Supplemental Information

Maturation of the Functional Mouse CRES Amyloid from Globular Form

Aveline Hewetson, Nazmul H. Khan, Matthew J. Dominguez, Hoa Quynh Do, R.E. Kusko, Collin G. Borcik, Daniel J. Rigden, Ronan M. Keegan, R. Bryan Sutton, Michael P. Latham, Benjamin J. Wylie, Gail A. Cornwall

SI Materials and Methods

Expression and purification of CRES. Tag-less CRES C48A (Cst8, Swiss-Pro P32766) was purified from the soluble fraction of bacteria using affinity, ion exchange, and gel filtration chromatography as previously described (1). The numbering of residues is based on the mature CRES protein. A tag-less CRES C48A protein with four alanine substitutions in the CRES loop (N94A, N101A, T102A, N104A) (loop mutant (LM)) was similarly purified from the soluble fraction of bacteria. A LM CRES C48A/pUC57 construct generated by Genewiz (South Plains, NJ) was cloned in-house into the pGEX-cs vector to allow for the production of a glutathione S-transferase fusion protein as done for wildtype CRES. The LM C48A CRES/pGEX-cs construct was verified by sequencing (Genewiz) to confirm alanine substitutions and in-frame cloning prior to transformation into Origami B *E. coli* for protein expression as described (1). Protein concentration was determined by using a Nanodrop Lite spectrophotometer (ThermoScientific, Rockford, IL) with an extinction coefficient of 16960 $M^{-1}cm^{-1}$ calculated based on CRES residues 20V-142V and a six residue N-terminal linker, GAMAHM.

MALDI. Sinapinic acid was used as matrix at a concentration of 50 mM in 50 % acetonitrile, 0.1 % TFA. 1 µL of protein sample (containing a mix of the two CRES forms) was mixed with 5 µL matrix solution, followed by spotting of 0.5 µL of mixture solution on the MALDI-plate and air-drying prior to analysis. The analysis was performed on MALDI-TOF/TOF 4800 mass spectrometer (Applied Biosystems). The mass spectra were acquired in the positive mode and a total of 1250 shots were accumulated per spectrum. The mass range was selected to be between *m/z* 5000-50,000 Da. SDS-PAGE purified CRES isoforms were subjected to tryptic digestion. Briefly, each sample was denatured at 90 °C for 10 min followed by adding 200 mM (1 µl) dithiothreitol (DTT) and incubating at 60°C for 45 min. The resulting reduced protein was then alkylated by adding 4µl iodoacetamide (IAA, 200 mM solution) and incubating at 37 °C in the dark

for 45 min. Excessive IAA was quenched by additional DTT and incubation at 37 °C for 30 min. Trypsin (Promega, Madison, WI) was added at a ratio of 1:25 (enzyme: protein, w/w) and incubated at 37 °C for 18 hours followed by addition of formic acid at a final concentration of 0.5% to quench the enzymatic reaction. Samples were centrifuged at 14,800 rpm for 10 min, and the resulting supernatant containing the digested peptides used for the MALDI analysis.

Crystallization and Structure Determination. CRES C48A in 25 mM MES, 250 mM NaCl, 1 mM EDTA, pH 6.0 was concentrated to 13.2 mg/ml using an Amicon Ultra-15 10K centrifugal filter (Millipore, Burlington, MA). Initial crystallization trials were performed using in house and commercially available screens to identify optimal growth conditions. CRES C48A crystals were grown at 10 °C in 2.1M $(NH_4)_2SO4$, 0.2M NaI, pH 5.8 using the hanging drop method. Specifically, 2 µl of protein + 2 µl mother liquor were spotted on a 22 x 22 mm plastic coverslip (ThermoScientific) inverted above 500 µl mother liquor in 24 well VDX Limbro culture plates (pre-greased, Hampton Research, 34 Journey, Aliso Viejo CA). After 6 months, crystals were captured into nylon loops, frozen in liquid nitrogen and shipped to SLAC (Stanford Linear Accelerator Center) beamline 7-1 for high resolution data collection. The wavelength of the final dataset was 1.195 Å, and the data were collected at 90 K. X-ray data were processed and scaled with HKL3000 (2). The X-ray crystal structure was initially solved in AMPLE as part of a CCP4 package (3). However, the final structure was solved using molecular replacement techniques (Phaser) (4) with 6ROA (5) as the search model and subsequently refined using Phenix (6) with manual building in Coot (7).

NMR sample preparation. Uniformly^{13}C-,^{15}N-labeled CRES C48A was expressed in *E. coli* grown in M9 medium containing ^{13}C-D-glucose (4 g/L) and ^{15}N-amonium chloride (2 g/L) (Cambridge Isotope Labs, Andover, MA) and supplemented with ^{13}C-,^{15}N labeled Cell Tone Base

powder (1 g/L) (Cambridge Isotope Labs, Andover, MA). CRES purification and amyloid assembly was performed as described previously (1).

Solution-state NMR. All solution state NMR data were recorded at 25 ºC on an 600 MHz (^{1}H Larmor frequency; 14 T) Agilent DD2 spectrometer equipped with a z-axis gradient room temperature HCN probe. Data were processed and analyzed with NMRPipe and associated programs and visualized with CCPN Analysis (8, 9). Backbone HN, N, C', Cα, and Cβ assignments were determined with a 1.1 mM uniformly ^{13}C,^{15}N-labeled CRES C48A sample in 25 mM MES, 250 mM NaCl, 1 mM EDTA, pH 6 buffer. Standard gradient-selected, sensitivity-enhanced triple resonance 3D HNCO, HNCACB, HN(CO)CACB, HNCA, and HN(CO)CA experiments were employed (10). We assigned ~99% of the ^{1}HN, ^{15}N, ^{13}C', ^{13}Cα, and ^{13}Cβ chemical shifts for the non-proline backbone atoms of residues A31 through V142. Secondary structure propensities and random coil index ^{15}N order parameters were calculated from the backbone assignments with TALOS-N (11) and 10,000 structures were generated, for comparison purposes, using the backbone chemical shifts and the CS-ROSETTA (12) webserver https://csrosetta.bmrb.wisc.edu/csrosetts/submit

To adjust pH for relaxation experiments, an equal volume of 50 mM HEPES, 100 mM NaCl, pH 8 was added to CRES C48A in pH 6 buffer with the final buffer consisting of 12.5 mM MES, 25 mM HEPES, 173 mM NaCl, and 0.5 mM EDTA, pH 7.5. To prepare CRES without NaCl, approximately 9 milligrams of CRES C48A were exchanged with 4 mM potassium phosphate buffer, pH 7.4 using an Amicon Ultra-15 10K centrifugal filter (Millipore, Burlington, MA). CRES C48A was then concentrated to 20 mg/mL using an Amicon Ultra-0.5 10K centrifugal filter.

For 2D ^{15}N,^{1}H HSQC and ^{15}N relaxation data, 1152 and 144 complex data points were collected over spectral widths of 8990 and 2000 Hz in the ^{1}H and ^{15}N dimensions, respectively. The carrier frequencies were placed on water and the center of the backbone amide region for ^{1}H

and ^{15}N, respectively. Amide chemical shift perturbations (CSPs), derived from peak positions in two-dimensional HSQC spectra, were calculated according to

$$CSPs = \sqrt{(\delta^N_{pH\ 6.5} - \delta^N_i)^2 / 25 + (\delta^H_{pH\ 6.5} - \delta^H_i)^2 / 2}$$

where δ^N and δ^H are the amide nitrogen and proton chemical shifts, respectively (13). CSPs were calculated for data collected in pH 6.5 / 250 mM NaCl buffer ($\delta_{pH\ 6.5}$) minus pH 7.5 / 173 mM NaCl or pH 7.5 / no salt (δ_i) buffers.

^{15}N longitudinal (R_1) and rotating-frame ($R_{1\rho}$) relaxation rates (14) were each calculated from six parametrically varied time points ranging from 10 – 500 msec (R_1) or 2 – 60 msec ($R_{1\rho}$). $R_{1\rho}$ values were converted to transverse (R_2) relaxation rates using the relation $R_{1\rho} = R_1 cos^2\theta + R_2 sin^2\theta$, where $\theta = \tan^{-1}(\nu_{SL} / \Omega)$ is the effective tilt angle of the rotating-frame field, ν_{SL} is the field strength of the rotating-frame field (~2000 Hz), and Ω is the offset of the peak from the ^{15}N carrier. Errors in the rates were calculated from the covariance matrix of the fit.

^{15}N CPMG relaxation dispersion profiles were recorded using the experiment described by Hansen *et al* with a 30 msec constant-time relaxation delay and 20 CPMG frequencies ranging between 33.3 and 1000 Hz (15). Two CPMG frequencies (66.6 and 866.6 Hz) were repeated for error analysis. Dispersion data were fitted to a two-site global exchange model by numerically solving the Bloch-McConnell (16, 17) equations using the program *chemex* (https://github.com/gbouvignies/ChemEx).

MAS solid-state NMR. The^{13}C-^{15}N-labeled CRES C48A amyloid sample was packed into a 3.2 mm pencil rotor (Agilent Technologies, Santa Clara, CA and Loveland, CO). SSNMR experiments were carried out on a 600 MHz Agilent DD2 three-channel spectrometer equipped with an HCN Balun probe (Agilent Technologies). Data were processed in NMRPipe and analyzed in CCPN Analysis. The magic-angle spinning (MAS) rate and the sample temperature were maintained at

13333 ± 2 Hz and at 0 ± 2 °C, respectively. ^{13}C chemical shifts were externally referenced using an adamantine sample. The downfield signal of adamantine was set to 40.48 ppm on the DSS scale. All SSNMR data were acquired using exponentially biased non-uniform sampling (NUS). Data was reconstructed using SMILE during processing in NMRPipe (18). All spectra were apodized with either 30 Hz of Gaussian line broadening or with a shifted sine-bell in all dimensions. The two-dimensional DARR spectra were acquired employing 1 ms of cross polarization (CP) ^{13}C and 73 kHz on ^{1}H. During CP a 15% tangent ramp was applied to ^{13}C. Data were acquired with 8 ms of indirect ^{13}C chemical shift evolution, 12, 25, and 50 ms of DARR mixing, and a 20 ms acquisition period in the direct dimension, all under high-powered ^{1}H decoupling (85 kHz ^{1}H field).

High-powered SPINAL-64 decoupling was also applied in all 3D spectra, including 7 ms of chemical shift evolution in each indirect dimension and 20 ms of direct acquisition. NCACX spectra were acquired with 1 ms ^{1}H to ^{15}N adiabatic cross polarization, employing a 15% tangent ramp on the ^{1}H channel and ^{1}H and ^{15}N fields of 60 kHz and 46 kHz respectively. Polarization is transferred from the backbone ^{15}N amide to ^{13}Cα with 6 ms of SPECIFIC CP with ^{15}N field set to 33 kHz and ^{13}C field set to 20 kHz. A 4% tangent ramp was set on the ^{13}C channel. Polarization was transferred from ^{13}Cα to neighboring ^{13}C nuclei via 25 ms of DARR mixing. NCO spectra were acquired in a similar fashion, but with 33 kHz on the ^{13}C' channel and 20 kHz on ^{15}N during SPECIFIC-CP. CANcoCA used the same CP matching conditions as the NCACX and NCOCX for the appropriate transfer, but the final polarization transfer step from ^{13}Cα to ^{13}C' was performed using radio frequency driven recoupling (RFDR).

ROSETTA Fold and Dock. SSNMR chemical shifts were submitted to the CS-ROSETTA server. The resulting monomeric fold exhibited the expected β-sheet secondary structure. The CRES model was generated by first selecting fragments using the CS-ROSETTA server and then

running ROSETTA Fold-and-Dock calculations using C2 symmetry (19). The lowest 10 out of 300 structures were selected. The structure in best agreement with the experimental chemical shifts was selected.

Negative stain TEM. CRES protein was spotted on to formvar/carbon coated 200 mesh nickel grids (Ted Pella, Redding, CA) as described (20).

Dynamic light scattering. Wildtype CRES and LM CRES freshly eluted of the gel filtration column were buffer exchanged into 4 mM potassium phosphate buffer pH 7.4 and concentrated to ~15 mg/ml using an Amicon Ultra-15 10 KDa filter followed by an Amicon Ultra-0.5 10 KDa filter with centrifugation at 2500 x g, 4 °C. Samples were immediately examined using a Zetasizer Nano ZS (ZEN3600, Malvern Instruments) equipped with a 633 nm red laser and 173° scattering angle as previously described (1). Samples were stored at 4 °C and examined over time.

Supplemental Text

Justification for conformational exchange of CRES to a larger oligomeric state. The overall rigidification of amide backbone (i.e., less internal dynamics) at pH 7.5 could lead to larger $R_{2,\,pH\,7.5}$ rates and the resulting negative deviation in the average ΔR_2; however, the fairly uniform change implies that CRES C48A tumbles more slowly at pH 7.5 than at pH 6. Numerical simulations suggest that an average increase in R_2 of 3.2 s^{-1} would arise from a ~4 - 5 ns larger global correlation time (for initial correlation times in the range of 6 – 24 ns, which is greater than the range of an expected correlation time for the 118 amino acid construct used here). At the low end of this correlation time range, a ~4 – 5 ns increase could correspond to at least doubling the size of the CRES unit observed at pH 6 / 250 mM NaCl.

We can place qualitative bounds on the populations and R_2 values within our model of fast exchange on the R_2 timescale between the observable, low molecular weight form of CRES (i.e., monomers) and an unobservable, high molecular weight amyloidogenic form. For example, assuming a large average $R_2^{oligomer}$ (50 s^{-1}) for the oligomeric state, such that peaks from this state are not detectable, a population of the oligomeric form (1 - $p_{monomer}$) of 10% would be present $\left(R_2^{obs} - R_2^{oligomer}\right) / \left(R_2^{monomer} - R_2^{oligomer}\right) = p_{monomer}$, where $R_2^{monomer}$ = 18.5 s^{-1} which is the average R_2 in the pH 6 / 250 mM salt sample). If the average $R_2^{oligomer}$ increases to 75 s^{-1}, then the population of the oligomeric form would be 6%. As outlined in the main text and below, it is unlikely that the millisecond timescale exchange observed in the CPMG relaxation dispersion data is the same conformation exchange process as that observed with the ΔR_2 values. However, if it were, the 1.2% minor state population detected in the CPMG experiment would result in an average $R_2^{oligomer}$ of 287 s^{-1}.

The deviation in R_2 observed when switching the buffer to pH 7.5 / no salt can also be interpreted as an increase in CRES tumbling, resulting from the stable or transient association of CRES protomers. However, numerical simulations suggest that the 9 s^{-1} increase in R_2 upon shifting the buffer to pH 7.5 and no salt would result in a 13 – 15 ns increase in the global tumbling time (for a range of initial tumbling times 6 – 15 ns). As the line widths of our spectra, which are directly proportional R_2, are not substantially greater at pH 7.5 / no salt (Figure 3A), the data suggest that CRES we observe is not tumbling more slowly in solution. Thus, monomeric CRES is rapidly exchanging with a larger oligomeric form and that this exchange is greater under the no salt condition. Again, we can use our exchange framework to place qualitative bounds on R_2 and populations. With an average $R_2^{oligomer}$ of 50 s^{-1}, the population of the oligomeric form would be 29%, whereas an average $R_2^{oligomer}$ of 75 s^{-1} gives a population of 16%. Finally, using the 3.1% population calculated from the CPMG data gives an average $R_2^{oligomer}$ of 312 s^{-1}. It is tempting to see the difference between the CPMG derived average $R_2^{oligomer}$ values at the two pH 7.5 conditions (287 s^{-1} *vs* 312 s^{-1} for 173 mM and no salt, respectively) as similar and indicating a common exchange process observed for the ΔR_2 and CPMG. Yet, the 25 s^{-1} difference in predicted R_2 is a significant difference equivalent to adding at least 2 – 4 CRES monomers to the oligomer; thus, the exchange processes (i.e., ΔR_2 and CPMG) are in fact different.

Justification for CRES amyloid containing antiparallel β-sheets. We know from the X-ray-derived coordinates that the β-sheets found in the globular state of CRES are antiparallel and that the solution-state NMR data are highly suggestive of a similar fold in solution. We observed few significant chemical shift perturbations between the solution-state and solid-state NMR data for a majority of the residues in β-strand regions once the structure was shifted toward amyloid. This suggests the local structure, including secondary structure and relative packing does not change significantly. Moreover, it is likely that β4 and β5 retain their antiparallel registry as the disulfide bond between these strands is preserved. Using the SSNMR chemical shift data, we found that backbone dihedral angles predicted by TALOS-N are closer to the idealized antiparallel dihedral angles (ϕ,ψ) which deviate from (ϕ,ψ) angles observed in parallel β-sheet structures. However, TALOS-N is naturally biased towards antiparallel β-sheets because these structures dominate the β-sheet structures found in the TALOS-N structural database. Our measured chemical shifts nevertheless contain clear clues to a largely antiparallel arrangement of β-strands. It was established that the isotropic chemical shift (δ_{iso}) of C' resonances is dominated by the geometries (i.e., length and angle) of its participating hydrogen bonds (21-24). Thus, the average C' chemical shift in an α-helix often is different from a β-sheet. It stands to reason that the differences in hydrogen bonding and registry between antiparallel and parallel β-sheets should also be reflected in the observed C' chemical shifts (Figure 6B). In fact, antiparallel β -sheets should have a slightly higher ^{13}C' chemical shift and narrower distribution (i.e., standard deviation) compared to parallel β-sheets because their characteristic hydrogen bond lengths are slightly shorter, their bonding angles are more linear, and their overall registry is more ordered. In addition, because the hydrogen bonding environment within parallel β-sheets is less regular, there should be a greater deviation within a population of parallel β-sheet C' chemical shifts. We pooled the C' chemical shifts from our CRES SSNMR chemical shift assignments for regions predicted by TALOS-N to be β-sheets. We then clustered C' chemical

shifts from deposited SSNMR structures of fibrillar proteins forming parallel intermolecular β-sheets. The fibrillar structures that we selected were α-synuclein, low-complexity domain of fused in sarcoma (FUS-LC), Aβ40, Aβ42, and HET-s (25-29). As depicted in Fig S12, all of the parallel β-sheet structures exhibit a greater standard deviation of the mean compared to CRES β-sheet C' resonances. In addition, all of the average C' chemical shifts from the parallel β-sheet structures are shifted slightly upfield from the CRES average. We then compared our CRES C' chemical shifts to well-established SSNMR-assigned proteins containing antiparallel β-sheets (Figure S13). In each case, including GB1, ubiquitin, and Crh, the C' chemical shifts exhibited a smaller deviation, a similar mean, and generally lay slightly downfield from C' resonances in the selected parallel amyloids (30-34). Thus, the statistical comparison of the average and standard deviation of ^{13}C' resonances found in CRES compared to a group of well-resolved parallel β-sheet amyloid structures and well-established antiparallel globular proteins suggest the majority of β -sheets found in the CRES amyloid are antiparallel.

Supplemental Tables

Table S1. X-ray crystallographic data collection and refinement statistics.	
CRES (Cystatin-Related Epididymal Spermatogenic)	
PDB ID	6UIO
Wavelength (Å)	1.195
Resolution Range (Å)	28.39 - 1.83 (1.895 - 1.83)
Space group	P 2_1 2_1 2_1
Unit cell (Å)	51.048, 85.924, 111.677, 90°, 90°, 90°
Total reflections	159751 (16251)
Unique reflections	39293 (4045)
Multiplicity	4.1 (4.0)
Completeness (%)	89.06 (93.68)
Mean I/σ(I)	17.22 (2.25)
Wilson B-factor (Å^2)	34.22
R-merge	0.0822 (0.8354)
R-meas	0.09578 (0.9626)
R-pim	0.04798 (0.4687)
CC1/2	0.996 (0.647)
CC*	0.999 (0.886)
Refinement Statistics	
Reflections used in refinement	39262 (4045)
Reflections used for R-free	3930 (405)
R-work	0.2243 (0.3187)
R-free	0.2623 (0.3465)
CC(work)	0.929 (0.461)
Number of non-hydrogen atoms	3947
Macromolecules	3664
Ligands	154
Solvent	129
Protein residues	444
RMS(bonds, Å)	0.016
RMS(angles, °)	2.06
Ramachandran favored (%)	97.94
Ramachandran allowed (%)	2.06
Rotamer outliers (%)	0
Rotamer outliers (%)	5.77
Clashscore	9.59
Average B-factor (Å^2)	45.98
Macromolecules (Å^2)	44.48
Ligands (Å^2)	83.20
Solvent (Å^2)	44.07
Statistics for the highest-resolution shell are shown in parentheses.	

Table S2. H-bond and salt bridge distances between CRES loop in chain A and β-strand 5 in chain B

Chain A	Chain B	Length (Å)	Type
Asn-94 [HD21]	Glu-59 [OE1]	2.1	H-bond
Lys-96 [H]	Lys-137 [O]	2.3	H-bond
Asn-101 [H]	Cys-139 [O]	2.4	H-bond
Gln-108 [HE22]	Leu-135 [O]	2.0	H-bond
Asn-94 [O]	Lys-137 [H]	2.1	H-bond
Asn-101 [OD1]	Asp-141 [H]	2.3	H-bond
Asn-104 [O]	Lys-137 [HZ3]	2.0	H-bond
Asn-104 [OD1]	Glu-61 [H]	2.3	H-bond
Lys-96 [NZ]	Glu-138 [OE1]	3.5	Salt Bridge
Lys-109 [NZ]	Glu-55 [OE2]	2.7	Salt Bridge

Table S3. ^{15}N chemical shift perturbations between pH 6 / 250 mM salt and pH 7.5 / 173 mM salt or pH 7.5 / no salt

Residue	pH 7.5 / 173 mM salt	pH 7.5 / no salt
F35	0.04	
I38	0.11	
N39	0.06	
I40	0.00	
N44		0.03
V45	0.01	0.04
K46	0.02	0.03
Q47	0.08	0.03
A48	0.01	0.03
V49	0.01	0.04
W50	0.02	0.06
F51	0.02	0.04
A52	0.03	0.01
M53	0.02	0.06
K54	0.02	0.04
E55	0.01	0.03
Y56	0.03	0.05
N57	0.00	0.05
K58	0.00	0.05
E59	0.04	0.07
S60	0.01	0.05
E61	0.03	0.04
D62	0.01	0.03
K63	0.03	0.05
Y64	0.00	0.02
V65	0.00	0.05
F66	0.01	0.03
L67	0.01	0.02
V68	0.01	0.06
D69	0.00	0.03
K70	0.02	0.02
I71	0.01	0.05
L72	0.07	0.06
H73	0.16	0.17
A74	0.02	0.02
K75	0.04	0.02

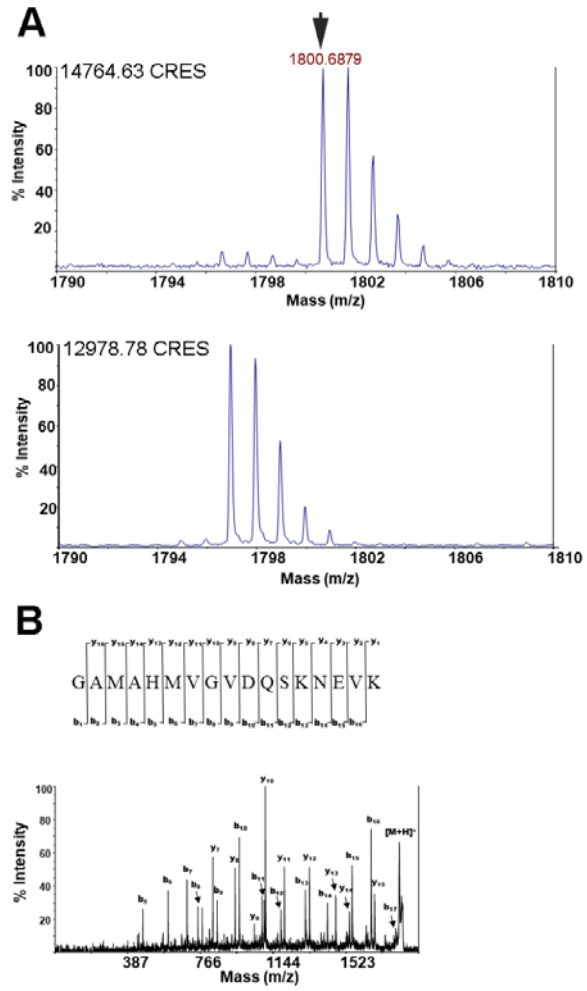
L76	0.05	0.07
Q77	0.02	0.02
I78	0.01	0.02
T79	0.02	0.02
D80	0.01	0.05
R81	0.01	0.01
M82	0.01	0.05
E83	0.01	0.04
Y84	0.01	0.03
Q85	0.05	0.03
I86	0.04	0.06
D87	0.08	0.10
V88	0.06	0.07
Q89	0.02	0.05
I90	0.01	0.03
S91	0.00	0.04
R92	0.00	0.04
S93	0.00	0.04
N94	0.00	0.04
C95	0.01	0.04
K96	0.03	0.02
K97	0.00	0.03
T102		0.05
E103	0.00	0.00
C105	0.00	0.01
I106	0.01	0.03
Q108	0.02	0.02
K109	0.05	0.09
K110	0.01	0.02
E112	0.00	0.00
L113	0.01	0.03
E114	0.02	0.04
K115	0.00	0.04
K116	0.00	0.02
M117	0.01	0.06
S118	0.01	0.02
C119	0.01	0.05
S120	0.02	0.06
F121	0.01	0.04
L122	0.01	0.03

V123	0.01	0.02
G124	0.01	0.03
A125	0.02	0.05
L126	0.01	0.04
W128	0.01	0.09
N129	0.01	0.07
G130	0.01	0.04
E131	0.01	0.02
F132	0.02	0.08
N133	0.01	0.05
L134	0.02	0.01
L135	0.01	0.05
S136	0.01	0.03
K137	0.01	0.02
E138	0.02	0.05
C139	0.08	0.11
K140	0.01	0.02
D141	0.02	0.03
V142	0.01	

Table S4. Absolute value of ^{15}N chemical shift differences between ground and excited states extracted from CPMG relaxation dispersion experiments

Residue	pH 7.5 / 173 mM salt	pH 7.5 / no salt
Q77	2.34 ± 0.38	1.91 ± 0.36
T79	1.88 ± 0.31	8.75 ± 0.33
D80	1.38 ± 0.48	1.05 ± 0.32
R81	2.09 ± 0.29	1.59 ± 0.32
M82	1.81 ± 0.30	1.80 ± 0.46
A125	4.96 ± 1.16	2.22 ± 0.54
L126	4.37 ± 0.87	1.76 ± 0.44
W128		1.17 ± 0.22
N129	2.84 ± 0.71	1.44 ± 0.54
G130		1.28 ± 0.33
E131	1.31 ± 0.21	1.11 ± 0.25
F132	2.40 ± 0.28	6.68 ± 0.17

Supplemental Figures and Legends

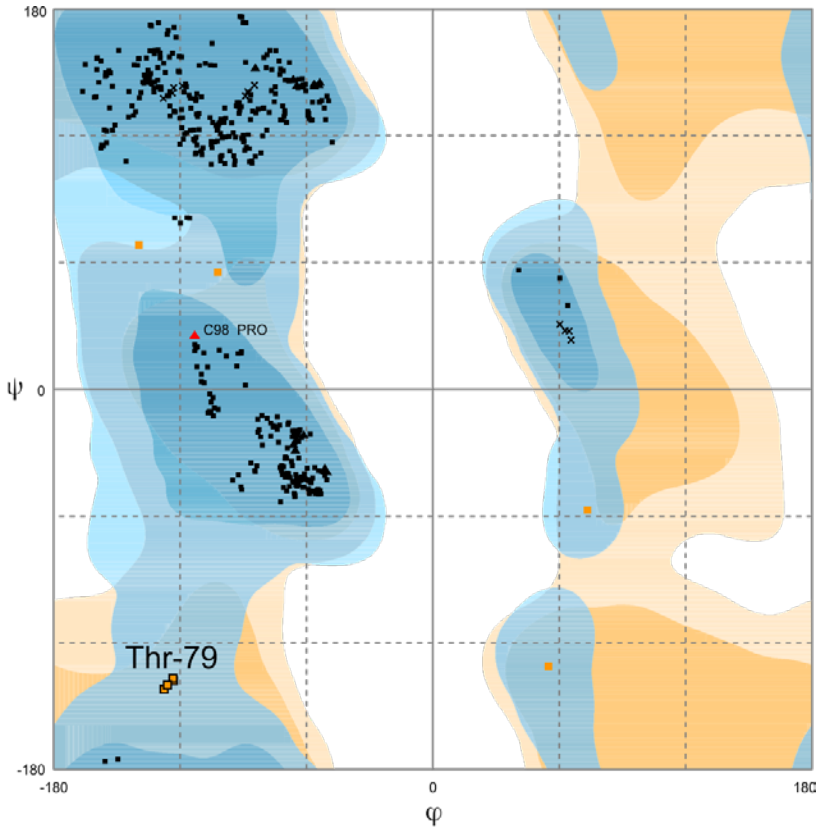


C

mass	CRES	Truncated CRES
theoretical	14762.98	12979.93
MALDI	14764.63	12978.78
# amino acids	129	112

Figure S1. Mass spectrophotometric analysis of CRES C48A. (A) MALDI of purified CRES C48A revealed two populations with masses of 14764.63 and 12978.78 Daltons. Tryptic digests of the CRES C48A isoforms showed that an 1800 Dalton peptide (arrowhead) was missing from the smaller CRES protein. (B) MS/MS spectrum for m/z= 1800 peptide indicated a 6 amino acid linker (GAMAHM) and the first 11 N-terminal amino acids of CRES (VGVDQSKNEVK). (C) The theoretical masses of full-length CRES and CRES lacking the N-terminal 11 amino acids and 6 amino acid linker match that observed by MALDI.

A



B

Chain	φ (°)	ψ (°)
A-Thr 79	-123.2	-137.9
B-Thr 79	-123.4	-136.8
C-Thr 79	-127.6	-142
D-Thr 79	-126	-140.1

C

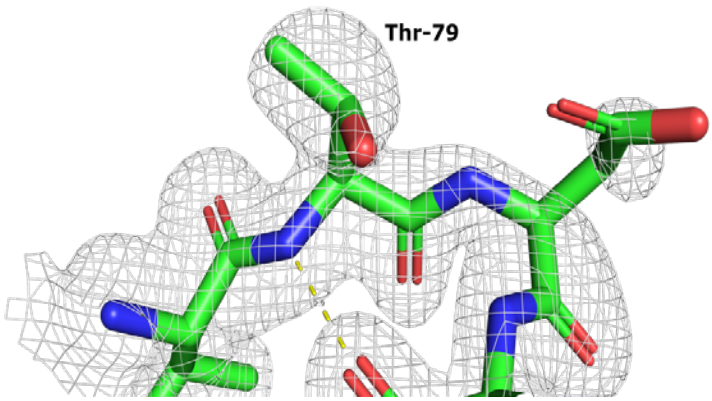


Figure S2. Structural assessment of the CRES X-ray crystal structure. (A) Ramachandran map of CRES residues from all four chains. The phi-psi angles for the energetically-constrained Thr-79 are highlighted. (B) Table of the phi-psi angles for Thr-79. (C) Representative electron density of Thr-79. The modeling and refinement of Thr-79 from all four chains is consistent with the electron density.

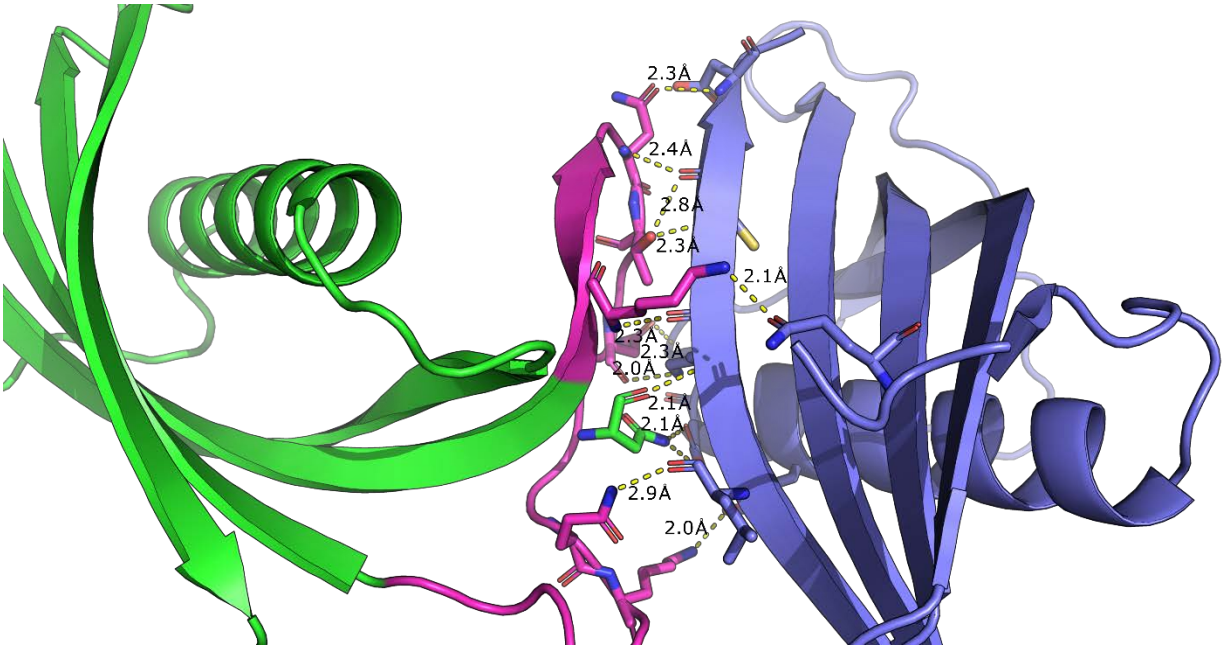


Figure S3. Inter-molecular interaction between two CRES molecules in the crystallographic asymmetric unit. The CRES loop (pink) of chain A forms a parallel β-sheet interaction with β5 of chain B through 8 hydrogen bonds and two salt bridges. Distances are indicated and are also presented in Table S2.

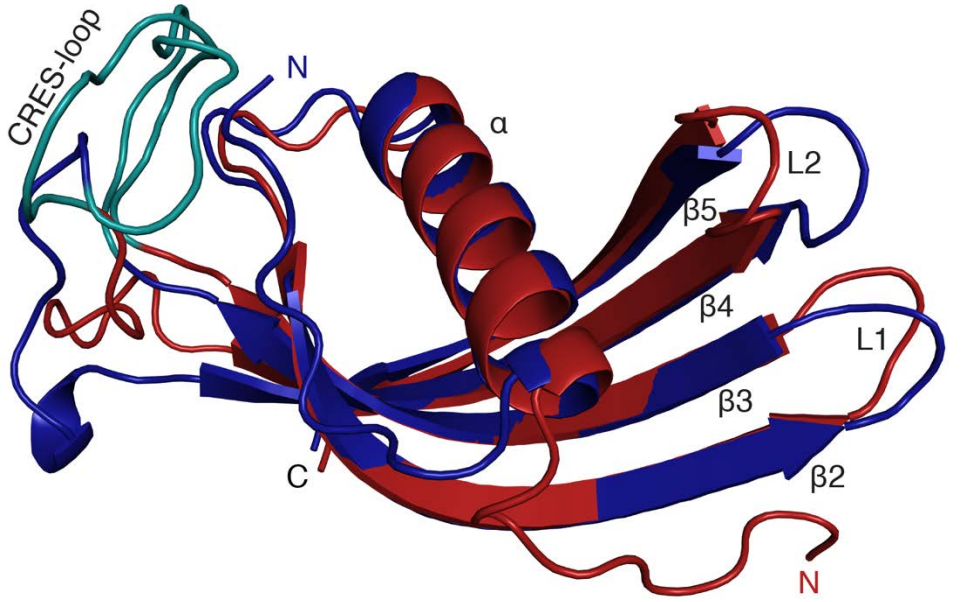


Figure S4. CRES X-ray and solution alignment. Overlay of CS-ROSETTA derived model of CRES (blue) and the X-ray crystal structure (Red), which highlights the similarity of the secondary structure and the overall fold.

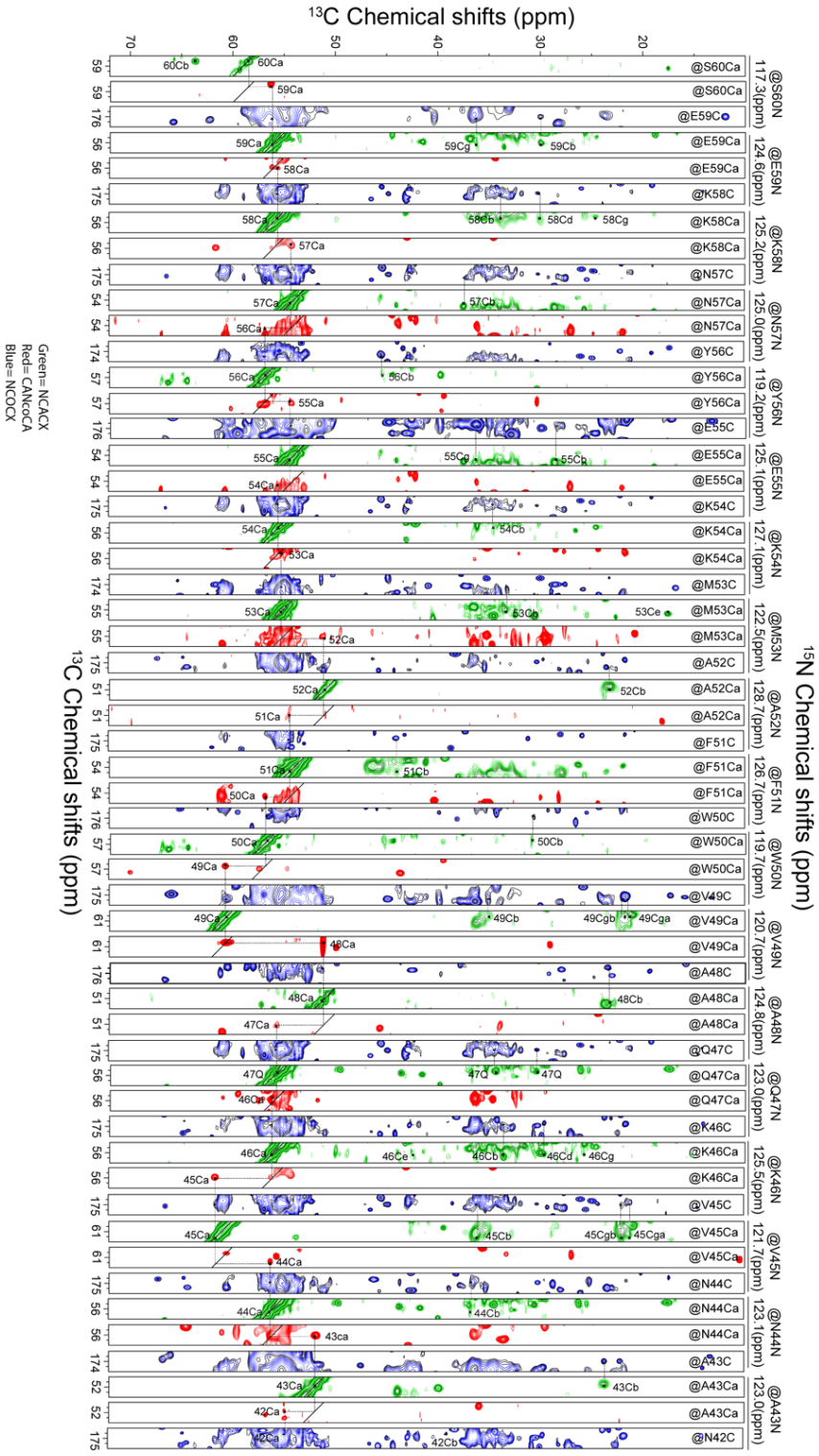


Figure S5. SSNMR backbone walk through the α-helical region of the X-ray crystal structure after CRES has transitioned to amyloid. Depicted are strip plots from three different 3D spectra: NCACX (green), CANcoCA (red), and NCOCX (blue). Each 2D ^{13}C-^{13}C slice is aligned using the connecting ^{15}N shift, which is indicated on the top of each strip. The observed chemical shifts are all consistent with this region of the protein converting into a β-strand.

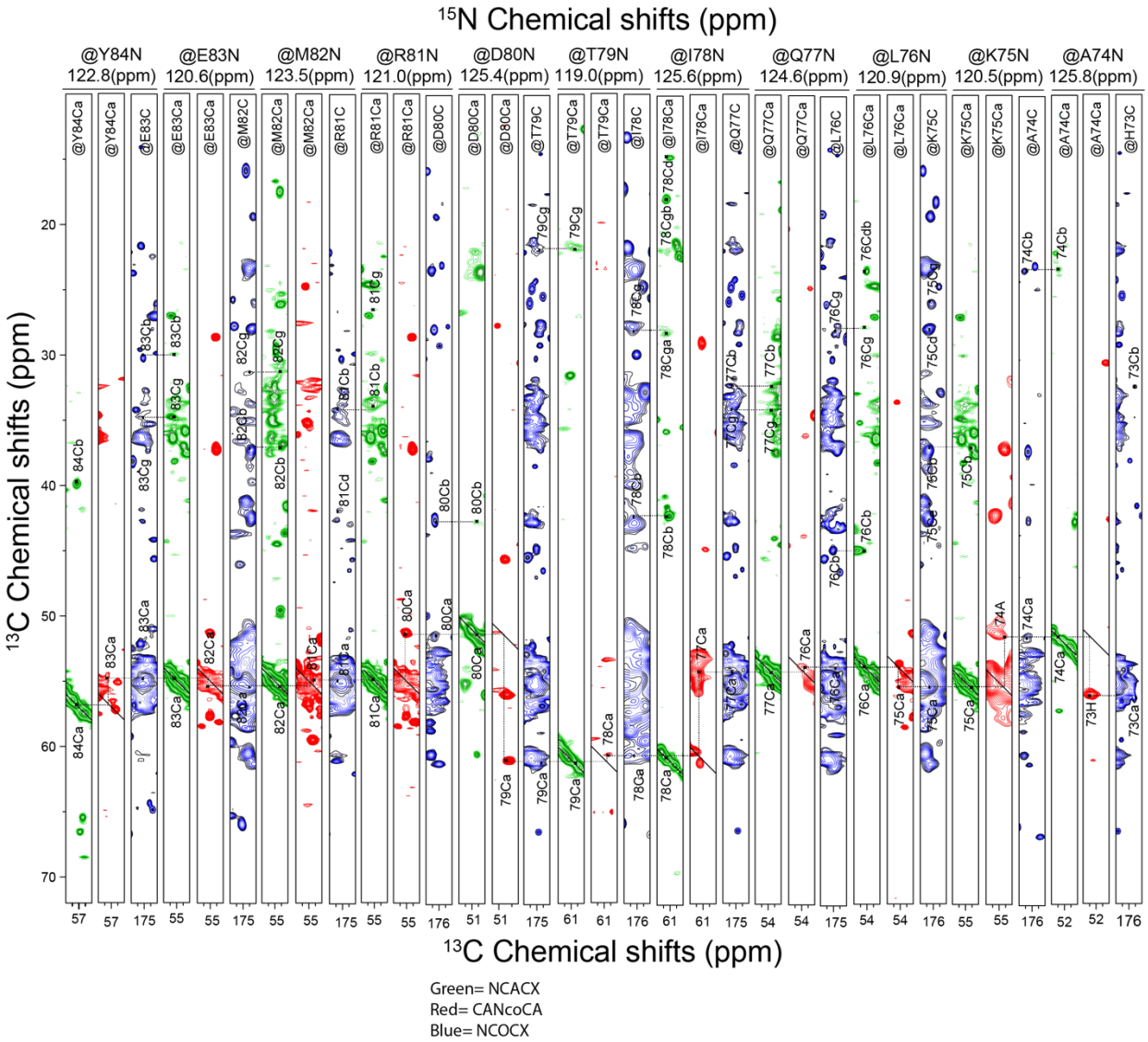


Figure S6. SSNMR backbone walk through the L1 loop, including regions of the β2 and β3 strands adjacent to this loop. Depicted are strip plots from three different 3D spectra: NCACX (green), CANcoCA (red), and NCOCX (blue). Each 2D ^{13}C-^{13}C slice is aligned using the connecting ^{15}N shift, which is indicated on the top of each strip. These chemical shifts indicate that this region has likely become a single β strand.

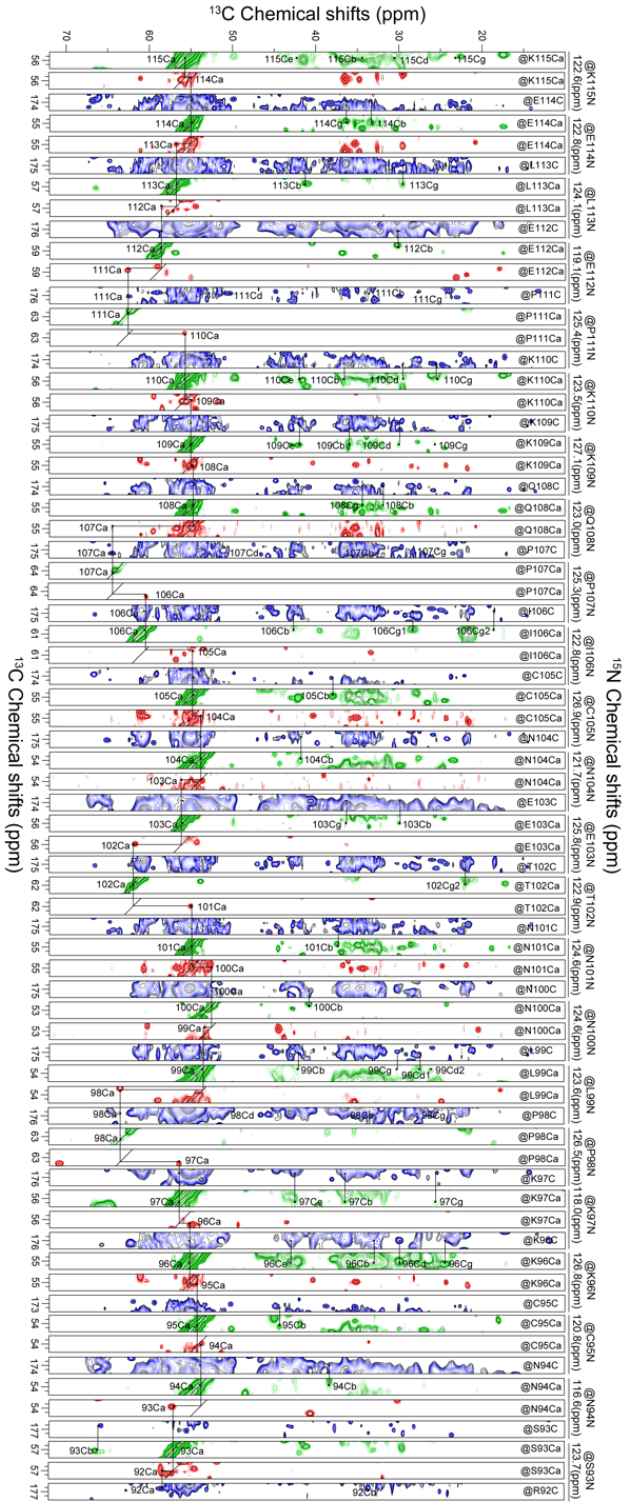


Figure S7. SSNMR backbone walk through the CRES Loop. Depicted are strip plots from three different 3D spectra: NCACX (green), CANcoCA (red), and NCOCX (blue). Each 2D ^{13}C-^{13}C slice is aligned using the connecting ^{15}N shift, which is indicated on the top of each strip. These chemical shifts indicate that this region has likely become two short β strand strands joined by a short coil.

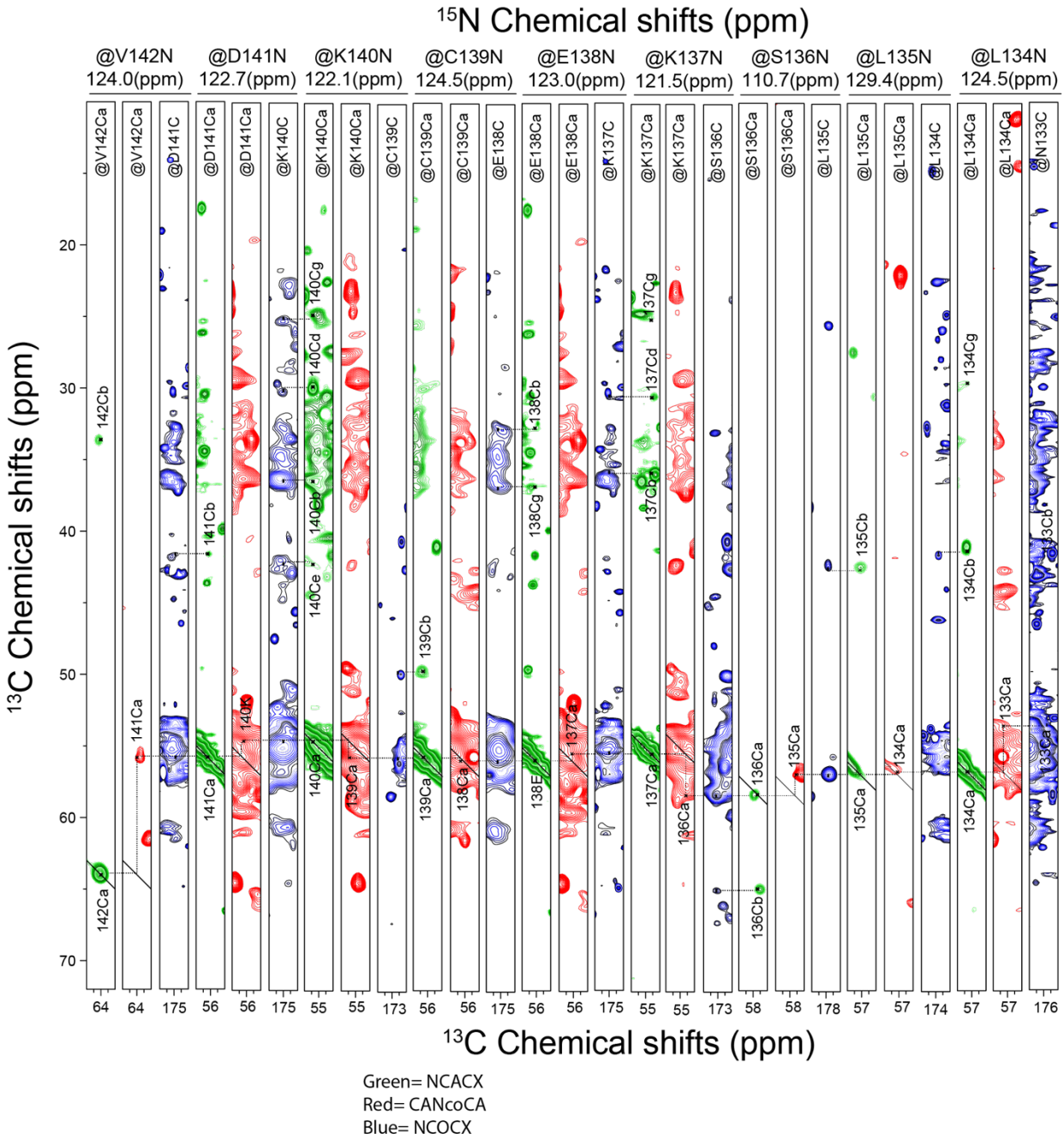


Figure S8. SSNMR backbone walk through parts of the C-terminus. Depicted are strip plots from three different 3D spectra: NCACX (green), CANcoCA (red), and NCOCX (blue). Each 2D ^{13}C-^{13}C slice is aligned using the connecting ^{15}N shift, which is indicated on the top of each strip. These chemical shifts indicate that this region retained its original β-strand architecture.

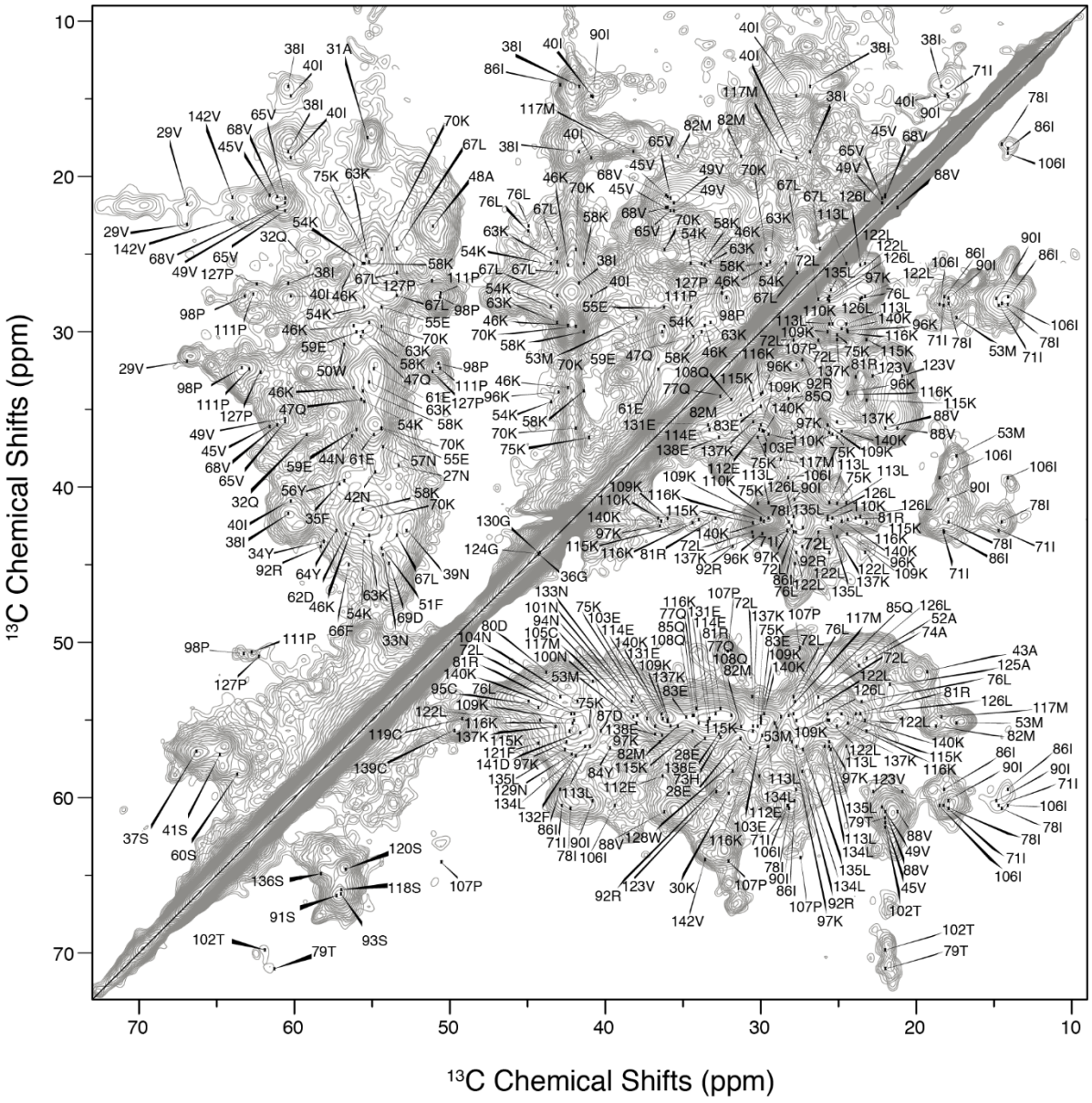


Figure S9. Aliphatic region of 2D ^{13}C-^{13}C DARR spectrum of CRES. The spectrum was recorded with a DARR mixing time of 25 ms at a magnetic field of 14 T. Each dimension is apodized with 30 Hz of Gaussian line broadening. The near-complete aliphatic ^{13}C chemical shift assignments are labeled for each cross-peak.

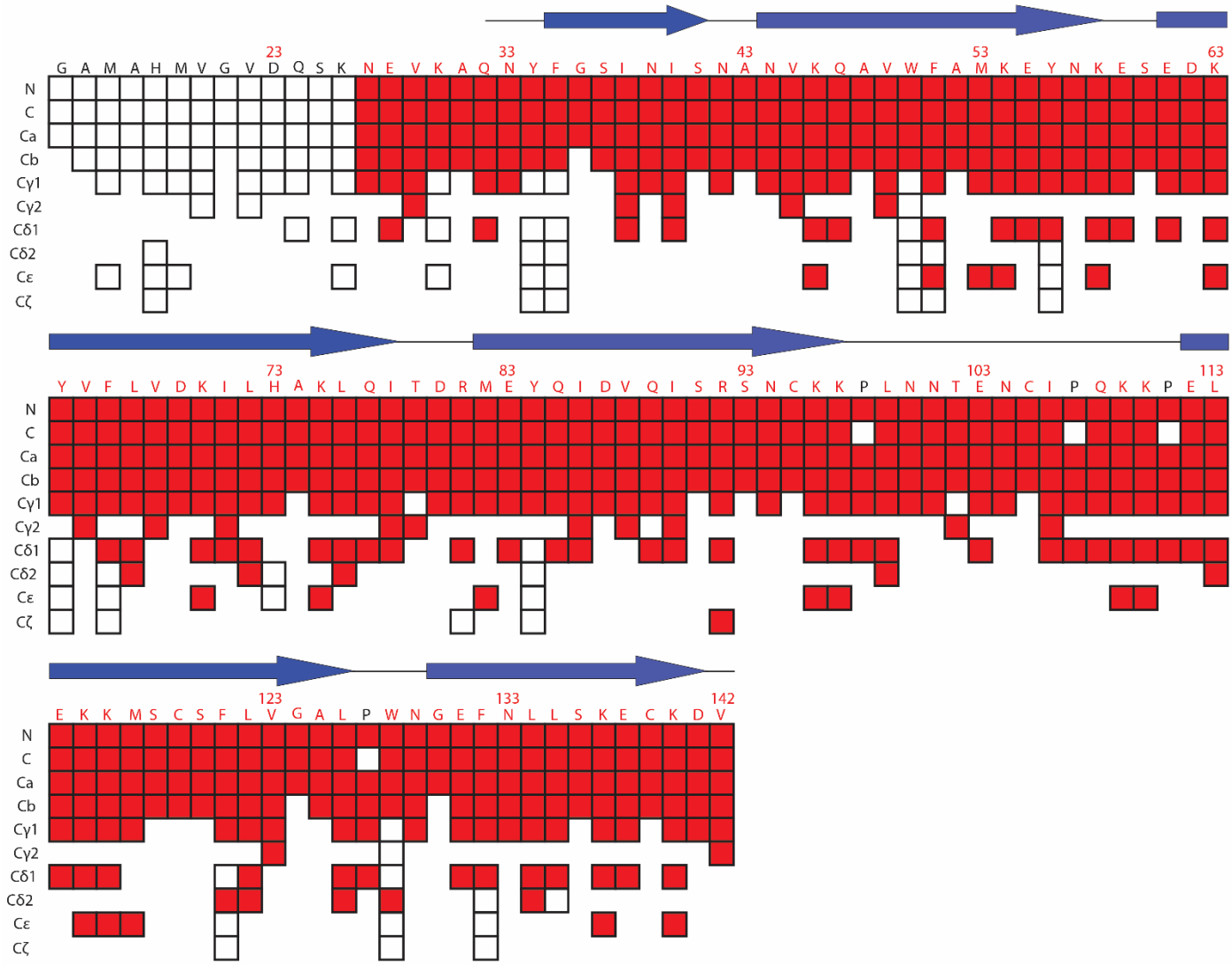


Figure S10. SSNMR assignment completeness. Red boxes correspond to assigned spin systems through use of NCACX, CANcoCA, and NCOCX 3D pulse sequences for fully ^{15}N-^{13}C labeled CRES.

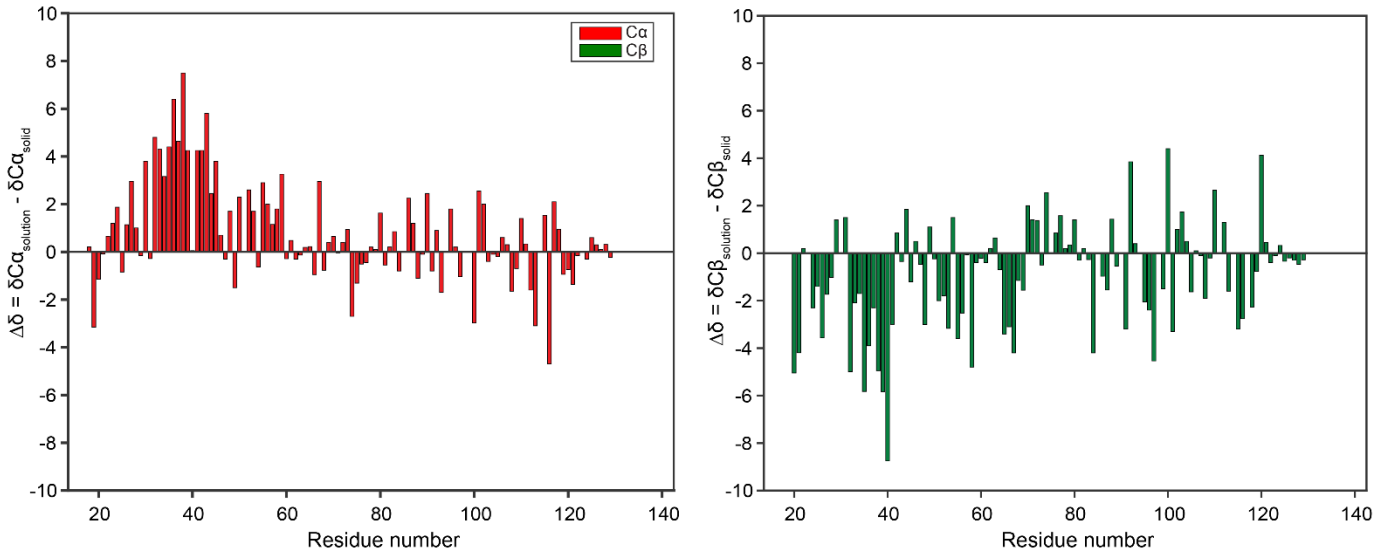


Figure S11. Bar plot showing the Cα (red) and Cβ (green) chemical shift perturbations between solution and solid-state assignments. The differences were calculated by using the equation $\Delta\delta = \delta C_{solution} - \delta C_{solid}$, where δ = chemical shifts in ppm and C = Cα or Cβ. These CSPs reflect the structural change as CRES transitions from the soluble monomer to the mature matrix.

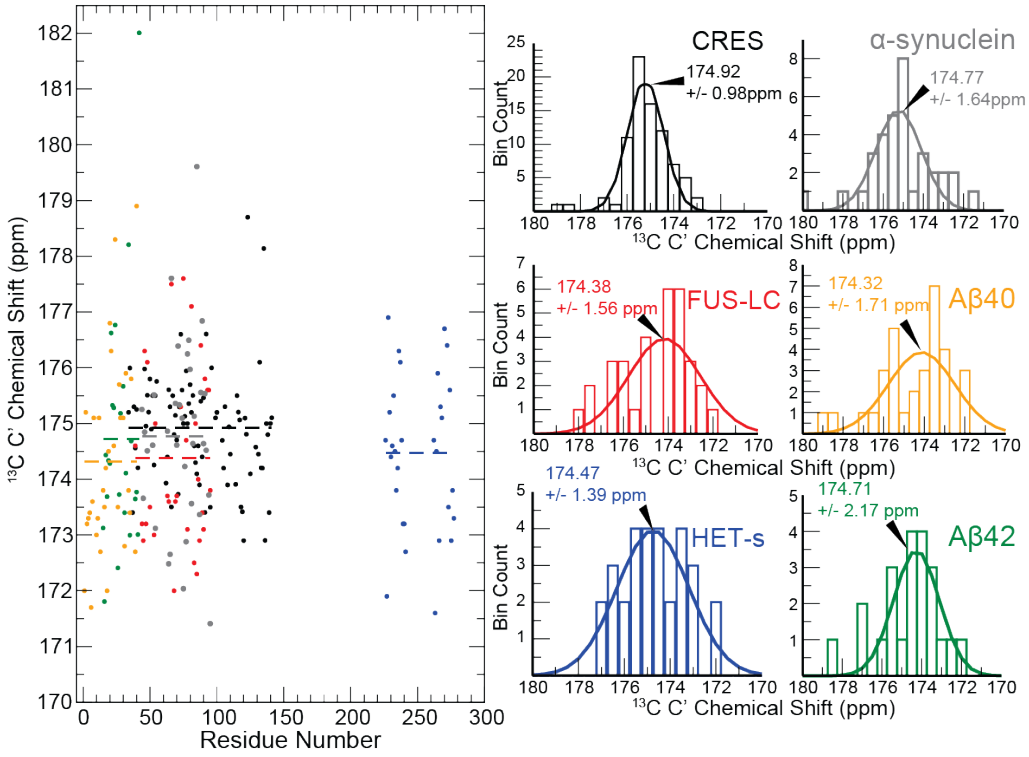


Figure S12. Statistical analysis of CRES ^{13}C' chemical shifts compared to well-known intermolecular parallel β-sheet structures. On the left is the distribution of all measured β-sheet ^{13}C' resonances colored by structure and organized by residue number: CRES (black), α-synuclein (grey) (25), FUS-LC (red) (26), Aβ40 (yellow) (27), HET-s (blue) (29), and Aβ42 (green) (28). On the right are the histograms of β-sheet ^{13}C' resonances for each protein. The mean and standard deviation are given for each.

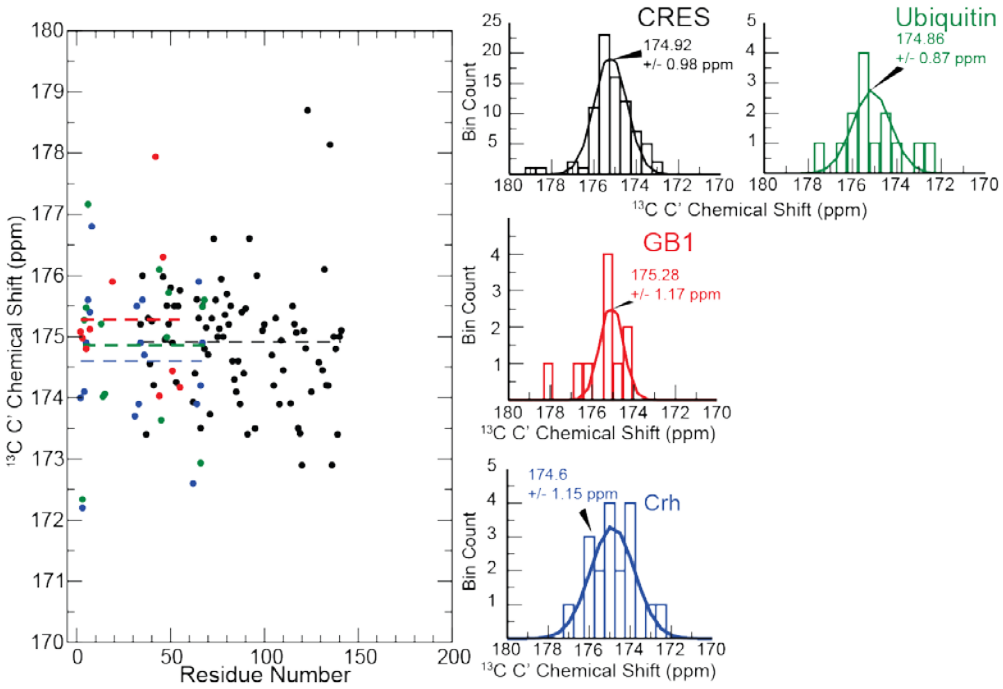


Figure S13. Statistical analysis of CRES ^{13}C' chemical shifts compared to well-known antiparallel β-sheet regions of globular proteins. On the left is the distribution of all measured β-sheet ^{13}C' resonances colored by structure and organized by residue number: CRES (black), Ubiquitin (green) (30-32), GB1 (red) (34), and Crh (blue) (33). On the right are the histograms of β-sheet ^{13}C' resonances for each protein. The mean and standard deviation are given for each.

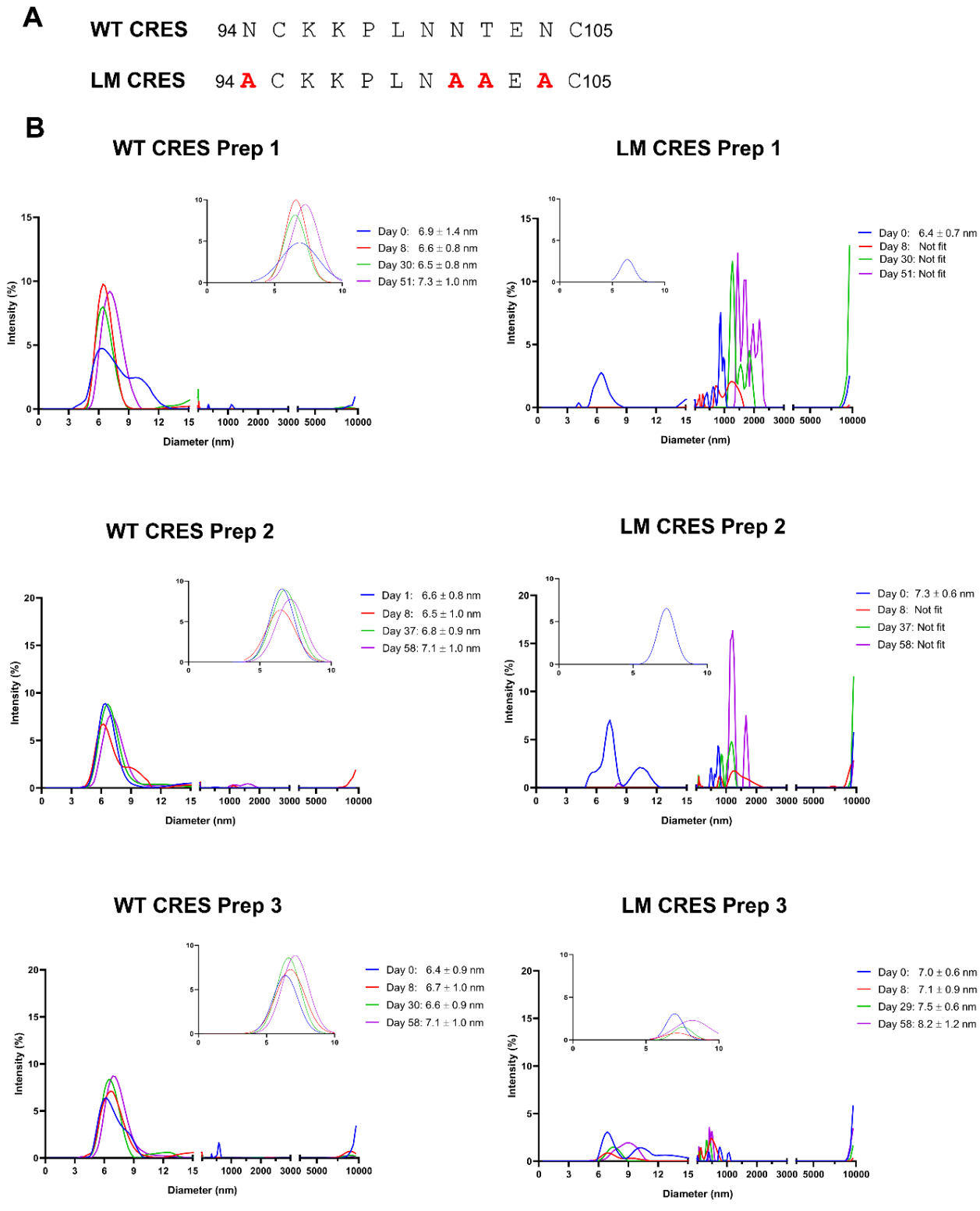


Figure S14. Loop mutant CRES does not form a stable oligomer and rapidly assembles into large aggregates. (A) Amino acid sequence of the CRES loop in wildtype (WT) CRES and

with alanine substitutions (red) in loop mutant (LM) CRES. (B) The aggregation of WT CRES and LM CRES was followed with time by dynamic light scattering (DLS). Proteins freshly eluted off the gel filtration column were buffer exchanged into 4 mM potassium phosphate, pH 7.4 and concentrated to ~15 mg/ml. Samples were examined by DLS within 24 hours and stored at 4 °C between reads. Inset, an average hydrodynamic radius for the particles between 4-9 nm was calculated from the fitted data and diameter ± SD is reported. Due to large variations in particle size we were unable to fit CRES particles larger than 100 nm. The data shown represent three different protein preparations for WT and LM CRES.

Supplemental references

1. Do HQ*, et al.* (2019) The Functional Mammalian CRES (Cystatin-Related Epididymal Spermatogenic) Amyloid is Antiparallel beta-Sheet Rich and Forms a Metastable Oligomer During Assembly. *Scientific Reports* 9.
2. Minor W, Cymborowski M, Otwinowski Z, & Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. *Acta Crystallographica Section D-Structural Biology* 62:859-866.
3. Bibby J, Keegan RM, Mayans O, Winn MD, & Rigden DJ (2012) AMPLE: a cluster-and-truncate approach to solve the crystal structures of small proteins using rapidly computed ab initio models. *Acta Crystallographica Section D-Biological Crystallography* 68:1622-1631.
4. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40:658-674.
5. Maszota-Zieleniak M*, et al.* (NMR and crystallographic structural studies of the extremely stable monomeric variant of human cystatin C with single amino acid substitution. *FEBS J.*
6. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Structural Biology* 66:213-221.
7. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Structural Biology* 60:2126-2132.
8. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6(3):277-293.
9. Vranken WF*, et al.* (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins-Structure Function and Bioinformatics* 59(4):687-696.
10. Muhandiram DR & Kay LE (1994) GRADIENT-ENHANCED TRIPLE-RESONANCE 3-DIMENSIONAL NMR EXPERIMENTS WITH IMPROVED SENSITIVITY. *Journal of Magnetic Resonance Series B* 103(3):203-216.
11. Shen Y & Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* 56(3):227-241.
12. Shen Y*, et al.* (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences of the United States of America* 105(12):4685-4690.
13. Grzesiek S*, et al.* (1996) The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat. Struct. Biol.* 3(4):340-345.
14. Kay LE, Nicholson LK, Delaglio F, Bax A, & Torchia DA (1992) PULSE SEQUENCES FOR REMOVAL OF THE EFFECTS OF CROSS-CORRELATION BETWEEN DIPOLAR AND CHEMICAL-SHIFT ANISOTROPY RELAXATION MECHANISM ON THE MEASUREMENT OF HETERONUCLEAR T1 AND T2 VALUES IN PROTEINS. *J. Magn. Reson.* 97(2):359-375.
15. Hansen DF, Vallurupalli P, & Kay LE (2008) An improved (15)N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *Journal of Physical Chemistry B* 112(19):5898-5904.
16. Korzhnev DM*, et al.* (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* 430(6999):586-590.
17. McConnell HM (1958) Reaction rates by nuclear magnetic resonance. *J Chem Phys* 28:430-431.

18. Ying JF, Delaglio F, Torchia DA, & Bax A (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* 68(2):101-118.
19. Das R*, et al.* (2009) Simultaneous prediction of protein folding and docking at high resolution. *Proceedings of the National Academy of Sciences of the United States of America* 106(45):18978-18983.
20. Whelly S*, et al.* (2012) Nonpathological extracellular amyloid is present during normal epididymal sperm maturation. *Plos One* 7(5).
21. Wylie BJ*, et al.* (2007) Chemical-shift anisotropy measurements of amide and carbonyl resonances in a microcrystalline protein with slow magic-angle spinning NMR spectroscopy. *J. Am. Chem. Soc.* 129(17):5318-+.
22. Wei YF, Lee DK, & Ramamoorthy A (2001) Solid-state C-13 NMR chemical shift anisotropy tensors of polypeptides. *J. Am. Chem. Soc.* 123(25):6118-6126.
23. Kameda T & Ando I (1997) The relationship between the helical conformation and C-13 NMR chemical shift of amino acid residue carbonyl carbons of polypeptides in the solid state. *J. Mol. Struct.* 412(3):197-203.
24. Kameda T*, et al.* (1996) Hydrogen-bonded structure and C-13 NMR chemical shift tensor of amino acid residue carbonyl carbons of peptides and polypeptides in the crystalline state .1. *J. Mol. Struct.* 384(1):17-23.
25. Tuttle MD*, et al.* (2016) Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein. *Nat. Struct. Mol. Biol.* 23(5):409-415.
26. Murray DT*, et al.* (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* 171(3):615-+.
27. Lu JX*, et al.* (2013) Molecular Structure of beta-Amyloid Fibrils in Alzheimer's Disease Brain Tissue. *Cell* 154(6):1257-1268.
28. Colvin MT*, et al.* (2016) Atomic Resolution Structure of Monomorphic A beta(42) Amyloid Fibrils. *J. Am. Chem. Soc.* 138(30):9663-9674.
29. Wasmer C*, et al.* (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319(5869):1523-1526.
30. Igumenova TI*, et al.* (2004) Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126(21):6720-6727.
31. Igumenova TI, Wand AJ, & McDermott AE (2004) Assignment of the backbone resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126(16):5323-5331.
32. Fasshuber HK*, et al.* (2015) Structural heterogeneity in microcrystalline ubiquitin studied by solid-state NMR. *Protein Sci.* 24(5):592-598.
33. Bockmann A*, et al.* (2003) Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR* 27(4):323-339.
34. Wylie BJ*, et al.* (2011) Ultrahigh resolution protein structures using NMR chemical shift tensors. *Proceedings of the National Academy of Sciences of the United States of America* 108(41):16974-16979.