CHEMBIOCHEM

Supporting Information

Propensity for *cis*-Proline Formation in Unfolded Proteins

T. Reid Alderson, Jung Ho Lee, Cyril Charlier, Jinfa Ying, and Ad Bax*[a]

cbic_201700548_sm_miscellaneous_information.pdf

Materials and Methods

Protein expression and purification

Codon-optimized human αS,[1] but lacking the post-translational amino-terminal acetylation, was expressed and purified from *E. coli* BL21(DE3) using standard methods. Briefly, αS expression was induced with 1 mM IPTG when the OD of 1 L of ^{13}C,^{15}N-M9 minimal medium at 37 °C reached 0.6-0.8 units. Protein expression proceeded for 3 hours, upon which cells were pelleted and stored at -80 °C until use. Cell pellets were resuspended in 15 mL of 20 mM Tris-HCl, 2 mM EDTA, 2 mM benzamidine, pH 7, and then placed in a water bath at 90 °C for 10 minutes. One protease inhibitor tablet (Roche) was included in the initial buffer, with only EDTA and benzamidine used in subsequent buffers. The resultant solution was centrifuged at 20,000 x g for 20 minutes, and the supernatant was loaded onto a HiTrap Q column equilibrated in the resuspension buffer. αS was eluted using a linear gradient of 1M NaCl from 7%-20%, with elution near 15% NaCl. All αS-containing fractions were pooled and then purified further by size exclusion chromatography and a Superdex S75 26/60 column (GE Healthcare) equilibrated in 20 mM NaH_2PO_4, pH 6, 2 mM EDTA, and 2 mM benzamidine as a broad spectrum protease inhibitor.

The C-terminal fragment of human αS (87-140), including an S87C mutation,[2] was purified from *E. coli* BL21(DE3) as above, except that 5 mM β-mercaptoethanol (BME) was included in all buffers to keep the Cys residue reduced.

The His_6-tagged α-crystallin domain (ACD) of human HSP27 (residues 84-171), including the C137S mutation, was expressed and purified from *E. coli* BL21(DE3) as described previously.[3] Protein expression was initiated upon the addition of 1 mM IPTG to ^{13}C,^{15}N-M9 minimal medium when the OD of the culture reached 0.6-0.8 units. Expression continued for 3 hours at 37 °C, and cells were then pelleted and stored at -80 °C until use. Cells were resuspended in 20 mM Tris-HCl, 2 mM benzamidine, 30 mM imidazole at pH 7, and lysed using a Cell Disruptor (Constant Systems LTD) with multiple passes at 2 kbar. The lysed cells were centrifuged at 20,000 x g for 20 minutes, and the supernatant was applied over a HisTrap column equilibrated in the resuspension buffer. Bound ACD was eluted from the column with 300 mM imidazole, and the ACD-containing fractions were digested with His_6-tagged TEV protease overnight at 4 °C in the resuspension buffer. The cleaved His_6 tag and His_6-tagged TEV protease were separated from the ACD by reverse Ni^{2+} affinity (HisTrap), and the ACD was subjected to size exclusion chromatography using a Superdex S75 26/60 column equilibrated in 30 mM NaH_2PO_4, pH 7, 2 mM EDTA, 2 mM benzamidine.

A plasmid encoding the amino acid sequence of human ubiquitin bearing the V17A and V26A mutations was purchased from ATUM and transformed into *E. coli* BL21(DE3) competent cells. The recombinant protein was expressed and purified, with uniform ^{2}H, ^{13}C, and ^{15}N labeling. Expression of ubiquitin was induced upon the addition of 1 mM IPTG to a culture of ^{2}H,^{13}C,^{15}N-M9 minimal medium containing 1 g of ^{2}H,^{13}C,^{15}N-ISOGRO (Sigma-Aldrich). After 3 hours of expression at 37 °C, cells were pelleted and lysed and centrifuged as above. The pH of the supernatant was then lowered to 3.3 using acetic acid, and this solution was centrifuged at 20,000 x g for 20 minutes at 4 °C. The resultant supernatant was subjected to size exclusion chromatography as above, except in 10 mM KPi at pH 6.8. Fractions that contained ubiquitin were additionally purified using reverse-phase HPLC with a Vydac 214TP C4 column. Ubiquitin V17A/V26A eluted near 35% acetonitrile with 0.1% trifluoroacetic acid, and this sample was dialyzed twice into four liters of 25 mM potassium phosphate buffer (pH 6.4), and concentrated.

Preparation of peptide solutions

Oligo-peptides used are listed in Table S1 and were kindly synthesized by Galina Abdoulaeva at the Center for Biologics Evaluation and Research (CBER) of the Food and Drug Administration (FDA). The N- and C-termini of αS peptides were blocked with acetylation and amidation, and the tri-peptides FPA, MPS, and APQ were synthesized with and without blocked termini. All peptides were dissolved at a concentration of 40-50 mM in 20 mM NaH_2PO_4 and the pH of the solution was adjusted to 6 using

concentrated NaOH. To match protein samples, the solution also contained 2 mM EDTA and 2 mM benzamidine.

NMR spectroscopy

All samples for NMR spectroscopy included 7% D_2O and NMR spectra were recorded at 288 K. ^{13}C,^{15}N-αS was concentrated to 0.9 mM in 20 mM NaH_2PO_4, pH 6, 2 mM EDTA, 2 mM benzamidine, with 290 µL loaded into a Shigemi micro NMR cell. A separate sample was prepared with 1.1 M NaCl, using a thick-walled, 2.2 mm ID NMR tube (New Era Enterprises, Inc.). Spectra were recorded on a Bruker Avance III 700 MHz spectrometer equipped with a cryogenically cooled probe.

^{13}C,^{15}N-ACD(C137S) was concentrated to 1.6 mM in 30 mM NaH_2PO_4, pH 7, 2 mM EDTA, 2 mM benzamidine. 280 µL of the sample was loaded into a high-pressure tube (Daedalus Innovations) and pressurized to 2.5 kbar. ^{2}H,^{13}C,^{15}N-ubiquitin(V17A/V26A) was concentrated to 0.7 mM in 25 mM potassium phosphate buffer at pH 6.4, loaded into a high-pressure tube, and pressurized to 2.5 kbar. Spectra were recorded on a Bruker Avance III 800 MHz spectrometer equipped with a cryogenically cooled probe.

3D HNCO and HNCA spectra were recorded to facilitate identification and assignment of the lowly populated *cis*-Pro conformations. When non-uniform sampling was used, the spectra were reconstructed using the SMILE program,[4] and all NMR spectra were processed and analyzed with NMRPipe[5] and Sparky.[6]

Chemical shift perturbations in Figure 3 were calculated using the following equation: $\Delta\delta_{eff} = \sqrt{(\Delta\delta H)^2 + (\Delta\delta N/5)^2}$.

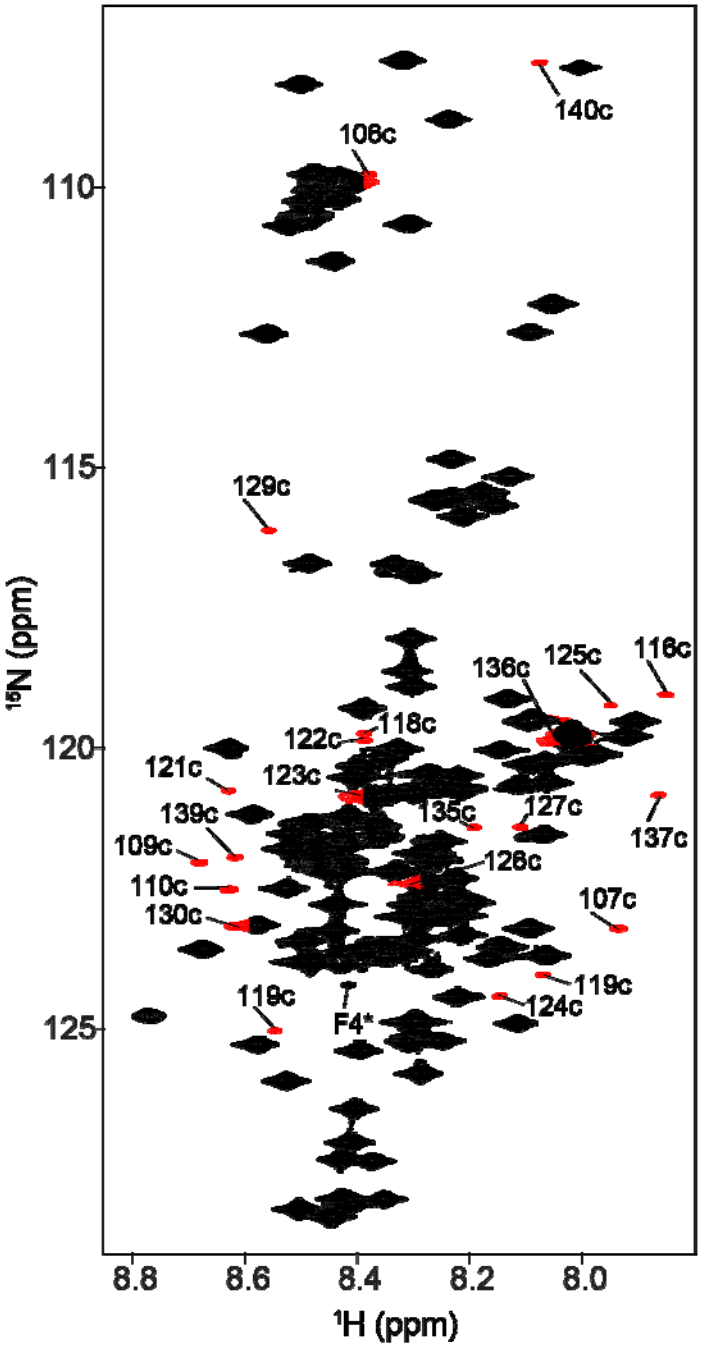

Figure S1. 2D ^{1}H-^{15}N HSQC NMR spectrum of αS, with resonance assignments marked for residues impacted by *cis*-Pro formation. The sample contained 0.9 mM αS in 20 mM NaH_2PO_4 at pH 6, 2 mM EDTA, 2 mM benzamidine, 288 K. Assignments for resonances affected by *cis*-Pro formation are colored *red* whereas resonances corresponding to *trans*-Pro are shown in *black*. For each of the five Pro residues, the fraction of *cis* was less than 5%, resulting in an effective NMR-observable concentration of ≤40 μM protein. A total of 27 out of 140 residues are visibly impacted by *cis*-Pro formation (including the Pro residues, which are not observable in the spectrum above). The peak labeled F4* corresponds to a ~3.5% fraction of protein that has been oxidized at M5.

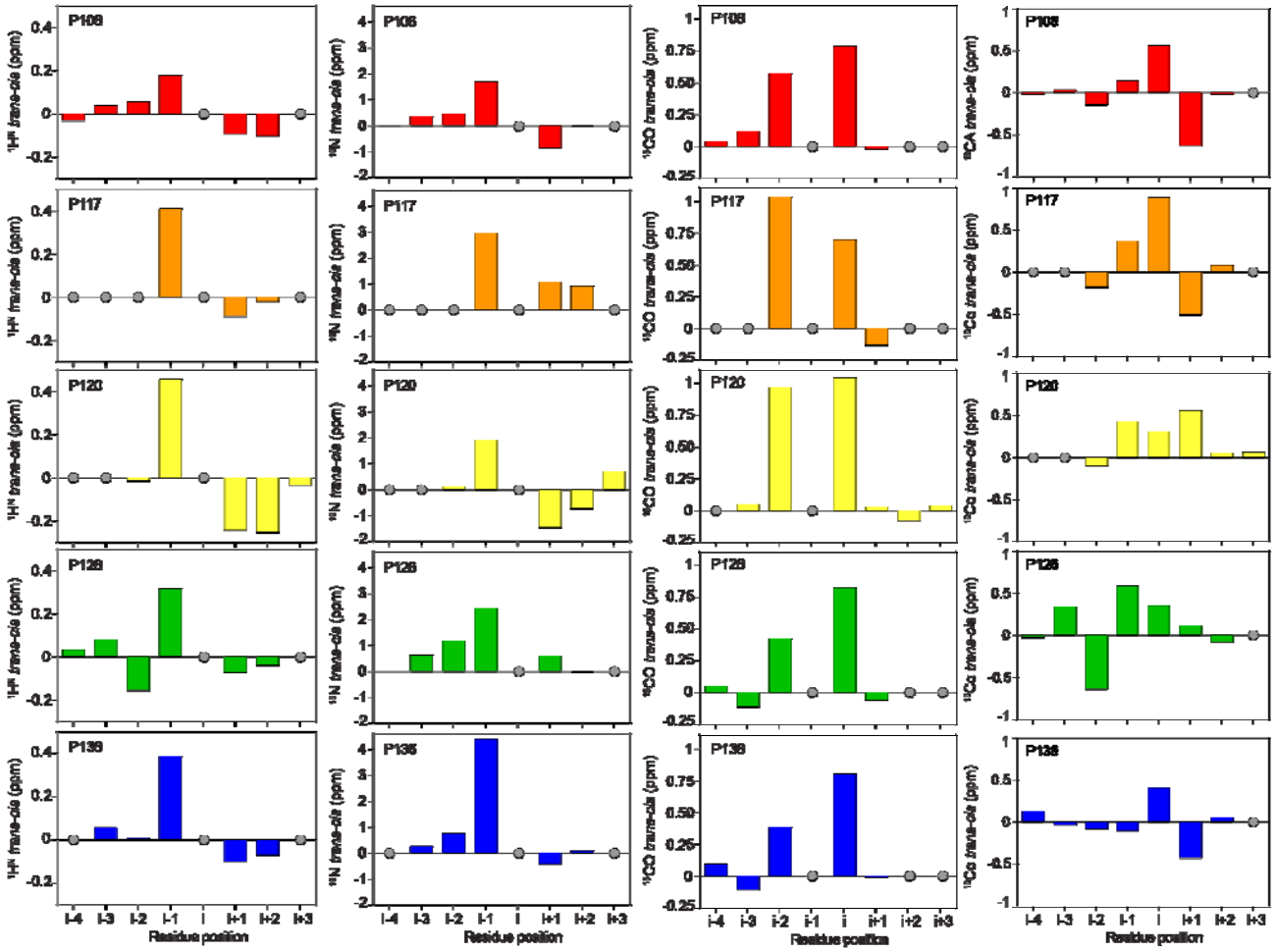

Figure S2. Chemical shift differences ($\delta_{trans} - \delta_{cis}$) between the *trans*- and *cis*-Pro-affected resonances in αS. Shown here are the differences for (i) $^{1}H^{N}$, (ii) ^{15}N, (iii) $^{13}C'$, and (iv) $^{13}C^{\alpha}$. Diagnostic chemical shift changes include $^{1}H^{N}$ and ^{15}N for the *i*-1 residue with respect to the isomerizing Pro, which show large positive values. Since these resonances are readily observed in 2D ^{1}H-^{15}N HSQC spectra, together with the constraint that the cis/trans intensity ratio must be the same for amides affected by a given Pro isomerization, can often be used for assigning the *cis*-isomer resonances to their nearby *trans* counterparts. $^{13}C'$ chemical shifts of *i* and *i*-2 residues show large upfield shifts for the *cis* isomer, and the same applies for the Pro $^{13}C^{\alpha}$.

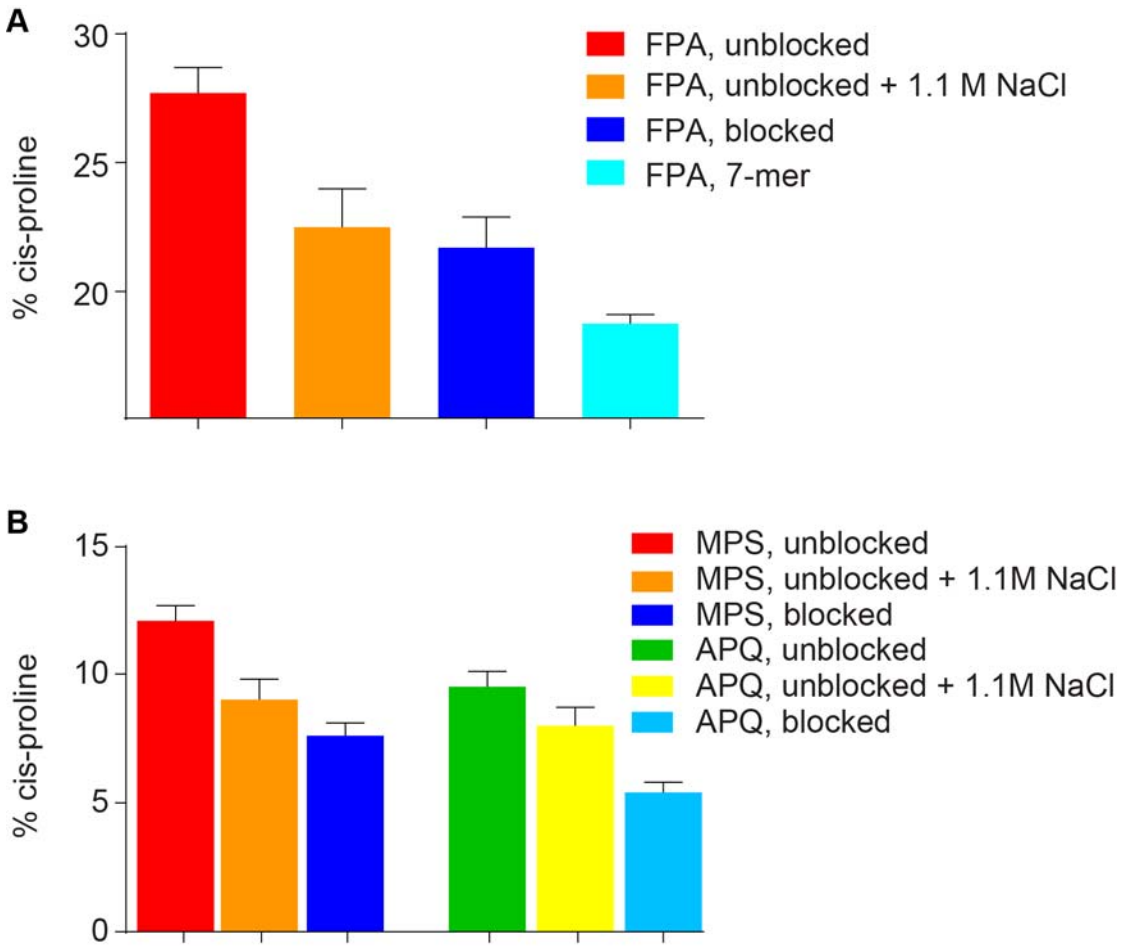

Figure S3. Impact of charged termini on *cis*-Pro formation. (A) Fraction of *cis*-Pro in a tri-peptide (FPA) without (*red*) or with (*blue*) acetylation and amidation of the N- and C-termini, respectively. Addition of 1.1 M NaCl to the unblocked FPA peptide (*orange*) significantly diminishes the fraction of *cis*-Pro, lowering it to levels essentially identical to the blocked peptide. Extending the peptide chain from a blocked tri-peptide to a blocked hepta-peptide (*cyan*) further decreases the fraction of *cis*-Pro. (B) Fraction of *cis*-Pro in αS tri-peptides MPS and APQ. The unblocked tri-peptides (*red*, *green*) show elevated *cis*-Pro fractions as compared to their blocked counterparts (*blue*, *cyan*). The addition of 1.1 M NaCl lowers the *cis*-Pro fraction of the unblocked MPS peptide to that of the blocked peptide.

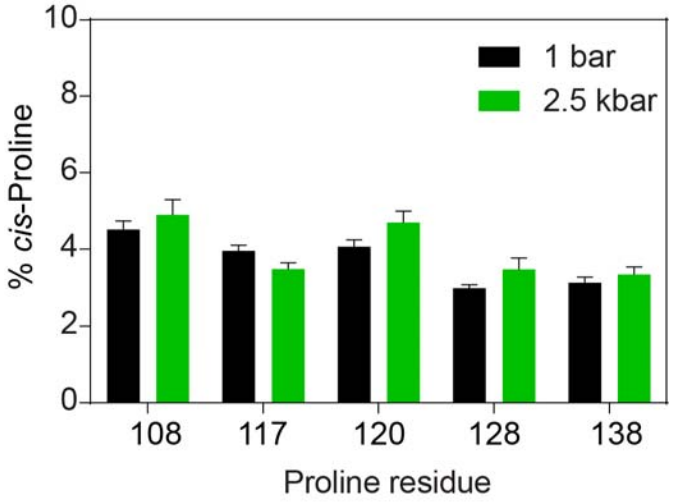

Figure S4. Impact of hydrostatic pressure on *cis*-Pro formation in αS. The fraction of *cis*-Pro was measured at 1 bar (*black*) and at 2.5 kbar (*green*). Agreement to within experimental uncertainty between the two datasets indicates that high pressures has no detectable impact on *cis*-Pro formation.

Table S1. Peptides used to quantify *cis*-Pro fractions. Ac and NH_2 refer to acetylation and amidation, respectively. The bolded proline numbers correspond to those in αS and are included for clarity.

Peptide sequence[a]	% *cis*-Pro
FPA	27.7 ± 1.1%
FPA + 1.1 M NaCl	22.5 ± 1.5%
Ac-FPA-NH_2	21.7 ± 1.2%
Ac-EGFPAEG-NH_2	18.7 ± 0.4%
P108: Ac-GAPQ-NH_2	5.6 ± 0.2%
GAPQ[b]	4.5 ± 0.2
Ac-GGPQ-NH_2	7.7 ± 0.2%
APQ	9.5 ± 0.7%
APQ + 1.1 M NaCl	8.0 ± 0.8%
Ac-APQ-NH_2	5.4 ± 0.4%
P117: Ac-DMPV-NH_2	7.6 ± 0.2%
DMPV[b]	3.9 ± 0.2%
P120: Ac-VDPD-NH_2	4.0 ± 0.1%
VDPD[b]	4.0 ± 0.2%
P128: Ac-EMPS-NH_2	7.1 ± 0.1%
EMPS[b]	3.0 ± 0.1%
MPS	12.1 ± 0.6%
MPS + 1.1 M NaCl	9.0 ± 0.8%
Ac-MPS-NH_2	7.6 ± 0.5%
Ac-YEMPSE-NH_2	4.5 ± 0.2%
Ac-AYEMPSEE-NH_2	4.1 ± 0.3%
P138: Ac-YEPE-NH_2	4.3 ± 0.3%
YEPE[b]	3.1 ± 0.2%
Ac-EAEPSDT-NH_2[c]	7.9 ± 0.4%

a. Unless otherwise indicated, all peptides were prepared in 20 mM sodium phosphate, 2 mM EDTA, 2 mM benzamidine, pH 6, and some contained 1.1M NaCl as noted.
b. Data obtained on full-length αS in the same buffer as above.
c. 25 mM potassium phosphate, pH 6.4. This peptide corresponds to P18 in human ubiquitin bearing the V17A mutation to match the protein used in this study (V17A/V26A ubiquitin).

Table S2. Amino acid sequences for the four proteins used in this study.

Human αS (Uniprot: P37840)

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAV
AQKTVEGAGSIAAATGFVKKDQLGKNEEGA**P**QEGILEDM**P**VD**P**DNEAYEM**P**SEEGYQDYE**P**EA

Human αS S87C-140

CIAAATGFVKKDQLGKNEEGA**P**QEGILEDM**P**VD**P**DNEAYEM**P**SEEGYQDYE**P**EA

cHSP27(C137S), the α-crystallin domain of human HSP27 (Uniprot: P04792)

GVSEIRHTADRWRVSLDVNHFA**P**DELTVKTKDGVVEITGKHEERQDEHGYISRSFTRKYTL**PP**GVD**P**TQVSSSLS**P**E
GTLTVEA**P**M**P**K

Human ubiquitin(V17A/V26A) (Uniprot: P0CG48)

MQIFVKTLTGKTITLEAE**P**SDTIENAKAKIQDKEGI**PP**DQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG

References

[1] M. Masuda, N. Dohmae, T. Nonaka, T. Oikawa, S. Hisanaga, M. Goedert, M. Hasegawa, *FEBS Lett.* **2006**, *580*, 1775–1779.
[2] Z. Jiang, F. Heinrich, R. P. McGlinchey, J. M. Gruschus, J. C. Lee, *J. Phys. Chem. Lett.* **2017**, *8*, 29–34.
[3] G. K. a. Hochberg, H. Ecroyd, C. Liu, D. Cox, D. Cascio, M. R. Sawaya, M. P. Collier, J. Stroud, J. A. Carver, A. J. Baldwin, et al., *Proc. Natl. Acad. Sci.* **2014**, *111*, E1562–E1570.
[4] J. Ying, F. Delaglio, D. A. Torchia, A. Bax, *J. Biomol. NMR* **2016**, DOI 10.1007/s10858-016-0072-7.
[5] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–93.
[6] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325–7.