relative to **3**, and both constructs retained their slowly exchanging protons. Given the sequence changes in these improved dimerizers, the increase in stability could be due to additional ion pairs between the two dimerizers.

In the context of an amino-terminal MG and carboxy-terminal GPP, the EFLIVKS dimerizer folds some but not all insert sequences into stable tertiary structures. Peptide **16** has a structure with similar stability to **3** and with less random coil, and peptides **15** and **14**, with a slightly modified second dimerizer, form stable tertiary structures. However, three epitope tag inserts flanked by multiple glycine residues, either had no slowly exchanging protons or had significant random-coil structure, as observed by CD. This could be due to the flexible nature of the insert sequence (containing eight and seven glycines, respectively) and the expected charge repulsion-driven extended structure of the flag sequence, which has five aspartates in a seven-residue stretch.

One possible concern about using peptides with high self-affinity is intermolecular aggregation. In theory, the high local concentration of dimerizing peptides, due to their tethering together by the insert sequence, should result in saturation of even moderate dimerizer self-binding constants forming 'intramolecular' dimers, and a monomeric constrained structure. This structure should predominate over hetero-oligomers when the intramolecular (local) concentration of dimerizers is greater than the solution concentration. No dimers or higher-order structures of peptide **3** were observed by MS (data not shown) or for peptides **6**, **7** and **18** by gel filtration in PBS.

To examine dimerizer-constrained peptides present after intracellular screening in mammalian cells, we synthesized and studied the structural features of peptides **14** and **15**. In **14**, the combination of 21 slowly exchanging protons, a T_m of 60°C with a broad melting transition, and a nearly pH-independent CD spectrum with little random coil, suggest that this peptide folds into a stable tertiary structure with multiple slowly exchanging, low-energy conformers. Human epidermal growth factor may use some residues from a four-residue surface aromatic cluster for receptor recognition [41], whereas a water-accessible cluster of five aromatic residues may be involved in ligand binding and receptor activation in the dopamine D2 receptor [42]. The 35-residue miniprotein HP35 has a cluster of three aromatic residues that pack together and may be important for folding [19,43]. Thus, the insert sequence in **14**, which has seven aromatic residues, may use these residues for the peptide's function, folding or stability.

For peptide **15**, the combination of 21 slowly exchanging protons, a T_m of 40.6°C with a narrow melting transition, and a nearly pH-independent CD spectrum with little random coil, suggests formation of a stable tertiary structure with relatively few slowly exchanging, low-energy conformers. Gel filtration experiments suggest the existence of some peptide dimer in a 59 μM solution at pH 7.3. The narrower melting curve may also reflect dimerization of this peptide. The insert in this peptide contains a repeat pattern of tryptophan and valine alternating with less hydrophobic residues, suggesting a run of β sheet that is consistent with the CD spectrum. As β sheets are known to aggregate, this short stretch may be involved in intermolecular dimerization of this construct. Neither peptide **14** nor **15** is very soluble. Beside the formation of a stable tertiary or quaternary structure, this limited solubility could be an additional mechanism for surviving intracellular screening and thus, presumably, for surviving cellular catabolism.

As different 18-mer inserts between the two terminal EFLIVKS peptides gave both folded and unfolded constructs, depending on the insert sequence, peptide libraries based on this scaffold will probably be a mixture of constrained and unconstrained peptides. A peptide library could contain stable folded structures such as peptides **14** or **15**, peptides with secondary, but not tertiary, structure such as **12** or **13**, or peptides with no stable structure, such as **17**. In view of the unfolding of peptide **3** by the fusion of additional amino-terminal or both amino- and carboxy-terminal sequences, libraries containing additional fusion sequences with this scaffold may be more biased to extended structures. The only short peptide insert tested, a nonamer, gave an unfolded peptide, perhaps because the overall length (23 residues) is too short to allow folding for most peptide sequences. Thus, to give more stable libraries, longer random inserts, such as 18-mers, appear to be preferable for intracellular screens.

Significance

Creating very small proteins without the disulfide bonds used in nature to stabilize small domains, such as conotoxins, is a challenge to protein engineers. A method to noncovalently constrain a small domain is important if the goal is to express a library of very small proteins in the cytosol of a cell for the purpose of intracellular screening. Current approaches to the creation of very small proteins include the careful construction of two or three contiguous segments of secondary structure or the inclusion of other constraints such as metal-ion binding sites. Here we have described a novel method to create a very small protein by attaching dimerizing peptides to the amino and carboxyl terminus of the intervening sequence. The overall structure is sensitive to changes in the dimerizers. By mutating individual residues to create potential ion pairs across the dimer interface, a construct with a T_m as high as 57ºC has been created. After several rounds of selection in a mammalian cell, two peptides using this scaffold appeared to have a stable structure, with T_m values of 41ºC and 60ºC. One has seven aromatic residues, suggesting that the random 18-mer insert can allow additional selection for a stable structure. The low solubility of these peptides, as well as their folding, may allow them to escape intracellular proteolysis. Although we have not determined the structure of these constructs, the use of terminal dimerizers appears to impart some unique structural elements to these minidomains.

Materials and methods

Synthesis and purification of peptides

Peptides were synthesized on a Symphony/ Multiplex peptide synthesizer (Protein Technologies Inc., Tucson, AZ, USA) using standard Fmoc (9-fluorenylmethyloxycarbonyl) chemistry [44], and purified by reversed-phase chromatography. Automated coupling methods and standard synthesis protocols were insufficient for some peptides, as outlined below. Since the sequences are enriched with β-branched amino acids as well as hydrophobic residues, syntheses of these test peptides were carried out in a semi-automated fashion after coupling the first 11 amino acids in an automated mode. Coupling of amino acids in general was carried out twice with a fivefold excess of Fmoc-protected α-amino acid for 2 h in a mixture of dimethylformamide (50%, [v/v]) and *N*-methyl pyrrolidone, using diisopropylcarbodiimide-hydroxybenzotriazole mediated couplings [45]. To ensure completion of the coupling reaction, an aliquot of the resin was tested using the ninhydrin reaction [46]. Segments of the test peptides that posed difficulty in completing either Fmoc-deprotection or coupling reaction steps are shown for three representative peptides. The synthesis of peptide **3** was performed on preloaded Fmoc-Ser(tBu)-Wang resin (Advanced ChemTech, Louisville, KY, USA), while the syntheses of peptides **14** and **15** used H-Pro-2-Cl-Trt-resin (Infinity Biotech, Upland, PA, USA). Deprotection was achieved with two 20 min reactions with 25% (v/v) piperidine in dimethyl formamide (DMF), except as noted.

Peptide solubility

Peptides were dissolved in phosphate-buffered saline at pH 7.5 by repeated sonication. The suspended solution was then centrifuged and the supernatant removed. A 100 µl aliquot was sent to the Protein Structure Lab at the University of California at Davis for amino acid analysis. The number of moles of the soluble tested peptide were then calculated from the known amino acid composition of the peptide.

Synthesis of dimerizer peptides

Peptide **3**. Difficult coupling steps in peptide **3** are in italics, and the reagents which were successfully used are listed below. EF*LI*VKSVG-TI*V*TMEY*R*ID*R*TRSFVEFLIVKS. Coupling of *R* to T used 2-(1-H-benzotriazole-1-yl)1,1,3,3-tetramethyluroniumhexafluorophosphate (HBTU), N-hydroxybenzotriazole (HOBt) and diisopropylethylamine [47]. Coupling of *R* to I used the pentafluorophenyl active ester of fmoc-arg. Coupling of *V* to T used *O*-(7-azabenzotriazole-1-yl)-1,1,3,3, tetramethyluronium hexafluorophosphate (HATU), 1-hydroxy-7-azabenzotriazole (HOAt), diisopropylethylamine [48]. Coupling of *I* to V and *L* to I utilized diisopropylcarbodiimide and HOBt at 55°C in 1:1:1 dichloromethane : dimethylformamide : *N*-methyl pyrrolidone with 1% Triton X-100 and 2 M ethylene carbonate [49]. Deprotection of Fmoc in the LIV sequence used two 20 min treatments with 2% 1, 8-diazabicyclo[5, 4, 0] undec-7-ene (DBU) in DMF with 5% (v/v) piperidine [50].

Peptide **14**. For this peptide, MGEF**LI**VKSGD*FNHFGNYLLDRRFFIAF*-GEFLIVKLGPP, all of the italicized amino acids were triple-coupled in a fivefold excess of the Fmoc-amino-acid using benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphoniumhexafluorophosphate (BOP), [51] - HOBt-diisopropylethylamine. The couplings of **I** to V and **L** to I (bold) used a three-hours-per-coupling triple coupling with HATU-HOAt-diisopropylethylamine [48]. Deprotections for the **L** to I and **I** to V couplings were performed using two 20 min reactions with 2% DBU in DMF containing 5% (v/v) piperidine [50].

Peptide **15**. For this peptide, MGEF**LI**VKSGHSS*GIPVGVGWCWN-SAGG*GEFLIVKSGPP, the italicized residues were triple-coupled as in peptide **14** above. **I** to V and **L** to I couplings (bold) and deprotections were as in peptide **14**.

Elastolysis of peptide constructs

Inhibition of elastase by synthetic peptides was measured with peptides (10 µM or as indicated) incubated with 50 nM elastase for 1 min followed by assay of the elastase at pH 7.88, 25°C, using 100 µM succinyl-Ala–Ala–Ala–*p*-nitroanilide as a substrate. The change in absorbance at 412 nm with time was observed, and the initial slope was calculated. Inhibition was calculated as:

1 – ([enzyme activity + peptide] – [enzyme activity + PMSF])/([enzyme activity + no addition] – [enzyme activity + PMSF]).

For examination of elastolysis by MS, purified synthetic peptides (5–10 µM final concentrations) were dissolved in 0.15 M tris buffer, pH 7.88, at 25°C. Elastase was added to 100 nM. At times of ~0, 15 min, and 1, 2, 3 and 24 h, an aliquot of the reaction mixture was injected onto a 0.1 × 25 cm C18 reverse-phase HPLC column (Vydac Inc., Hesperia, CA, USA) for peptides **1** and **2**, or onto a 0.1 × 25 cm betabasic cyano reversed-phase column (Keystone Scientific, Bellefonte, PA, USA) for peptides **3**–**6**. For peptides **1** and **2**, the reaction mixture was eluted at 100 µl/min. using 100% A (99.88% H_2O, 0.1% [v/v] acetic acid, 0.02% trifluoroacetic acid) for 4 min, followed by a 1%/min increasing gradient of B (99.88% acetonitrile, 0.1% acetic acid, 0.02% trifluoroacetic acid) to 40% B, and a 6%/min gradient of B to 100%. For peptides **3**–**6**, elution was with 10% B for 4 min followed by a 1%/min gradient increasing in B for 60 min, and a 6.7%/min increasing gradient in B for 6 min. Peptides were examined by direct elution from the column into the electrospray source of a Finnigan LCQ ion trap mass spectrometer. Peptide masses were scanned from 300–2000 amu, and identified by searching their mass with that of different fragments of the full-length peptide, or comparing their mass with different masses of expected elastolytic fragments in the case of the cyclic peptide, using MacBioSpec (obtained courtesy of PE-Sciex, Foster City, CA, USA).

Proteolysis of the reduced peptide CGTIVTMEYRIDRTRSFC was carried out in the presence of 2 mM DTT (Sigma Chemical Co.,

St. Louis, MO, USA). Cleavage products of the oxidized peptide were either directly chromatographed without reduction, or chromatographed after an aliquot was treated first with 1 mM PMSF for 1 h and then with 30 mM DTT for 10 min.

Gel filtration of peptides

Gel filtrations experiments to obtain the apparent molecular weight of selected peptide constructs in saturated solutions were run on a 1×30 cm Pharmacia Superdex Peptide column equilibrated with Dulbecco's phosphate buffered saline (pH 7.3) (Mediatech Inc., Herndon, VA, USA). Column elution was carried out using a Hewlett-Packard 1100 high-performance liquid chromatography instrument at 0.35 ml/min at 23°C. The apparent molecular weights of the constructs were calculated from a standard curve containing eight proteins ranging in molecular weights from 21.5–3.5 kDa and 16 peptides ranging in molecular weights from 0.4–5 kDa.

Deuterium exchange measurements

Deuterium exchange experiments were carried out by dissolving the peptide of interest in water at pH 5, and diluting the peptide ten-fold into D_2O at $t = 0$. For the initial constructs tested, the peptide concentrations after dilution in D_2O were in the range of 10 μM. For other time points, an aliquot of the peptide solution was quenched by addition of a 2.5-fold volume excess of 1:1 H_2O : MeCN with 1% formic acid at 0°C or 25°C and immediately infused into the mass spectrometer. This acidic pH jump slows the rate of amide-bond hydrogen exchange with solvent. For selected time points, the mass derived from the first two minutes of the infusion was compared with that of later 2 min blocks to assess the significance of back-exchange, which was usually one proton or less. The total number of exchangeable protons was derived by (a) initially dissolving the peptide in DMSO, diluting it directly into D_2O before quenching, and measuring the new mass of the peptide several minutes later; (b) diluting a peptide dissolved in 5% dimethylsulfoxide (DMSO) tenfold into D_2O; or (c) heating the solution of peptide diluted into D_2O at 100°C for 15 min. DMSO was included since, in preliminary experiments, low levels added to aqueous peptides appeared to greatly accelerate proton exchange. When all three methods were used with peptide **3** they gave the same results. Calculation of the total protons exchanged included correction for the 10% by volume of H_2O present after dilution in D_2O. For peptides which were soluble in the 1 mM range, samples from the tenfold D_2O solution at pH 5 were directly infused into the mass spectrometer without quenching. Rate constants and amplitudes for deuterium exchange were derived by fitting the timecourse of the gain in mass above the fully protonated form to a single exponential function.

Circular dichroism measurements

Circular dichroisin spectra were recorded on an AVIV 62A DS CD spectropolarimeter (Lakewood, NJ, USA) equipped with a Peltier cooler. The temperature of the instrument was maintained constantly below 20°C using a Neslab CFT-33 refrigerated recirculator waterbath. The instrument was periodically calibrated with the ammonium salt of (+)-10-camphorsulfonic acid, according to the manufacturer's recommendations. Spectra were recorded between 250 and 195 nm at 0.2 nm intervals with a time constant of 1 s at 25°C. Data were collected from five separate scans and averaged using an IBM PS/2 computer . A rectangular quartz cell of path length 0.1 cm was used for the spectral range with the peptide concentration in the range of 0.02–0.05 mM as determined by amino acid analysis. Peptide stock solutions (up to 1 mM) were made in 10 mM KPO_4 buffer containing 100 mM KF at pH 7.5, and were diluted in the same buffer to the final concentration. For pH titration experiments, the pH of the buffer was carefully adjusted to the desired value using either 0.1 M HCl or 0.1 M NaOH, before adding the above peptide stock solution. Raw data were converted to ASCII format and plotted using Microsoft Excel [52]. Thermal denaturation data were acquired on samples containing 20 μM peptide in 10 mM KPO_4 buffer containing 100 mM KF at pH 7.5. The thermal denaturation was measured at the appropriate wavelength over a range of 4–98°C, with a temperature step of 2°C, a 2 min equilibration time and a 60 s signal averaging time. The data were fitted to a logistic sigmoid equation using the Levenberg-Marquardt algorithm in Ultrafit (Biosoft, Cambridge, UK) or the apparent T_m was calculated as the maximum of the first derivative of the CD signal with respect to temperature. Both methods of T_m calculation agreed well. CD spectra were deconvoluted with the program CDNN (CD neural network, [53]) downloaded from http://bioinformatik.biochemtech.uni-halle.de/cdnn/index.html.

Supplementary material

Supplementary material including a table of elastolytic fragments of peptides **1**–**3** and a table of chemical shift values for peptide **11** at pH4.0 is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We thank Richard Scheller (Stanford University) for use of the AVIV circular dichroism spectrometer, Michael Levine (SUNY, Buffalo) for providing NMR time, Karla Blonsky for a careful reading of the manuscript, and Rigel Inc. for support of this work.

References

1. Goodman, M., *et al.*, & Reisine, T. (1992). Topochemical design of bioactive peptides and peptidomimetics. *Bioorg. Khim.* **18**, 1375-1393.
2. Cunningham, B.C. & Wells, J.A. (1997). Minimized proteins. *Curr. Opin. Struct. Biol.* **7**, 457-462.
3. Souroujon, M.C. & Mochly-Rosen, D. (1998). Peptide modulators of protein–protein interactions in intracellular signaling. *Nat. Biotechnol.* **16**, 919-924.
4. Goldberg, A.L., Akopian, T.N., Kisselev, A.F., Lee, D.H. & Rohrwild, M. (1997). New insights into the mechanisms and importance of the proteasome in intracellular protein degradation. *Biol. Chem.* **378**, 131-140.
5. Hilt, W., & Wolf, D.H. (1996). Proteasomes: destruction as a programme. *Trends. Biochem. Sci.* **21**, 96-102
6. Lee, D.H. & Goldberg, A.L. (1998). Proteasome inhibitors: valuable new tools for cell biologists. *Trends Cell Biol.* **8**, 397-402.
7. Belich, M.P., & Trowsdale, J. (1995). Proteasome and class I antigen processing and presentation. *Mol. Biol. Rep.* **21**, 53-56.
8. Meister, A. (1988). On the discovery of glutathione. *Trends Biol. Sci.* **13**, 185-188.
9. Uhlig, S. & Wendel, A. (1992). The physiological consequences of glutathione variations. *Life Sci.* **51**, 1083-1094.
10. Derman, A.I., Prinz, W.A., Belin, D. & Beckwith, J. (1993). Mutations that allow disulfide bond formation in the cytoplasm of *E. coli*. *Science* **262**, 1744-1747.
11. Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. & Brent, R. (1996). Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* **380**, 548-550.
12. Abedi, M., Caponigro, G. & Kamb, A. (1998). Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res.* **26**, 623-630.
13. Scott, C., Abel-Santos, E., Wall, M., Wahnon, D. & Benkovic, S. (1999). Production of cyclic peptides and proteins *in vivo*. *Proc. Natl Acad. Sci. USA* **96**, 13638-13643.
14. Bodenmuller, H., Schilling, W., Zachmann, B. & Schaller, H. (1986). The neuropeptide head activator loses its biological activity by dimerization. *EMBO J.* **5**, 1825-1829.
15. McPhalen, C.A. & James, M.N. (1987). Crystal and molecular structure of the serine proteinase inhibitor CI-2. *Biochemistry* **26**, 261-269.
16. Leatherbarrow, R.J. & Salacinski, H.J. (1991). Design of a small peptide-based proteinase inhibitor by modeling the active-site region of barley chymotrypsin inhibitor 2. *Biochemistry* **30**, 10717-10721.
17. Struthers, M., Cheng, R. & Imperiali, B. (1996). Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science* **271**, 342-345.
18. Braisted, A. & Wells, J. (1996). Minimizing a binding domain from protein A. *Proc. Natl Acad. Sci. USA* **93**, 5688-5692.
19. McKnight, C., Doering, D., Matsudaira, P. & Kim, P. (1996). A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **12**, 126-134.
20. Dahiyat, B. & Mayo, S. (1997). *De novo* protein design: fully automated sequence selection. *Science* **278**, 82-87.
21. Spector, S., Young, P. & Raleigh, D.P. (1999). Native-like structure and stability in a truncation mutant of a protein minidomain: the peripheral subunit-binding domain. *Biochemistry* **38**, 4128-4136.
22. Englander, S.W., Mayne, L., Bai, Y. & Sosnick, T.R. (1997). Hydrogen

exchange: the modern legacy of Linderstrom-Lang. *Protein Sci.* **6**, 1101-1109.
23. Woodward, C. (1999). Advances in protein hydrogen exchange by mass spectrometry. *J. Am. Soc. Mass Spectrom.* **10**, 672-674.
24. Chung, E.D., *et al.*, & Robinson, C. (1997). Hydrogen exchange properties of proteins in native and denatured states monitored by mass spectrometry. *Protein Sci.* **6**, 1316-1324.
25. Smith, D.L., Deng, Y. & Zhang, Z. (1997). Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* **32**, 135-146.
26. Bai, Y., Milne, J., Mayne, L. & Englander, S. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **67**, 75-86.
27. Gariani, T. & Leatherbarrow, R. (1997). Stability of protease inhibitors based on the Bowman-Birk reactive site loop to hydrolysis by proteases. *J. Peptide Res.* **49**, 467-475.
28. Bloemendal, M. & Johnson, W.C. (1995). Structural information on proteins from circular dichroism spectroscopy possibilities and limitations. *Pharm. Biotechnol.* **7**, 65-100.
29. Woody, R.W. (1995). Circular dichroism. *Methods Enzymol.* **246**, 34-71.
30. Greenfield, N. (1996). Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.* **235**, 1-10.
31. Greenfield, N. & Fasman, G. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**, 4108-4116.
32. Perczel, A., Hollosi, M., Sandor, P. & Fasman, G.D. (1993). The evaluation of type I and type II β-turn mixtures. Circular dichroism, NMR and molecular dynamics studies. *Int. J. Peptide Prot. Res.* **41**, 223-236.
33. Johnson, W.C. (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins Struct. Funct. Genet.* **7**, 205-214.
34. Rance, M., Sorenson, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. & Wüthrich, K. (1983). Improved spectral resolution in cosy ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Comm.* **117**, 479-485.
35. Pardi, A., Billeter, M. & Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C alpha proton coupling constants, $^{3}J_{HN\alpha}$, in a globular protein. Use of $^{3}J_{HN\alpha}$ for identification of helical secondary structure. *J. Mol. Biol.* **180**, 741-751.
36. Gururaja, T.L., Levine, J.H., Tran, D.T., Naganagowda, G.A., Ramalingam, K. & Levine, M.J. (1999). Candidacidal activity prompted by N-terminus histatin-like domain of human salivary mucin (MUC7). *Biochem. Biophys. Acta* **1431**, 107-119.
37. Brahms, S. & Brahms, J. (1980). Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* **138**, 149-178.
38. Martsev, S.P., Vlasov, A.P. & Arosio, P. (1998). Distinct stability of recombinant L and H subunits of human ferritin: calorimetric and ANS binding studies. *Protein Eng.* **11**, 377-81.
39. Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D. & Scharpe, S. (1995). Proline motifs in peptides and their biological processing. *FASEB J.* **9**, 736-744.
40. Imperiali, B. & Ottesen, J. (1998). Design strategies for the construction of independently folded polypeptide motifs. *Biopolymers* **47**, 23-29.
41. Murray, M., *et al.*, & Niyogi, S. (1998). Structure–function analysis of a conserved aromatic cluster in the N-terminal domain of human epidermal growth factor. *Protein Eng.* **11**, 1041-1050.
42. Javitch, J., Ballesteros, J., Weinstein, H. & Chen, J. (1998). A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* **37**, 998-1006.
43. Frank, B., Buckley, D. & McKnight, C. (1998). Hydrophobic core mutations in the mini-protein HP35: a key phenylalanine specifies the 3D fold. *Protein Sci.* **7(suppl 1)**, 264S.
44. Fields, G. & Noble, R. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxy-carbonyl amino acids. *Int. J. Pept. Protein Res.* **35**, 161-214.
45. Hudson, D. (1988). Methological implications of simultaneous solid-phase peptide synthesis. I. Comparison of different coupling procedures. *J. Org. Chem.* **53**, 617-624.
46. Kaiser, E., Colescott, R., Bossinger, C. & Cook, P. (1970). Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34**, 595-598.
47. Knorr, R., Trazeciak, A. & Bannworth, W. (1989). New coupling reagents in peptide chemistry. *Tetrahedron Lett.* **30**, 1927-1930.
48. Carpino, L.A. (1993). 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* **115**, 4397-4398.
49. Zhang, L. (1994). In *Innovation and perspectives in solid phase synthesis. 3rd International Symposium.* (Epton R. ed.), pp. 711, Mayflower Scientific Ltd., Birmingham.
50. Wade, J., Bedford, J., Sheppard, R. & Tregear G. (1991). DBU as an N alpha-deprotecting reagent for the fluorenylmethoxycarbonyl group in continuous flow solid-phase peptide synthesis. *Peptide Res.* **4**, 194-199.
51. Castro, B., Dormoy, J.R., Evin, G. & Selve, C. (1975). Peptide coupling reagents. N-[Oxytris(dimethylamino)phosphonium]benzotriazole hexafluorophosphate. *Tetrahedron Lett.* **14**, 1219-1222
52. Gururaja, T.L. & Levine, M.J. (1996). Solid-phase synthesis and characterization of human salivary statherin: a tyrosine-rich phosphoprotein inhibitor of calcium phosphate precipitation. *Peptide Res.* **9**, 283-289.
53. Bohm, G., Muhr, R. & Jaenicke, R. (1992). Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **5**, 191-195.