Maturation of the Functional Mouse CRES Amyloid from Globular Form

Aveline Hewetson[1#], Nazmul H. Khan[2#], Matthew J. Dominguez [3], Hoa Quynh Do[1], R.E. Kusko[1], Collin G. Borcik[2], Daniel J. Rigden[4], Ronan M. Keegan[4, 5], R. Bryan Sutton[3,6], Michael P. Latham[2,6*], Benjamin J. Wylie[2,6*], Gail A. Cornwall[1*]

[1]Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX, 79430. [2] Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, 79409, [3]Department of Cell Physiology and Molecular Biophysics, Texas Tech University Health Sciences Center, Lubbock, TX, 79430, [4]Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, U.K. , [5]Science and Technology Facilities Council, Rutherford Appleton Laboratory, Oxfordshire, OX11 0QX, U.K., [6]Center for Membrane Protein Research, Texas Tech University Health Sciences Center, Lubbock, TX, 79430.

Email: Aveline.hewetson@ttuhsc.edu, Nazmul.khan@ttu.edu, Matthew.Dominguez@ttuhsc.edu, Quynh-hoa.do@ttuhsc.edu, rebecca.kusko@ttuhsc.edu, collin.borcik@ttu.edu, drigden@liverpool.ac.uk, ronan.keegan@stfc.ac.uk, Roger.b.sutton@ttuhsc.edu, Michael.Latham@ttu.edu, Benjamin.J.Wylie@ttu.edu, Gail.Cornwall@ttuhsc.edu

ORCID: 0000-0001-8741-2443 (AH), 0000-0001-7361-6066 (HQD), 0000-0001-8393-5239 (RBS), 0000-0002-2209-5798 (MPL), 0000-0001-8183-2762 (BJW), 0000-0002-2903-1690 (GAC)

*Corresponding authors: Gail A. Cornwall, Benjamin J. Wylie, Michael P. Latham.

To whom correspondence should be addressed: Gail A. Cornwall, Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, 3601 4th Street, Lubbock, TX 79430, 806-743-2687, gail.cornwall@ttuhsc.edu

# contributed equally to this work.

Classification: Biological Sciences/Biophysics and Computational Biology

Author contributions: RBS, MPL, BJW, GAC designed research; AH, NHK, HQD, MJD, RBS, REK, MPL, BJW performed research; MJD, HQD, DJR, REK, RMK, CGB, RBS, MPL, BJW, GAC analyzed data; NHK, RBS, MPL, BJW, GAC wrote the paper.

Abstract: The epididymal lumen contains a complex cystatin-rich nonpathological amyloid matrix with putative roles in sperm maturation and sperm protection. Given our growing understanding for the biological function of this and other functional amyloids, the problem still remains: how functional amyloids assemble including their initial transition to early oligomeric forms. To examine this, we developed a protocol for the purification of nondenatured mouse CRES, a component of the epididymal amyloid matrix, allowing us to examine its assembly to amyloid under conditions that may mimic those *in vivo*. Herein we use X-ray crystallography, solution-state NMR, and solid-state NMR to follow at the atomic level the assembly of the CRES amyloidogenic precursor as it progressed from monomeric folded protein to an advanced amyloid. We show the CRES monomer has a typical cystatin fold that assembles into highly branched amyloid matrices, comparable to those *in vivo,* by forming β-sheet assemblies that our data suggest occur via two distinct mechanisms: a unique conformational switch of a highly flexible disulfide-anchored loop to a rigid β-strand and by traditional cystatin domain swapping. Our results provide key insight into our understanding of functional amyloid assembly by revealing the earliest structural transitions from monomer to oligomer and by showing that some functional amyloid structures may be built by multiple and distinctive assembly mechanisms.

Keywords: cystatin, amyloid, X-ray crystallography, NMR spectroscopy, domain swapping, loop

Significance Statement: Some amyloids are functional in the normal physiology of cells. Of these, several have been found in complex amyloid matrices that are composed of multiple amyloidogenic precursors. Here we trace the lifecycle of the mouse CRES (cystatin-related epididymal spermatogenic) protein, a functional amyloidogenic precursor and component of a complex amyloid matrix in the epididymis, as it progresses from monomeric folded protein to an advanced amyloid. We discovered that CRES amyloid formation may involve β-sheet assemblies generated by two distinct mechanisms: those formed by a flexible CRES loop and those formed by domain swapping. These mechanisms would provide multiple contact surfaces for further β-sheet assemblies with neighboring proteins resulting in sophisticated quaternary structures which may be integral for biological function.

Introduction

Amyloids, including fibrils, films and matrices, are formed by the self-association of proteins into higher ordered cross β-sheet structures. Although traditionally viewed solely as pathological oligomers or plaques integral to neurodegenerative diseases and prionopathies, a growing body of evidence now suggests that many proteins form amyloid which perform important biological roles. These functional amyloids have been implicated in processes including melanin synthesis; germline specification; formation of long-term memory; sperm maturation, clearance, and fertilization; and hormone storage (1-8). Functional amyloids follow similar aggregation pathways as pathological forms, assembling from monomer to intermediate oligomers to advanced amyloids. Despite forming similar structures, functional amyloids are distinct from pathological forms since they are produced *in vivo* without significant cytotoxicity and often reversibly disassemble (6, 8, 9). Further, some functional amyloids are elaborate structures composed of multiple amyloidogenic precursors adding another level of complexity to their assembly (6, 7, 10). However, the mechanism(s) by which functional amyloids assemble and disassemble is poorly understood. This lack of knowledge is due, in part, to the difficulty in purifying amyloidogenic precursors under nondenaturing conditions which would then allow for examination of the structural transitions required for the assembly of amyloid from natively folded protein. Indeed, although structural biology techniques including cryoEM and NMR spectroscopy have revealed atomic resolution of protofibrils and highly ordered amyloid fibrils, this work has primarily focused on pathological amyloids such as Aβ and/or discrete amyloidogenic domains (11-17). Elucidation of the structural and dynamic transitions of whole amyloidogenic precursors along their assembly pathway is therefore needed to reveal possible mechanisms that initiate and control amyloidogenesis and to identify differences between functional and pathogenic amyloid assembly and structure.

In previous studies, we demonstrated that a nonpathological amyloid matrix with proposed roles in sperm maturation and sperm protection is present throughout the normal mouse epididymal lumen (4). Further, the amyloid matrix changes along the length of the epididymal tubule transitioning from an anti-A11 immunoreactive amyloid composed of matrices and films in the proximal epididymis to one rich in fibrils in the distal epididymis (18) . The epididymal amyloid matrix contains several family 2 cystatins of cysteine protease inhibitors including cystatin C and four members of the CRES (cystatin-related epididymal spermatogenic) subgroup, a reproductive subgroup of family 2 cystatins (10) . We further showed the amyloid forms of CRES subgroup members are present in the epididymal amyloid matrix and their recombinant forms readily form amyloid *in vitro* (10). These results suggest CRES subgroup members are functional amyloidogenic precursors that together contribute to the formation of a complex extracellular amyloid structure.

To determine how a functional amyloidogenic precursor assembles into amyloid, we previously showed that native CRES assembled from a monomer with mixed secondary structure into a metastable antiparallel β-sheet rich oligomer, that eventually assembled into a higher-ordered amyloid (19). In the present study we now follow the assembly of CRES from monomer to amyloid at the atomic level. We used X-ray crystallography and solution state NMR to determine the structure and dynamics of the CRES monomer and its earliest oligomeric forms. Solid-state NMR then revealed structural details of the advanced CRES amyloid. Together, our data suggest that CRES amyloid assembly occurs by two distinct mechanisms which may allow the controlled formation of complex quaternary structures, including highly branched amyloid matrices, that are characteristic of the functional amyloid structure in the lumen of the proximal mouse epididymis.

Results and Discussion

Crystal structure of CRES reveals a monomeric cystatin fold. Mature mouse CRES lacking the signal sequence (residues 1-19) was modified with a cysteine 48 to alanine mutation (C48A) to prevent inappropriate disulfide bond formation. The crystallization experiments were carried out using the hanging-drop vapor-diffusion method, and diffraction quality crystals were obtained after 6 months via a robotic screen using Molecular Dimensions screening conditions. Crystals of CRES C48A were grown in 2.1 M ammonium sulfate, 0.2 M Na iodide, pH 5.8 at 10°C. The crystals were in space group $P2_12_12_1$ and diffracted to 1.8 Å resolution. We modeled 22 iodide (I^-) and 2 molecular iodine (I_2) ligands in the structure. The iodine positions were confirmed by anomalous difference maps. Although we were unable to solve the structure with traditional replacement techniques using the stabilized cystatin C monomeric structure (3GAX) (20), we were able to compute a partial solution using AMPLE (21). AMPLE was able to place a significantly trimmed ensemble of eight distantly homologous cystatin-fold structures. Interestingly, we were able to solve the CRES structure using the recently published crystal structure of V57G mutant of human cystatin C (6ROA) (22). There were four CRES molecules in the asymmetric unit. The crystallographic data are summarized in the SI Appendix,Table S1.

The CRES C48A structure is shown in Figure 1A. The overall structural similarity of CRES to other family 2 cystatins was high, as expected, as the primary sequence alignment between cystatin C and CRES is 47.6% similar (24% identical). However, unlike other wild type cystatins which crystallized as domain swapped dimers, CRES C48A remained a monomer and exhibited a typical cystatin fold characterized by an unstructured N-terminus leading into a long α-helix sitting atop a five-stranded antiparallel β-sheet. Similar to crystal structures of other cystatins, no electron density was observed for the first eleven N-terminal residues (22). In cystatin C this is thought to be due to the inherent flexibility of the N-terminal region (23). MALDI mass spectrometry analysis of C48A CRES, however, revealed two populations with masses of

14764.63 and 12978.78 Daltons suggesting a proportion of CRES was processed during purification. Tryptic digests indicated that the first eleven residues from the N-terminus were missing from the smaller CRES C48A protein (SI Appendix, Figure S1). Although the mechanism by which this processing occurred is unknown, the N-terminal truncation of CRES may help initiate its aggregation. Indeed, the loss of the first ten to eleven N-terminal amino acids is a characteristic of several family 2 cystatins and, as shown for cystatin C and cystatin E, facilitated their self-assembly (22, 24-26). Both 14 kDa and 12 kDa CRES forms are present in the mouse epididymal lumen suggesting the N-terminal processing of CRES is biologically relevant (4). The electron density of all four CRES molecules in the asymmetric unit began after the cleavage site suggesting only N-terminally truncated molecules in the crystal.

Like other family 2 cystatins, CRES C48A contained two defined loop structures, L1 which connects β-strands 2 and 3 and L2 which connects β-strands 4 and 5 (Figure 1A). Although L1 and L2 contribute to the papain (C1 cysteine protease) binding site in typical family 2 cystatin monomers, CRES does not inhibit papain *in vitro*, likely due to the lack of the highly conserved Q-X-V-X-G site for inhibition (27). Notably, CRES C48A contains a threonine (Thr) instead of a valine (Val) residue at this position in L1 (...QITDR...). In cystatin C the conformational strain around Val-57 in L1 creates a hinge within the loop that causes it to domain swap. In fact, mutating this residue to asparagine (N) or glycine (G) resulted in monomeric cystatin C (22, 25). The absence of this critical residue in L1 may be why CRES C48A did not undergo domain swapping during crystallization. Indeed, Val has the lowest probability of being in a β-turn, hence the structural instability, while Thr has an average probability (28). The electron density in the CRES L1 loop shows that the Thr-79 in L1 is well-ordered (SI Appendix, Figure S2). While the φ-ψ angles of Thr-79 map to a similar region of Ramachandran space as Val-57 in cystatin C, the backbone angles for Thr-79 in CRES are not technically forbidden (SI Appendix, Figure S2). However, this site in CRES, like other cystatins, is likely structurally strained for similar reasons.

An interaction in the crystal structure suggests a path to amyloid. Opposite to the papain inhibitory region in the CRES C48A monomer is a disordered disulfide-anchored loop connecting β3 and β4, which we refer to as the CRES loop (Figure 1A). This disordered region is also present in other cystatins and has been called an “appending structure” (AS) or legumain exosite loop (LEL) because of its interaction with the C13 cysteine protease legumain (25, 26) . It is not known if CRES, like cystatins C, E, and F, inhibits legumain. An overlay of the four CRES C48A molecules in the asymmetric unit showed the only variability between the four molecules was in the flexible CRES loop (Figure 1B). Unexpectedly, CRES molecules A and B interacted with each other via an asparagine-rich region of the CRES loop of molecule A and the main chain atoms along β-strand 5 of molecule B, forming what may be an extended pseudo parallel β-sheet interaction (Figure 1C, SI Appendix, Figure S3). In fact, DSSP, which computes secondary structure, made the CRES loop into an elongated β-strand 3 of chain A upon contact with β-strand 5 of CRES chain B (29). The parallel β-strand formed by the interaction between the two molecules was held together by 2 salt bridges and 8 H-bonds with distances typical of those present in β-sheet interactions (SI Appendix, Figure S3, Table S2). The interface formed between these two molecules buried approximately 1295 Å^2 as measured using PISA (30), suggesting a biologically meaningful interaction (Figure 1D). This interaction could represent an early assembly state of CRES as it transitions to a higher ordered amyloid via a mechanism other than that of domain swapping.

Solution-State NMR confirms the cystatin fold of CRES. To better understand the transition of the cystatin C-like CRES structure observed in the crystal to the form found in the amyloid, we turned to solution-state NMR. Uniformly ^{13}C and ^{15}N-labeled CRES C48A in 25 mM MES, 250 mM NaCl, 1 mM EDTA buffer, pH 6 was used for standard triple-resonance backbone assignment experiments that were performed immediately following protein purification (Figure 2A). We previously showed by dynamic light scattering that, under these buffer conditions, the majority of particles were of a size expected for a CRES monomer with occasionally slightly larger particles detected (19). TALOS-N was then used to calculate secondary structure propensities from the backbone assignments (31). As shown in Figure 2B, a good agreement was observed between the secondary structural elements observed in the crystal structure and the solution state NMR data confirming the cystatin C-like fold of CRES C48A in solution. TALOS-N was also used to derive random coil index ^{15}N order parameters (S^2), which are a measure of backbone flexibility (32) . A good inverse correlation was observed between the average crystallographic b-factors (Figure 2C, green points) and the predicted ^{15}N S^2 (blue points), where amides with low b-factors have high S^2, indicating a rigid backbone, and amides with high b-factors have low S^2, indicating a flexible backbone (33). We note that several residues (E103 – I106) in the CRES loop connecting the β3 and β4 strands had elevated S^2 values (decreased b-factors) designating decreased motion in this section of the loop in both the crystal and solution states. CS-ROSETTA confirmed nearly the same fold as the crystal structure (SI Appendix, Figure S4).

Monitoring the CRES transition to amyloid. To characterize the initial steps of CRES C48A into its early oligomeric amyloid form, we measured solution state NMR ^{15}N relaxation experiments sensitive to the pico-to-nanosecond and millisecond timescales. Data were initially collected in 25 mM MES, 250 mM NaCl, 1 mM EDTA, pH 6 buffer conditions where we predict CRES is predominantly a monomer. The pH of the buffer was then adjusted to pH 7.5, by the addition of 50 mM HEPES, pH 8, which also reduced the salt concentration to 173 mM, to promote amyloidogenesis (19). Initial comparison of the 2D ^{15}N,^{1}H heteronuclear single-quantum coherence (HSQC) spectra recorded in the two buffer conditions revealed a series of chemical shift perturbations (CSPs; average of all CSPs = 0.02, SI Appendix,Table S3), which are mapped on the structure of CRES C48A in Figure 3A. The largest CSPs upon increasing the pH occurred in strands β2 and β3 and to the adjacent region of α-helix. Notably, the largest CSP involved 73H in strand β2: this perturbation and the neighboring effects likely reflect deprotonation of the imidazole side chain in the histidine residue upon shifting the pH to more basic conditions. In addition to the CSPs highlighted in Figure 3A, several resonances in the CRES loop (99L, 100N, 101N, 102T, and 104N) experience extreme line broadening and are undetectable upon shifting the pH to 7.5. A set of peaks in the homologous region were also missing in the NMR data for monomeric V57G cystatin C mutant, which was collected at pH 7.4 and 50 mM NaCl (22). This commonality suggests that the AS structure in cystatin C, which is equivalent to the CRES loop here, also changes structure under physiological conditions. The disappearance of peaks in an HSQC spectrum, as seen here, is indicative of 'intermediate' conformational exchange on the millisecond timescale between the observable 'ground' state and a lowly and transiently populated excited state. Thus, our ^{15}N,^{1}H HSQC data suggested that motions within CRES, including the flexible CRES loop, are present at physiological pH.

We next compared the difference in the ^{15}N transverse relaxation (R_2) rates from the initial and pH 7.5 high salt conditions (250 vs 173 mM) (Figure 3B, blue points). The ^{15}N R_2 is predominantly sensitive to the nanosecond timescale of global macromolecular tumbling as well as the amplitude and picosecond timescale of local amide bond motion, although millisecond timescale motions can have an additive effect as well (see below) (34). Therefore, differences in oligomeric state or protein dynamics upon shifting the solution conditions to pH 7.5 could be manifested as a change in ^{15}N R_2 (12). Generally, we observed an average difference in R_2 ($\Delta R_2 = R_{2,\text{ initial}} - R_{2,\text{ pH 7.5}}$) of -3.22 s^{-1} (blue dashed line in Figure 3B). Although a reason for this difference could be from overall rigidification and/or slower tumbling of the observed CRES (see SI Appendix), another explanation for the larger R_2 values at pH 7.5 is that monomeric CRES is rapidly exchanging with a much larger oligomeric form of CRES, which would have much larger R_2 rates. If this exchange process is faster than R_2, then $R_{2,\text{ pH 7.5}}$ would be the population weighted average of the rates from the monomeric and oligomeric forms (12). We hypothesize that this larger tumbling time was the result of the early stages of amyloidogenesis with CRES C48A subunits associating – either stably or transiently. We also observed larger differences in

ΔR_2 (both positive and negative) for several residues (Figure 3B). These ΔR_2 values likely reflect changes in amide dynamics that accompany differences in structure.

We also collected ^{15}N Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments, which are a powerful method to quantify conformational exchange processes that occur on the millisecond timescale. Transitions to these lowly populated ‘excited’ states are important for a range of biological functions including enzyme catalysis and protein folding (35). Under the initial buffer conditions, CRES displayed flat dispersion curves indicating that an exchange process on this timescale is not present (Figure 3C, black points). However, when the solution conditions were shifted to pH 7.5, typical dispersion curves, with $R_{2,eff}$ decreasing with increasing ν_{CPMG} as shown in Figure 3C (blue points), were obtained. These profiles were only observed for residues Q77 – M82 and A125, L126, N129, E131 and F132, which are in the L1 and L2 loops respectively. Since these residues have little or no CSPs upon switching to the low salt buffer and are on the other side of the protein from the CRES loop residues that experience extreme line broadening, we suggest that this exchange phenomenon is separate from the process(es) causing CSP, line broadening, and the observed ΔR_2. Relaxation dispersion data collected at one static magnetic field strength, corresponding to ^1H field of 600 MHz, were fit to a global two-site exchange model to derive $k_{ex} = 840 \pm 180\ s^{-1}$ and $p_b = 1.2\ \% \pm 0.2\%$ (Table S4). Together, the comparison of solution state NMR data at 250 mM NaCl and pH 6, where CRES exists as a stable monomer, and 173 mM NaCl and pH 7.5, which may promote the earliest stages amyloidogenesis, suggests that the initial steps of amyloid formation are characterized by changes in the structure and dynamics of the α-helix, the CRES loop connecting strands β3 and β4, and the opposing edge of the β-sheet (i.e., L1 and L2).

We next sought to drive the formation of CRES amyloid further; therefore, we prepared a ^{15}N-labeled CRES sample in no salt and pH 7.5 buffer, which we have previously shown rapidly accelerates the formation of CRES oligomers (19), and measured the same suite of ^{15}N

relaxation experiments. Initial characterization of the $^{15}N,^{1}H$ HSQC spectra (Figure 3A) revealed that, in addition to the CSPs observed upon shifting the pH to 7.5 and 173 mM NaCl, more perturbations occurred when salt was omitted from the buffer (average of all CSPs = 0.04, SI Appendix, Table S3). These CSPs now include the majority of the residues in CRES, including perturbations in the β-sheet. Notably, peaks in the CRES-loop between strands β3 and β4 were still missing under these conditions (residues L99, N100, N101, and N104), indicating that the millisecond timescale dynamics observed at 173 mM NaCl for this region of the protein were still present in the no salt buffer condition. We next compared the ^{15}N R_2 relaxation rates of the amide groups under the original and pH 7.5 /no salt buffer conditions (Figure 3B, red points). Again, a generally uniform decrease in the R_2 rates were observed with an average ΔR_2 ($R_{2,\ \mathrm{initial}} - R_{2,\ \mathrm{pH\ 7.5\ /\ no\ salt}}$; red dashed line in Figure 3B) of $-9.09\ s^{-1}$. These data suggest that monomeric CRES is rapidly exchanging with a larger oligomeric form and that this exchange is greater under these conditions (see SI Appendix). Finally, we also measured ^{15}N relaxation dispersion data on the pH 7.5 / no salt sample (Figure 3C, red points). Here, we observed relaxation dispersion profiles in the L2 loop connecting strands β4 and β5 (A125, L126, W128, G130) and in the L1 loop. Fitting these data to a two-state exchange model resulted in a $k_{ex} = 1220 \pm 290\ s^{-1}$ and $p_b = 3.1\ \% \pm 0.7\%$ (SI Appendix, Table S4), which are consistent with the notion that monomeric CRES is undergoing more exchange in the pH 7.5 / no salt buffer condition.

In total, our solution state NMR relaxation data present a picture of CRES sampling an oligomeric conformation using interactions on either end of the β-sheet (i.e., L1/L2 loops and CRES loop). Although the role of the L1 loop in domain swapping and oligomerization is well established for several cystatins, the flexible CRES loop, that is also present in all family 2 cystatins, has largely been ignored. Interestingly, the intermediate exchange line broadening for CRES loop resonances observed in the HSQC spectra and the CPMG data observed for L1/L2 residues imply that these two mechanisms for amyloidogenesis are distinct and complementary processes, as these two NMR observations could be on different timescales (see SI Appendix).

Thus, our data suggest the CRES loop may also participate in oligomerization but by an assembly mechanism distinct from domain swapping.

SSNMR chemical shift assignments. A sample of ^{15}N, ^{13}C-labeled CRES C48A was sonicated into a stable amyloid form and packed into an SSNMR rotor. This sample formed a complex branched amyloid matrix similar to that in the proximal epididymal lumen (see below, Figure 6D) (4, 18). As previously reported, β-sheet and random coil ^{13}C chemical shifts dominated the SSNMR spectra of the advanced amyloid state of CRES (19) . The abundance of antiparallel β-sheet secondary structure produced significant chemical shift overlap, which was worsened by natural ^{13}C linewidths of 0.5 ppm (75 Hz). These broad lines are likely due, in part, to sample polymorphism as discussed below. To address the issue of peak overlap, in addition to 2D dipole-assisted rotational resonance (DARR; 12 ms, 25 ms and 50 ms mixing times) and NcaCX and NcoCX spectra, we acquired three dimensional (3D) CANcoCA (Cα-N-c'-$C_{\alpha i-1}$), NCACX (N_i-$C\alpha_i$-$C_{i,sidechain}$), and NCOCX (N_i-C'_{i-1}-$C_{i-1,sidechain}$) spectra. The signal-to-noise of these spectra was improved using double exponential weighted sampling and following protocols we and others previously reported (36, 37). We assigned all backbone ^{15}N and ^{13}C and most sidechain resonances for 112 residues stretching from Q32 through V142, with limited ambiguity. This process required high-resolution processing of data sets to resolve correlations along the backbone, with processing optimized for signal-to-noise to identify long side chain spin systems. Proline backbone resonances were confidently assigned using the CANcoCA spectrum. We depict backbone walks through key regions of the protein, including the residues found in the α-helix of the X-ray crystal structure (SI Appendix, Figure S5), L1 loop region (SI Appendix, Figure S6), CRES loop (SI Appendix, Figure S7) and the C-terminus (SI Appendix, Figure S8). The CANcoCA spectrum was essential to identification of unambiguous chemical shift correlations. The completeness of the SSNMR assignments is depicted in an extensively assigned DARR spectrum (Figure 4, SI Appendix, Figure S9), strip plots (SI Appendix, Figures S5-S8), and is explicitly illustrated in SI Appendix, Figure S10. All observed cysteine Cβ chemical shifts were indicative of oxidized sidechains, signifying that disulfide bonds formed in the monomeric globular state are preserved throughout amyloidogenesis. The DARR spectra further revealed the relative rigidity of the amyloid structure, as multiple Cα-Cα cross peaks (correlations between nearest neighboring Cα residues) appear at relatively short (25 ms) DARR mixing times, indicating little motional attenuation of ^{13}C-^{13}C through-space couplings. The completeness of our chemical shift assignments allowed us to perform a detailed analysis of the differences between the globular monomeric structure and the amyloid form, as we describe below.

We calculated the changes in Cα and Cβ chemical shifts between solution and solid-state NMR assignments ($CSP = \delta_{C_{solution}} - \delta_{C_{solid}}$, where δ_C is the Cα or Cβ chemical shift in ppm) to determine the change in local structure (SI Appendix, Figure S11). We observed 34 significant (1.5 ppm) Cα CSPs and mapped the absolute values onto the liquid-state NMR structure (Figure 5A). Among these 34 residues, 14 were located in the α-helix found in the X-ray crystal structure revealing changes in the protein fold. The other 20 Cα CSPs were distributed throughout the rest of the protein but predominated in the L1, L2, and CRES loops. The appearance of multiple CSPs in these loops implied there may be some reorientation of the β-strands in the amyloid complex, while the strands themselves largely retained their original structure. The data further indicate the CRES loop may indeed participate in amyloid assembly because the chemical shifts now appear to have β-sheet secondary structure.

We then calculated the secondary structure of the protein using TALOS-N. This analysis confirmed that the globular β-sheet was largely conserved. However, the dihedral angles predicted from SSNMR chemical shifts deviated substantially in other regions of CRES (Figure 5B) from the secondary structure propensities derived from the liquid-state NMR chemical shifts which were in good agreement with the X-ray crystallographic structure (Figure 2B). Here, we observed dramatic changes for residues A43 – E59, situated in the α-helix of the X-ray crystal structure, the L1 loop, and the CRES loop. TALOS-N predicted the long α-helix transitioned to a β-strand. The L1 loop also shifted toward β-strand secondary structure. As shown in Figure 5B, TALOS-N predicted a continuous β-strand was formed by β-strands 2 and 3 and the L1 loop. This observation is bolstered by the large CSPs observed in the L1 loop, which all shifted toward canonical $^{13}C\alpha$ and $^{13}C\beta$ β-sheet chemical shifts (Figure 5A). The formation of a long β-strand is consistent with domain swapped structures observed in other cystatin proteins (26, 38), thus this observation is not unexpected. Additionally, the chemical shifts also indicate the CRES loop can form two short β-strands. We further established the increased rigidity of the protein by calculating the random coil index S^2 with TALOS-N. The predicted S^2 from solid-state chemical shifts were generally higher compared to the globular form of the protein (Figure 5C). This was especially true for the CRES loop region and the L1 loop, which may coincide with the formation of extended secondary structure.

To further our chemical shift analysis, we used AmylPred2, a web tool which uses the consensus from 11 different methods, to predict the amyloid prone regions of CRES (39). In Figure 5A (right), residues colored in grey have no predicted amyloid propensity, whereas other residues are shaded from green to purple depending upon the number of AmylPred2 predictors that selected that residue as amyloid prone. Four residues found in the α-helix were predicted to be amyloidogenic by all eleven components, and 88 residues overall were picked to have some amyloid propensity. It should be noted that even though AmylPred2 did not predict that residues within the CRES loop were amyloid prone, the shift toward β-sheet secondary structure (Figure 5B) implies that these residues could still play a role in amyloid formation. Together, TALOS-N and AmylPred2 suggest that the α-helix, L1 loop, CRES loop, and multiple regions of the preexisting, preassembled β-strands are pivotal regions for the transition into the amyloid fold. The many amyloid-prone sites found in the β-sheet indicate the CRES amyloid may come partially assembled within the globular protein. This could further help to explain the unusually fast rate (days vs weeks) of amyloid formation in CRES proteins compared to pathological amyloidogenic proteins. This also could help to explain why previous FTIR and CD measurements of CRES aggregates exhibited much higher antiparallel vs. parallel β-sheet content, which runs contrary to commonly observed amyloid structures (19). Many of the globularly-native intramolecular antiparallel β-sheets are conserved within the amyloid matrix and in turn nucleate intermolecular antiparallel β-sheets.

Determination of a CRES model and implications for the complexity of CRES amyloid quaternary structure. Our experimental chemical shifts strongly suggest β-strands 2 and 3 and the L1 loop form an elongated β-strand within the mature amyloid matrix. This longer β-stand may indicate a domain-swapped structure as observed in other cystatins or could indicate this region stacks with other proteins analogously to many other reported amyloids. Our data also suggest β-sheet assemblies may occur through rigidification of the CRES loop. To generate possible structural models for the CRES amyloid, fragments were selected using the SSNMR chemical shift assignments and CS-ROSETTA and these were employed along with C2 symmetry in the ROSETTA Fold-and-Dock algorithm (40). While this model is dramatically underdetermined, it reveals key features which could help to explain the web-like, rather than fibrillar, structures observed via EM (Figure 6D).

The lowest energy structural model in best agreement with the SSNMR chemical shifts is shown in Figure 6A (in the top panel each monomer is colored blue or yellow, below this all residues predicted to be β-sheet by TALOS-N are highlighted in green). It is composed of six intra- and intermolecular β-strands as compared to five intramolecular β-strands found in the X-ray and solution NMR structures of monomeric CRES. In agreement with the TALOS-N analysis, CS-ROSETTA predicted the long α-helix is converted to a β-strand (residues A43 – E59) in the amyloid fold of the protein, and β2-L1-β3 form a long intermolecular strand, with this region of the first protomer forming an antiparallel sheet with the same region of the second (possibly reflective of domain swapping) (Figure 6A). Another key feature predicted by this structure is that the amyloid matrix is composed of mostly antiparallel β-sheets rather than parallel β-sheets observed in many pathogenic amyloid structures. We previously reported CD and FTIR measurements indicative of antiparallel β-sheets in the early CRES aggregates (19). The SSNMR ^{13}C' chemical shifts reported here further support this conclusion in the advanced CRES amyloid. Backbone ^{13}C' chemical shifts are known to be highly sensitive to hydrogen bond length and geometry (41-44). Antiparallel β -sheets should have a slightly higher ^{13}C' chemical shift and narrower distribution (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(45). 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 (14, 15, 46-48) and well-established antiparallel globular proteins (49-52) (Figure 6B,SI Appendix, Figures S12,13), suggest the majority of β -sheets found in the CRES amyloid are antiparallel. Further justification is provided in the SI Appendix. Lastly, the CRES loop in these CS-ROSETTA models is transitioning from its original random coil secondary structure towards a β-strand as predicted by TALOS-N. It does not transition completely to β-strand because within the limitations of the modelling software, which relies upon molecular symmetry, there is no packing partner.

It is unlikely the structure of the CRES amyloid matrix has any intrinsic symmetry. Although SSNMR suggests there may be a basic unit, which could be a domain-swapped core, the extensive β-sheet assemblies that may form via the CRES loop and other extremities of the model are likely not detected because they are less frequent and are therefore not detectable in this SSNMR data. Indeed, there are multiple amyloid-prone contact surfaces branching out from the core of our model, implying extensive possibilities for the formation of intermolecular β-sheets. For example, the N-terminus, CRES loop, β4 and β5 are capable of branching out from the core to form a complex matrix. In addition, multiple repeats of a β2-L1-β3 domain-swapped β-sheet could stack in a similar manner to fibrillar amyloids. It is thus possible that within the quaternary structure of the amyloid matrix multiple intermolecular β-sheets could form along surfaces on multiple neighboring proteins. Consequently, as CRES aggregates, many of these β-strands intermingle, producing a far more complicated tertiary and quaternary structure. This β-sheet rich structure is further supported by X-ray fiber diffraction of the SSNMR sample (Figure 6C), where the indicative 4.8 Å and 10.9 Å β-strand and β-sheet spacings are detected. The observation of multiple possible branch points is supported by negative stain TEM images of the SSNMR sample (Figure 6D). Similar to previously reported TEM images of both *in vivo* and *in vitro* samples, our CRES sample forms a webbed matrix (4, 18). Together, these data support the formation of an aggregated protein containing the β-sheet stacking characteristic of many amyloids, but uniquely prone toward antiparallel β-sheet secondary structures and a more complicated (non-fibrillar) quaternary organization.

Functional amyloid assembly. Our studies presented herein show at the atomic level the assembly of a functional mammalian amyloidogenic monomer to advanced amyloid. Using a combination of three powerful structural biology techniques, we not only reveal the structure of the CRES monomer and its advanced amyloid form but provide critical knowledge of how a biological amyloidogenic precursor initiates its assembly into early oligomeric forms under physiological conditions. Together, our data suggest that CRES amyloid formation is complex and likely utilizes several mechanisms including a unique interaction driven by changes in the structure of a CRES loop from a flexible linker in the monomer to a β-strand conformation in the advanced amyloid as well as traditional domain swapping typical of other cystatins. We speculate that one role of the CRES loop may be to guide and control the building of highly branched and elaborate amyloid matrices whose β-sheet assemblies with neighboring proteins, including other amyloidogenic precursors, may be essential for biological function. Indeed, early mutagenesis studies in which three asparagine and one threonine residue in the CRES loop were replaced with alanine residues revealed a monomer that no longer polymerized into a metastable oligomer typical of wildtype CRES but rather which rapidly progressed to large aggregates suggestive of an uncontrolled amyloid assembly process (19)(SI Appendix, Figure S14). Structural heterogeneity, including lateral attachments, from a basic structural unit has also been observed in the fungal hydrophobin $EAS_{\Delta 15}$ amyloid, a functional truncated variant which forms an extended amphipathic monolayer at hydrophobic:hydrophilic interfaces (53). Based on mutagenesis experiments and docking models, a highly mobile, disulfide-anchored and asparagine-rich loop was proposed to be central to the conformational transition of EAS to its antiparallel β-sheet-rich amyloid form (54). The intriguing similarity of this model with our demonstrated role of the CRES loop in mammalian amyloid assembly suggests that rigidification of a flexible loop may be an evolutionarily conserved mechanism for the controlled assembly of some functional amyloids. The use of multiple assembly mechanisms could also allow functional amyloids to adopt very different geometries if one mechanism dominated over the other, which could be mediated by local environmental conditions. Further, the use of distinctive mechanisms to assemble amyloid within a common structure could also be a means to selectively disassemble specific β-sheet interactions allowing a highly dynamic and plastic functional amyloid structure.

Materials and Methods

Protein expression and purification. Tag-less CRES C48A was purified from the soluble fraction of bacteria using affinity, ion exchange, and gel filtration chromatography as previously described (19). Mass spectra of intact and tryptic digests of CRES isoforms were acquired using a MALDI-TOF/TOF 4800 mass spectrometer (Applied Biosystems) as detailed in the SI Appendix.

X-ray crystallization. CRES C48A crystals were grown using the hanging drop method and shipped to Stanford Linear Accelerator Center beamline 7-1 for high resolution data collection. The structure was solved as detailed in the SI Appendix.

Solution-state and MAS solid-state NMR. All solution-state NMR data were recorded on a 600 MHz (^{1}H Larmor frequency; 14 T) Agilent DD2 spectrometer equipped with a z-axis gradient room temperature HCN probe. Solid-state NMR experiments were carried out on a 600 MHz Agilent DD2 three-channel spectrometer equipped with an HCN Baun probe. Details are provided in SI Appendix.

Negative Stain TEM. CRES C48A was spotted on to formvar/carbon coated 200 mesh nickel grids (Ted Pella, Redding, CA) as described (4).

Data availability. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 6UIO). Solution-state NMR data ID 50273 and solid-state NMR data ID 50275 were deposited in the Biological Magnetic Resonance Bank.

Acknowledgements: The authors would like to thank Masoud Zabet, Ph.D., of the TTU Proteomics core for the mass spectrophotometric analysis of CRES isoforms and the preparation of Fig S1B. The authors acknowledge the High Performance Computing Center (HPCC) at Texas Tech University at Lubbock for providing HPC resources that have contributed to the research results reported within this paper (http://www.depts.ttu.edu/hpcc/). Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. Supported by TTUHSC, The Wilson Foundation, and The Newby Family (GAC), NIH R01AR063634 (RBS), NIH 1R35GM128906 (MPL), NIH 1R35GM124979 (BJW).

References

1. Fowler DM, *et al.* (2006) Functional amyloid formation within mammalian tissue. *PLoS Biol.* 4(1):100-107.
2. Boke E, *et al.* (2016) Amyloid-like self-assembly of a cellular compartment. *Cell* 166(3):637-650.
3. Raveendra BL, *et al.* (2013) Characterization of prion-like conformational changes of the neuronal isoform of Aplysia CPEB. *Nat. Struct. Mol. Biol.* 20(4):495-+.
4. Whelly S, *et al.* (2012) Nonpathological extracellular amyloid is present during normal epididymal sperm maturation. *Plos One* 7(5).
5. Roan NR, *et al.* (2017) Semen amyloids participate in spermatozoa selection and clearance. *Elife* 6.
6. Guyonnet B, Egge N, & Cornwall GA (2014) Functional amyloids in the mouse sperm acrosome. *Mol. Cell. Biol.* 34(14):2624-2634.
7. Egge N, Muthusubramanian A, & Cornwall GA (2015) Amyloid properties of the mouse egg zona pellucida. *Plos One* 10(6).
8. Maji SK, *et al.* (2009) Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325(5938):328-332.
9. Pfefferkorn CM, McGlinchey RP, & Lee JC (2010) Effects of pH on aggregation kinetics of the repeat domain of a functional amyloid, Pmel17. *Proceedings of the National Academy of Sciences of the United States of America* 107(50):21447-21452.
10. Whelly S, *et al.* (2016) Cystatin-related epididymal spermatogenic subgroup members are part of an amyloid matrix and associated with extracellular vesicles in the mouse epididymal lumen. *Mol. Human Reprod.* 22(11):729-744.
11. Fitzpatrick AWP, *et al.* (2017) Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* 547(7662):185-+.
12. Fawzi NL, Ying JF, Ghirlando R, Torchia DA, & Clore GM (2011) Atomic-resolution dynamics on the surface of amyloid-beta protofibrils probed by solution NMR. *Nature* 480(7376):268-U161.
13. Walti MA, *et al.* (2016) Atomic-resolution structure of a disease-relevant A beta(1-42) amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* 113(34):E4976-E4984.
14. Colvin MT, *et al.* (2016) Atomic Resolution Structure of Monomorphic A beta(42) Amyloid Fibrils. *J. Am. Chem. Soc.* 138(30):9663-9674.
15. 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-+.
16. Mompean M, *et al.* (2018) The Structure of the Necrosome RIPK1-RIPK3 Core, a Human Hetero-Amyloid Signaling Complex. *Cell* 173(5):1244-+.
17. Van Melckebeke H, *et al.* (2010) Atomic-Resolution Three-Dimensional Structure of HET-s(218-289) Amyloid Fibrils by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132(39):13765-13775.
18. Cornwall GA, Do HQ, Hewetson A, Muthusubramanian A, & Myers C (2019) The epididymal amyloid matrix: structure and putative functions. *Andrology*:1-7.
19. 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.
20. Kolodziejczyk R, *et al.* (2010) Crystal structure of human cystatin C stabilized against amyloid formation. *FEBS J.* 277(7):1726-1737.
21. 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.

22. 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.*
23. Janowski R, Kozak M, Abrahamson M, Grubb A, & Jaskolski M (2005) 3D domain-swapped human cystatin c with amyloidlike intermolecular beta-sheets. *Proteins-Structure Function and Bioinformatics* 61(3):570-578.
24. Ghiso J, Jensson O, & Frangione B (1986) Amyloid fibrils ini hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variaant of gamma trace basic protein (cystatin C). *Proc. Natl. Acad. Sci. USA* 83:2974-2978.
25. Orlikowska M, Jankowska E, Kolodziejczyk R, Jaskolski M, & Szymanska A (2011) Hinge-loop mutation can be used to control 3D domain swapping and amyloidogenesis of human cystatin C. *Journal of Structural Biology* 173(2):406-413.
26. Dall E*, et al.* (2018) Structural and functional analysis of cystatin E reveals enzymologically relevant dimer and amyloid fibril states. *J. Biol. Chem.* 293(34):13151-13165.
27. Cornwall GA*, et al.* (2003) The cystatin-related epididymal spermatogenic protein inhibits the serine protease prohormone convertase 2. *Endocrinology* 144(3):901-908.
28. Wilmot CM, Thornton JM (1988) Analysis and prediction of the different types of beta-turn in proteins. *J. Mol. Biol.* 203(1):221-232.
29. Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22(12):2577-2637.
30. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372(3):774-797.
31. 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.
32. Berjanskii MV & Wishart DS (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* 127(43):14970-14971.
33. Powers R, Clore GM, Garrett DS, & Gronenborn AM (1993) RELATIONSHIPS BETWEEN THE PRECISION OF HIGH-RESOLUTION PROTEIN NMR STRUCTURES, SOLUTION-ORDER PARAMETERS, AND CRYSTALLOGRAPHIC B-FACTORS. *Journal of Magnetic Resonance Series B* 101(3):325-327.
34. Jarymowycz VA & Stone MJ (2006) Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chemical Reviews* 106(5):1624-1671.
35. Sekhar A & Kay LE (2013) NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proceedings of the National Academy of Sciences of the United States of America* 110(32):12867-12874.
36. Palmer MR*, et al.* (2015) Sensitivity of Nonuniform Sampling NMR. *Journal of Physical Chemistry B* 119(22):6502-6515.
37. Amani R*, et al.* (2020) Conformational changes upon gating of KirBac1.1 into an open-activated state revealed by solid-state NMR and functional assays. *Proceedings of the National Academy of Sciences of the United States of America* 117(6):2938-2947.
38. Janowski R*, et al.* (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat. Struct. Biol.* 8(4):316-320.
39. Tsolis AC, Papandreou NC, Iconomidou VA, & Hamodrakas SJ (2013) A Consensus Method for the Prediction of 'Aggregation-Prone' Peptides in Globular Proteins. *Plos One* 8(1).

40. 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.
41. 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-+.
42. Wei Y, Lee D, & Ramamoorthy A (2001) Solid-state C-13 NMR chemical shift anisotropy tensors of polypeptides. *Journal of the American Chemical Society* 123(25):6118-6126.
43. 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.
44. 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.
45. Pauling L and Corey RB (1951) The pleated sheet, a new layer of configuration of polypeptide chains. *Proceedings of the National Academy of Sciences* 37:251-256.
46. 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.
47. Lu JX*, et al.* (2013) Molecular Structure of beta-Amyloid Fibrils in Alzheimer's Disease Brain Tissue. *Cell* 154(6):1257-1268.
48. 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.
49. 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.
50. Igumenova TI*, et al.* (2004) Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126(21):6720-6727.
51. Igumenova TI, Wand AJ, & McDermott AE (2004) Assignment of the backbone resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126(16):5323-5331.
52. Fasshuber HK*, et al.* (2015) Structural heterogeneity in microcrystalline ubiquitin studied by solid-state NMR. *Protein Sci.* 24(5):592-598.
53. Morris VK*, et al.* (2012) Solid-State NMR Spectroscopy of Functional Amyloid from a Fungal Hydrophobin: A Well-Ordered beta-Sheet Core Amidst Structural Heterogeneity. *Angewandte Chemie-International Edition* 51(50):12621-12625.
54. Macindoe I*, et al.* (2012) Self-assembly of functional, amphipathic amyloid monolayers by the fungal hydrophobin EAS. *Proceedings of the National Academy of Sciences of the United States of America* 109(14):E804-E811.

Figure 1. X-ray crystal structure of CRES. (A) Ribbon diagram of the CRES monomer highlighting the CRES loop in red and the two disulfide bonds as yellow sticks. **(B)** Superposition of the four CRES molecules in the crystallographic asymmetric unit (PDB code 6UIO). **(C)** Inter-molecular interaction of two CRES molecules in the crystallographic asymmetric unit. Highlighted in pink is the CRES loop and β3 of one molecule forming a parallel β-sheet interaction with β5 of an adjacent molecule. **(D)** Space fill model of the two interacting CRES molecules.

Figure 2. Solution state NMR on monomeric CRES. (A) 2D ^{15}N,^{1}H HSQC spectrum of 1.4 mM CRES in 25 mM MES, 250 mM NaCl, 1 mM EDTA, pH 6 at 25 °C. Resonance assignments are given as the residue number and amino acid single letter code. **(B)** Secondary structure propensity of solution state CRES derived from backbone chemical shifts and TALOS-N. The secondary structure observed in the crystal structure is diagrammed above. **(C)** Comparison of the average b-factors from the four CRES monomers in the crystallographic asymmetric unit (green points) and the predicted random coil index (RCI) order parameter (S^2) (blue points), which were also calculated from the backbone chemical shifts using TALOS-N. The standard deviations of the crystallographic b-factors of the four monomers are given as grey error bars.

Figure 3. Changes in the structure and dynamics of CRES upon amyloidogenesis. (A) Left, overlay of a region from the 2D ^{15}N,^{1}H HSQC spectra of 1.4 mM CRES at pH 6 / 250 mM NaCl (black contours), pH 7.5 / 173 mM NaCl (blue contours), and pH 7.5 / 0 mM NaCl (red contours). Right, calculated chemical shift perturbations (CSPs) mapped onto the crystal structure of CRES. The CSPs between pH 6 / 250 mM NaCl and pH 7.5 / 173 mM NaCl are given as a grey-to-blue gradient, and the CSPs between pH 6 / 250 mM NaCl and pH 7.5 / 0 mM NaCl are given as a grey-to-red gradient. **(B)** Plot of the ΔR_2 versus residue number. Blue points are the differences between pH 6 / 250 mM NaCl and pH 7.5 / 173 mM NaCl, and the red points are the differences between pH 6 / 250 mM NaCl and pH 7.5 / 0 mM NaCl. The error in

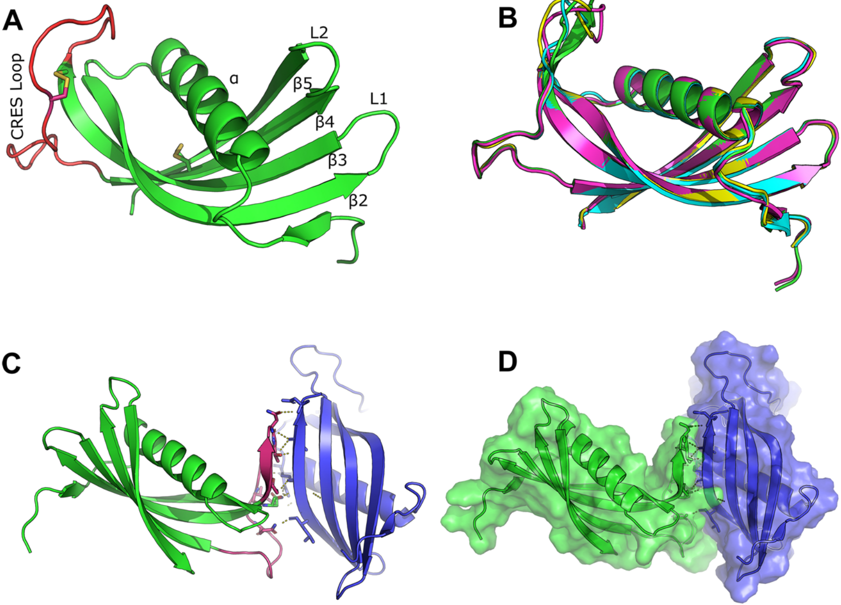
the values, propagated from the covariance matrix of the fits, are shown as grey bars. The blue and red dashed lines highlight the average difference in ΔR_2 for the two conditions. **(C)** Representative relaxation dispersion CPMG profiles are shown for CRES at pH 6 / 250 mM NaCl (black circles), pH 7.5 / 173 mM NaCl (blue triangles), and pH 7.5 / 0 mM NaCl (red squares) collected at 600 MHz and 25 C. Error bars are derived from duplicate measurements of $R_{2,eff}$ at two CPMG fields. Solid lines are fits to the Bloch-McConnell equations describing two-site exchange. Below the plots, the residues for which chemical exchange phenomenon are observed are shown as sticks on the crystal structure.

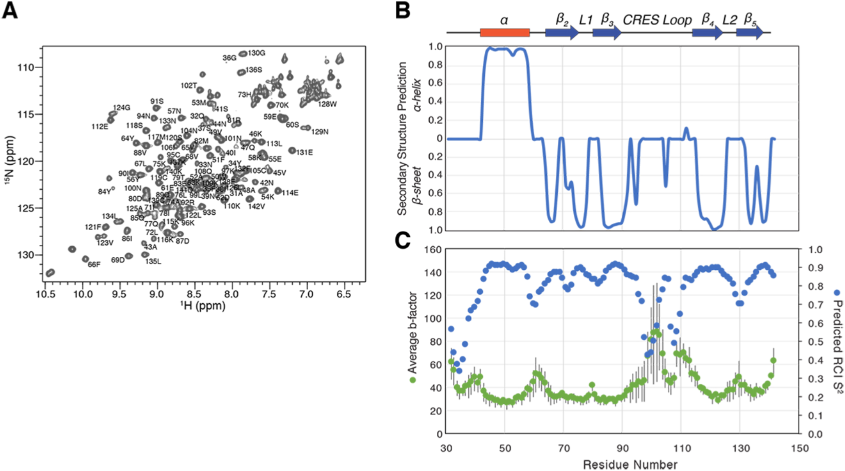
Figure 4. SSNMR assignments of an advanced amyloid form of CRES. 2D ^{13}C,^{13}C DARR spectrum of CRES. The spectrum was recorded with a mixing time of 25 ms at a magnetic field of 14 T. All assigned correlations are labeled with residue number.

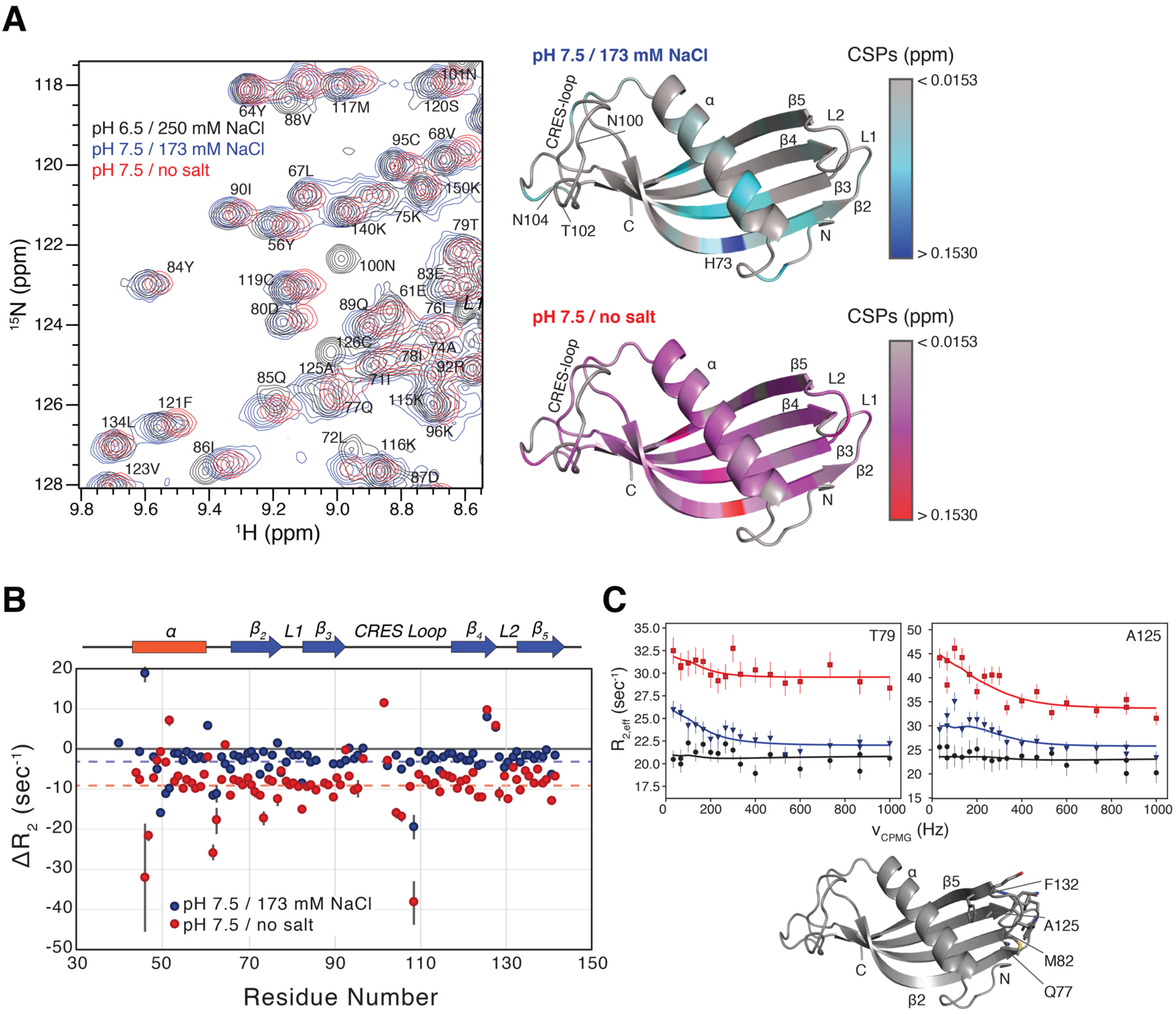
Figure 5. Structural analysis of CRES using SSNMR chemical shift assignments. (A) The absolute value of CSPs between solid-state and solution state NMR mapped onto solution-state NMR structure (shaded from grey to red) and AmylPred2 predicted amyloidogenic sites of the protein (shaded from grey (no prediction) to green (at least 1 prediction) to purple (unanimous prediction). **(B)** Protein secondary structure predicted from SSNMR assigned chemical shifts using TALOS-N. Secondary structure components of CRES X-ray crystal structure and solution NMR are presented in red rectangles for α-helices and blue arrows for β-sheets. **(C)** Dot plot of predicted RCI S^2 from SSNMR (red) and solution NMR (green) chemical shifts.

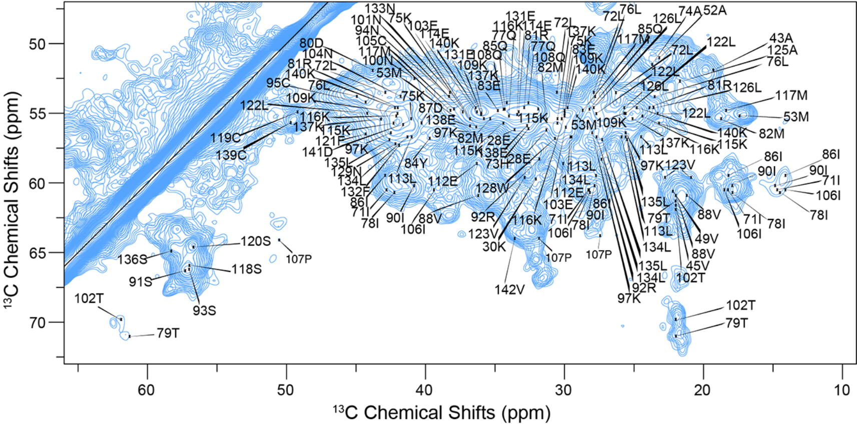
Figure. 6. Structural model of the CRES amyloid and X-ray diffraction and TEM images of advanced CRES amyloid matrix formed by the SSNMR sample. (A) Domain-swapped structures with monomeric units colored in blue and yellow (top) and regions predicted to form β-sheets shaded in green (bottom). Numbering of secondary structure is based on Figure 1. **(B)** Comparison of average and standard deviation of $^{13}C'$ chemical shifts measured in advanced

CRES amyloid compared to average and standard deviation of well-defined parallel and antiparallel-sheet structures. The dark horizontal line in each box signifies the average, the box signifies the 68% percent confidence interval, and the vertical line is the full spread of values. Color-coded dots are actual values reported for different proteins listed on the right. **(C)** X-ray fiber diffraction of CRES amyloid matrix. X-ray fiber diffraction exhibited reflections at 4.8 Å and 10.9 Å, indicative of the spacing between β-strands and β-sheets, respectively. **(D)** Negative stain TEM image reveals a complicated web-like matrix formed by the CRES protein studied by SSNMR.

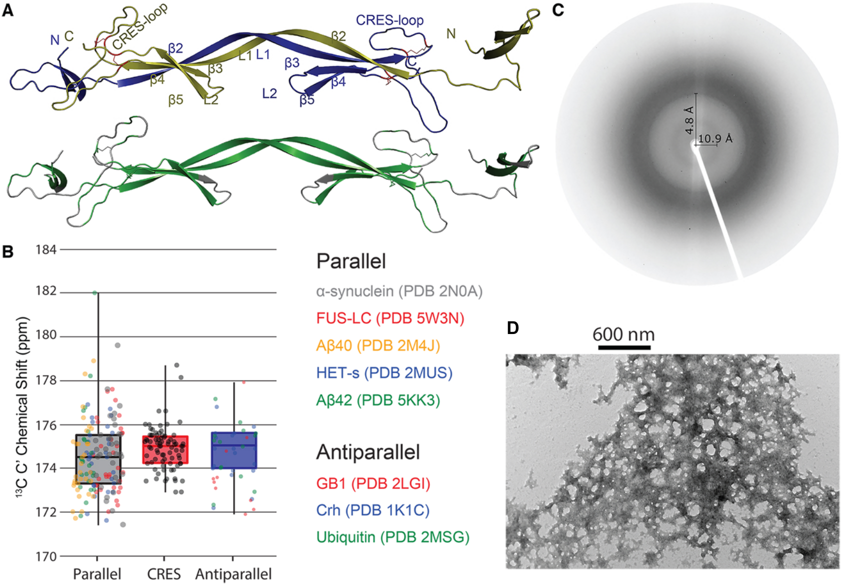

A
CRES Loop
L2
α
β5
β4
L1
β3
β2
B
C
D


A
B
C
¹⁵N (ppm)
¹H (ppm)
110
115
120
125
130
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
36G
130G
136S
102T
73H
53M
41S
70K
128W
124G
91S
57N
94N
133N
32Q
44N
81R
59E
60S
129N
112E
118S
104N
37S
49V
46K
117M
120S
64Y
82M
101N
113L
106I
65V
47Q
88V
95C
68V
40I
131E
51F
58K
55E
67L
75K
137K
33N
34Y
132F
105C
45V
90I
140K
108Q
97K
56Y
119C
79T
52A
50W
138E
42N
100N
83E
63K
109K
84Y
61E
141D
35F
38I
126L
48A
114E
89Q
76L
99L
39N
62D
31A
54K
80D
139C
74A
92R
110K
142V
125A
71I
93S
78I
122L
134L
85Q
77Q
115K
96K
121F
72L
86I
123V
116K
87D
43A
69D
135L
66F
α
β2
L1
β3
CRES Loop
β4
L2
β5
Secondary Structure Prediction
α-helix
β-sheet
1.0
0.8
0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1.0
Average b-factor
Predicted RCI S²
160
140
120
100
80
60
40
20
0
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
30
50
70
90
110
130
150
Residue Number


A
pH 6.5 / 250 mM NaCl
pH 7.5 / 173 mM NaCl
pH 7.5 / no salt
^{15}N (ppm)
^{1}H (ppm)
118
120
122
124
126
128
9.8
9.6
9.4
9.2
9.0
8.8
8.6
64Y
88V
117M
120S
101N
95C
68V
67L
90I
150K
75K
140K
56Y
79T
100N
84Y
119C
83E
61E
80D
89Q
76L
126C
74A
78I
92R
85Q
125A
71I
77Q
115K
121F
134L
96K
72L
116K
86I
123V
87D
pH 7.5 / 173 mM NaCl
CSPs (ppm)
< 0.0153
> 0.1530
CRES-loop
N100
N104
T102
C
N
H73
α
β5
β4
β3
β2
L2
L1
pH 7.5 / no salt
CSPs (ppm)
< 0.0153
> 0.1530
CRES-loop
α
β5
β4
β3
β2
L2
L1
C
N
B
α
β2
L1
β3
CRES Loop
β4
L2
β5
ΔR$_2$ (sec^{-1})
20
10
0
-10
-20
-30
-40
-50
30
50
70
90
110
130
150
Residue Number
pH 7.5 / 173 mM NaCl
pH 7.5 / no salt
C
T79
A125
R$_{2,eff}$ (sec^{-1})
35.0
32.5
30.0
27.5
25.0
22.5
20.0
17.5
50
45
40
35
30
25
20
15
0
200
400
600
800
1000
ν$_{CPMG}$ (Hz)
α
β5
F132
A125
M82
Q77
C
N
β2



133N
101N
75K
103E
94N
114E
105C
140K
80D
117M
131E
108Q
104N
100N
109K
81R
72L
53M
137K
140K
83E
95C
76L
75K
109K
87D
122L
116K
138E
137K
115K
97K
82M
119C
121F
28E
141D
97K
115K
138E
139C
135L
84Y
73H
28E
129N
134L
113L
112E
132F
128W
86I
71I
78I
90I
88V
106I
131E
116K
114E
72L
77Q
81R
137K
85Q
77Q
75K
108Q
83E
109K
82M
140K
115K
53M
109K
113L
134L
112E
103E
86I
71I
90I
78I
106I
92R
116K
123V
30K
107P
107P
142V
76L
72L
126L
74A
52A
85Q
117M
122L
72L
43A
125A
76L
122L
126L
81R
126L
117M
122L
53M
140K
82M
137K
116K
115K
113L
86I
90I
71I
106I
78I
97K
123V
135L
88V
79T
113L
49V
134L
88V
135L
45V
134L
102T
92R
97K
102T
79T
120S
136S
107P
118S
91S
93S
102T
79T
50
55
60
65
70
60
50
40
30
20
10
^{13}C Chemical Shifts (ppm)
^{13}C Chemical Shifts (ppm)


A
N
C
CRES-loop
β2
β3
β4
β5
L1
L1
L2
L2
β2
β3
β4
β5
CRES-loop
C
N
B
184
182
180
178
176
174
172
170
¹³C C' Chemical Shift (ppm)
Parallel
CRES
Antiparallel
Parallel
α-synuclein (PDB 2N0A)
FUS-LC (PDB 5W3N)
Aβ40 (PDB 2M4J)
HET-s (PDB 2MUS)
Aβ42 (PDB 5KK3)
Antiparallel
GB1 (PDB 2LGI)
Crh (PDB 1K1C)
Ubiquitin (PDB 2MSG)
C
4.8 Å
10.9 Å
D
600 nm