Peptidomimetics

DOI: 10.1002/anie.201105119

Protein α-Turns Recreated in Structurally Stable Small Molecules**

*Huy N. Hoang, Russell W. Driver, Renée L. Beyer, Alpeshkumar K. Malde, Giang T. Le, Giovanni Abbenante, Alan E. Mark, and David P. Fairlie**

About 30% of protein structure is composed of peptide helices. An ideal α-helix consists of repeating (classical) α-turns, each of 3.6 amino acids, with identical backbone dihedral angles (ϕ −58°, ψ −47°), and a 13-membered hydrogen-bonded ring (Figure 1).[1] However, α-turns in proteins are not identical due to variations in consecutive (ϕ, ψ) angles that subtly alter peptide backbone shape, side chain orientation and helical pitch (Figure 1).[2a,b] For example, a study of 460 proteins identified 9 distinctly different α-turn types, with very different (ϕ, ψ) angles.[2c] Of 7548 helices in 1085 proteins,[3] the most common helix length was just 4 amino acids or one α-turn. Corresponding 4-residue synthetic peptides are not thermodynamically stable structures in water, making it hard to study properties for α-turn motifs independent of packing influences in proteins. Here we present a solution to this problem, constraining tetrapeptides to adopt extremely water-stable α-turns of two types. We use 2D ^{1}H NMR spectroscopy and molecular dynamics simulations to prove that the resulting three-dimensional structures are quite rigid and closely match two α-turn types that dictate structure at important sites in 20 different protein environments.

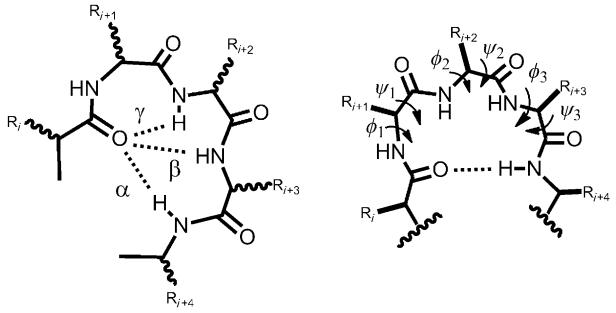


Figure 1. Left: Turns (α,β,γ) in proteins. Right: The α-turn can vary in dihedral angles of three consecutive residues, $i+1$ (ϕ_1, ψ_1), $i+2$ (ϕ_2, ψ_2), $i+3$ (ϕ_3, ψ_3), and may show a 13-membered H-bonded ring.

We assemble evidence (Table S1 in the Supporting Information) from protein structures for two α-turn types that occur at important locations in proteins—in enzyme active sites, in metal-binding domains, at ends of structural motifs, in kinks of helices, or in isolation away from α-helices. Despite key structural influences of these α-turns in proteins,[4] no small molecules have yet been developed to mimic them, although approaches have been established to stabilize α-helices.[5–9] Our attempts to trap the first water-stable "non-classical" α-turns in small molecules addressed the question of whether 3-residue units could recapitulate the backbone geometries of non-classical α-turn types within cyclic tetrapeptides (Figure 2), by linking the side-chain of residue i to the C-terminus of residue $i+3$ (e.g. cyclo-(1,4)-Ac-[DapARA], **1**) or the side-chain of residue $i+3$ to the N-terminus of residue i (e.g. cyclo-(1,4)-[ARAE]-NH_2, **2**). These connections are too short to induce conventional classical α-helical backbone geometry (ϕ −58°, ψ −47°), so we expected more planar conformations. For comparison, we also synthesized cyclic pentapeptide, cyclo-(1,5)-Ac[KARAD]-NH_2 (**3**), which we have shown to be an α-helical turn (ϕ −58°, ψ −47°).[6] Four peptides (**1**, **2**, Ala-Arg-Ala-Glu-NH_2, Ac-Dap-Ala-Arg-Ala-OH) were synthesized on solid phase using Fmoc/HBTU, purified by semi-preparative HPLC using ACN/H_2O gradients, and characterized by analytical rp-HPLC, high-resolution MALDI-TOF and ^{1}H NMR spectra (Supporting Information). Peptides were cyclized with diphenyl-phosphoroazide/DIPEA (diisopropylethylamine) at low μM concentrations in DMF at 4°C in low yields (15% **1**; 12% **2**).

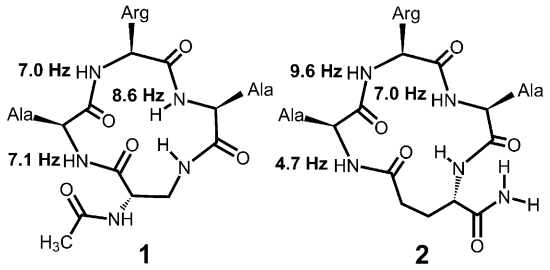


Figure 2. Cyclic peptides **1** and **2** are 13- and 14-membered ring structures, with amide coupling constants ($^3J_{NH\text{-}CH\alpha}$) shown in Hz.

α-Helical peptides with > 25 amino acids typically display circular dichroism (CD) spectra with two molar ellipticity minima ($\lambda = 222$, 208 nm; ratio 0.9 ± 0.2) corresponding to n–π* and π–π* electronic transitions, and an ellipticity max-

[*] Dr. H. N. Hoang, Dr. R. W. Driver, Dr. R. L. Beyer, Dr. G. T. Le, Dr. G. Abbenante, Prof. A. E. Mark, Prof. D. P. Fairlie
Division of Chemistry and Structural Biology
Institute for Molecular Bioscience
The University of Queensland, Brisbane, Qld 4072 (Australia)
E-mail: d.fairlie@imb.uq.edu.au
Dr. A. K. Malde, Prof. A. E. Mark
School of Chemistry and Molecular Biosciences (SCMB)
The University of Queensland, Brisbane, Qld 4072 (Australia)

[**] We thank the Australian Research Council for research support and for a Federation Fellowship to D.P.F. (FF0668733 and DP1096290).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201105119.

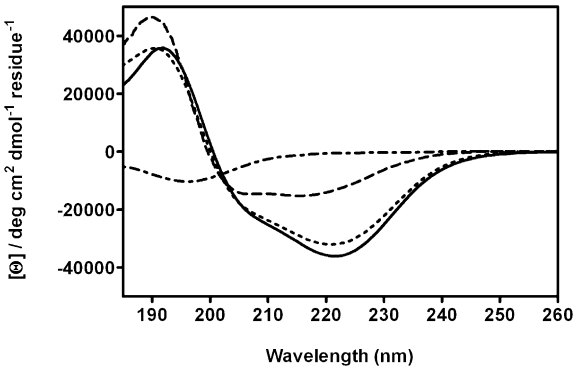
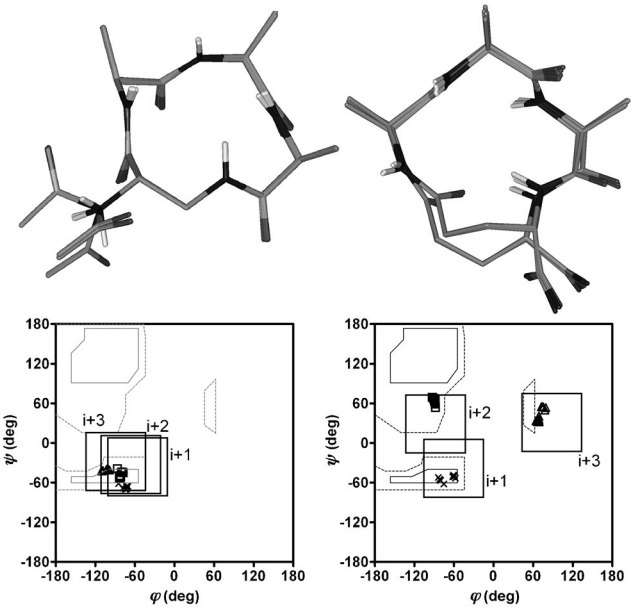


Figure 3. CD spectra for **1** (solid line), **2** (dotted line), **3** (dashed line) vs. acyclic tetrapeptide Ac-DapARA-OH (dot–dashed line) in aqueous 10 mM phosphate buffer (pH 7.4, 298 K).

imum (192 nm).[10] CD spectra for **1**–**3** in 10 mM phosphate buffer pH 7.4, 298 K (Figure 3) are like those often attributed to peptide helices, while the CD spectra of the uncyclized analogues were characteristic of unstructured peptides (Figure 3, dot–dashed line). Adding 2,2,2-trifluoroethanol or varying concentrations (50 μM–7.7 mM) of **1**–**3** did not change spectral lineshapes or intensities, suggesting that there was neither conformational averaging nor concentration-dependent aggregation (Figures S3, S15, S16). In addition, NMR parameters for **1** and **2** (e.g. ^{1}H,^{13}C chemical shifts, $^{3}J_{\text{NH-CH}\alpha}$, and amide proton temperature coefficients (Figures S8, S9, S10, S11, S17, and S18) are sequence-dependent and therefore support well-defined rigid structures in **1** and **2**. A key difference between **1** or **2** and the α-helical **3** is the ratio of $[\Theta]_{222}$:$[\Theta]_{208}$ (ca. 1.5:1 vs ca. 1:1) in CD spectra, which have not previously been reported for α-turn types reported herein.

^{1}H NMR spectra for **1** and **2** revealed amide coupling constants ($^{3}J_{\text{NH-CH}\alpha} > 6$ Hz, Figure 2) not consistent with an α-helix (< 6.0 Hz). Structures **1** and **2** are clearly different, based on $^{3}J_{\text{NH-CH}\alpha}$ being 7.1 vs 4.7 Hz for Ala1, 7.0 vs 9.6 Hz for Arg2, and 8.6 vs 7.0 Hz for Ala3. Coupling constants between 6–8 Hz in proteins often imply a lack of a defined structure and conformational averaging but, in these small-molecule cases, well-defined specific structures are clearly observed (see ahead). There were no clear intramolecular hydrogen bonds in the cycles **1** and **2**, since variable-temperature ^{1}H NMR spectra showed temperature-dependent amide NH chemical shifts, except possibly for the "$i+4$" amide ($\Delta\delta/T$ 5.4 ppb/deg (**1**), 5.5 ppb/deg (**2**)). The latter temperature coefficients are inconsistent with strong hydrogen bonds,[11] but may be weak hydrogen bonds that would define 6- and 7-membered rings, respectively, with hydrogen bonds connecting the CO corresponding to residue "i" to the NH of residue "$i+4$", each fused to a common 13-membered hydrogen-bonded ring. A weak hydrogen-bonded "macrocycle" is sometimes a feature in α-turns in protein crystal structures.

Solution structures for **1** and **2** were determined in H_2O/D_2O (9:1) at 298 K using ROESY 2D ^{1}H NMR spectra. There were no *cis*-amide bonds. The structure of **1** was calculated from 23 ROE distance restraints, 1 backbone ϕ-dihedral angle restraint derived from $^{3}J_{\text{NH-CH}\alpha}$ ($-120 \pm 30°$), and no hydrogen bond restraints. Structure **2** was calculated from 25 ROE distance restraints, 2 backbone ϕ-dihedral angles derived from $^{3}J_{\text{NH-CH}\alpha}$ ($-120 \pm 30°$, $-60 \pm 30°$) with or without a H-bond constraint that had no effect on structure (Figures S10 and S11). Structures were calculated in XPLOR-NIH[12] using a dynamic simulated annealing protocol in a geometric force field, and energy-minimized using CHARMm force field.[13] The 20 lowest-energy structures (Figure 4) for **1** and **2** had no



	φ_{i+1}	ψ_{i+1}	φ_{i+2}	ψ_{i+2}	φ_{i+3}	ψ_{i+3}	d(Å)
1	-72±1	-67±1	-81±1	-49±3	-103±3	-41±1	5.62±0.02
2	-67±10	-53±4	-89±1	63±4	70±4	41±10	5.55±0.06
3	-55±5	-40±4	-65±5	-46±3	-59±3	-40±4	5.62±0.08

Figure 4. Twenty lowest-energy NMR solution structures (top), Ramachandran plots (middle), and dihedral angles (bottom, in degree) for **1** and **2**. (ϕ, ψ) regions defining α-turn types I-α_{RS} (middle left) and II-α_{LU} (middle right) in proteins are demarcated by squares (45° long, average (ϕ, ψ) angles[3]). Corresponding ϕ and ψ angles Ala1 (crosses), Arg2 (squares), Ala3 (triangles) for **1** (left) and **2** (right) map to regions consistent with assigned α-turn types I-α_{RS} and II-α_{LU}, respectively. d: distance from C$\alpha_{(i+1)}$ to C$\alpha_{(i+3)}$.

distance (≥ 0.2 Å) or dihedral angle ($\geq 5°$) violations and are quite rigid, convergent structures (average pairwise backbone RMSD 0.04 Å and 0.26 Å for **1** and **2**, respectively). The (ϕ, ψ) angles for residues Ala1, Arg2, Ala3 of **1** (Figure 4) are within α-helix space, but ϕ is more negative than $-58°$. For **2** there is greater deviation from α-helicity; the $(i+3)$ residue is in left-handed α-helix (ϕ, ψ) space, while the $(i+2)$ residue occupies β-structure (ϕ, ψ) space. Plots of (ϕ, ψ) of the 20 lowest-energy structures of **1** and **2** fit within regions previously defined for α-turn types I-α_{RS} and II-α_{LU} (Figure 4) in proteins.[3] The (ϕ, ψ) angles for **1** and **2** are distinctly different from **3** (Figure 4), and from known β-turns (Figures S13 and S14), consistent with their unique fit to α-turns.

Small cyclic peptides tend not to exhibit as many NOEs as proteins, so for NMR-derived structures it is sometimes necessary to distinguish between small rigid structures and a

time-averaged conformation that may not reflect the conformational ensemble populated by the peptide.[14] To assess the degree of rigidity in the NMR-derived structures for peptides **1** and **2**, molecular dynamics (MD) simulations were carried out (Figure 5) using the GROMOS simulation package[15] with the GROMOS 53A6[16] parameter set. The net charge on each peptide was +1e. MD simulations were started from the NMR-calculated structures that were allowed to equilibrate for 100 ns at 298 K in a water environment (1038 H_2O molecules per peptide). Backbone atom-positional RMSD were calculated after translational superposition of centers of mass and least-squares rotational fitting of atomic positions. Cluster analysis (RMSD cutoff 0.05 nm) was performed using structures extracted from the trajectory every 0.01 ns. Figure 5 shows that both peptide **1** and **2** were almost exclusively in the NMR-derived rigid conformation during the simulations. The average inter-proton ROE distances and $^3J_{NHH\alpha}$ calculated from the simulations (Figure 5; Tables S6–S9) agreed well with the experimental NMR data. The proposed structures for **1** and **2** in water were also the dominant conformations sampled during the simulations. Thus, the results from the MD simulations clearly validate the NMR-derived solution structures of peptides **1** and **2**.

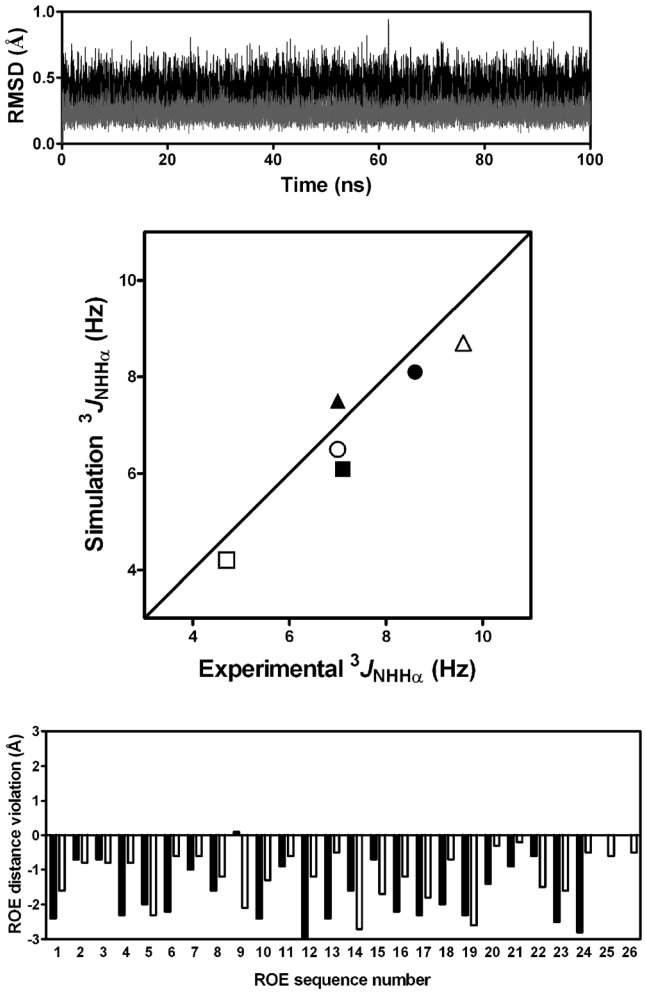
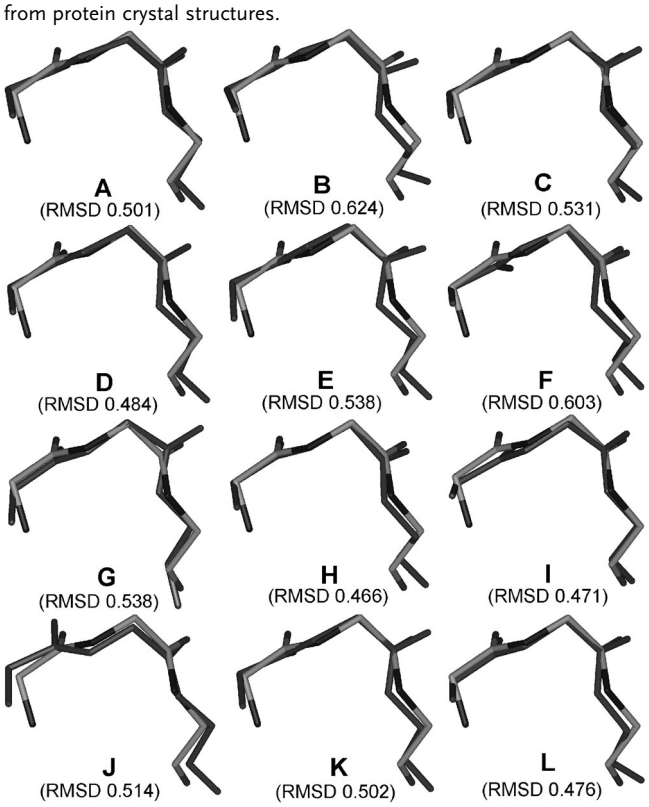


Figure 5. Molecular dynamics simulations for **1** (black) and **2** (gray, unfilled). Above: Backbone RMSD variations over 100 ns. Middle: Comparison of $^3J_{NHH\alpha}$ derived for three residues: Ala(i+1) (square), Arg(i+2) (triangle), and Ala(i+3) (circle) from NMR experimental versus MD simulation studies. Bottom: Difference of averaged MD simulated and NMR ROE upper-bound distances, showing no positive distance violations (>0.1 Å). The sequence of ROEs are in Tables S8 and S9.

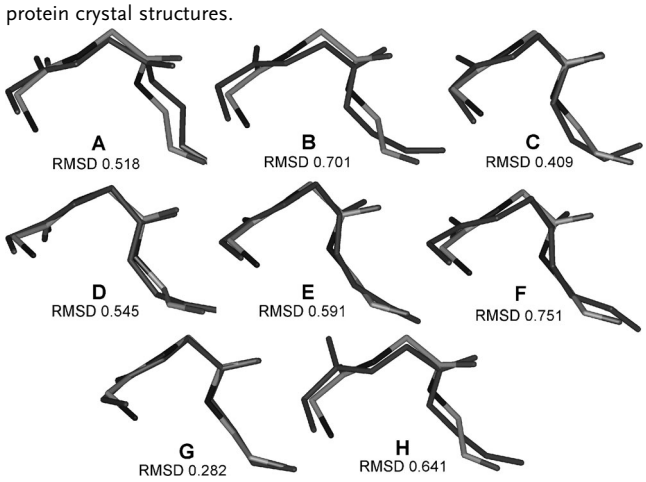
Table 1: Average backbone solution structure of **1** (gray; residues: A, R, A) superimposed on the three middle, consecutive residues of α-turns from protein crystal structures.



	Protein	Turn type	Residues	Location	PDB
A	human lysozyme	I-α_{RS}	46–50	active-site cleft	1lz1
B	ferredoxin I	I-α_{RS}	35–39	ligated to the 8 Fe ferredoxin	1fd2
C	blue copper protein azurin	I-α_{RS}	117–121	His117 and Met121 coordinate copper	2aza
D	cholesterol oxidase	I-α_{RS}	115–119	contacts FAD	1cox
E	T-cell surface glycoprotein CD4	I-α_{RS}	51–55	HIV gp120 binding site	2CD4
F	bilin-binding protein	I-α_{RS}	19–23	connected to β-strand	1bbp
G	[2Fe-2S] ferredoxin I	I-α_{RS}	75–79	Ile74 in β-barrel	1fxi
H	glyceraldehyde-3 phosphate dehydrogenase	I-α_{RS}	192–196	Arg195 in active site	1gd1
I	photosynthetic reaction center	I-α_{RS}	87–91	heme-binding	1prC
J	α-lytic protease	I-α_{RS}	140–156	adjacent to active site	1P02
K	*S. griseus* trypsin	I-α_{RS}	177–181	coordinates Ca^{2+}	1SGT
L	*p*-hydroxybenzoate hydroxylate	I-α_{RS}	285–289	connects two β-strands	1phh

Gratifyingly, the structures of **1** and **2**, calculated based on NMR results and validated by MD simulation, closely match two α-turn types I-α_{RS} and II-α_{LU} in 20 different environments in protein crystal structures examined (Tables 1 and 2). These non-classical α-turns clearly play important structural roles within human, mammalian, and bacterial proteins of very diverse functions (Table 1 and 2) and the capacity to recreate these important structural motifs in small molecules allows their study independent of packing influences in protein structures.

Table 2: Average backbone solution structure of **2** (gray; residues: A, R, A) superimposed on three middle, consecutive residues of α-turns from protein crystal structures.



	Protein	Turn type	Residues	Location	PDB
A	selenoenzyme glutathione peroxidase	II-α_{LU}	75–79	connects β-sheet and α-helix	1gp1
B	isolectins 1 and 2 of wheat germ agglutinin	II-α_{LU}	148–152	Trp150 and Ser152 interact with sugar	2wgc
C	hirudin human α-thrombin complex	II-α_{LU}	14I-14M A chain	Interacts with fibrinogen exosite of thrombin	4htc
D	*S. griseus* trypsin	II-α_{LU}	33–37	connects two β-sheets	1sgt
E	proteinase K	II-α_{LU}	149–153	connects β-sheet and α-helix	2prk
F	photoreversible cinnamate	II-α_{LU}	35–39	connects two β-sheets	3gch
G	electron transport protein	II-α_{LU}	109–113	C-terminal cap for α-helical domain	3c2c
H	serine protease thermitase	II-α_{LU}	239–243	connects two α-helices	1tec

In summary, we have devised small molecules which closely recapitulate two important non-classical α-turns that are influential structural motifs in important structural sites in proteins. The cyclic tetrapeptides, cyclo-(1,4)-Ac[DapARA] (**1**) and cyclo-(1,4)-[ARAE]-NH_2 (**2**), closely mimic two distinct protein α-turn types, I-α_{RS} and II-α_{LU}, respectively. Compared with the known conventional α-helical turn in pentapeptide (**3**), these new structures differ significantly in (ϕ, ψ) angles, $^3J_{NH\text{-}CH\alpha}$ coupling constants, NOE data, and CD spectra. Peptides **1** and **2** are pseudo-planar rather than having a helical pitch as in **3**. We conclude that the particular side-chain to main-chain connections in the cyclic tetrapeptides **1** and **2** have successfully constrained the peptide backbone of the intervening tripeptide fragment to fold into distinct "non-classical" α-turn types, and a similar approach may enable creation of small molecules that mimic other α-turn types that occur in proteins. These studies contribute to a better understanding of the properties of structural motifs found in proteins and teach us how to translate those motifs into small molecules of well-defined structure.

Received: July 21, 2011
Published online: September 29, 2011

Keywords: α-turn · circulardichroism · helices · NMR spectroscopy · peptidomimetics

[1] a) D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1988**, *201*, 601; b) T. E. Creighton in *Proteins: Structures and Molecular Properties*, 2nd ed., Freeman, San Francisco, **1993**.
[2] a) D. V. Nataraj, N. Srinivasan, R. Sowdhamini, C. Ramakrishnan, *Curr. Sci.* **1995**, *69*, 434; b) K. C. Chou, *Anal. Biochem.* **2000**, *286*, 1; c) V. Pavone, G. Gaeta, A. Lombardi, F. Nastri, O. Maglio, C. Isernia, M. Saviano, *Biopolymers* **1996**, *38*, 705.
[3] U. Pal, P. Chakrabarti, G. Basu, *J. Mol. Biol.* **2003**, *326*, 273.
[4] a) K. C. Chou, *Biopolymers* **1997**, *42*, 837; b) Ref. [2b]; c) B. Dasgupta, L. Pal, G. Basu, P. Chakrabarti, *Proteins Struct. Funct. Bioinf.* **2004**, *55*, 305.
[5] Classical α-turns, however, have been mimicked: a) M. J. I. Andrews, A. B. Tabor, *Tetrahedron* **1999**, *55*, 11711; b) M. J. Kelso, H. N. Hoang, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2000**, *122*, 10488; c) J. W. Taylor, *Biopolymers* **2002**, *66*, 49; d) M. J. Kelso, R. L. Beyer, H. N. Hoang, A. S. Lakdawala, J. P. Snyder, W. V. Oliver, T. A. Robertson, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2004**, *126*, 4828; e) R. S. Harrison, N. E. Shepherd, H. N. Hoang, G. Ruiz-Gomez, T. A. Hill, R. W. Driver, V. S. Desai, P. R. Young, G. Abbenante, D. P. Fairlie, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11686.
[6] a) N. E. Shepherd, G. Abbenante, D. P. Fairlie, *Angew. Chem.* **2004**, *116*, 2741; *Angew. Chem. Int. Ed.* **2004**, *43*, 2687; b) N. E. Shepherd, H. N. Hoang, G. Abbenante, D. P. Fairlie, *J. Am. Chem. Soc.* **2005**, *127*, 2974.
[7] a) C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem. Soc.* **2000**, *122*, 5891; b) G. L. Verdine, L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, S. J. Korsmeyer, *Science* **2004**, *305*, 1466; c) E. I. Shakhnovich, P. S. Kutchukian, J. S. Yang, G. L. Verdine, *J. Am. Chem. Soc.* **2009**, *131*, 4622; d) G. L. Verdine, Y. W. Kim, T. N. Grossmann, *Nat. Protoc.* **2011**, *6*, 761.
[8] a) P. S. Arora, R. N. Chapman, G. Dimartino, *J. Am. Chem. Soc.* **2004**, *126*, 12252; b) A. Patgiri, A. L. Jochim, P. S. Arora, *Acc. Chem. Res.* **2008**, *41*, 1289; c) A. Patgiri, K. K. Yadav, P. S. Arora, D. Bar-Sagi, *Nat. Chem. Biol.* **2011**, *7*, 585.
[9] a) L. D. Walensky, K. Pitter, J. Morash, K. J. Oh, S. Barbuto, J. Fisher, E. Smith, G. L. Verdine, S. J. Korsmeyer, *Mol. Cell* **2006**, *24*, 199; b) L. D. Walensky, M. L. Stewart, E. Fire, A. E. Keating, *Nat. Chem. Biol.* **2010**, *6*, 595; c) L. D. Walensky, G. H. Bird, N. Madani, A. F. Perry, A. M. Princiotto, J. G. Supko, X. Y. He, E. Gavathiotis, J. G. Sodroski, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14093.
[10] a) *Circular Dichroism* (Eds.: K. Kakanishi, N. Berova, R. Woody), VCH, New York, **1994**; b) *Circular Dichroism and*

Conformational Analysis of Biomolecules (Ed.: G. D. Fasman), Plenum, New York, **1996**.
[11] D. S. Wishart, B. D. Sykes, F. M. Richards, *Biochemistry* **1992**, *31*, 1647.
[12] a) A. T. Brunger, *X-PLOR Manual Version 3.1*, Yale University Press, New Haven, **1992**; b) M. Nilges, A. M. Gronenborn, A. T. Brunger, G. M. Clore, *Protein Eng.* **1988**, *2*, 27.
[13] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187.
[14] M. P. Williamson, J. P. Waltho, *Chem. Soc. Rev.* **1992**, *21*, 227.
[15] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, A. E. Torda, T. Huber, P. Krüger, W. F. van Gunsteren, *J. Phys. Chem. A* **1999**, *103*, 3596.
[16] C. Oostenbrink, A. Vila, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.

