Spider Silk

DOI: 10.1002/anie.201003795

pH-Dependent Dimerization and Salt-Dependent Stabilization of the N-terminal Domain of Spider Dragline Silk—Implications for Fiber Formation**

*Franz Hagn, Christopher Thamm, Thomas Scheibel, and Horst Kessler**

The formation of spider dragline silk is controlled by the relatively small C- and N-terminal domains of the spidroins.[1–6] The formidable and unrivaled mechanical tensile strength of spider silk fibers is a result of the carefully matched assembly of polyalanine (polyA) or poly(glycine-alanine) (polyGA) repeat sequences separated by GGX or GPGXX repeats, which are thought to confer elasticity to the thread.[4,5,7] The correct alignment of polyA/polyGA sequences to form microcrystalline structures is controlled by the pH value, salt concentration, and shear-force-induced partial unfolding of the disulfide-bridged dimeric C-terminal domain.[8] The N-terminal domain was also shown to be important for the pH-dependent assembly of fiber.[9] Here, we use NMR spectroscopy and light-scattering techniques to show that the N-terminal domain of the major ampullate spider silk from *Latrodectus hesperus* (black widow spider) is mainly monomeric at neutral pH, as found in the spinning gland. The slight tendency to dimerize disappears under high salt conditions, as found in the gland. However, the N-terminal domain will dimerize at the lower pH value found in the spinning duct. Hence, acidification mainly controls the assembly of the N terminus, which is important for the formation of silk fiber, while high ionic strength stabilizes the monomeric N-terminal structure. The crystal structure of the N-terminal domain shows a homodimer with an antiparallel orientation of the subunit. In addition to this picture, our NMR data provide further evidence for the regulation and functional role of this domain in forming elongated silk threads.

Spider dragline silk threads consist of two different proteins, with highly conserved N-terminal domains (see Figure 1 in the Supporting Information). We chose the N termini of the major ampullate silk proteins of *Latrodectus hesperus* for our structural analysis (called N1 and N2; see Table 1 in the Supporting Information). Consistent with an earlier report,[10] circular dichroism (CD) spectroscopy indicates a high α-helix content and reasonable thermal stability (see Figure 2 in the Supporting Information) for both constructs. N1 showed a slightly higher stability than N2. Therefore, an NMR spectroscopic characterization was conducted on N1.

By using a set of triple resonance experiments, 97% (128 out of 132) of all the amino acid residues could be assigned (Figure 1a) and, furthermore, an almost complete assignment of the amino acid side chains could be achieved. An evaluation of the secondary structure content in N1 clearly confirmed the presence of five α helices within the protein (Figure 1b). The previously reported dimeric structure of an N-terminal domain from the major ampullate spidroin from *Euprosthenops australis* (pdb code: 3lr2[9]) was first used to

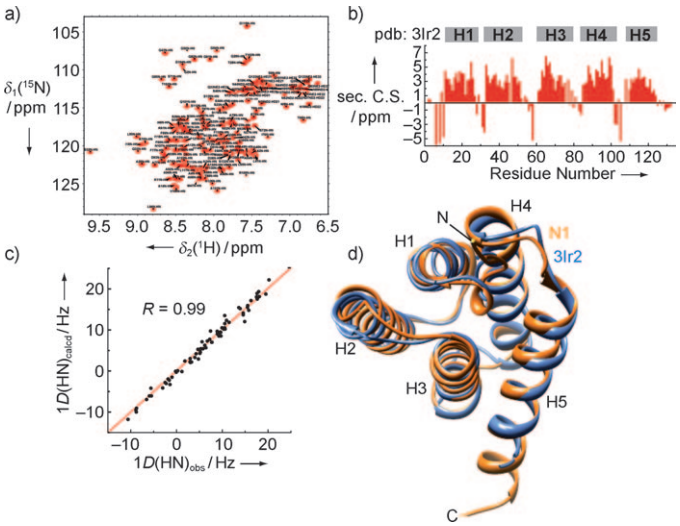

Figure 1. NMR analysis of N1. a) ^{1}H-^{15}N HSQC ananlysis of N1 with the assigned resonances labeled. b) Secondary chemical shifts of N1 indicate five α helices. c) Experimental HN-RDCs of N1 in 20 mg mL^{-1} Pf1 phage medium ($^{1}D(HN)_{obs}$) plotted against HN-RDCs back-calculated from the refined structure of N1 ($^{1}D(HN)_{calcd}$). d) Overlay between the N-terminal domain of *Euprosthenops australis* (pdb: 3Lr2, blue) and the refined structure of N1 (orange), with an rmsd of 0.8 Å for backbone atoms. Sec.C.S.: secondary C^{α},C^{β} chemical shift ($\Delta\delta(^{13}C^{\alpha})-\Delta\delta(^{13}C^{\beta})$).

[*] Dr. F. Hagn, Prof. Dr. H. Kessler
Technische Universität München, Institute for Advanced Study and Center for Integrated Protein Science
Lichtenbergstrasse 4, 85747 Garching (Germany)
Fax: (+49) 89-289-13210
E-mail: horst.kessler@ch.tum.de
Homepage: http://www.org.chemie.tu-muenchen.de
C. Thamm, Prof. Dr. T. Scheibel
Universität Bayreuth, Chair of Biomaterials
Fakultät für Angewandte Naturwissenschaften
Universitätsstrasse 30, 95440 Bayreuth (Germany)

[**] This work was supported by CIPSM and the Deutsche Forschungsgemeinschaft (to H.K.), the Elitenetzwerk Bayern, CompInt (to F.H.), and BMBF grant 13N9736 (to T.S.). We want to thank Lukas Eisoldt for providing plasmids of N1 and N2, and Dr. Martin Humenik for support with the SEC-MALS experiments.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201003795.

generate a monomeric homology model for the protein used in this study. The location of the secondary structure elements in this model agrees perfectly with the chemical shift data in the NMR spectra recorded in solution (gray bars in Figure 1b). Only the C-terminal helix 5 is four residues longer in N1 than in the model. The domain borders could additionally be confirmed by dynamics measurements in the ns to ps time scale as obtained by heteronuclear $\{^1H\}^{15}N$ NOE experiments (see Figure 3 in the Supporting Information).

For refinement of the structural model of N1, we used a combination of chemical shift derived backbone angles (obtained using TALOS[11]), a few NOE interactions (mostly H^N-H^N contacts), and HN residual dipolar couplings (HN-RDCs). The experimental HN-RDCs agreed well with back-calculated values (Figure 1c), and the refined structure showed a root-mean-square deviation (rmsd) of 0.8 Å to the reference structure (Figure 1d). These results clearly show that the structure of the N-terminal domain is conserved between different spider species.

To test the effect of solvent conditions on the structure of the N-terminal domain we first used far-UV CD spectroscopy to monitor changes in the secondary structure. No significant change in the spectra could be observed upon altering the salt concentration or the pH value (see Figure 4 in the Supporting Information). We additionally measured near-UV CD spectra of N1 (monitoring the environment of Trp 8) at two different pH values and salt conditions to obtain information on changes in the tertiary structure. As can be seen in Figure 2a, there are significant changes in the spectra upon lowering the pH value from 7.2 to 6.0. At pH 7.2, the pH value at which silk proteins are stored in the lumen of the spinning gland, the presence of salt leads to no significant changes in the spectrum, whereas at pH 6, which is present during fiber assembly, salt has a more pronounced effect, thus indicating that salt has an influencing role on the tertiary structure at this pH value.

The structural stability of N1 is higher during fiber assembly (pH ≈ 6) than under storage conditions (pH ≈ 7.2; Figure 2b). This effect is most likely caused by protonation of carboxylate moieties of either glutamic or aspartic acid side chains. A similar but rather more pronounced behavior can be observed on increasing the salt concentration from 0 to 800 mM at pH 7.2. An increase in the thermal stability of almost 20 °C was measured for N1 under high salt conditions (Figure 2c), and experiments on urea-induced unfolding indicates a more than doubled stability in the presence of 500 mM sodium chloride (Figure 2d). This effect of the pH value and salt on the stability of N1 seems to be caused by the local proximity of equal charges on the protein surface and reflects the destabilization of proteins by such clusters.[12] Indeed, in the N-terminal domain, clusters of negative and positive charges can be observed (Figure 3a) located at the interface between the two monomers in the dimeric structure.

Chemical shift perturbation (CSP) NMR experiments (see Figure 5 in the Supporting Information) enabled the effect of salt and pH to be characterized at a per-residue resolution. The addition of salt leads to large changes in the chemical shifts of resonances within the ^1H-^{15}N HSQC spectrum, particularly of residues located at the dimerization

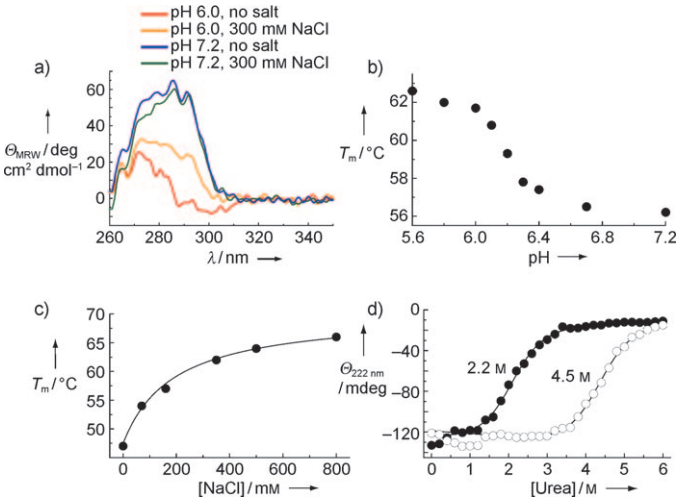

Figure 2. Effect of pH value and salt concentration on the structure and stability of N1. a) The near-UV CD spectra of 100 μM N1 at pH 6 and pH 7.2 show significant differences. b) and c) Thermal stability of the secondary structure of N1 at different pH values (100 mM NaCl) and increasing sodium chloride concentrations at pH 7.2. d) Urea-induced unfolding of N1 without salt (●) and in the presence of 500 mM sodium chloride (○).

site (Figure 3b). The same region is affected by lowering the pH value from 7.2 to pH 6.

Many residues were not visible at pH 6, most likely because of an exchange process in the ms to μs time scale as a result of a monomer–dimer equilibrium (Figure 3c). We additionally observed chemical shift perturbations within N1 at increasing protein concentrations, and these occurred exactly at the residues that the crystal structure[9] had shown to be at the dimerization site (Figure 4a). The line widths in the ^1H-^{15}N HSQC experiment are significantly smaller in the presence of sodium chloride (Figure 4b), which indicates a stabilization of the N1 monomer by salt at pH 7.2.

To address this issue we performed size-exclusion chromatography at pH 8.0, 7.2, and 6.0 as well as in the absence and presence of salt. At pH 6.0 N1 formed a stable dimer irrespective of the presence of salt. The monomer was clearly

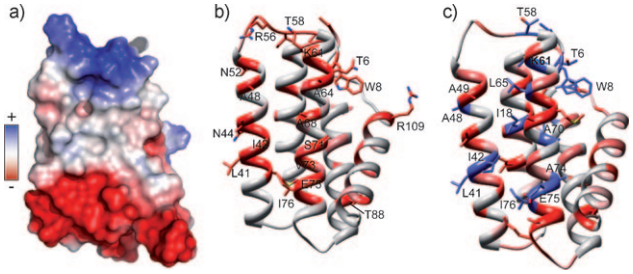

Figure 3. NMR spectroscopic characterization of the effects of pH value and salt concentration. a) The electrostatic potential of N1 shows there are clusters of positive and negative charges on the surface. b) The color intensity in the structure corresponds to CSPs upon the addition of 300 mM NaCl. c) CSPs upon a change in the pH value from 7.2 to 6.0. Resonances of the residues in blue disappeared during the pH titration.

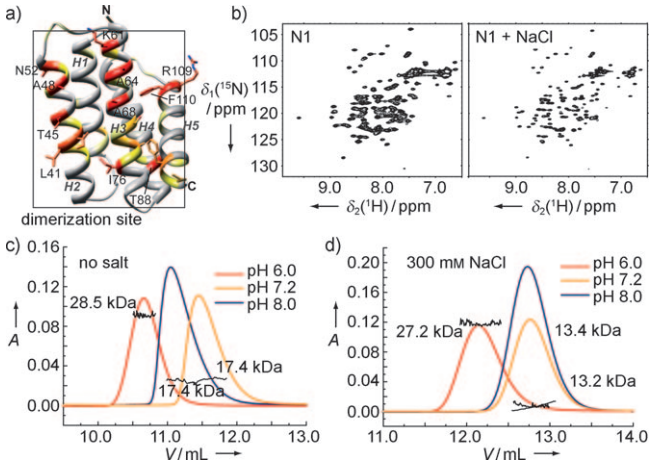

Figure 4. Monomer/dimer analysis of N1. a) CSPs of N1 at increasing protein concentrations. b) Line widths in the ^1H-^{15}N HSQC spectra of 400 μM N1 at pH 7.2 in the absence and presence of salt (300 mM NaCl) indicate a shift to the formation of the monomer by the addition of salt. c,d) Size-exclusion chromatograms of N1 at different pH values in the c) absence and d) presence of salt, including the calculated molecular weights by MALS.

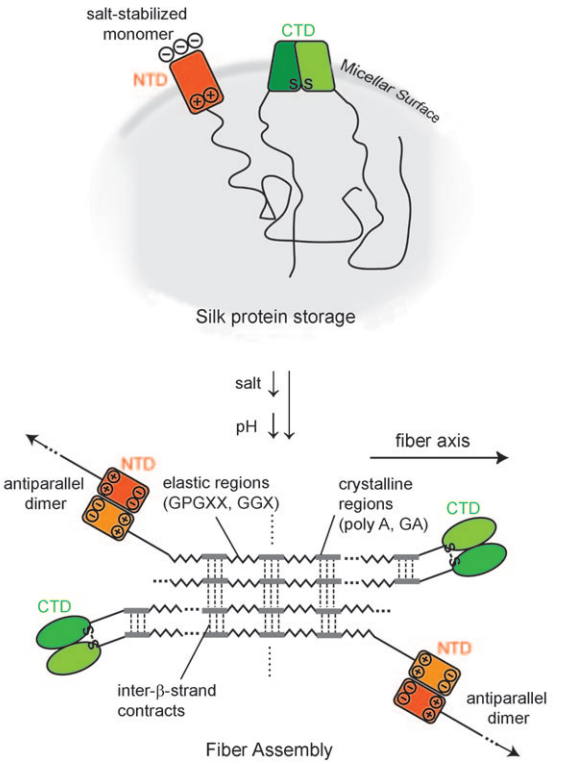

Figure 5. Mechanism of the initiation of fiber assembly, including the pH- and salt-dependent role of the N-terminal domain of spider dragline silk. During storage of silk protein at high protein concentrations both the N- and C-terminal domains are most likely located at the surface of the formed protein micelles. At fiber-forming conditions (lower pH value, less salt), the N-terminal domains are able to dimerize. The microcrystalline regions formed by the repetitive polyA/GA sequence elements[17] (gray boxes) are responsible for further noncovalent interactions between chains (dashed lines). The elastic regions are indicated by a zig-zag motive. CTD: C-terminal domain; NTD: N-terminal domain. Dotted lines indicate that the repetitive elements are much longer than shown in the picture. This model does not imply the real dimensions of the crystalline regions. The positions of the termini are chosen arbitrarily.

stabilized at neutral pH in the presence of salt, while in the absence of salt a slight tendency to dimerize could be detected, as seen by the asymmetric elution peak and the slightly increased mean molecular mass (MW of the monomer is 13.7 kDa; Figure 4c,d).

NMR diffusion measurements at pH 7.2 showed an increase in the diffusion coefficient of 16% on increasing the sodium chloride concentration from 0 to 300 mM (see Figure 6 in the Supporting Information), which is in agreement with the shifted dimer to monomer equilibrium induced by salt at pH 7.2. The charged clusters on the surface of N1 facilitate antiparallel dimerization, but render this dimeric state sensitive to the pH value and slightly to salt. Highly charged surface regions have been known to be of functional relevance for several decades.[13] Thus, it is not surprising that such a region is the key mediator of protein dimerization in the case of N1.

Together with our previous results on the C-terminal domain,[8] the previous data of the N-terminal domain of *E. australis* MA spidroin,[9] and the presented data of the N-terminal domain of *L. hesperus*, a detailed molecular picture of the initiation process of spider dragline fiber assembly is now evident (Figure 5). During storage of the silk protein at high protein concentrations, neutral pH, and high salt concentrations[14] in the spinning gland, the silk proteins form a supramolecular micelle-like structure[3,5,8,15]—an efficient storage form at concentrations of up to 40% (w/v)—in which the N-terminal domain is most likely present as a salt-stabilized monomer (as suggested by the NMR data at high protein concentrations; Figure 4b). During passage through the spinning duct the NaCl concentration is lowered[14,16] and shear forces are applied which lead to alignment of the protein chains parallel to the long axis of the fiber.[5,17] This process is induced by the C-terminal domain providing the correct alignment of the repetitive elements.[8] Both polyA and polyGA blocks are able to form β-sheet-rich microcrystallites.[17] The function of the N terminus in the initiation process of fiber assembly is to generate the salt- and pH-dependent interaction between already-existing supramolecular structures by dimerization with a further N-terminal domain. As reported recently, the N termini are able to form antiparallel dimers,[9] which is in agreement with the surface charges shown in this study (Figure 3a). Together with the crystalline regions, which consist of the repeat sequences, the solvent-dependent multivalent anchoring[18] of the N-terminal domains enforce controlled interaction between protein chains and chain elongation.

In summary, a model of the mechanism of how assembly of spider silk fiber is initiated has been established. Storage of spider silk proteins at high concentrations in aqueous solution (as found in the gland) is possible in (reversible) micelles in which the folded polar ends (N- and C-terminal domains) are located at the surface, whereas the unfolded repeat sequences are inside. During passage of this solution through the duct,

the chains align through shear forces. Changes in both the salt concentration and composition in the duct partially destabilize the C-terminal domains, thus allowing alignment of the associating repeat sequences to form the initial β-sheet-rich structures that potentially act as seeds for forming the final fibrillar structures. In addition, the lower pH value induces antiparallel dimerization of the N-terminal domains to yield head-to-tail dimers of the N-terminal domains, which result in a multivalent network connecting the microcrystalline β sheets. The presence of both terminal domains in one spider silk protein ensures the possibility to endlessly assembly these proteins into stable fibers. Hence, it is the careful balance of the solvent conditions as well as shear forces which are key for initiating the fiber assembly process of spider silk proteins.

Experimental Section

The genes encoding the N-terminal domains of *Latrodectus hesperus* major ampullate spidroins 1 and 2 (residues without the putative signal sequences) were synthesized (GeneArt, Regensburg, Germany), cloned into a pET28a expression vector (Novagen), and expressed in the *E. coli* strain BL21 (DE3) at 20 °C for 16 h. For labeling with NMR-active isotopes, 1 g [^{15}N]ammonium chloride and 2 g [^{13}C]glucose per liter of M9 medium were used. Protein purification was performed by nickel-NTA chromatography and size-exclusion chromatography.

CD measurements were made with a Jasco J-715 spectropolarimeter (Jasco, Gross-Umstadt, Germany). For far-UV spectra and thermal transitions (10 μM protein concentration) a 0.1 cm, for urea-induced unfolding (2 μM protein concentration) a 1 cm, path length cuvette was used. Near-UV CD spectra were measured at 100 μM protein concentration and in a 1 cm path length cuvette. The response was set to 2 s and the bandwidth to 5 nm.

Thermal transitions were measured by CD at 222 nm and a heating rate of 60 °C h^{-1}. Chemical unfolding was achieved by adding increasing amounts of urea until a concentration of 6 M was reached. Samples (2 μM protein) were incubated at 4 °C over night and the CD signal at 222 nm was measured. The data were evaluated with a two-state folding model as described previously.[19]

Size-exclusion chromatography was performed on an Agilent 1100 system equipped with a Superdex 75 10/300 GL column (GE Healthcare, München, Germany), with UV detection at 280 nm and a flow rate of 0.4 mL min^{-1}. 100 μL of a 1 mg mL^{-1} protein solution were injected and the column was equilibrated with the respective buffer before each run. The chromatography system was in line with a multiangle light-scattering (MALS) device, a quasielastic light scattering (QELS) device at 99°, and differential refractive index detection. DAWN EOS, WYATT QELS detectors (Wyatt Technology Europe, Dernbach, Germany), and an RI detector (Shodex RI71, Techlab, Erkerode, Germany) were connected in series. Data acquisition and processing were performed using Wyatt's ASTRA software program (5.3.4.14).

NMR spectroscopy was conducted on 600 and 900 MHz instruments (Bruker Biospin, Rheinstetten, Germany). Resonance assignment was achieved by conventional 3D-heteronuclear NMR experiments using a U-[^{13}C,^{15}N]-labeled sample of 800 μM N1 in 10 mM sodium phosphate at pH 7.2, 300 mM NaCl.[20] The assignment process was facilitated by the program PASTA.[21] NOESY spectra were performed as described.[8,22] HN-RDCs were measured with IPAP experiments[23] (in 10 mM sodium phosphate at pH 7.2, 300 mM NaCl) and protein alignment was induced by the addition of 20 mg mL^{-1} Pf1 phage medium[24] (Helios, Regensburg, Germany). Structure calculation and RDC refinement was done with Xplor-NIH[25] using standard scripts. For {^{1}H},^{15}N heteronuclear NOE measurements a 2 s proton irradiation time (and recycle delay) was used. CSP-NMR experiments were evaluated as the weighted mean of ^{1}H and ^{15}N chemical shift deviations as described previously.[26] Homology modeling was done with the program Swiss Model.[27] Stimulated echo NMR diffusion experiments were performed using a 350 ms diffusion time and two bipolar dephasing gradients of 2 ms duration. 32 data points were recorded at increasing gradient strengths ranging from 0.7 to 32.4 G cm^{-1}. The assignment of the NMR resonances has been deposited at the BMRB data bank under accession code 17131.

Received: June 21, 2010
Revised: August 20, 2010
Published online: November 9, 2010

Keywords: biopolymers · circular dichroism · NMR spectroscopy · protein folding · protein structure

[1] D. Motriuk-Smith, A. Smith, C. Y. Hayashi, R. V. Lewis, *Biomacromolecules* **2005**, *6*, 3152.
[2] A. Sponner, W. Vater, W. Rommerskirch, F. Vollrath, E. Unger, F. Grosse, K. Weisshart, *Biochem. Biophys. Res. Commun.* **2005**, *338*, 897.
[3] J. H. Exler, D. Hümmerich, T. Scheibel, *Angew. Chem.* **2007**, *119*, 3629; *Angew. Chem. Int. Ed.* **2007**, *46*, 3559.
[4] M. Heim, D. Keerl, T. Scheibel, *Angew. Chem.* **2009**, *121*, 3638; *Angew. Chem. Int. Ed.* **2009**, *48*, 3584.
[5] L. Eisoldt, J. G. Hardy, M. Heim, T. R. Scheibel, *J. Struct. Biol.* **2010**, *170*, 413.
[6] S. Rammensee, U. Slotta, T. Scheibel, A. R. Bausch, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6590.
[7] J. M. Gosline, P. A. Guerette, C. S. Ortlepp, K. N. Savage, *J. Exp. Biol.* **1999**, *202*, 3295.
[8] F. Hagn, L. Eisoldt, J. G. Hardy, C. Vendrely, M. Coles, T. Scheibel, H. Kessler, *Nature* **2010**, *465*, 239.
[9] G. Askarieh, M. Hedhammar, K. Nordling, A. Saenz, C. Casals, A. Rising, J. Johansson, S. D. Knight, *Nature* **2010**, *465*, 236.
[10] A. Rising, G. Hjalm, W. Engstrom, J. Johansson, *Biomacromolecules* **2006**, *7*, 3120.
[11] G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **1999**, *13*, 289.
[12] M. Akke, S. Forsen, *Proteins Struct. Funct. Genet.* **1990**, *8*, 23.
[13] E. D. Getzoff, J. A. Tainer, P. K. Weiner, P. A. Kollman, J. S. Richardson, D. C. Richardson, *Nature* **1983**, *306*, 287.
[14] E. K. Tillinghast, S. F. Chase, M. A. Townley, *J. Insect Physiol.* **1984**, *30*, 591.
[15] Z. Lin, W. Huang, J. Zhang, J. S. Fan, D. Yang, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8906.
[16] D. P. Knight, F. Vollrath, *Naturwissenschaften* **2001**, *88*, 179.
[17] J. D. van Beek, S. Hess, F. Vollrath, B. H. Meier, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10266.
[18] M. Mammen, S. K. Choi, G. M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908; *Angew. Chem. Int. Ed.* **1998**, *37*, 2754.
[19] P. L. Privalov, *Adv. Protein Chem.* **1979**, *33*, 167.
[20] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93.
[21] M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker, H. Kessler, *J. Biomol. NMR* **1998**, *11*, 31.
[22] T. Diercks, M. Coles, H. Kessler, *J. Biomol. NMR* **1999**, *15*, 177.
[23] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, *131*, 373.
[24] M. R. Hansen, L. Mueller, A. Pardi, *Nat. Struct. Biol.* **1998**, *5*, 1065.
[25] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65.
[26] F. Hagn, C. Klein, O. Demmer, N. Marchenko, A. Vaseva, U. M. Moll, H. Kessler, *J. Biol. Chem.* **2010**, *285*, 3439.
[27] N. Guex, M. C. Peitsch, *Electrophoresis* **1997**, *18*, 2714.

