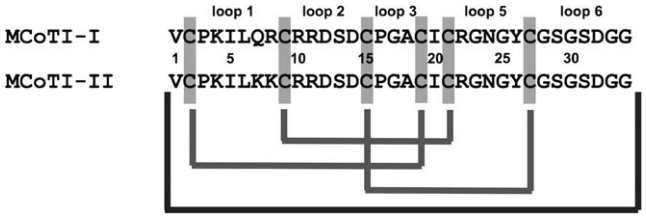
Molecular Dynamics

DOI: 10.1002/anie.201002906

Backbone Dynamics of Cyclotide MCoTI-I Free and Complexed with Trypsin**

*Shadakshara S. Puttamadappa, Krishnappa Jagadish, Alexander Shekhtman, and Julio A. Camarero**

Cyclotides are a new emerging family of large plant-derived backbone-cyclized polypeptides (about 28–37 amino acids long) that share a disulfide-stabilized core (three disulfide bonds) characterized by an unusual knotted arrangement.[1] Cyclotides contrast with other circular polypeptides in that they have a well-defined three-dimensional structure, and despite their small size can be considered as microproteins. Their unique circular backbone topology and knotted arrangement of three disulfide bonds makes them exceptionally stable to thermal and enzymatic degradation (Scheme 1).



Scheme 1. Primary structure and disulfide connectivities of MCoTI cyclotides. Dark gray and light gray connectors represent peptide and disulfide bonds, respectively.

Furthermore, their well-defined structures have been associated with a wide range of biological functions.[2,3] Cyclotides MCoTI-I/II are powerful trypsin inhibitors ($K_i \approx 20$–30 pM) that have been recently isolated from the dormant seeds of *Momordica cochinchinensis*, a plant member of the cucurbitaceae family.[4] Although MCoTI cyclotides do not share significant sequence homology with other cyclotides beyond the presence of the three cystine bridges, structural analysis by NMR spectroscopy has shown that they adopt a similar backbone-cyclic cystine-knot topology.[5,6] MCoTI cyclotides, however, show high sequence homology with related cystine-knot squash trypsin inhibitors,[4] and therefore represent interesting molecular scaffolds for drug design.[7–10]

Determination of the backbone dynamics of these fascinating microproteins is key for understanding their physical and biological properties. Internal motions of a protein on different timescales, extending from picoseconds to a second, have been suggested to play an important role in its biological function.[11] A better understanding of the backbone dynamics of the cyclotide scaffold will be extremely helpful for evaluating its utility as a scaffold for peptide-based drug discovery. Such insight will help in the design of optimal focused libraries that can be used for the discovery of new cyclotide sequences with novel biological activities.[12,13]

Herein, we report for the first time the determination of the internal dynamics of the cyclotide MCoTI-I in the free state and complexed with trypsin. Uniformly ^{15}N-labeled natively folded cyclotide MCoTI-I was recombinantly produced in *Escherichia coli* growing in minimal M9 medium containing $^{15}NH_4Cl$ as the only source of nitrogen. Concomitant backbone cyclization and folding were accomplished by using intramolecular native chemical ligation[14,15] in combination with a modified protein splicing unit (Figure S1, Supporting Information).[16–18] The internal dynamics of cyclotide MCoTI-I was obtained from ^{15}N spin–lattice and spin–spin relaxation times and $^{15}N\{^1H\}$ heteronuclear Overhauser effect (NOE) enhancements.[11] The backbone flexibility was characterized by the square of the generalized order parameter, S^2, which reveals the dynamics of backbone NH groups on the pico- to nanosecond timescale.[19,20] The order parameter satisfies the inequality $0 \le S^2 \le 1$, in which lower values indicate larger amplitudes of intramolecular motions. Motions on the milli- to microsecond timescale were assessed by the presence of the chemical exchange terms in the spin–spin relaxation.

The NMR spectrum and S^2 values, derived from the ^{15}N relaxation data of free MCoTI-I, are shown in Figure 1 a and d, respectively. Residues Ile5 and Gly23 of free MCoTI-I were excluded from the backbone dynamics analysis since the relaxation data could not be fitted to a monoexponential function, possibly as a result of chemical exchange.[21] Gln7 was not assigned because of broadening of the NMR signal, presumably caused by fast exchange with water. The S^2 values for free MCoTI-I show that most of the NH groups of the cyclotide backbone are highly constrained with S^2 values ≥ 0.8, thus resembling those found in well-folded globular proteins (Table 1). The average S^2 value, $\langle S^2 \rangle$, for free MCoTI-I was 0.83 ± 0.03. This value is similar to that found

[*] Dr. K. Jagadish, Dr. J. A. Camarero
Department of Pharmaceutical Sciences and Pharmacology
University of Southern California
Los Angeles, CA 90033 (USA)
Fax: (+1) 323-224-7473
E-mail: jcamarer@usc.edu
S. S. Puttamadappa, Dr. A. Shekhtman
Department of Chemistry, State University of New York
Albany, NY 12222 (USA)

[**] This work was supported by funding from the School of Pharmacy at the University of Southern California, by National Institute of Health award GM090323-01 to J.A.C., and by American Diabetes Association award 1-06-CD-23 and National Institute of Health award GM090323-01 to A.S.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201002906.

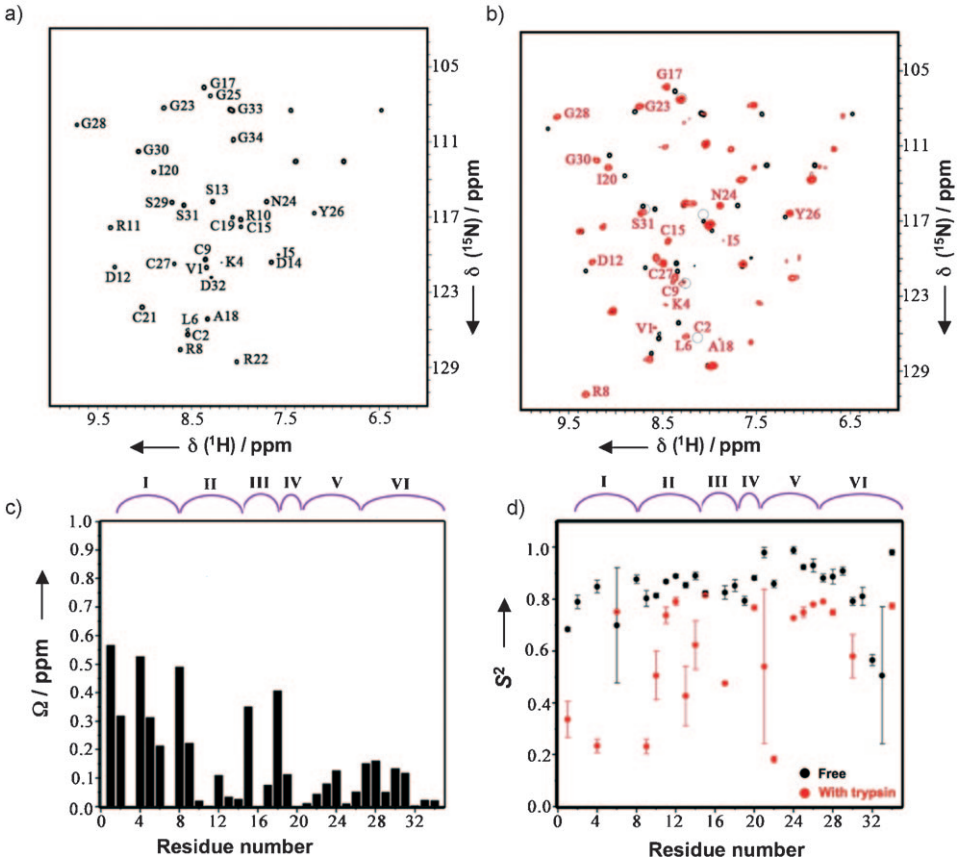


Figure 1. NMR analysis of the backbone dynamics of free and trypsin-bound MCoTI-I. a) $\{^{15}N,^{1}H\}$ NMR heteronuclear single quantum correlation (HSQC) spectrum of free MCoTI-I. Chemical shift assignments of the backbone amides are indicated. b) Overlay of the $\{^{15}N,^{1}H\}$ HSQC spectra of free (black) and trypsin-bound MCoTI-I (red). Residues with large average amide chemical shift differences between two different states (>0.3 ppm) are indicated. Peaks that are broadened in trypsin-bound MCoTI-I are indicated by gray circles. c) Average amide chemical shift difference for all the assigned residues in free and trypsin-bound MCoTI-I. The chemical shift difference was calculated as: $\Delta\Omega = [(\Delta\Omega^2_{NH} + 0.04\,\Omega^2_N)/2]^{1/2}$, where $\Delta\Omega_{NH}$ and $\Delta\Omega_N$ are the changes in the amide proton and nitrogen chemical shifts (ppm), respectively. d) Order parameter, S^2, for free (black) and trypsin-bound MCoTI-I (red). The S^2 value is a measure of backbone flexibility and represents the degree of angular restriction of the N–H vector in the molecular frame. The MCoTI-I loops are shown at the top of (c) and (d). Small unassigned peaks in the spectra of both free and trypsin-bound MCoTI-I are from a minor conformation of the protein, and result from a known isomerization of the backbone at an Asp–Gly sequence in loop 6 of MCoTI-I.

Table 1: Average order parameters of structural elements in MCoTI-I in the free state and bound to trypsin.

Structural element	Sequence	$<S^2>$ [a] Free MCoTI-I	$<S^2>$ [b] Trypsin–MCoTI-I
loop 1	3–8	0.81 ± 0.01	0.49 ± 0.05
loop 2	10–14	0.81 ± 0.01	0.62 ± 0.07
loop 3	16–18	0.84 ± 0.02	0.48[c]
loop 4	20	0.88[c]	0.76[c]
loop 5	22–26	0.92 ± 0.02	0.61 ± 0.01
loop 6	28–34	0.76 ± 0.05	0.61 ± 0.05
cystine knot	2,10,15,19,21,27	0.84 ± 0.02	0.60 ± 0.08

[a] S^2 values for residues 5 and 23 from free MCoTI-I are not included in the average because the relaxation data could not be fitted to a monoexponential function. [b] S^2 values for residues 2, 5, 8, 18, 19, 23, 29, 31, 32, and 33 from trypsin-bound MCoTI-I are not included in the average because of the lack of signal intensity or because the relaxation data could not be fitted to a monoexponential function. [c] $<S^2>$ contains the S^2 value for a single residue.

for the six cystine residues involved in the cystine knot ($<S^2> = 0.84 \pm 0.02$) and is considerably larger than those found for other linear squash trypsin inhibitors ($<S^2> = 0.71$ for trypsin inhibitor from *Cucurbita maxima* (CMTI-III, 78% homology with MCoTI-I)),[22] thus indicating the importance of the backbone cyclization to rigidifying the overall structure. Loops 2 through 5 in free MCoTI-I showed $<S^2>$ values ≥ 0.8. In particular, loop 5 showed an $<S^2>$ value of 0.92 ± 0.02, well above the average for the molecule and the cystine knot. In contrast, loops 1 and 6 showed $<S^2>$ values below the average for the molecule. Thus, loop 6, which is believed to act as a very flexible linker to allow cyclization,[23] had an $<S^2>$ value of 0.76 ± 0.17 with only two residues, Asp32 and Gly33, having values below 0.6. Despite this small $<S^2>$ value, residues in loop 6 did not require significant chemical exchange terms (Figure S2 and Table S1, Supporting Information), which suggests that the mobility observed arises mostly from local vibrations.

The $<S^2>$ value for loop 1, which is responsible for binding trypsin, was 0.81 ± 0.07. This value is $\approx 90\%$ of the average value for free MCoTI-I. Residue Leu6 in loop 1 also required chemical exchange terms to be considered, thus indicating the existence of intramolecular conformational exchange on the micro- to millisecond timescale. The mobility observed in loop 1 at both nano- to picosecond and millisecond timescales has also been described in other trypsin inhibitors,[22,24,25] and it has been suggested to play an important role in receptor–ligand binding.[11]

To explore whether that was the case in the MCoTI cyclotides, we next studied the effect of ligand binding on the backbone dynamics of MCoTI-I (Figures 1 and 2). To exclude the possibility that trypsin could cleave or scramble the disulfide bonds of MCoTI-I upon complex formation, we used a competition experiment of trypsin–[^{15}N]MCoTI-I with unlabeled MCoTI-I. The results indicated that the structure of MCoTI-I is unaltered upon trypsin binding (Figure S3, Supporting Information). Trypsin binding led to large (>0.3 ppm) and specific changes in the chemical shifts of

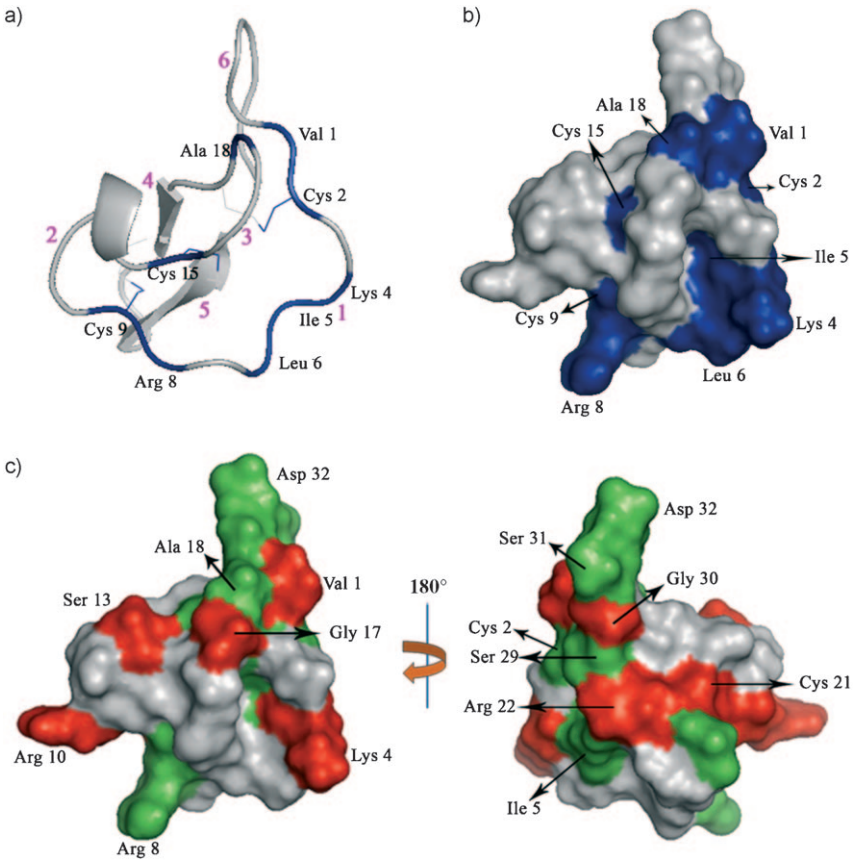


Figure 2. Trypsin binding to MCoTI-I affects the MCoTI-I backbone dynamics. a) Ribbon and b) surface diagrams of the trypsin–MCoTI-I interaction map. Red numbers indicate the positions of the MCoTI-I loops. The MCoTI-I residues with a large chemical shift difference (>0.3 ppm) are in blue. c) Changes in the MCoTI-I order parameter as a result of binding to trypsin. Residues with $S_f^2-S_b^2>0.2$, where S_f^2 and S_b^2 are the order parameters of the free and trypsin-bound MCoTI-I, respectively, are depicted in red. MCoTI-I residues that were broadened in $\{^{15}N,^1H\}$ HSQC because of binding to trypsin are shown in green. The structure of free MCoTI/II (PDB code: 1IB9)[6] was used to illustrate the changes of MCoTI-I dynamics arising from trypsin binding.

the residues located in loop 1 (Cys2, Lys4, Ile5, Arg8), loop 3 (Cys15 and Ala18), and loop 6 (Val1) (Figures 1c and 2b and Table S2, Supporting Information). NMR signals of Cys2, Ile5, Cys19, Ser29, Ser31, Asp32, and Gly33 were significantly broadened, presumably because of intramolecular chemical exchanges in the trypsin–MCoTI-I complex. Arg8, Ala18, and Gly23 were excluded from backbone dynamics analysis because their peaks were broadened in the $^{15}N\{^1H\}$ NOE spectra. Similar findings have already been reported for other biomolecular interactions.[26]

We used these changes to construct the trypsin–MCoTI-I interaction surface. The binding surface is contiguous and spans 46% of the total molecular area of MCoTI-I (Figure 2b). As expected, the major difference in the backbone dynamics was observed in the binding loop (Table 1), where the mobility in the nano- to picosecond timescale was increased in MCoTI once bound to trypsin. Loop 1 showed $<S^2>=0.49\pm0.02$, which is much lower than the value for the rest of the molecule ($<S^2>=0.65\pm0.07$). Several residues in loop 2 (Cys9, Arg10, Ser13, and Asp14), loop 3 (Gly17), and loop 5 (Cys21 and Arg22) also showed significantly lower values of S^2 upon complex formation (Figures 1d and 2c). It is likely that the increase in mobility observed in these loops may help to accommodate the increased flexibility of the binding loop (Figure 2c).

Since our data clearly show that backbone flexibility of cyclotide MCoTI-I increases significantly upon binding to trypsin, we decided to estimate the contribution of these motions to the overall Gibbs free energy of binding (ΔG). The energetic benefit of this increase in backbone flexibility can be estimated from the experimental relaxation data, by using the experimentally measured order parameters S^2.[27] The estimated ΔG value was approximately $-62\,kJ\,mol^{-1}$ at 298 K. This value is almost identical to the calculated value from the trypsin inhibitory constant of MCoTI-I ($K_i\approx20$ pM,[28] $\Delta G\approx-61\,kJ\,mol^{-1}$). The calculated entropic contribution ($-T\Delta S$) at the same temperature was approximately $-46\,kJ\,mol^{-1}$. These results highlight the importance of the backbone entropic term to the formation of the trypsin–MCoTI-I complex, although a more detailed thermodynamic analysis that also includes the side-chain motions may be required.

In summary, we have reported the backbone dynamics of the cyclotide MCoTI-I in the free state and complexed to its binding partner trypsin in solution. To our knowledge this is the first time the backbone dynamics of a natively folded cyclotide has been reported. This has been possible because of the use of modified protein splicing units for the heterologous expression of folded cyclotides using bacterial expression systems[17,18,29] to incorporate NMR-active nuclei such as ^{15}N. Our results on the backbone dynamics of free cyclotide MCoTI-I confirm that MCoTI-I adopts a well-folded and highly compact structure with an $<S^2>$ value of 0.83. This value is similar to those found in the regions of well-folded proteins with restricted backbone dynamics.

The results also indicate that the trypsin-binding loop (loop 1) has a smaller S^2 value than the average value for the whole molecule, thus indicating a higher mobility of this region in the pico- to nanosecond timescale. This region also showed significant conformational exchange motions in the micro- to millisecond timescale. Loop 6 also possesses a higher mobility in the pico- to nanosecond timescale than the averaged value for MCoTI-I, although no significant conformational exchange motions were detected in the micro- to millisecond timescale. This result is intriguing since this loop contains a potentially flexible Gly–Ser-rich sequence that is mostly absent among other linear trypsin squash inhibitors,

and therefore it was thought to be a highly flexible linker to allow cyclization. More surprising, however, was the fact that the backbone of MCoTI-I, and especially loop 1, increased the pico- to nanosecond mobility when bound to trypsin. This interesting result has already been observed in other high-affinity protein–protein interactions.[30,31]

The thermodynamic analysis of the backbone contribution to the formation of the trypsin–MCoTI-I complex by using measured S^2 values also revealed the importance of the backbone entropic term in the formation of the complex. Similar findings have also been found in other protease inhibitors.[32] This increment in backbone mobility may help to minimize the entropic penalties required for binding. Hence, we also observed in the HSQC spectrum of the trypsin–MCoTI-I complex the appearance of a signal corresponding to the ε-NH_3^+ of Lys4 located in loop 1, which suggests that the ammonium group is protected and more rigid when forming the complex (Figure 3). This is the only

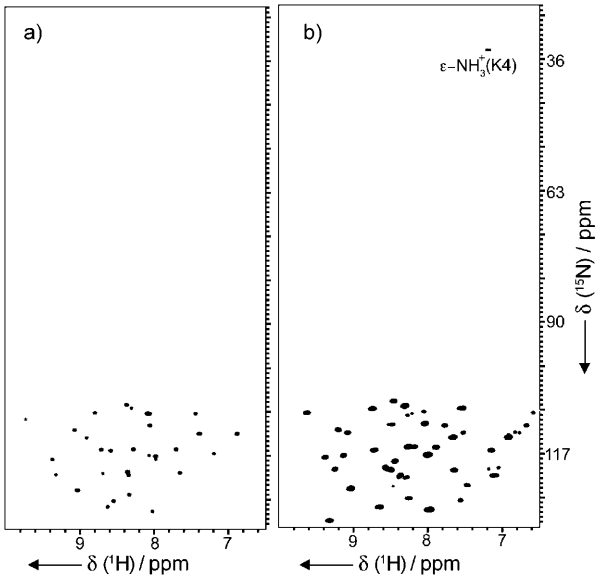


Figure 3. ε-NH_3^+ of Lys4 is protected from fast exchange with the solvent in trypsin-bound MCoTI-I. {^{15}N,^{1}H}-HSQC spectra of free (a) and trypsin-bound MCoTI-I (b) were collected at room temperature with the ^{15}N-carrier position at 82 ppm and ^{15}N radio-frequency field strengths of 5.2 kHz for 90° and 180° pulses and 1.2 kHz for composite decoupling during acquisition.

Lys residue present in the sequence of MCoTI-I (Scheme 1) and therefore it can be unambiguously assigned. This residue is key for binding to trypsin[29] and is responsible for binding to the specificity pocket of trypsin. This cross-peak was totally absent in the free MCoTI-I sample, which indicates that the ε-NH_3^+ of Lys4 is less rigid and rapidly exchanging with solvent (Figure 3a).

Similarly, the broadening of aliphatic resonances for Arg side chains with essentially rigid guanidinium groups (that is, $^{\varepsilon}$N–H bond vectors) has also been described for protein–peptide complexes.[26] Palmer and co-workers have recently suggested that this dynamic decoupling between the side-chain terminus from the rest of the aliphatic part of the side chain may be a general biophysical strategy for maximizing residual side-chain and potentially backbone conformational entropy in proteins and their complexes,[33] which is in agreement with our observations regarding the increase in MCoTI-I backbone mobility upon complex formation.

We have also mapped the binding surface of MCoTI-I once bound to trypsin. Major changes in chemical shifts were observed for the solvent-exposed residues located in loops 1 (Lys4, Ile5, Arg8), 3 (Ala18), and 6 (Val1) (Figures 1c and 2b). In agreement with these results, we have recently shown that the introduction of nonconservative mutations in these positions has a negative effect on the affinity for trypsin,[29] thus indicating that they may be in close contact with the protease at the binding interface of the molecular complex.

Cyclotides present several characteristics that make them appear as promising leads or frameworks for peptide drug design.[7,8] Investigation of the backbone dynamics is crucial for a better understanding of the dynamic structural properties of the cyclotide scaffold and how it affects the mode of binding of these interesting molecules. The reported data will help in the design of cyclotide-based libraries for molecular screening and the selection of de novo sequences with new biological activities, or the development of grafted analogues for use as peptide-based drugs.[9,10]

Received: May 14, 2010
Published online: August 16, 2010

Keywords: cyclotides · molecular dynamics · NMR spectroscopy · protein–protein interactions · structural biology

[1] N. L. Daly, K. J. Rosengren, D. J. Craik, *Adv. Drug Delivery Rev.* **2009**, *61*, 918.
[2] D. J. Craik, S. Simonsen, N. L. Daly, *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 251.
[3] D. J. Craik, M. Cemazar, C. K. Wang, N. L. Daly, *Biopolymers* **2006**, *84*, 250.
[4] J. F. Hernandez, J. Gagnon, L. Chiche, T. M. Nguyen, J. P. Andrieu, A. Heitz, T. Trinh Hong, T. T. Pham, D. Le Nguyen, *Biochemistry* **2000**, *39*, 5722.
[5] A. Heitz, J. F. Hernandez, J. Gagnon, T. Trinh Hong, T. T. Pham, T. M. Nguyen, D. Le-Nguyen, L. Chiche, *Biochemistry* **2001**, *40*, 7973.
[6] M. E. Felizmenio-Quimio, N. L. Daly, D. J. Craik, *J. Biol. Chem.* **2001**, *276*, 22875.
[7] R. J. Clark, N. L. Daly, D. J. Craik, *Biochem. J.* **2006**, *394*, 85.
[8] D. J. Craik, M. Cemazar, N. L. Daly, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 251.
[9] D. J. Craik, N. L. Daly, J. Mulvenna, M. R. Plan, M. Trabi, *Curr. Protein Pept. Sci.* **2004**, *5*, 297.
[10] P. Thongyoo, N. Roque-Rosell, R. J. Leatherbarrow, E. W. Tate, *Org. Biomol. Chem.* **2008**, *6*, 1462.
[11] V. A. Jarymowycz, M. J. Stone, *Chem. Rev.* **2006**, *106*, 1624.
[12] H. Sancheti, J. A. Camarero, *Adv. Drug Delivery Rev.* **2009**, *61*, 908.
[13] D. C. Ireland, M. L. Colgrave, N. L. Daly, D. J. Craik, *Adv. Exp. Med. Biol.* **2009**, *611*, 477.
[14] J. A. Camarero, T. W. Muir, *Chem. Commun.* **1997**, 1369.
[15] J. A. Camarero, J. Pavel, T. W. Muir, *Angew. Chem.* **1998**, *110*, 361; *Angew. Chem. Int. Ed.* **1998**, *37*, 347.
[16] J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 5597.

[17] R. H. Kimura, A. T. Tran, J. A. Camarero, *Angew. Chem.* **2006**, *118*, 987; *Angew. Chem. Int. Ed.* **2006**, *45*, 973.
[18] J. A. Camarero, R. H. Kimura, Y. H. Woo, A. Shekhtman, J. Cantor, *ChemBioChem* **2007**, *8*, 1363.
[19] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4546.
[20] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4559.
[21] A. Horská, J. Horsky, R. G. S. Spencer, *J. Magn. Reson. Ser. A* **1994**, *110*, 82.
[22] J. Liu, Y. Gong, O. Prakash, L. Wen, I. Lee, J. K. Huang, R. Krishnamoorthi, *Protein Sci.* **1998**, *7*, 132.
[23] A. Heitz, O. Avrutina, D. Le-Nguyen, U. Diederichsen, J. F. Hernandez, J. Gracy, H. Kolmar, L. Chiche, *BMC Struct. Biol.* **2008**, *8*, 54.
[24] J. Liu, O. Prakash, M. Cai, Y. Gong, Y. Huang, L. Wen, J. J. Wen, J. K. Huang, R. Krishnamoorthi, *Biochemistry* **1996**, *35*, 1516.
[25] B. Szenthe, Z. Gaspari, A. Nagy, A. Perczel, L. Graf, *Biochemistry* **2004**, *43*, 3376.
[26] S. M. Pascal, T. Yamazaki, A. U. Singer, L. E. Kay, J. D. Forman-Kay, *Biochemistry* **1995**, *34*, 11353.
[27] A. G. Palmer III, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 129.
[28] O. Avrutina, H. U. Schmoldt, D. Gabrijelcic-Geiger, D. Le Nguyen, C. P. Sommerhoff, U. Diederichsen, H. Kolmar, *Biol. Chem.* **2005**, *386*, 1301.
[29] J. Austin, W. Wang, S. Puttamadappa, A. Shekhtman, J. A. Camarero, *ChemBioChem* **2009**, *10*, 2663.
[30] K. Dutta, H. Shi, E. R. Cruz-Chu, K. Kami, R. Ghose, *Biochemistry* **2004**, *43*, 8094.
[31] R. Ghose, A. Shekhtman, M. J. Goger, H. Ji, D. Cowburn, *Nat. Struct. Biol.* **2001**, *8*, 998.
[32] S. Arumugam, G. Gao, B. L. Patton, V. Semenchenko, K. Brew, S. R. Van Doren, *J. Mol. Biol.* **2003**, *327*, 719.
[33] N. Trbovic, J. H. Cho, R. Abel, R. A. Friesner, M. Rance, A. G. Palmer III, *J. Am. Chem. Soc.* **2009**, *131*, 615.

