Supplementary Information for

Unravelling the effect of N(ε)-(carboxyethyl)lysine on the conformation, dynamics and aggregation propensity of α-synuclein

Laura Mariño, Rafael Ramis, Rodrigo Casasnovas, Joaquín Ortega-Castro, Bartolomé Vilanova, Juan Frau and Miquel Adrover*

Institut Universitari d'Investigació en Ciències de la Salut (IUNICS). Institut d'Investigació Sanitària Illes Balears (IdISBa). Departament de Química, Universitat de les Illes Balears, Ctra. Valldemossa km 7.5, E-07122 Palma de Mallorca, Spain.

*Correspondence to: Miquel Adrover, University of Balearic Islands,
Phone: +34 971 173491; Fax +34 971 173426;
e-mail: miquel.adrover@uib.es

EXPERIMENTAL PROCEDURES

Chemicals and reagents

Pyruvic acid, guanidine hydrochloride (Gdn-HCl) and N^{α}-Ac-Lys were purchased from Sigma-Aldrich, whereas $NaBH_3CN$ was supplied by Across Organics. Sodium phosphate was purchased from Scharlau. All buffer reagents were ACS grade and all solutions were prepared by using milli-Q water.

α-*Synuclein expression and purification*

E. coli BL21(DE3) cells were transformed with the pT7-7 plasmid containing the gene encoding for human αS (kindly provided by Dr. Daniel Otzen [Aarhus University]). Transformed cells were grown in sterilized Luria Bertani media (LB) (25g/l) containing ampicillin (100µg/ml) at 37°C and 180rpm. Cells were also grown in sterilized M9 media supplied with $^{15}NH_4Cl$ and $^{13}C_6$-glucose as the only sources of nitrogen and carbon, respectively. This allowed us to obtain ^{15}N- and ^{13}C-labelled αS. At OD_{600nm}=0.6-0.8 αS expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG) (1mM) and further incubated for additional 4h at 37°C and 180rpm. Afterwards, cells were centrifuged (at 4°C and 4000rpm for 15min) and the resulting pellet was resuspended in lysis buffer (10mM Tris-HCl, 1mM EDTA, 1mM PMSF, pH 8.0) and stirred for 1h at 4°C. Cells were then lysed by three cycles of 2min sonication step. Cellular debris was removed by centrifugation (20min at 10000 x *g* and at 4°C). Nucleic acids were removed from the lysate by adding streptomycin sulfate (1% w/v) and stirring for 1h at 4°C, followed by centrifugation during 30min (13,500 x *g* at 4°C). The obtained supernatant was then supplied by the slow addition of ammonium sulfate (up to 0.295 g/ml), and additionally stirred for 1h at 4°C to induce the precipitation of αS. Thereafter, the pellet was collected by centrifugation (13,500 x *g* at 4°C for 30min), dissolved in 10mM Tris-HCl at pH 7.4 ($1/20^{th}$ of the LB culture media) and filtered through a 0.22µm filter.

The obtained solution was directly loaded onto an anion exchange column (GE Healthcare RESOURCE™ Q; 6ml) using a GE ÄKTA Start FPLC. αS was eluted with a NaCl gradient (0-600mM) using two different solutions (A: 10mM Tris-HCl at pH 7.4; B: 10mM Tris-HCl and 600mM NaCl at pH 7.4) at a flow rate of 4ml/min. The αS fractions were pooled together and dialyzed extensively at 4°C against different phosphate buffers. The purity of the obtained αS was checked by using MALDI-TOF/TOF (**Fig. S2A**) and SDS-PAGE electrophoresis (**Fig. S2B**) (>96%). αS concentration was measured by UV-Vis spectroscopy using a molar extinction coefficient estimated on the basis of its amino acid content: $\varepsilon_{\alpha S\_280nm}$=5960$M^{-1}\cdot cm^{-1}$.[1]

Chemical synthesis of N^{ε}-(carboxyethyl)lysine (CEL) on α-synuclein

Unlabeled or ^{15}N,^{13}C-labeled αS (100-200µM) was incubated in the presence of pyruvic acid (50mM) in 150mM sodium phosphate buffer (pH 7.4) at 50°C for 48h. The

reaction was carried out in the presence of 75mM $NaBH_3CN$, a reagent that selectively reduces the imine groups at neutral pH[2] (**Fig. S3A**). After 48h of incubation, the reaction mixtures were dialyzed against phosphate buffer to remove the excess of the non-protein reagents.

Size exclusion chromatography (SEC)

The monomeric forms of αS and αS-CEL were further purified by Size Exclusion Chromatography (SEC). Purification was carried out at room temperature using a Superdex-75 HR 10/300 column (GE Healthcare) equilibrated with 20mM sodium phosphate and 150mM NaCl at pH 7.4. Aliquots (500μL) were injected into an Äkta purifier FPLC system at a flow rate of 0.5 mL/min coupled to an UV/Vis detector (280nm) (**Fig. S3B**). The fractions corresponding to the main peaks were pooled together, and the purity was assessed using MALDI-TOF/TOF and SDS-PAGE electrophoresis. The concentrations of the αS-CEL solutions were measured by UV-Vis spectroscopy. The addition of CEL moieties on αS does not involve the formation of additional chromophores; thus they do not change the UV-Vis spectrum profile of native αS (**Fig. S4**). Hence, the αS-CEL concentration was measured using the $\varepsilon_{\alpha S\_280nm}$=5960$M^{-1}\cdot cm^{-1}$ estimated for native αS[1].

Sodium Dodecyl Sulfate (SDS)-PAGE Electrophoresis

Samples containing monomeric αS and monomeric αS-CEL were subjected to SDS-PAGE analysis by using 4–20% Mini- Protean TGX precast gels (Bio-Rad). 15μL were mixed with 15μL of Laemmli sample buffer (Bio-Rad) and then loaded onto the gel. Proteins were visualized with PAGE-Blue Staining Solution (Thermo Scientific).

Mass spectrometry study of α-synuclein modified with N^{ε}-(carboxyethyl)lysine (αS-CEL)

Solutions of purified monomeric αS and αS-CEL were subjected to MALDI-TOF/TOF analysis. Before analysis, the samples were dialyzed against milli-Q water to remove salts and thereafter, they were combined in a 1:1 ratio with the matrix solution (10μg of α-cyano-4-hydroxycinnamic acid in a (1:1) water:acetonitrile solution containing 0.1% TFA). The samples were then spotted onto a steel target plate (MTP 384), air-dried, and analyzed. Mass spectra were acquired in a Bruker Autoflex III MALDI-TOF spectrometer equipped with a 200-Hz smart-beam pulsed N_2 laser (λ 337nm). The IS1 and IS2 voltages were 19kV and 16.65kV respectively, and the lens voltage was 8.6kV. Measurements were performed using a positive reflector mode with matrix suppression below 400Da. The spectra were calibrated externally using a protein calibration standard (3600-17000Da) from Bruker. The experiments were done in duplicate.

In addition, a reaction mixture containing native αS (45μM) and MG (50mM) was prepared in 200mM sodium phosphate buffer at pH 7.4. The sample was then incubated at 37ºC, and several aliquots were collected at different incubation times. They were diluted 1:10 in milli-Q water and afterwards, combined in a 1:1 ratio with the matrix solution. The MALDI-TOF/TOF spectra of the resulting solutions were obtained as described before. The MG was purchased from Sigma-Aldrich as a 40% solution and purified by steam-distillation before use, as described earlier[3].

Circular dichroism (CD) measurements

Buffer solutions containing monomeric αS or αS-CEL were subjected to CD measurements. Before analysis, the aliquots were diluted in 200mM phosphate buffer (pH 7.4) to a final protein concentration of 8μM. The CD spectra of the different samples were recorded at 25ºC between 260 and 190nm in a J-715 spectropolarimeter (JASCO) using a 1mm quartz cell. The scan speed was 50 nm/min with a response time of 1s and a band width of 1.0nm, while 15 scans were accumulated. Data were baseline corrected by subtracting the buffer contribution and the average value of the first 15nm of each spectrum to the entire curve. The different CD spectra were used to derive the protein secondary structure content using the BeStSel online platform (http://bestsel.elte.hu/)[4].

NMR spectroscopy measurements

^{15}N- and ^{13}C- double labeled αS (220μM) and αS-CEL (160μM) solutions were used for NMR studies. These solutions were prepared in 20mM sodium phosphate at pH 6.5 in the presence of 150mM NaCl and 10% (v/v) D_2O. All NMR experiments were recorded at 12.5ºC on a Bruker Avance III operating at a ^{1}H resonance frequency of 600.1 MHz, and equipped with a 5-m ^{13}C,^{15}N,^{1}H triple resonance cryoprobe. In all experiments, water suppression was achieved by the watergate pulse sequence[5] and proton chemical shifts were referenced to the water signal fixed at 4.89ppm. ^{13}C and ^{15}N chemical shifts were referenced indirectly using the ^{1}H,X frequency ratios of the zero-point[6]. The spectra were processed using the software packages NMRPipe/NMRDraw[7] and Topspin (Bruker), whereas the data were analyzed using Xeasy/Cara, Sparky and Protein Dynamics Center software.

To determine whether the CEL-induced chemical shift variations arise from a conformational rearrangement of αS or on the contrary, from a CEL-induced inductive effect, we compared the ^{1}H and ^{13}C NMR chemical shifts of N^{α}-Ac-Lys with those obtained from a CEL-modified N^{α}-Ac-Lys. The 1D-^{1}H and ^{13}C-HSQC spectra were adquired for a solution containing N^{α}-Ac-Lys (10mM) prepared in 150mM phosphate buffer (pH 7.4) in the presence of 150mM NaCl and 10% (v/v) D_2O. These spectra were

adquired at 12.5ºC in the absence and in the presence of pyruvic acid (50mM) and $NaBH_3CN$ (75mM) (afther incubation for 48h at 50ºC).

NMR assignment of αS and αS-CEL

The sequence-specific backbone assignments of αS and of αS-CEL, as well as the assignments of their C_β were achieved using different 2D and 3D NMR experiments: ^1H,^{15}N-HSQC, ^{15}N-NOESY-HSQC (250ms), HNCACB, CACB(CO)HN, HNCO, HNHA, HN(CO)CA and HN(CA)CO. The acquisition of the HCCH-TOCSY spectra allowed us to obtain the assignment of the side chain protons and carbons for Lys and CEL residues.

The assignment of native αS was transferred from that deposited for αS (BMRB code 6968)[8]. However, it was additionally confirmed using triple resonance experiments and then deposited in the BMRB data base (27796). The chemical shift assignment of αS-CEL was achieved using the described NMR experiments, and it was deposited in the BMRB under the accession code 27797.

The chemical shift assignments of αS and αS-CEL were used to estimate the secondary structure content at residue level. This was carried out using different algorithms: i) the secondary structure propensity (SSP) approach[9]; ii) the neighbor-corrected structure propensity calculator (ncSPC)[10], which bases its calculation on the ncIDP random coil library and adds an additional weighting procedure that accounts for the backbone conformational sensitivity of each amino acid type; and iii) the TALOS+ program[11a], which uses the chemical shifts and the sequence information to make quantitative predictions of the secondary structural content.

NMR measurement of $^3J_{HNH\alpha}$ and $^1J_{C\alpha C\beta}$ coupling constants for αS and αS-CEL

The HNHA spectra obtained for αS and αS-CEL were used to determine the $^3J_{HNH\alpha}$ coupling constants for each residue in both proteins. The $^3J_{HNH\alpha}$ coupling constants were calculated based on the intensity ratio S_{cross}/S_{diag} using the Eq.1[12],

$$S_{cross}/S_{diag} = -tan^2(2\pi \cdot J_{HNH\alpha} \cdot \zeta) \qquad (1)$$

where S_{cross} is the intensity of the HN-H_α cross-peak of a given residue, S_{diag} is the intensity of the H_N-H_N diagonal cross-peak of the same residue, and ζ is a delay time in HAHN pulse sequence, which was set to 13.05ms.

The one-bond $^1J_{C\alpha C\beta}$ coupling constants were determined from the HN(CO)CA spectrum of αS and from that of αS-CEL. The $^1J_{C\alpha C\beta}$ values were calculated for each residue from the splitting observed in the C dimension for the C_α cross-peak[13].

Estimation of the phi (ϕ) and psi (ψ) backbone torsion angles of αS and αS-CEL

The phi (ϕ) dihedral angles for αS and αS-CEL were calculated from their experimental $^{3}J_{HNH\alpha}$ coupling constants by using the following Eq. 2[12],

$$^{3}J_{HNH\alpha} = A\ \cos^2(\phi - 60) + B\cos(\phi - 60) + C \qquad (2)$$

where A (7.66), B (-1.00) and C (0.37) are the averaged values obtained from the best fit between the experimental $^{3}J_{HNH\alpha}$ values determined for αS[14a] and the ϕ values estimated for the averaged conformational population of αS[15]. The error of each ϕ value was determined through the application of the propagation of uncertainty method considering: i) the error of the A, B and C constants used in the Eq. 2; and ii) the errors of our $^{3}J_{HNH\alpha}$ values.

The experimental ϕ dihedral angles and the $^{1}J_{C\alpha C\beta}$ coupling constants were then jointly used to estimate the psi (ψ) dihedrals for αS and αS-CEL. The ψ dihedrals were obtained through the application of the following Karplus equation (Eq. 3)[13],

$$^{1}J_{C_\alpha C_\beta} = 1.3 + 0.6\cos(\psi - 61°) + 2.2cos[2(\psi - 61°)] - 0.9cos[2(\phi + 20°)] + {}^{1}J_{C_\alpha C_{\beta\_res}} \qquad (3)$$

where the $^{1}J_{C\alpha C\beta\_res}$ constants are the averaged amino acid-specific $^{1}J_{C\alpha C\beta}$ values obtained from six different proteins (i.e. flavodoxin, xylanase, RNase T1, frataxin, ubiquitin and DFPase)[16]. The error of each ψ value was determined through the application of the propagation of uncertainty method considering: i) the error of our ϕ values; ii) the error of our $^{1}J_{C\alpha C\beta}$ coupling constants; iii) and the error of the residue-specific $^{1}J_{C\alpha C\beta\_res}$ constants[16].

MERA (Maximum Entropy Ramachandran map Analysis) approach was used to estimate the propensity of each residue to explore the distinct regions of Ramachandran space with a resolution of 15° x 15° (ϕ,ψ) voxels[14]. As input, we used the N, Cα, and CO chemical shifts, and the $^{3}J_{HNH\alpha}$ coupling constants. This data was given for all residues, except for Gly since their Hα resonances have identical chemical shifts, thus they were stereospecifically undistinguishable. In addition, for the residues for which it was possible, we also included the $d_{NN}(i,i+1)$, the $d_{\alpha N}(i,i)$ and the $d_{\alpha N}(i,i+1)$ NOEs as inputs. MERA provided a residue-by-residue Ramachandran map distribution for disordered proteins or disordered regions in folded proteins in terms of populations, and the agreement between each input parameter and its distribution-derived value. The minimum RMSD (χ) between the experimental input data and the calculated values increases when the weight (θ) of the entropy term (S; defined in reference[14b]) increases. Here, we have chosen a θ=0.8, which yielded χ^2 values ≤1.5 for all residues in αS and in αS-CEL.

NMR diffusion experiments

The relative diffusion coefficients (*D*) of αS and αS-CEL were measured from the diffusion-ordered spectroscopy (DOSY) spectra adquired by using the pulse field fradient spin echo (PGSE) using a standard ledbpgp2s experiment[17a]. Experiments were carried out at 12.5ºC on samples containing 188µM αS or αS-CEL prepared in 20mM phosphate buffer (pH 6.5) in the presence of 150mM NaCl, 80µM DSS and 10% D_2O. To have statistically significant *D* values we acquired a set of six different DOSY spectra for each sample. These experiments were collected using different diffusion times (from 0.35 to 5s) and different lengths of the gradient pulse (from 3 to 5ms), and each experiment was acquired in duplicate. For each spectrum we determined the *D* values for the ^{1}H-NMR peaks appearing at 6.96ppm, 6.69ppm, 2.86ppm, 1.09ppm, 0.8ppm (all these signals belong to the protein) and 0ppm (a signal that belongs to the DSS, which was used as internal standard). The *D* value for each signal was obtained by fitting the intensity decays versus the gradient strength as described elsewhere[17b]. From the *D* values we calculated the hydrodynamic radius (R_h) of αS and αS-CEL using the Stokes-Einstein equation (Eq. 4),

$$R_h = k_B T / 6\pi\eta D \qquad (4)$$

where k_B is the Boltzman constant, T is the temperature (285.65K) and η is the viscosity of H_2O at 285.65K ($1.24 \cdot 10^{-3}$ N·s/m^2).

NMR relaxation measurements

Transverse (R_2) relaxation measurements were acquired for αS and αS-CEL by using a series of three 3D-HNCO experiments[18] with relaxation delays of 60, 360 and 570ms, respectively. This experiment is especially suitable for proteins exhibiting extensive spectral overlap in the ^{1}H,^{15}N-HSQC spectrum. R_2 measurements were carried out adopting the methodology developed by Yuwen and Skrynnikov, which allows the increase of the relaxation delays avoiding the cryoprobe heating[19]. The R_2 values were determined by fitting the peak heights (*I*) as a function of the relaxation delay (*t*) to a two-parameter (I_0 and R_2) exponential decay function: $I(t) = I_0 \cdot e^{-R_2 \cdot t}$. The average uncertainty values of R_2 were determined from the three data sets curve fitting. The ^{15}N HET-NOE measurements were performed by using the standard pulse sequence[20]. ^{1}H saturation was achieved for a 2s period by the application of 120º ^{1}H pulses separated by 5ms. The spectral width was 9615Hz over 2048 complex points in the ϖ2 (^{1}H) dimension and 2432Hz over 64 complex points in the ϖ1 (^{15}N) dimension. The steady-state NOE values and errors were calculated from the ratios of the peak intensities with and without presaturation by using the software Protein Dynamics Center (Bruker Biospin).

Carr-Purcell-Meiboom-Gill relaxation dispersion (CPMG-RD) measurements[21] were acquired at 285.5 K. The effective transversal relaxation rates (R_2^{eff}) at 14 different

CPMG frequencies (υ_{CPMG}), from 31.25 to 1000Hz, were calculated from the intensities of resonance in the ^{1}H-^{15}N correlation spectra using Eq. 5:

$$R_2^{eff} = (-1/T_{relax}) \cdot ln\,(I\upsilon_{CPMG}/I_0) \quad (5)$$

where T is constant-time delay (32ms) and I_0 and $I\upsilon_{CPMG}$ are the intensities of resonance obtained from experiments without and with CPMG pulse block. The difference between R_2^{eff} at υ_{CPMG} of 31.25Hz and 1000Hz ($\Delta R_2^{eff} = R_2^{eff}\upsilon_{CPMG=31.25Hz} - R_2^{eff}\upsilon_{CPMG=1000Hz}$) was plotted as a function of the αS sequence to study the contribution of the conformational exchange motions to the residue dynamics.

Calculations of conformational ensembles for αS and αS-CEL

Several bioinformatics tools were used to generate conformational ensembles compatible with the experimental chemical shifts. Initially, the TraDES-2[22] was used to generate 10,000 structures that sampled a random coil distribution. These structures were then clustered with the gromos method3[23] implemented in Gromacs 2016.4[24]. The cluster algorithm employed the RMSD of the C_α after proper alignment. Different cluster analyses were performed by increasing the cut-off value from 1.5 to 2.0 nm in steps of 0.1 nm. The cluster analyses that yielded less than 1000 clusters (i.e. those from cut-offs 1.7, 1.8, 1.9 and 2.0 nm) were considered for ensemble generation. Each group of clusters was used independently to generate one ensemble for αS and one for αS-CEL. Ensembles were generated by using the Mollack software[25], which tries to find the best linear combination of structures that reproduce some fed experimental data. We specifically used the NMR chemical shifts (N, H_N, C_α, C_β, H_α and CO) of each residue. For each structure in the group of resulting clusters, the chemical shifts were calculated with the software SPARTA+[11b]. The weighted ensemble average properties were calculated by taking into account those structures with weights higher than 10^{-3}.

Small Angle X-ray scattering (SAXS) measurements

SAXS experiments were performed on a Xeuss 2.0 instrument (Xenocs, France) equipped with a microfocus Cu K_α source (λ 1.54Å) collimated with scatterless slits. The scattering was measured using a Pilatus 300k detector with a pixel size of 0.172µm x 0.172µm (Dectris, Switzerland). The distance between the detector and the sample was calibrated using silver behenate and it was set at 1m. The X-ray scattering curves were obtained at 12.5ºC for solutions containing 0.3mM αS or αS-CEL in Tris buffer (10mM) at pH 7.4 in the presence of 150mM NaCl. The measurements were carried out for 10min under vacuum using a temperature-controlled low noise flow cell. The SAXS data was reduced to absolute units, avegared and solvent substrated using the RAW 1.5 software[26]. All data processing and analysis was also carried out

using the RAW 1.5 software. The magnitude of the scattering vector (Q) was calculated by Eq. 6,

$$Q=(4\pi\sin\theta)/\lambda \qquad (6)$$

where 2θ is the scattering angle and λ is the wavelength of X-ray, and futher used to determine the radius radius of gyration (R_g) from the Guinier approximation (Eq. 7)

$$\ln I(Q) = \ln I(0) - R_g^2 Q^2/3 \qquad (7)$$

where Q is the scattering vector, $I(Q)$ is the scattered intensity as a function of the scattering vector, and $I(0)$ is the forward scattering intensity, which is proportional to the molecular mass of the scattering profile.

α-Synuclein fibril formation

Monomeric αS and αS-CEL isolated from SEC were diluted in 20mM sodium phosphate buffer at pH 7.4 in the presence of 150mM of NaCl to a final concentration of 70μM. Then, each sample was filtered through a Millex-LG hydrophilic PTFE membrane with a 0.20μm pore size and degassed prior to being incubated at 37°C for several days while shaking at 1000rpm. In the case of αS-CEL, the incubation was carried out alone or in the presence of 150μM of $FeCl_3$, $AlCl_3$ or $CuCl_2$. Additionally, preformed amyloid fibrils of monomeric αS were dialyzed in 150mM sodium phosphate (pH 7.4) and then incubated in the presence of 75mM $NaBH_3CN$ and 50mM of pyruvic acid at 50°C for 48h.

ThT-fluorescent measurements

Aliquots of solutions containing monomeric αS or αS-CEL prepared in 20mM phosphate buffer (pH 7.4) in the presence of NaCl (150mM), were taken at different incubation times and diluted in milli-Q water to a final concentration of 10μM. In addition, solutions containing αS amyloid fibrils or CEL-modified amyloid fibrils (CEL was chemically synthesized on preformed αS amyloid fibrils) were subjected to ThT assays. These samples were then mixed thoroughly with a ThT stock solution to a final ThT concentration of 50μM. Afterwards, the fluorescence spectra were measured between 460 and 600nm (λ_{exc} 440 nm) at room temperature on a PerkinElmer LS55 Luminescence Spectrometer. The scan speed was 200nm/min with an excitation and emission slit of 2.5nm, while 5 scans were accumulated. In addition, the depolymerization of αS amyloid fibrils or CEL-modified amyloid fibrils was studied using Gdn-HCl as a chaotropic agent. The ThT fluorescent spectra of those fibrils (λ_{exc} 440nm) was adquired at room temperature between 460 and 660nm at diffent Gdn-HCl concentrations, ranging from 0 to 4.5M. The fluorescence emision intensity obtained at 487nm was plotted against the Gdn-HCl concentration.

Atomic force microscopy (AFM)

Aliquots (70µl) from the samples used to study αS and αS-CEL fibril formation were taken at different incubation times and placed onto a mica surface. Then, they were incubated for 5min at room temperature before drying with N_2 gas. The mica was rinsed 5 times with 1ml milli-Q water and dried with N_2 gas before observation under a Veeco Multimode atomic force microscope equipped with a NanoScope IV controller. The particle dimensions were measured using the NanoScope SPM v5 software.

Dynamic Light Scattering (DLS)

Aliquots (20µL) from the samples containing monomeric αS or αS-CEL (70µM), which were prepared in 20mM sodium phosphate (pH 7.4) in the presence of 150mM of NaCl, were taken at different incubation times (37ºC), degassed and placed in disposable cuvettes. Afterwards, they were subjected to DLS analysis at 37ºC by using a DynoPro NanoStar 467-DPN (Wyatt Technology) equipped with a 660.2nm laser. 10 acquisitions of 5 or/and 10s were collected in the range of 0.5µs to 1s. Those 5 or 10s accumulations with abnormally high SOS function were removed and the remaining ones were averaged. The data was analyzed using the Dynamics software (Wyatt Technology, v.7.1.9.3).

Coarse-grained molecular dynamics (CG-MD) calculations

CG-MD calculations carried out on αS and αS-CEL were performed using the replica exchange with solute scaling (REST2) simulations with a modified version of the coarse-grained force field SIRAH. The complete description of the computational details of this methodological approach has been recently published in a preliminary work of our group[27]. There, we proved that the molecular ensembles arising from the use of this computational approach -when using a factor *f* of 1.3 multiplying the standard ε Lennard-Jones parameter associated to the protein-solvent atom pairs- are able to reproduce most of the experimental descriptors of αS. Simulations were started from the central structure (9AAC-522.pdb) of the αS ensemble deposited in the Protein Ensemble Database (pE-DB)[25]. Simulations carried out on αS-CEL were performed from the same initial structure (9AAC-522.pdb) but replacing its 15 Lys by CEL (**Fig. 1A**).

All-atom molecular dynamics simulations on native and CEL-modified αS amyloid fibrils

Two fibril models were simulated, one formed by αS monomers and the other by αS-CEL monomers. The coordinates to build these two models were taken from the cryo-EM structure of the αS amyloid fibril (PDB code 6A6B)[29]. The CEL-modified Lys were modeled on all native monomers of the fibril with the Pymol software. The two models were placed in truncated dodecahedron boxes whose closest edges were at 2.0nm of

any of the proteins forming the fibrils. Then, the boxes were filled with water molecules. In each system, the required number of Na^+ and Cl^- ions were added to achieve electroneutrality and to reproduce our experimental ionic strength.

The Amber ff03 force field[30] was used for the protein, whereas the TIP3P force field[31] was used to model the water molecules. The Joung and Cheatham[32] parameters for using ions in combination with Amber and TIP3P force fields were used to model the Na^+ and Cl^- ions. Recently, we described the parametrization of the CEL-modified Lys consistent with the Amber ff03 force field[33].

The Gromacs 2016.4[24] software was employed for all the all-atom simulations. Initially, the geometries of αS and αS-CEL fibrils were relaxed with the steepest descent algorithm until the maximum force was <100.0 $kJ{\cdot}mol^{-1}{\cdot}nm^{-1}$. Then, 1ns-long NVT simulations were performed to equilibrate the temperature at 310 K. The thermostat of Bussi et al.[34] was applied separately to the protein and to the water plus ions with time couplings of 0.1 ps in both cases. Another 1ns of simulations was performed to equilibrate the densities of the systems for an external isotropic pressure of 1.0bar with the Parrinello-Rahman[35] barostat with a time coupling of 2.0ps and a compressibility of $4.5{\cdot}10^{-5}$ bar^{-1}. During the temperature and pressure equilibrations, position restrains of 1000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-2}$ were applied to the heavy atoms of all the monomers. Then, the restraints were kept only for the first monomer of the protofilament in one of the extremes. This was to maintain the integrity of the short fibril during the simulation. After the release of the restraints, 10ns of simulation at the equilibrated temperature and pressure were performed to allow the relaxation of the monomers. Finally, a production run of 60ns was performed.

Steered molecular dynamics of native and CEL-modified αS amyloid fibrils

Steered Molecular Dynamics (SMD) simulations[36] were performed to estimate the stability of αS and αS-CEL monomers at the ending of the fibril. All the calculations were performed with the Gromacs 2016.4[24] software patched with Plumed 2.4.0. In SMD, an external force is applied along a coordinate or collective variable (CV) to drive the system over free energy barriers. Then, the work performed by the external force in the non-equilibrium simulation can be used to estimate the ΔG of the corresponding equilibrium process by means of the Jarzynski equality[37]. Application of the Jarzynski equation (Eq. 8) allows the calculation of the equilibrium free energy from the non-equilibrium simulations work as

$$\Delta G = -\frac{1}{\beta}\ln\langle\exp(-\beta W_i)\rangle \qquad (8)$$

where β is the product of Boltzmann's constant (k_B) and the temperature, and W_i the work carried out to detach the monomer from the fibril.

Fifty 1ns long SMD simulations were performed for each of the two systems (i.e. αS and αS-CEL protofibrils). Each simulation was started from the previously equilibrated structure, but with different initial random velocities. The monomer pulled from the fibril was at the ending of the fibril while the restrained monomers were at the beginning. The CV for the pulling was defined as the minimum distance between the ending monomer and its two closest monomer neighbors in the fibril. One closest neighbor is below the pulled monomer in the same protofilament. The other closest neighbor is at the top of the second protofilament. The minimum distance between monomers was calculated as the minimum distance between the C_α of L38, T44, H50, A56, Q62, G68, V74, K80, G86, T92 and D98 of the pulled monomer and the equivalent atoms in the closest neighbors. To make the variable continuous, the following function of the distances (d_i) was used:

$$min = -\frac{1}{C}\log\sum_i \exp(-Cd_i) \qquad (9)$$

where C is a constant that in this case was chosen as $10nm^{-1}$. The pulling force was modeled as a moving harmonic potential acting on the CV. The harmonic constant of such potential was 4000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-2}$. Initially, the minimum of the potential was set at the equilibrium value of the CV (i.e. ~0.27 nm). Then, the minimum of the potential was augmented uniformly to a value of 3.27nm during the SMD simulations.

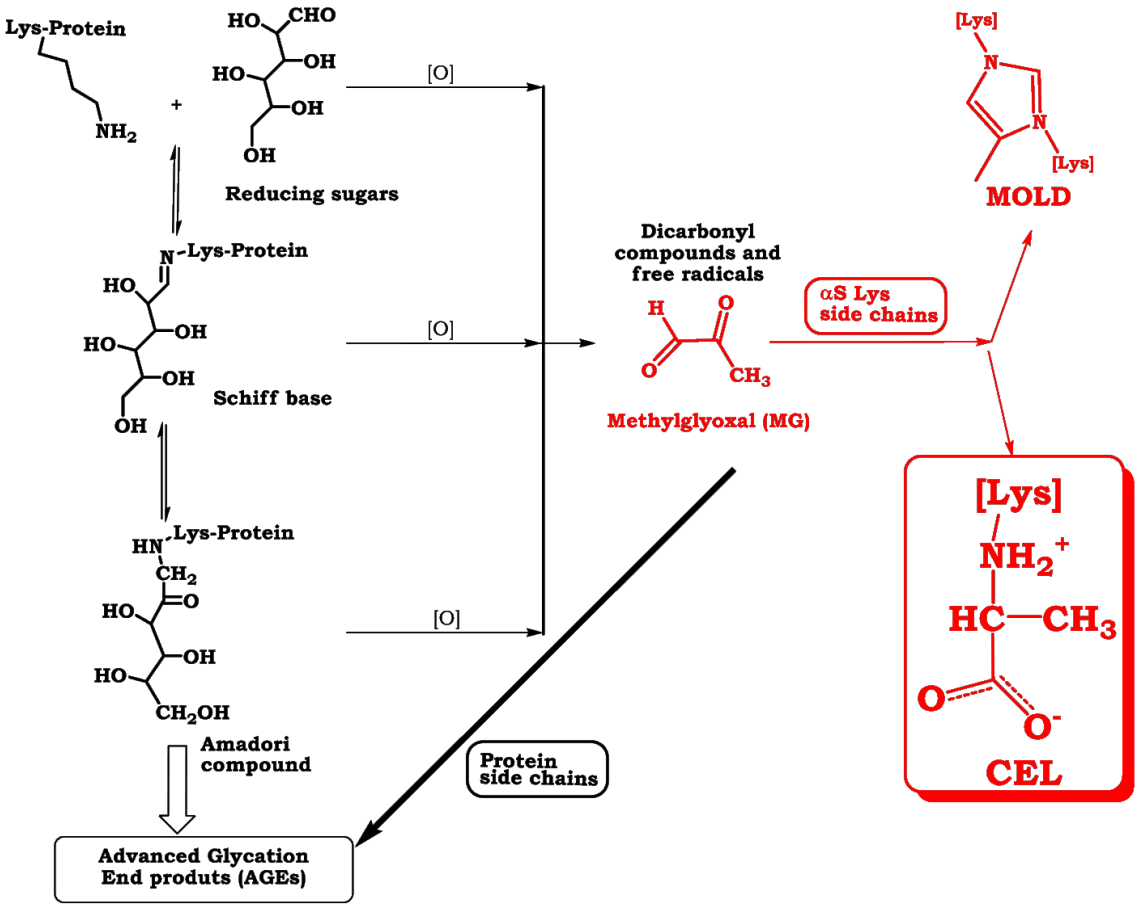

Figure S1. The overall mechanism of protein glycation and the formation of methylglyoxal-derived AGEs. Protein glycation starts with the chemical reaction of reducing carbohydrates (mainly glucose) with primary amino groups of proteins. This encompasses the reversible formation of a Schiff base that converts into an Amadori product, which can then further rearrange to yield the advanced glycation end products (AGEs). Although the formation of the Schiff base and the Amadori compound constitute the central pathways along the protein glycation mechanism, the whole process become much more complex as a result of collateral oxidative reactions of reducing sugars, Schiff bases and Amadori compounds[38]. These reactions yield highly reactive carbonyl species such as methylglyoxal (MG), which can further react with other amino acid side chains contributing to AGEs formation. MG is the most relevant glycating compound inside the neurons[39], and it is able to modify αS *in vivo* inducing the formation of MOLD and *N*ε-(carboxyethyl)lysine (CEL) (*see its chemical structure drawn in the red box*)[40] on its Lys side chains.

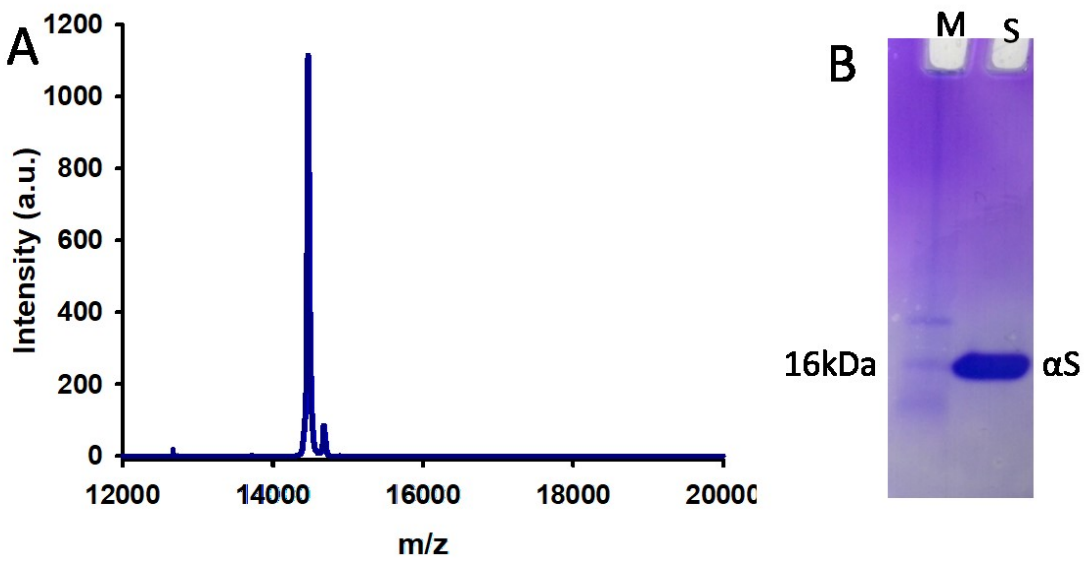

Figure S2. Characterization of purified wild type αS. **(A)** MALDI-TOF/TOF spectrum of purified recombinant αS. A single signal was obtained, whose m/z value matched the theoretical molecular weight of αS. **(B)** SDS-PAGE/Coomassie brilliant blue staining of purified recombinant αS (S) and of the marker (M; Sigma-Aldrich S8445). A single band with a molecular weight similar to that of αS was obtained.

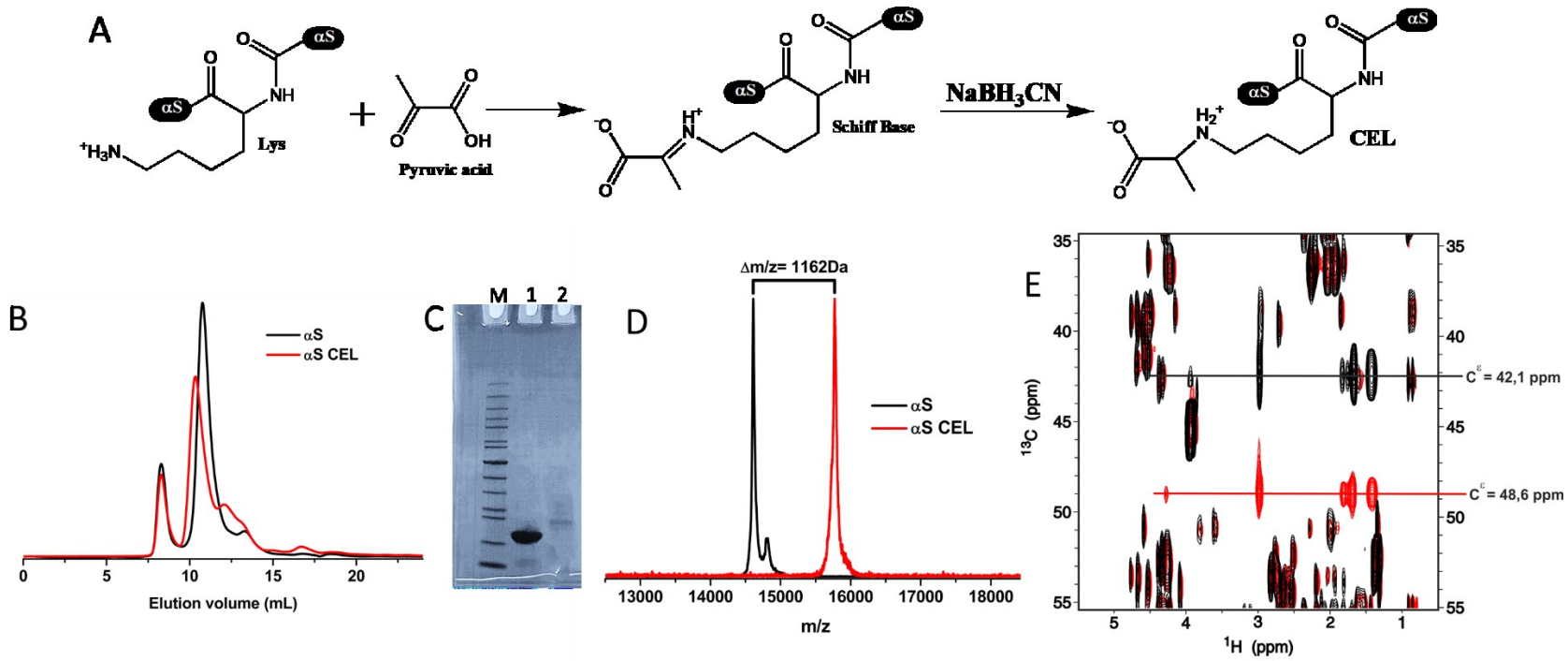

Figure S3. Synthesis and characterization of αS-CEL. **(A)** Schematic representation of the chemical synthesis of CEL on αS. The protein was incubated with pyruvic acid in 150mM sodium phosphate buffer (pH 7.4) for 48h at 50°C in the presence of $NaBH_3CN$. The chemical structures of Lys side chain, Schiff base and CEL have been drawn considering their main protonated states at neutral pH according to their pKas. **(B)** Size-exclusion chromatograms (Superdex-75 HR 10/300 column) of native αS and an αS incubated with pyruvic acid in the presence of $NaBH_3CN$ (αS-CEL). The sample containing αS-CEL was loaded into the column after a dialysis step in 150mM sodium phosphate buffer (pH 7.4). Fractions of the main peak, which corresponded to the monomeric αS-CEL, were pooled together and used for further analysis. **(C)** SDS-PAGE/PAGE-blue staining analysis of native αS (1) and purified αS-CEL (2). A marker (M; Sigma-Aldrich S8445) was used as reference. The sample containing monomeric αS-CEL only displayed a single band with a molecular weight slightly higher than that corresponding to the native αS. This would result from the covalent addition of different carboxyethyl groups on the Lys groups of αS. **(D)** Overlapping of the MALDI-TOF/TOF spectrum of the native αS with that corresponding to the synthetic αS-CEL. The intensity of the peaks was normalized for comparison purposes. As suggested the SDS-PAGE analysis, the signal corresponding to the αS-CEL shifted towards higher molecular weights as compared with that of native αS, which confirms the covalent addition of CEL moieties on αS. **(E)** Overlapping of the projections corresponding to the HCCH-TOCSY spectra of αS (*black*) and of the αS-CEL (*red*). The chemical shifts corresponding to the Lys-C^{ε} downfield shifted as a result of CEL formation.

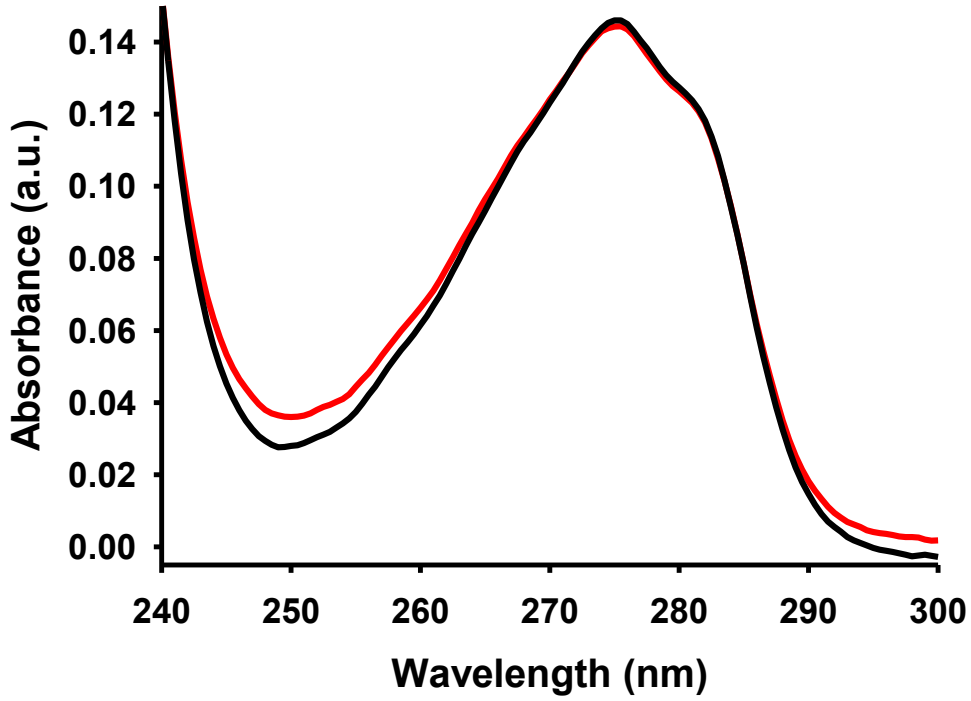

Figure S4. Overlapping of the UV-Vis spectrum of native αS (*black*) with that of αS-CEL (*red*). The spectra were normalized at 276nm for comparison purposes. The formation of CEL on αS does neither change the profile nor the intensity of the UV-Vis spectrum of αS.

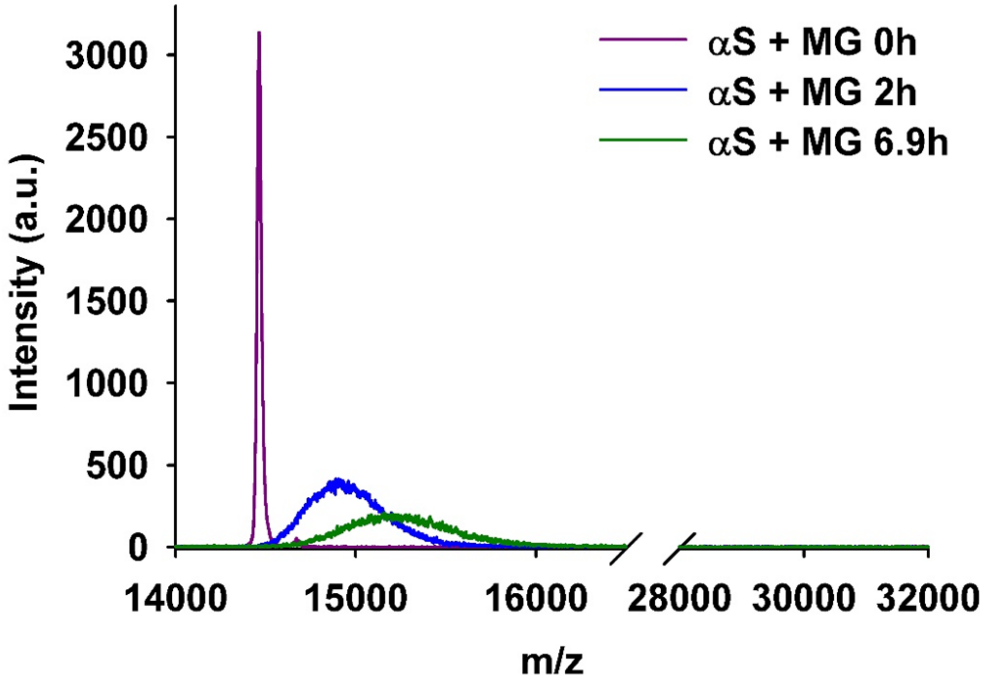

Figure S5. MALDI-TOF/TOF spectra of αS (40μM) in the presence of MG (50mM). MG in solution (40%) was purchased from Sigma-Aldrich (M0252) and additionally purified by steam distillation. Different distilled fractions were collected and their pH was adjusted to 7.0 with 0.1M NaOH before determining their concentration. The resulting solutions were mixed with H_2O_2 and the concentration of pure MG was indirectly quantified by titrating the remaining H_2O_2 with $KMnO_4$, according to the Firedemann's protocol[41]. The pure MG concentration was estimated to be ~0.94M and the fractions were frozen until use. Afterwards, a reaction mixture containing αS and MG was incubated at 37ºC in 200mM phosphate buffer at pH 7.4. Aliquots were taken at different times and then diluted 5 times in mili-Q water to reduce the phosphate concentration before the MALDI-TOF/TOF analysis. The initial peak corresponding to αS shifted towards higher molecular weights, which proves the covalent addition of different MG moieties during the incubation. However, the protein peak also became lower and broader upon incubation, which proves the formation of a heterogeneous mixture of αS molecules with different glycation degree. The incubation of αS with MG did not result into the formation of cross-linked oligomers, as proved by the absence of MALDI-TOF/TOF peaks at a m/z ratio higher than 16kDa.

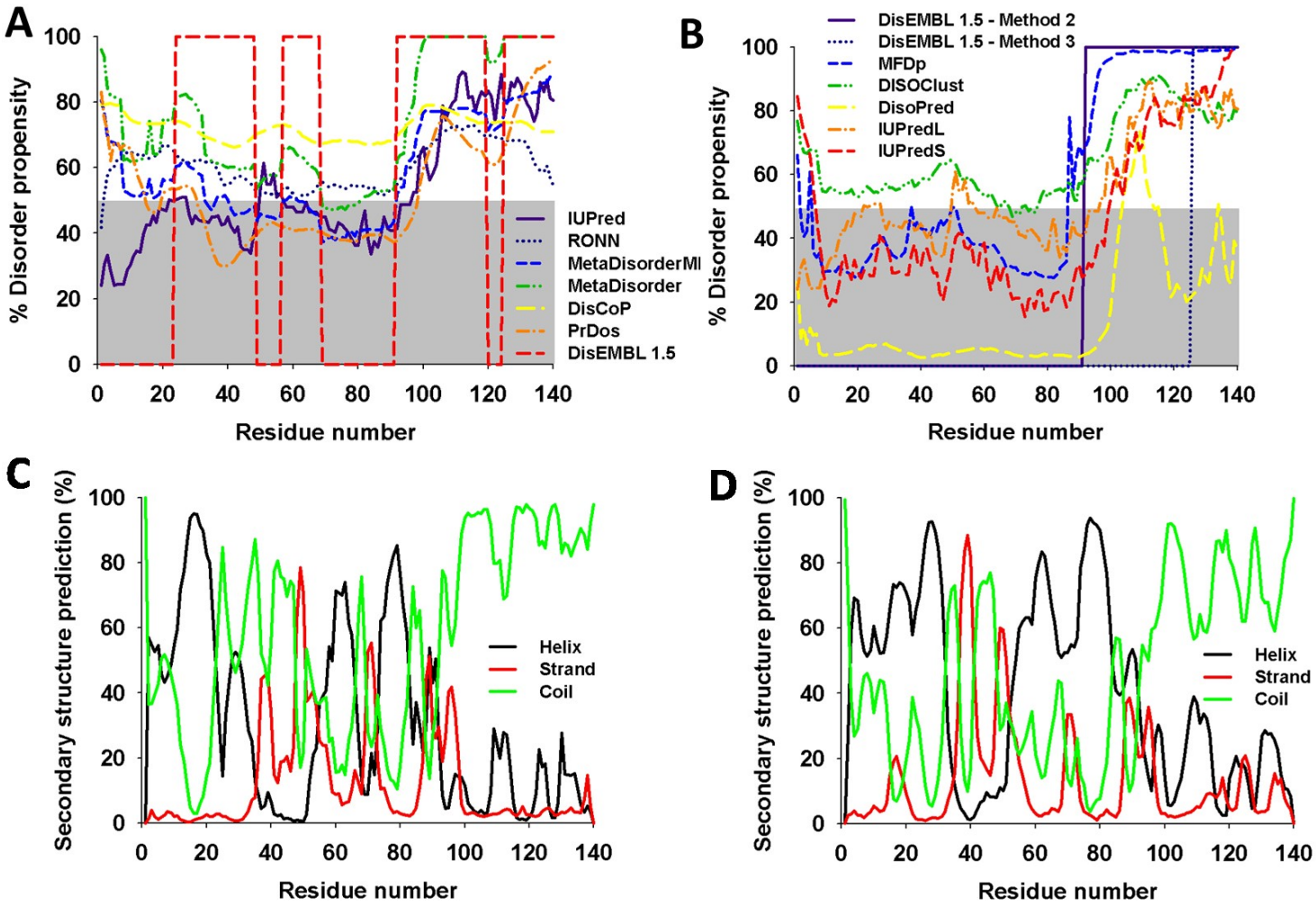

Figure S6. Sequence-based analysis of the structuration level of αS. **(A-B)** The disorder propensity prediction of αS was carried out using the following algorithms: IUPred[42a], RONN[42b], MetaDisorder MI, MetaDisorder[42c], DisCoP[42d], PrDos[42e], DisEMBL 1.5, DisEMBL 1.5 – Method 2, DisEMBL 1.5 – Method 3[42f], MFDp[42g], DISOClust[42h], DisoPred[42i], IUPredL[42a] and IUPredS[42a]. The predictions have been plotted in two different representations (**A** and **B**) to have a better view of them. All algorithms used the same disorder scale: residues with values between 0 and 50 are considered to retain a certain degree of structuration (*grey background*), whereas those with a disorder propensity between 50 and 100 are considered as disordered (*white background*). All predictions have been done using the αS sequence. **(C-D)** Secondary structure content prediction for each residue of αS. Predictions were carried out using the PSIPRED algorithm[43a] **(C)** and the Cspritz algorithm[43b] **(D)**.

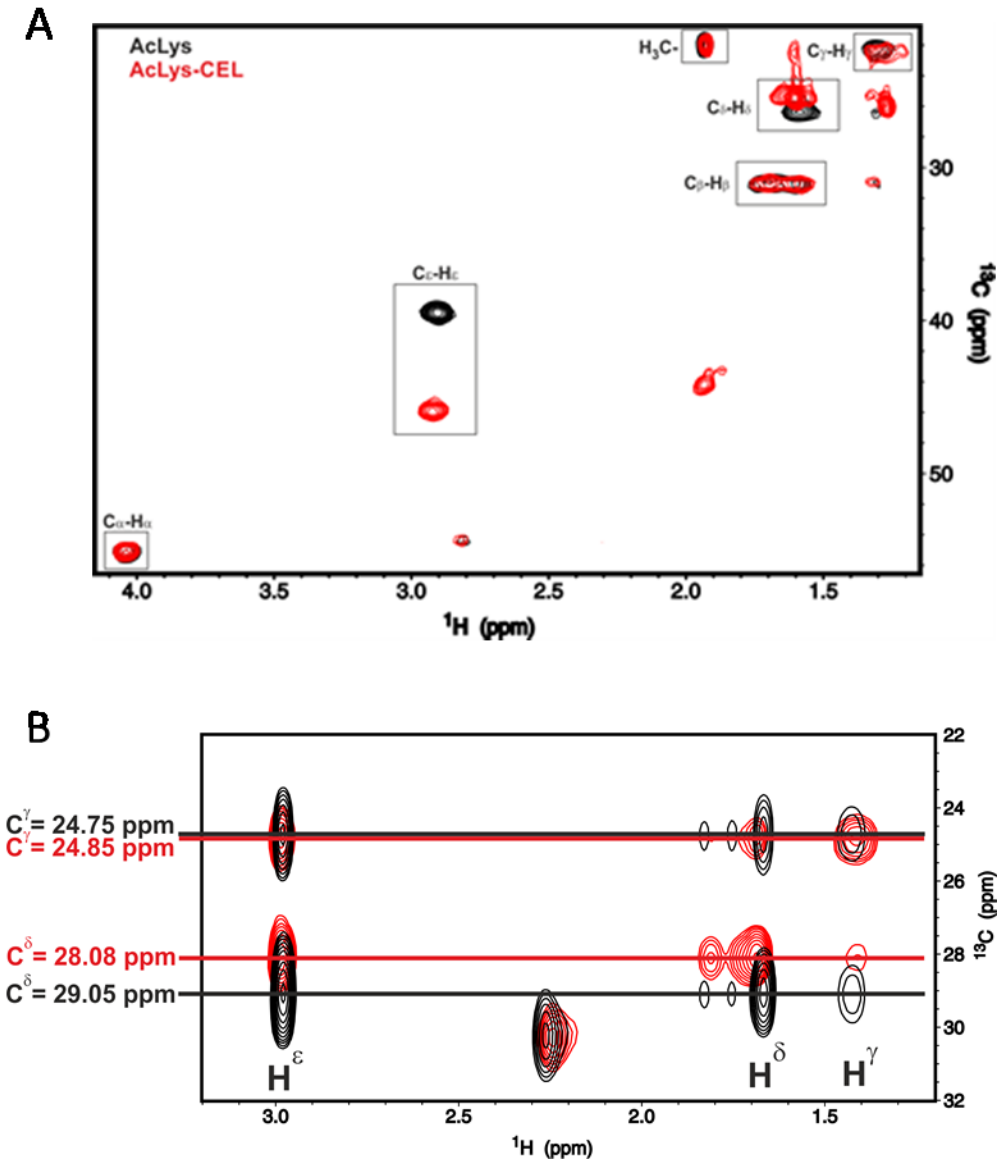

Figure S7. Study of the inductive effect of CEL formation on the ^{1}H anc ^{13}C chemical shifts: controls and evidences. **(A)** Overlapping of the aliphatic region of the ^{13}C-HSQC spectra of native N^{α}-Ac-Lys (*black*) on that corresponding to the CEL-modified N^{α}-Ac-Lys (*red*). The cross-peaks corresponding to the H_{α}/C_{α}, H_{β}/C_{β}, H_{γ}/C_{γ}, H_{δ}/C_{δ}, and $H_{\varepsilon}/C_{\varepsilon}$ are labelled and squared. **(B)** Overlapping of the projections corresponding to the HCCH-TOCSY spectra of αS (*black*) and of the αS-CEL (*red*). The Lys-H_{γ}/C_{γ} and Lys-H_{δ}/C_{δ} cross-preaks have been labelled to better compare the differences between αS and αS-CEL. Their chemical shifts are also indicated.

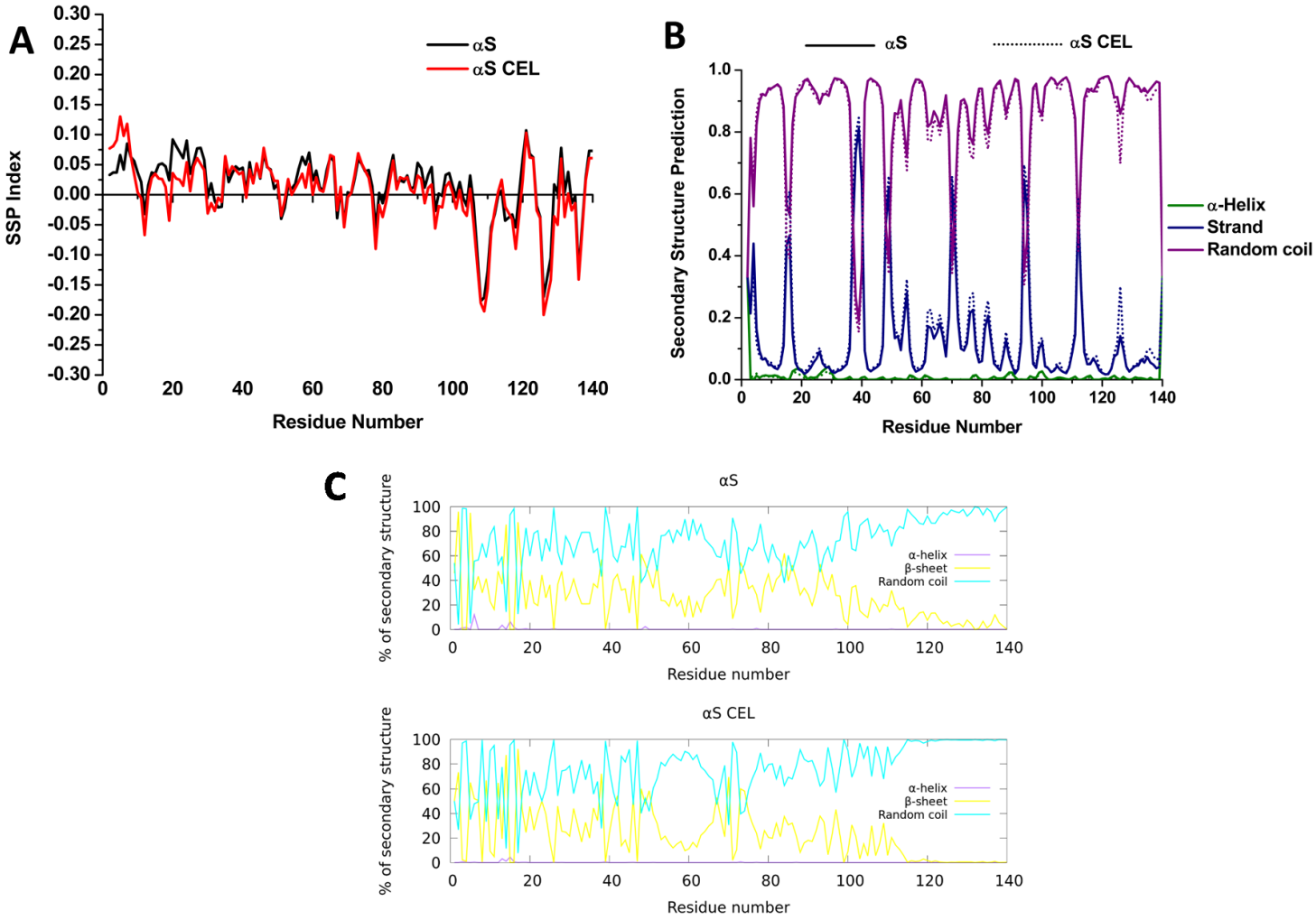

Figure S8. Secondary structure content predictions for αS and αS-CEL. **(A)** Secondary structure propensity (SSP) scores obtained for αS (*black*) and αS-CEL (*red*). The SSP values were calculated using the H_α, C_α and C_β chemical shifts for each assigned residue, which is the recommended procedure when studying intrinsically disordered proteins[44]. +1 indicates a fully formed α-helix; "-1" indicates a fully formed β-sheet; and "0" indicates disorder. **(B)** Secondary structure predictions obtained from TALOS+ using the H_N, N, H_α, C_α, C_β, and CO chemical shifts obtained for αS and αS-CEL at 12.5°C and at pH 6.5. **(C)** Average secondary structure predictions obtained for αS (*top*) and αS-CEL (*bottom*) using CG-MD after 1400ns of simulation. The simulations were carried out using replica exchange with solute scaling (REST2) simulations with the coarse-grained force field SIRAH. The simulations were carried out using a factor *f* of 1.3 multiplying the standard ε Lennard-Jones parameter associated to the protein-solvent atom pairs[27].

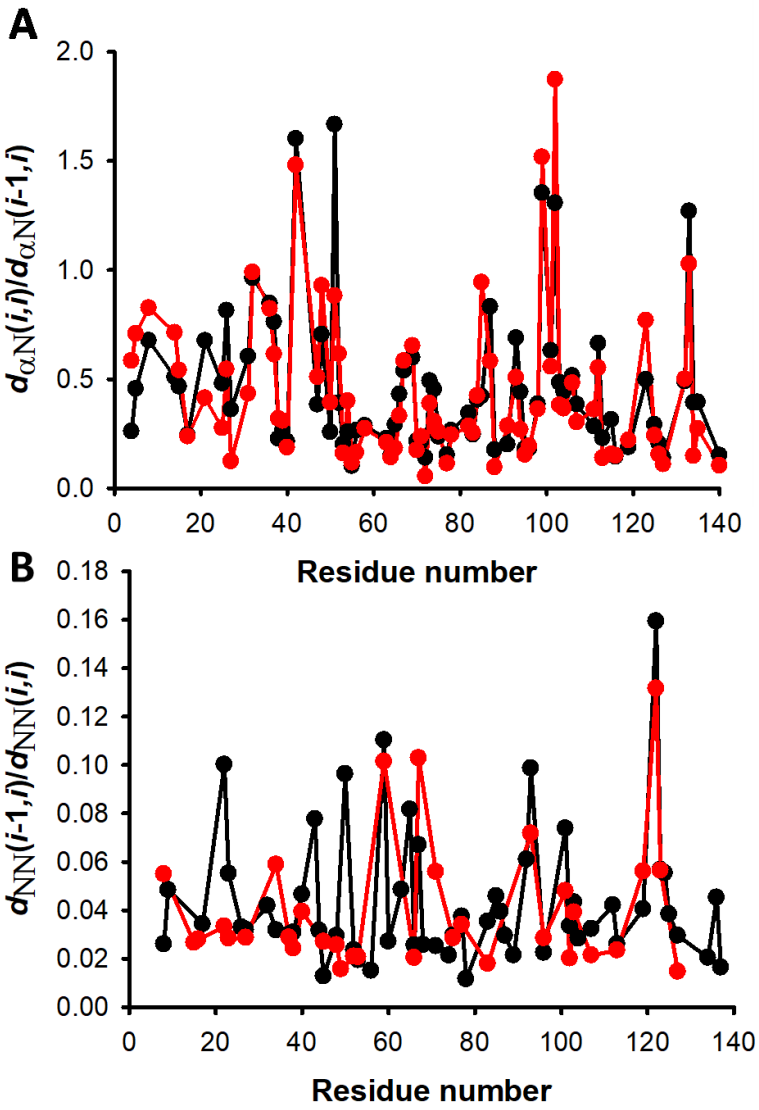

Figure S9. Sequential NOE intensity ratios for αS (*black*) and αS-CEL (*red*). **(A)** Ratios of intraresidue to sequential H_α-H_N NOE intensities in αS and αS-CEL. For comparison purposes, the plot only displays the values for those residues whose $d_{\alpha N}(i,i)/d_{\alpha N}(i-1,i)$ values could be determined for αS and for αS-CEL. The ratios corresponding to Gly residues were divided by 2, to correct for the presence of two H_α atoms. **(B)** Ratios of intraresidue to sequential H_N-H_N NOE intensities in αS and αS-CEL. For comparison purposes, the plot only displays the values for those residues whose $d_{NN}(i,i)/d_{NN}(i-1,i)$ values could be determined for αS and for αS-CEL.

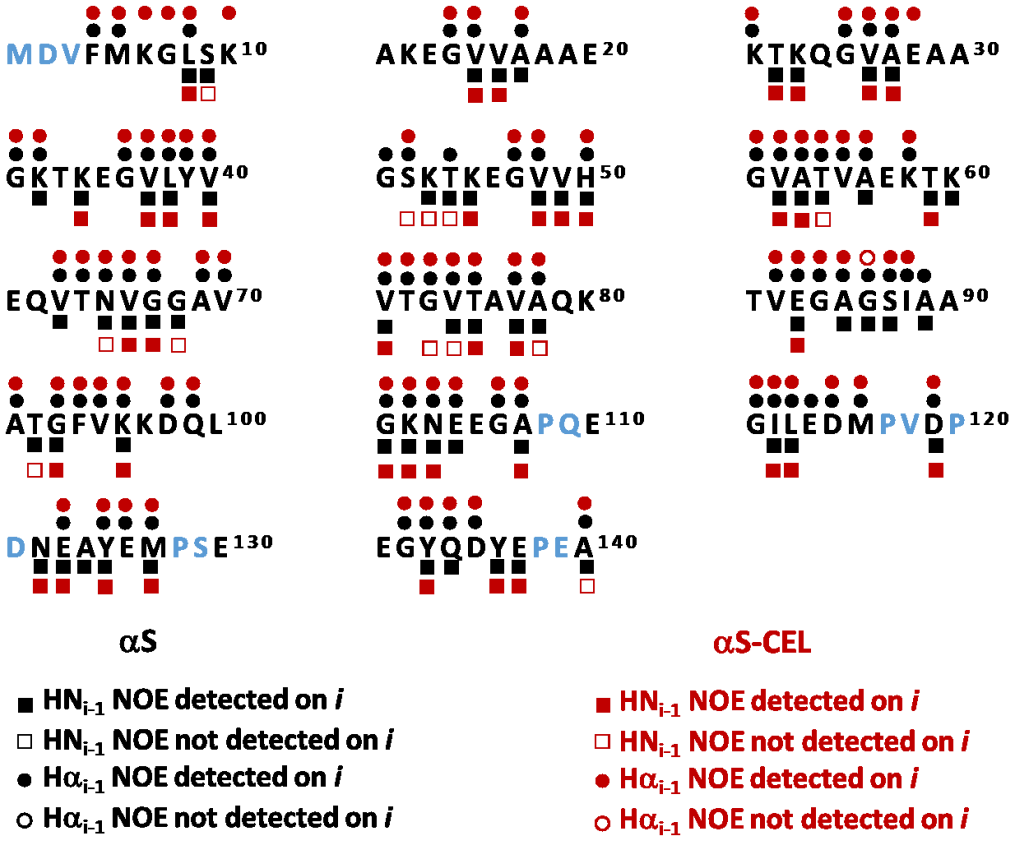

Figure S10. Primary sequence of αS complemented with NOE patters observed for native αS (*black*) and αS-CEL (*red*), which are displayed above ($H_{\alpha i-1}/HN_i$; *circles*) and below (HN_{i-1}/HN_i; *squares*) the sequence. The NOE intensities for each residue were obtained from the corresponding $^1H,^1H$-NOESY spectra. Amino acids colored in blue were not included in the analysis of the sequential NOEs. Filled symbols display the residues for which the *i*-1 NOE have been detected; empty symbols represent those residues for which the *i*-1 NOE was not observed; and the absence of a symbol indicates that it was not possible to determine whether it was a NOE or not since the signal of the *i*-1 residue overlapped with that of the *i* residue.

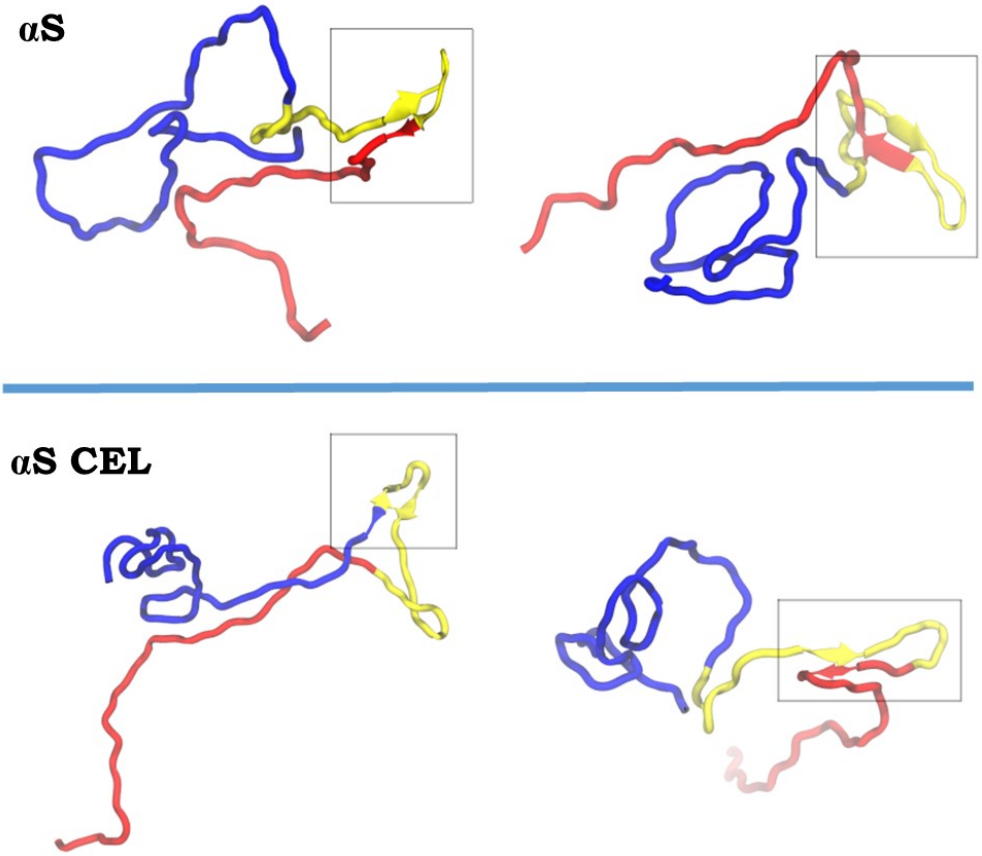

Figure S11. Structural conformations representative of αS (*top*) and αS-CEL (*bottom*). These structures were obtained from microsecond-long replica exchange with solute scaling (REST2) simulations with the coarse-grained force field SIRAH. The simulations were carried out using a factor *f* of 1.3 multiplying the standard ε Lennard-Jones parameter associated to the protein-solvent atom pairs, which is the one that provides conformations of αS that better reproduce experimental parameters such as chemical shifts or R_g values[27]. All the conformations characteristic of αS or of αS-CEL, displayed transient β-hairpins (framed in grey squares) between the N-terminal (*blue*) and the NAC (*yellow*) domains, and between the NAC (*yellow*) and the C-terminal (*red*) domains.

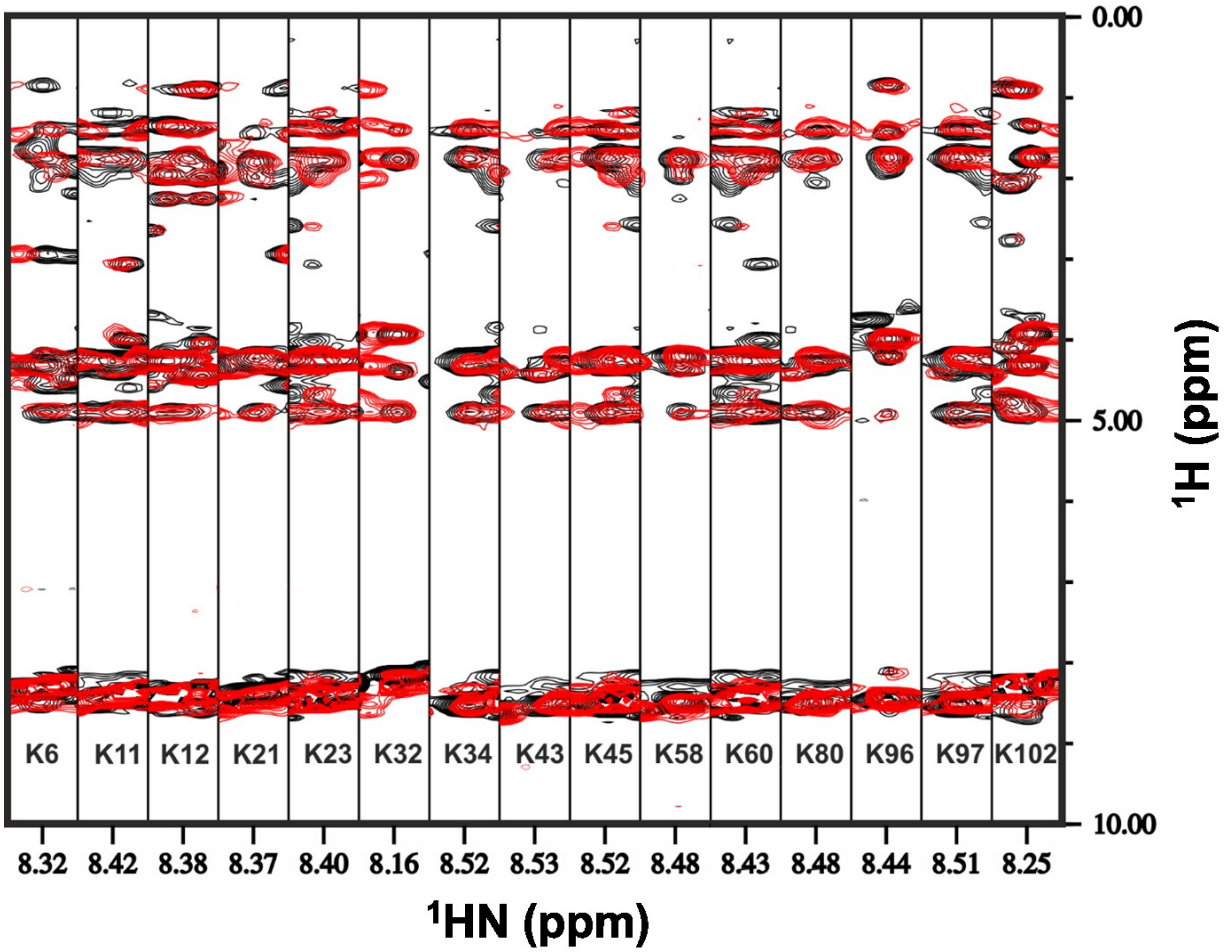

Figure S12. ^{15}N-NOESY-HSQC strips corresponding to each one of the fifteen Lys residues of the αS sequence (at the bottom, each strip is labelled with its residue number). The strips corresponding to αS (*black*) are overlapped with those obtained from αS-CEL (*red*), where all Lys residues have been replaced by CEL.

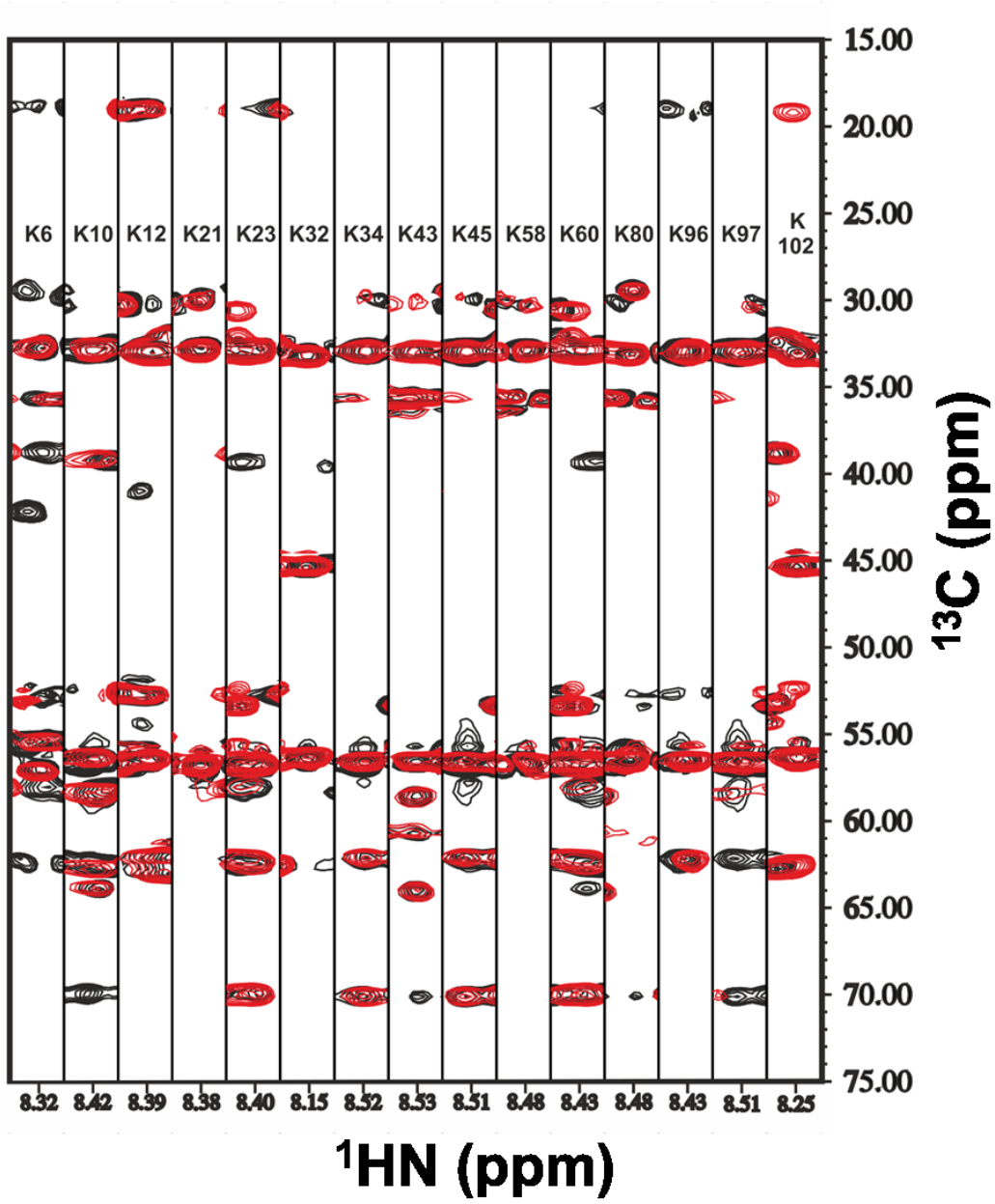

Figure S13. HNCACB strips obtained from the ^{15}N plane corresponding to each one of the fifteen Lys residues of the αS sequence (each strip is labelled with its residue number). The strips corresponding to αS (*black*) are overlapped with those obtained from αS-CEL (*red*), where all Lys residues have been replaced by CEL.

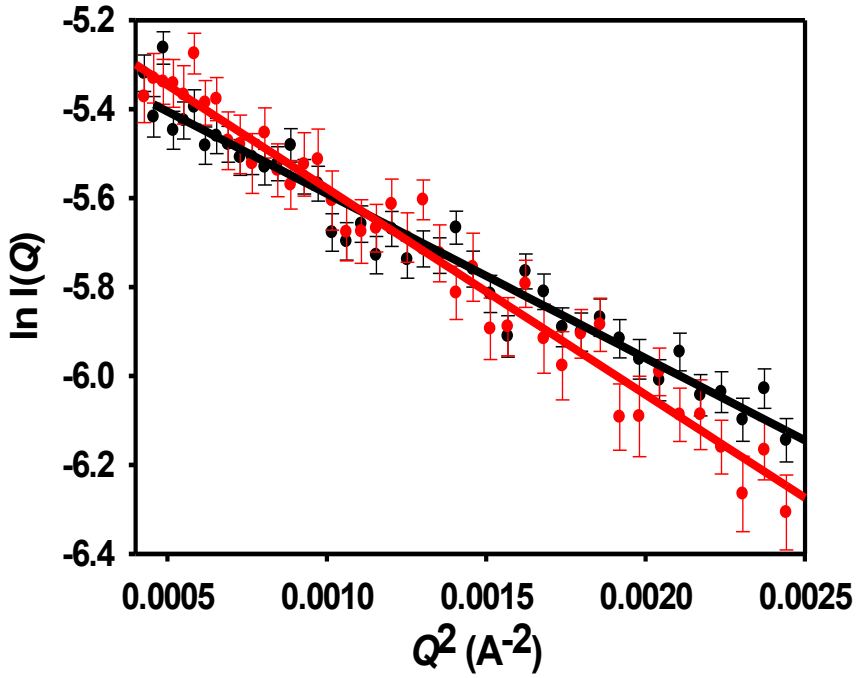

Figure S14. Guinier plots of the X-ray scattering curves corresponding to the low *Q* region of monomeric αS (*black*) and of the αS-CEL (*red*). The scattering curves were collected at 12.5°C using a 180μM protein concentration in 10mM Tris buffer (pH 7.4) in the presence of 150mM NaCl.

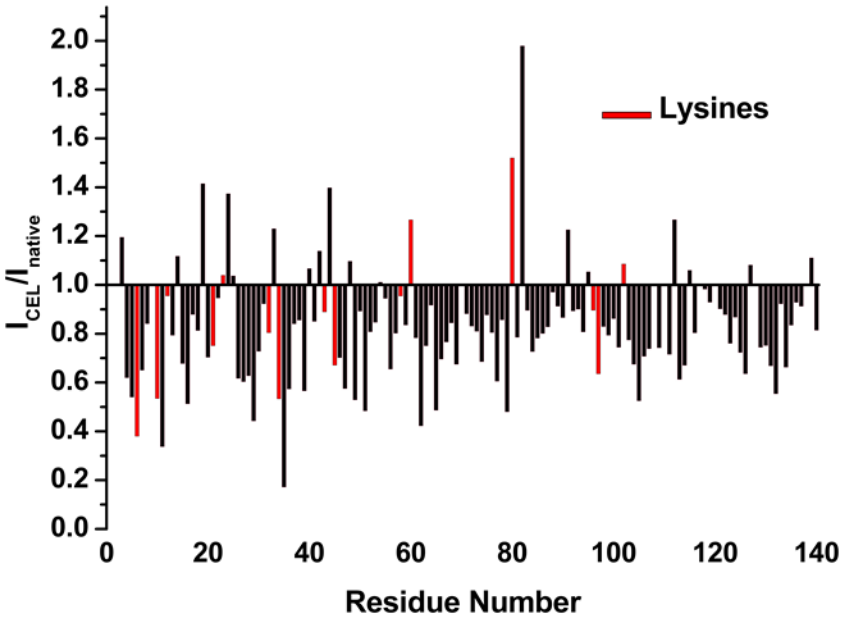

Figure S15. Ratios between the ^{15}N-HSQC peak heights of αS-CEL with those obtained for αS. The peak height of each residue in αS or αS-CEL was normalized using the height of an Asn side chain cross peak, which was used as internal standard. Data corresponding to the Lys residues are colored in red.

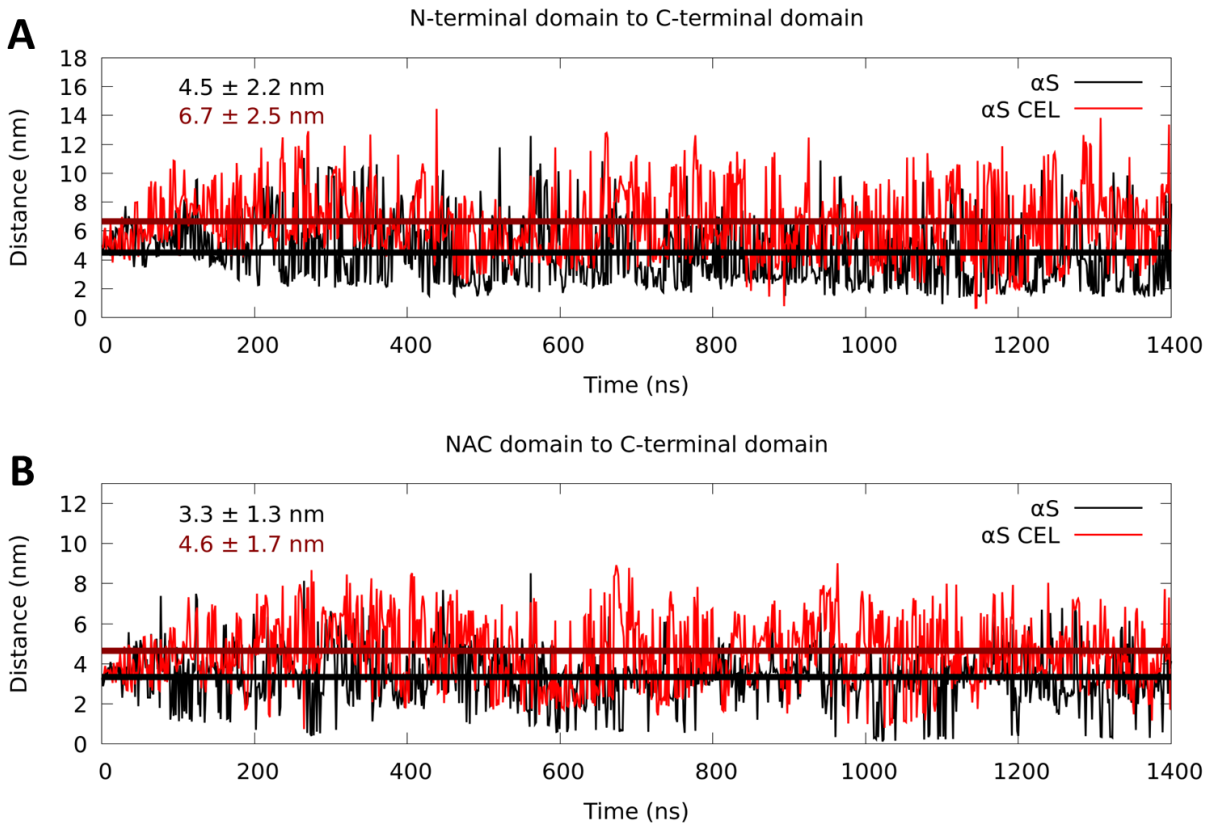

Figure S16. Distances between the centers of geometry of the N-terminal **(A)** or the NAC **(B)** domains to the centre of geometry of the C-terminal domain as a function of the simulation time for αS (*black*) and for αS-CEL (*red*). These distances were determined along 1400ns CG-MD simulations using replica exchange with solute scaling (REST2) simulations with the coarse-grained force field SIRAH. The horizontal lines represent the average distances in each case.

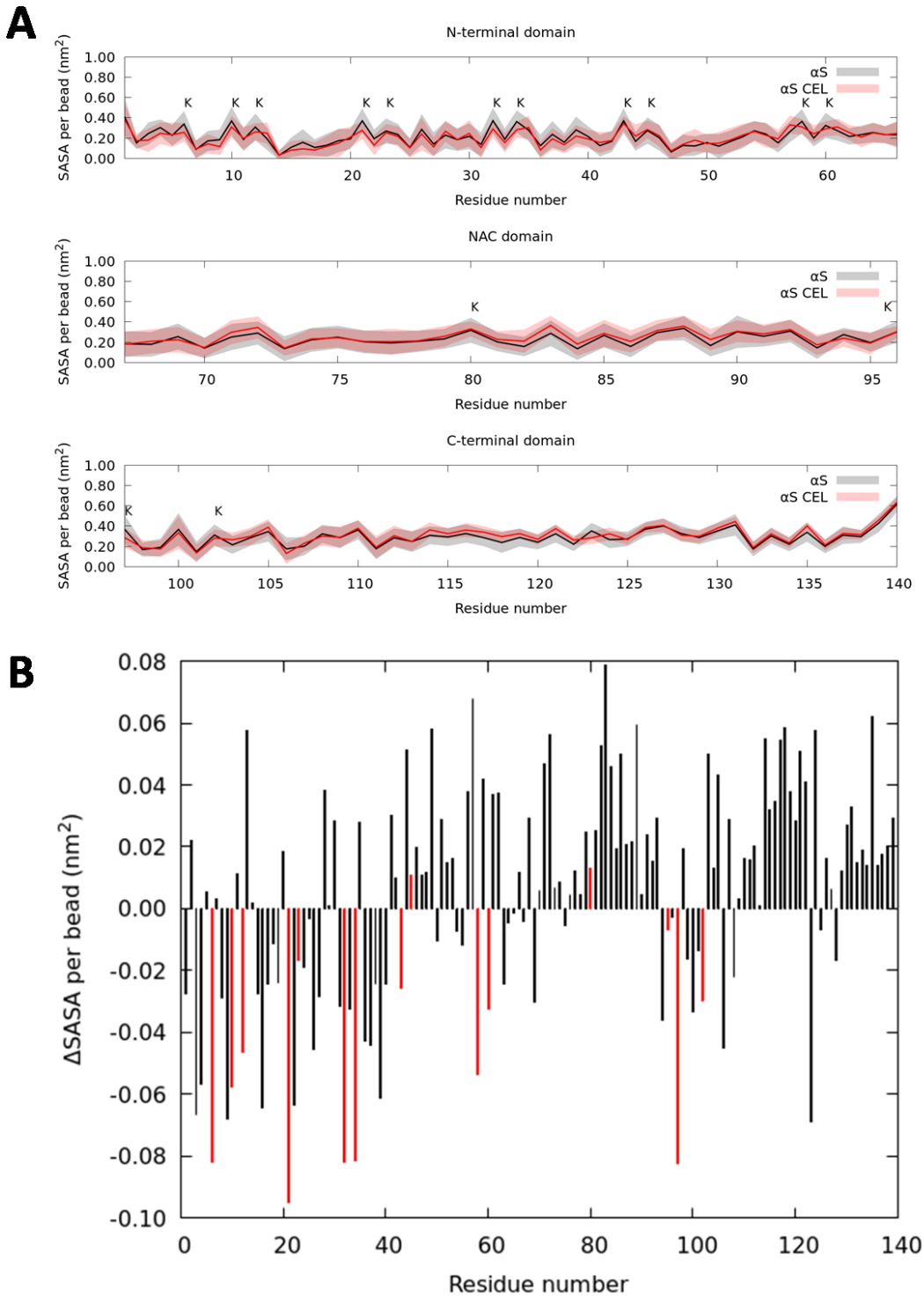

Figure S17. Effect of CEL on the solvent-accessible surface area of αS. **(A)** Average solvent-accessible surface area (SASA) per coarse-grained bead for each residue in the N-terminal (*top*), NAC (*middle*) and C-terminal (*bottom*) domains of αS (*black*) and αS-CEL (*red*). Lys/CEL locations along the sequence are marked as "K". SASA were calculated with the *sasa tool* in GROMACS for the most populated clusters in αS and in αS-CEL obtained from CG-MD simulations (*f*=1.3). The solid lines indicate the average over the trajectory and the shaded regions represent the standard deviations. **(B)** Differences in SASA per coarse-grained bead between αS-CEL and αS as a function of the residue number. Data corresponding to the Lys residues are colored in red.

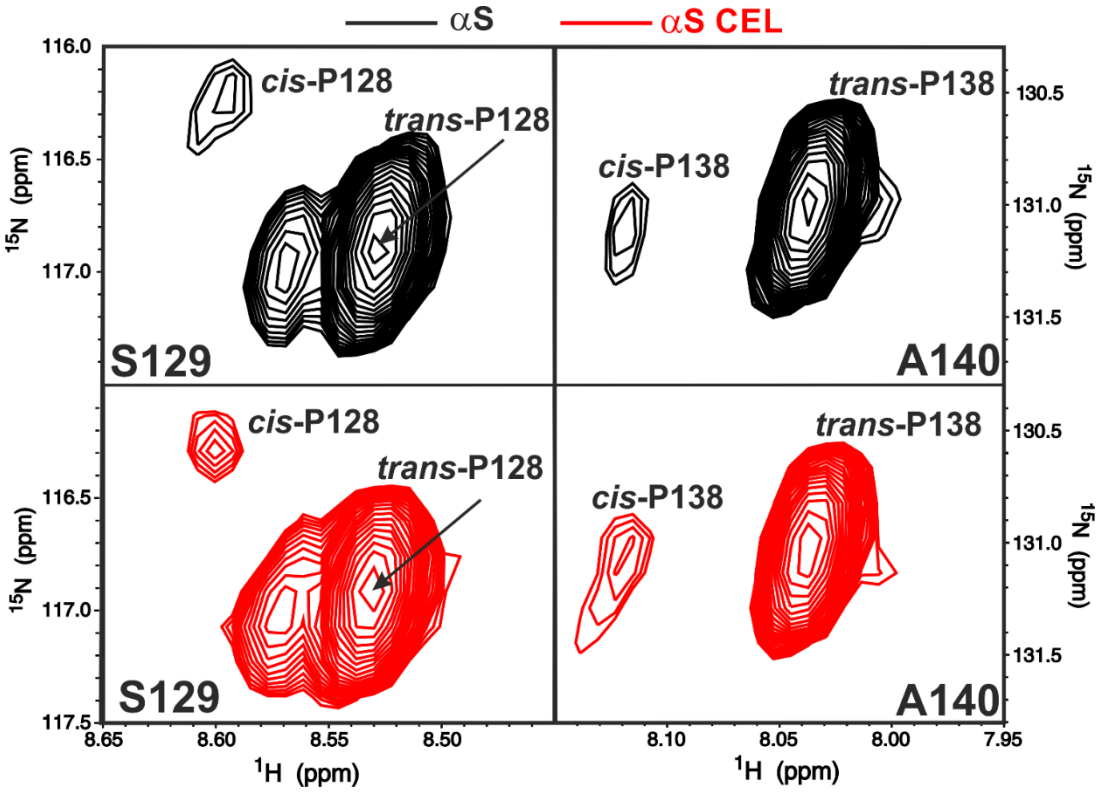

Figure S18. Zoomed-in regions corresponding to ^{15}N-HSQC peaks of S129 and A140 in αS (*black*) and in αS-CEL (*red*). The peaks of S129 and A140 are split in two, due to the effect of a neighbouring Pro, which includes a major *trans* and a minor *cis*-Pro peaks.

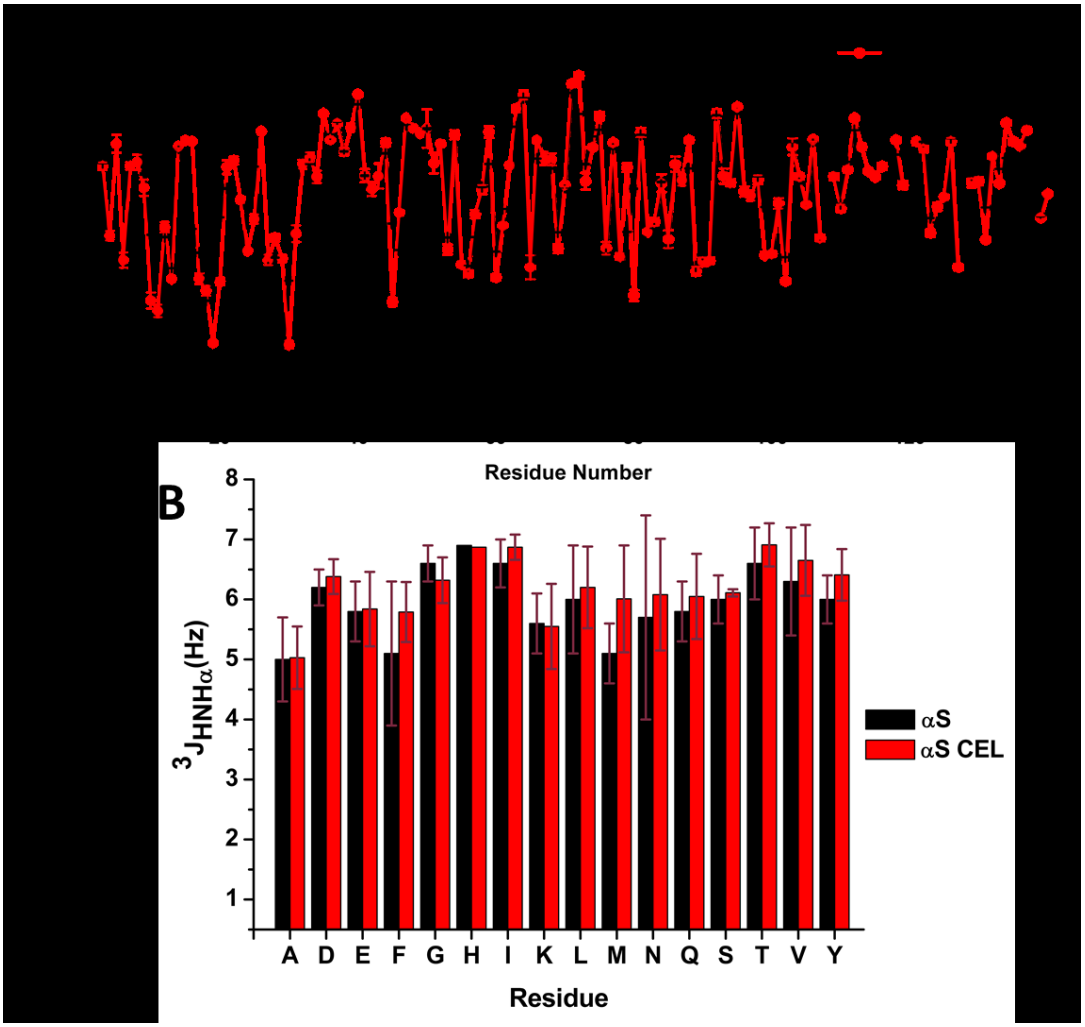

Figure S19. Effect of CEL formation on the $^3J_{HN\text{-}H\alpha}$ coupling constants of αS. **(A)** Sequence-dependent variation of the $^3J_{HN\text{-}H\alpha}$ coupling constants measured for αS (*black*) and for αS-CEL (*red*) in 20mM sodium phosphate buffer (pH 6.5) in the presence of 150mM NaCl. Error bar for each point in the coupling constant were calculated from the peak intensities and base plane noise levels from the HNHA spectra. **(B)** Distribution of $^3J_{HN\text{-}H\alpha}$ values grouped by amino-acid type. Each bar represents the averaged value for each residue type, whereas the errors represent the standard deviation for each residue type.

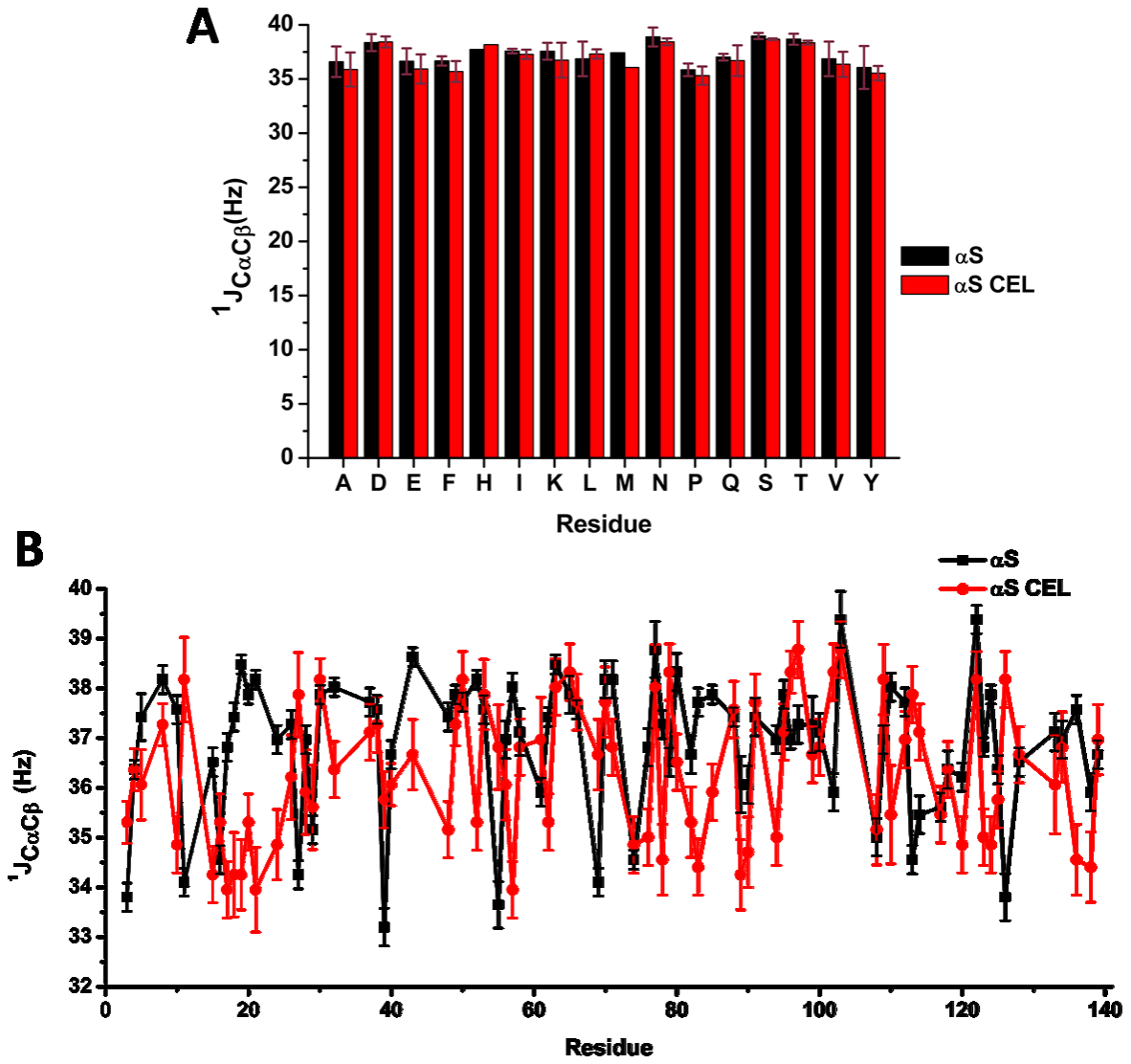

Figure S20. Effect of CEL formation on the $^1J_{C\alpha C\beta}$ coupling constants of αS. **(A)** Distribution of the $^1J_{C\alpha C\beta}$ values grouped by amino-acid type. Each bar represents the averaged value for each residue type, whereas the errors represent the standard deviation for each residue type. The values of Gly residues are not plotted since the $^1J_{C\alpha C\beta}$ is non-existent. **(B)** Sequence-dependent variation of the $^1J_{C\alpha C\beta}$ coupling constants measured for αS (*black*) and for αS-CEL (*red*) in 20mM sodium phosphate buffer (pH 6.5) in the presence of 150mM NaCl. The values of Gly are not plotted since $^1J_{C\alpha C\beta}$ does not exist for this residue. Moreover, the $^1J_{C\alpha C\beta}$ values corresponding to the Asp, Asn, Ser and Thr residues are excluded from the plots, since their higher $^1J_{C\alpha C\beta}$ values (*see panel A*) might induce the misinterpretation of the structural data[16]. Error bar for each point in the coupling constant was determined from the uncertainty in the determination of the maxium of each peak arising from the splitting of the C_α cross-peak in the HN(CO)CA spectra.

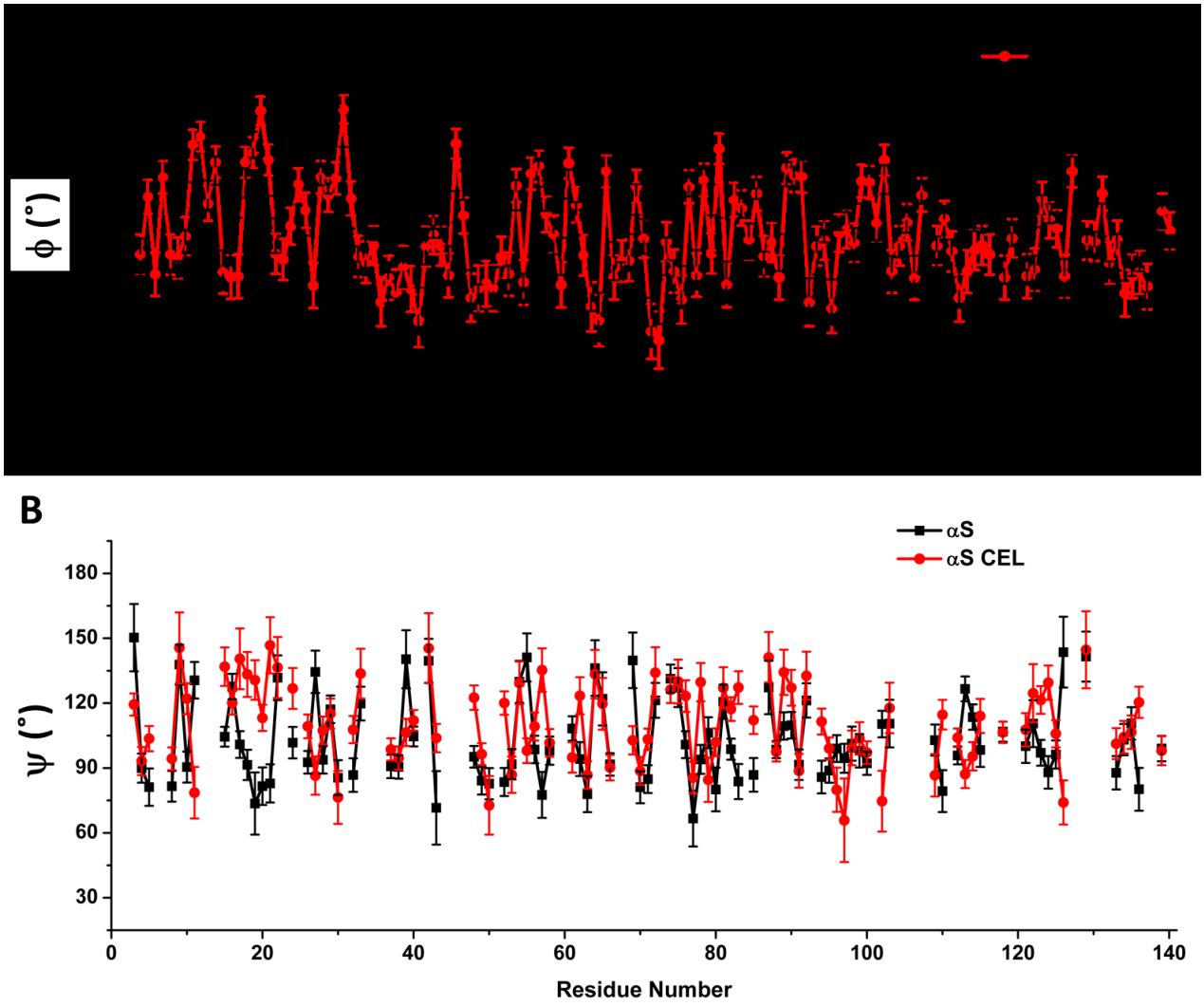

Figure S21. Values of the dihedral angles of αS and αS-CEL. **(A)** Sequence-dependent variation of the phi (ϕ) dihedral angles of αS (*black*) and αS-CEL (*red*). The ϕ dihedral angles and their errors were estimated from the $^{3}J_{HN\text{-}H\alpha}$ coupling constants as described in the materials and methods section. **(B)** Sequence-dependent variation of the psi (ψ) dihedral angles of αS (*black*) and αS-CEL (*red*). The ψ dihedral angles and their errors were estimated from the $^{1}J_{C\alpha C\beta}$ coupling constants and the ϕ dihedral angles, as described in the materials and methods section.

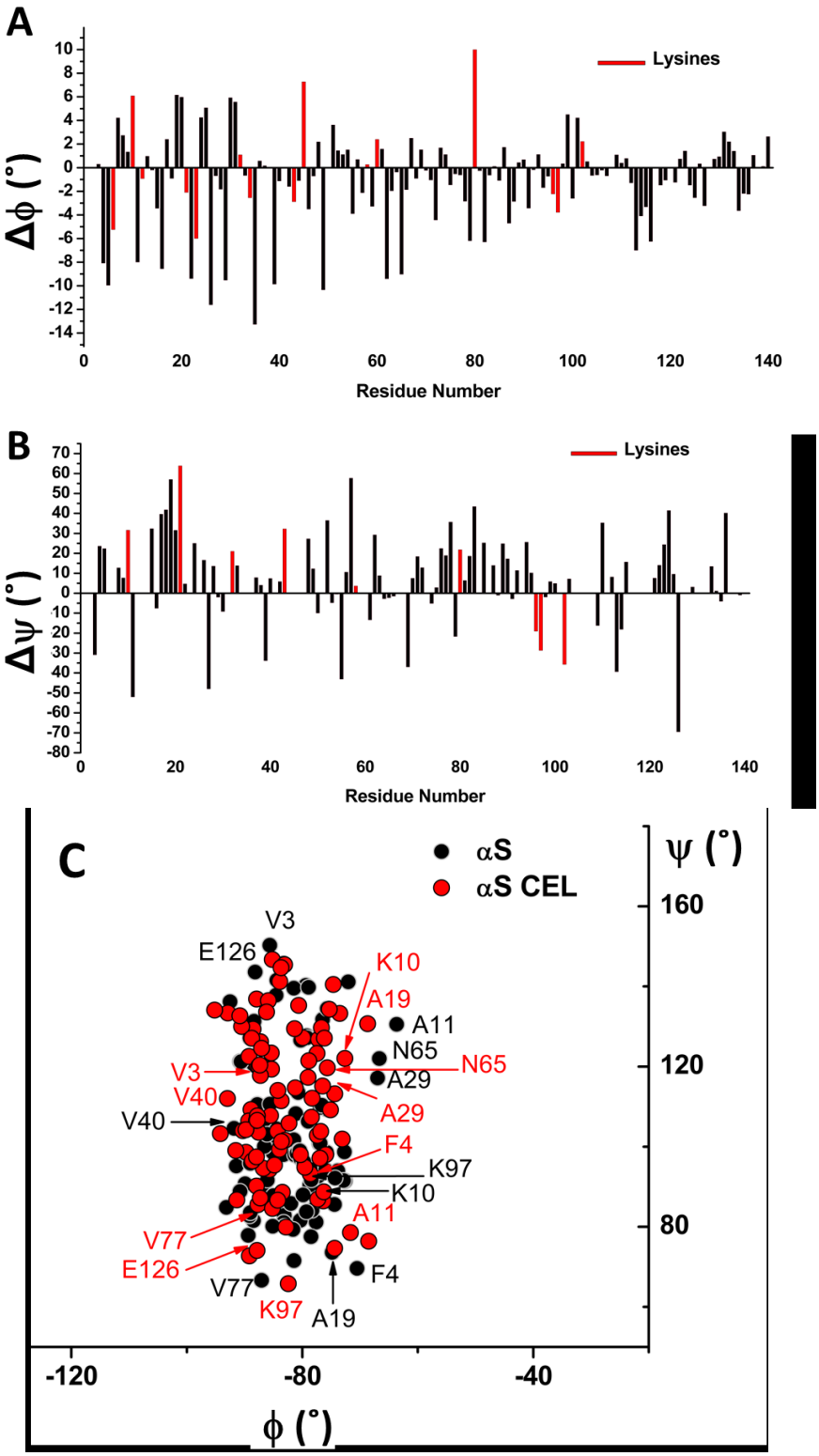

Figure S22. Effect of CEL formation on the dihedral angles of αS. **(A)** Sequence-dependent variations of the Δϕ between αS and αS-CEL ($\Delta\phi=\phi_{\_\alpha S\text{-}CEL}-\phi_{\_\alpha S}$). Data corresponding to the Lys residues are colored in red. **(B)** Sequence-dependent variations of the Δψ between αS and αS-CEL ($\Delta\psi=\psi_{\_\alpha S\text{-}CEL}-\psi_{\_\alpha S}$). Data corresponding to the Lys residues are colored in red. In panels *B* and *C*, we only plotted the Δϕ and Δψ values of those residues for which we had both angles in αS and in αS-CEL. **(C)** Zoom on the Ramachandran plot displayed in **Fig. 4C.** This plot only shows the region containing values of ψ>0 and values of ϕ<0. The signals corresponding to the ϕ/ψ angles of each amino acid are shown in *black* for αS and in *red* for αS-CEL. The signals of those residues whose Δϕ and/or Δψ between αS and αS-CEL are larger than the average, are labelled with the residue number.

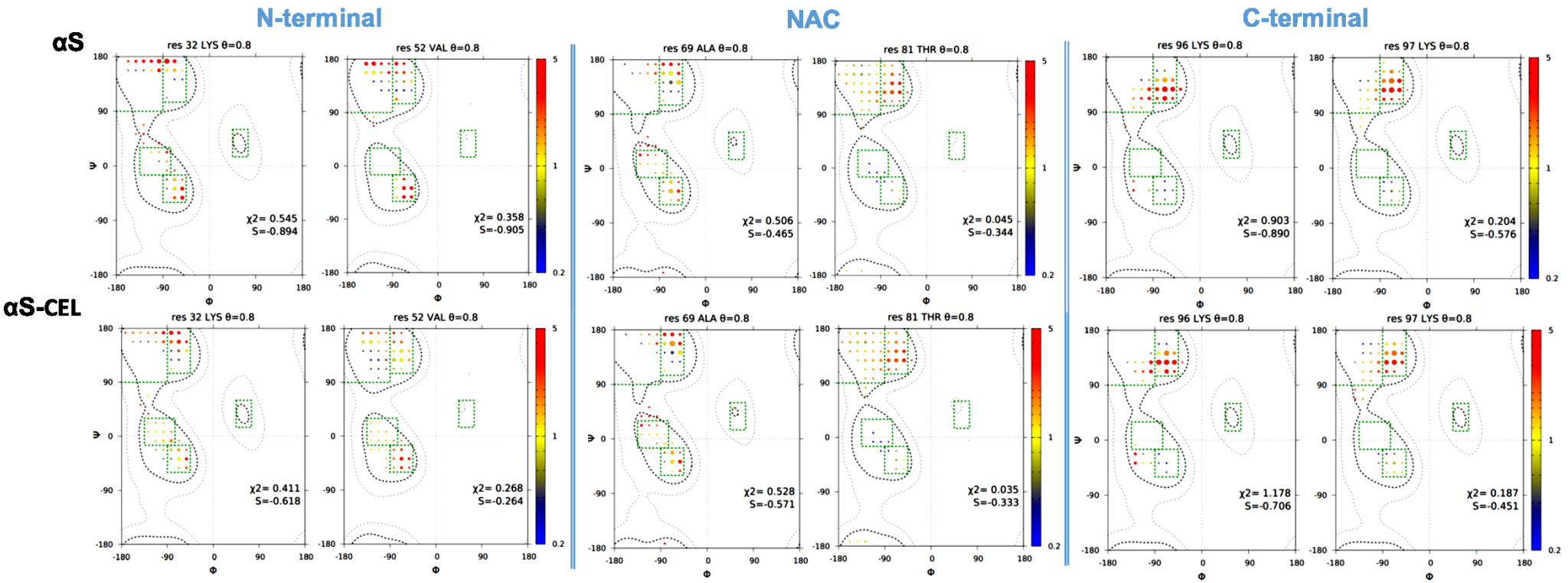

Figure S23. ϕ/ψ distributions derived from MERA calculations for αS (*top*) and αS-CEL (*bottom*). As examples, we have chosen the Ramachandran distribution plots for two residues representative of the N-terminal domain (i.e. K32 and V52; *left*), of the NAC domain (i.e. A69 and T81; *middle*), and of the C-terminal domain (i.e. K96 and K97; *right*). The surface area of each circle is proportional to the population of its 15°x15° voxel, and its color represents the ratio relative to that of the population seen in the coil database for that residue type, from 0.2 (*blue*) to 5 (*red*). An entropy weight factor of 0.8 was used. Each Ramachandran plot includes its χ^2 and S values.

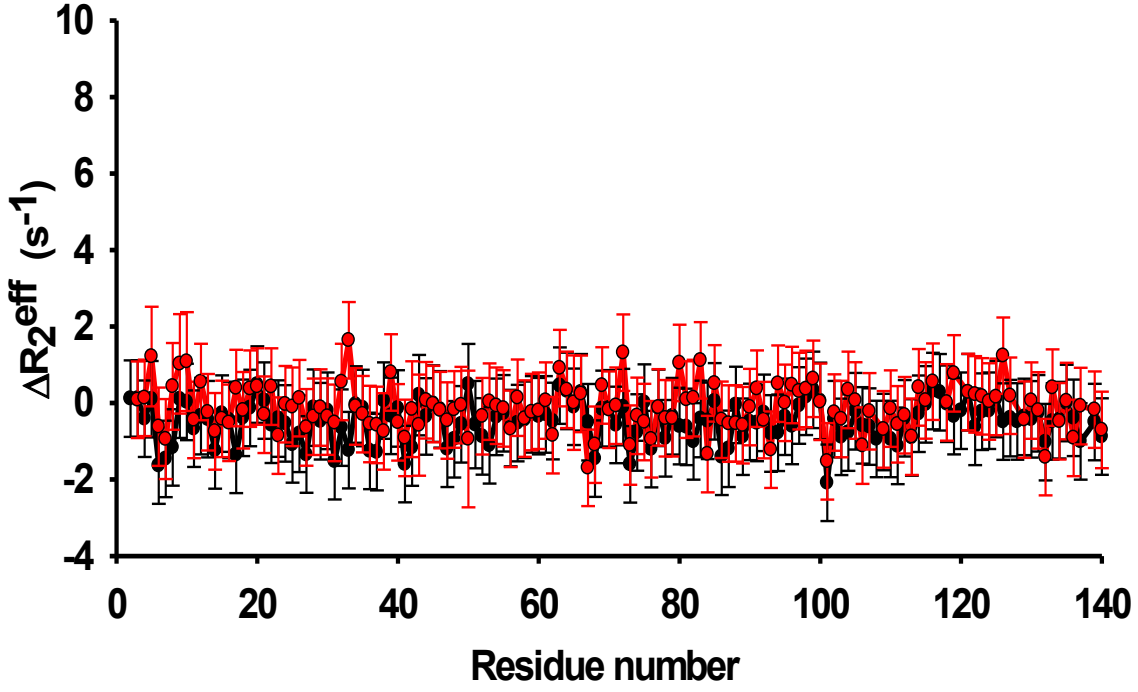

Figure S24. CPMG RD data showing the difference between the effective R_2 rates (R_2^{eff}) at low (31.25 Hz) and high (1kHz) CPMG frequencies obtained at 600MHz for ^{15}N-αS (black) and αS-CEL (red) at 12.5°C.

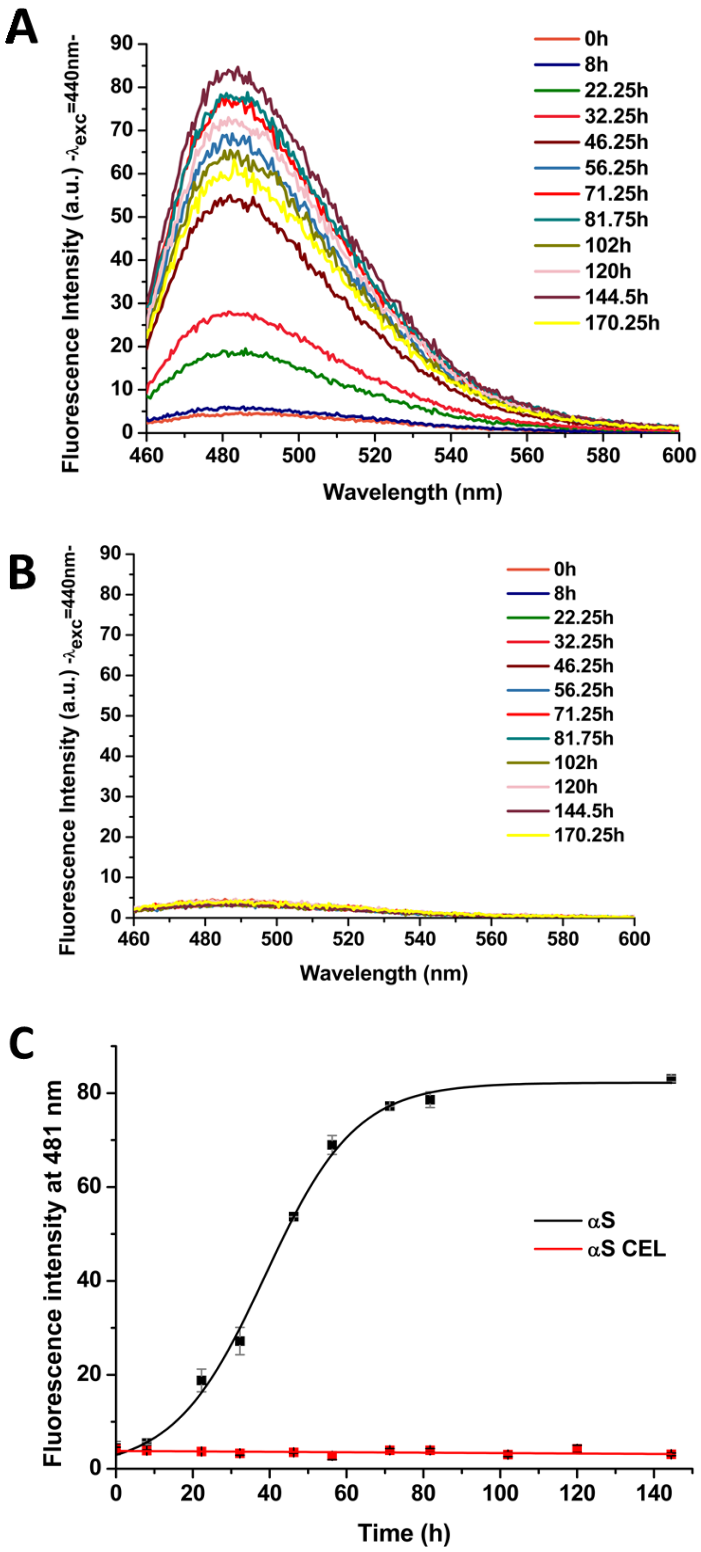

Figure S25. Study of amyloid fibril formation from αS and αS-CEL. **(A,B)** ThT fluorescence spectra of solutions containing αS **(A)** or αS-CEL **(B)** (at 10μM protein concentration), which were prepared in 20mM phosphate buffer (pH 7.4) in the presence of NaCl (150mM). The spectra were acquired at different incubation times (37°C; see figure legends) in the presence of ThT (50μM). **(C)** Changes in the ThT fluorescence intensity (λ_{exc} 440 nm) as a function of the incubation time (at 37°C while shaking at 1000rpm) of solutions containing αS (black) or αS-CEL (red) at 10μM protein concentration. Solutions were prepared in 20mM phosphate buffer (pH 7.4) and in the presence of NaCl (150mM). ThT (50μM) was added before measurements.

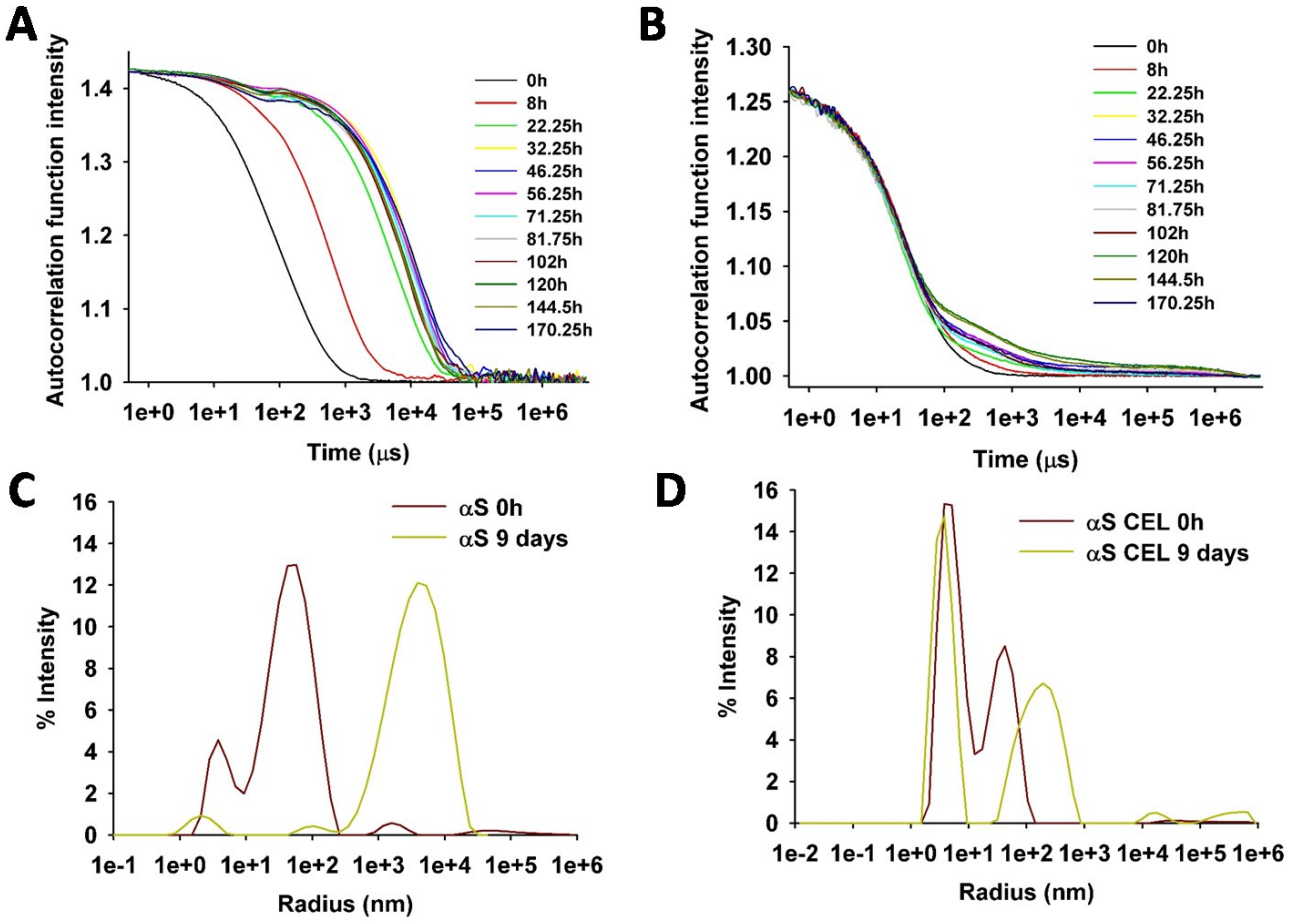

Figure S26. DLS study of the aggregation process of αS and αS-CEL. **(A-B)** DLS autocorrelation functions obtained at different incubation times (*see figure legends*) for solutions containing monomeric αS **(A)** or monomeric αS-CEL **(B)** at 70μM protein concentration. Solutions were incubated at 37°C in 20mM phosphate buffer (pH 7.4) and in the presence of 150mM NaCl, while shaking at 1000rpm. **(C-D)** Intensity-weighted DLS size distributions obtained for monomeric αS **(C)** and monomeric αS-CEL **(D)** after 0 and 9d of incubation at 37°C in 20mM phosphate buffer (pH 7.4) and in the presence of 150mM NaCl, while shaking at 1000rpm.

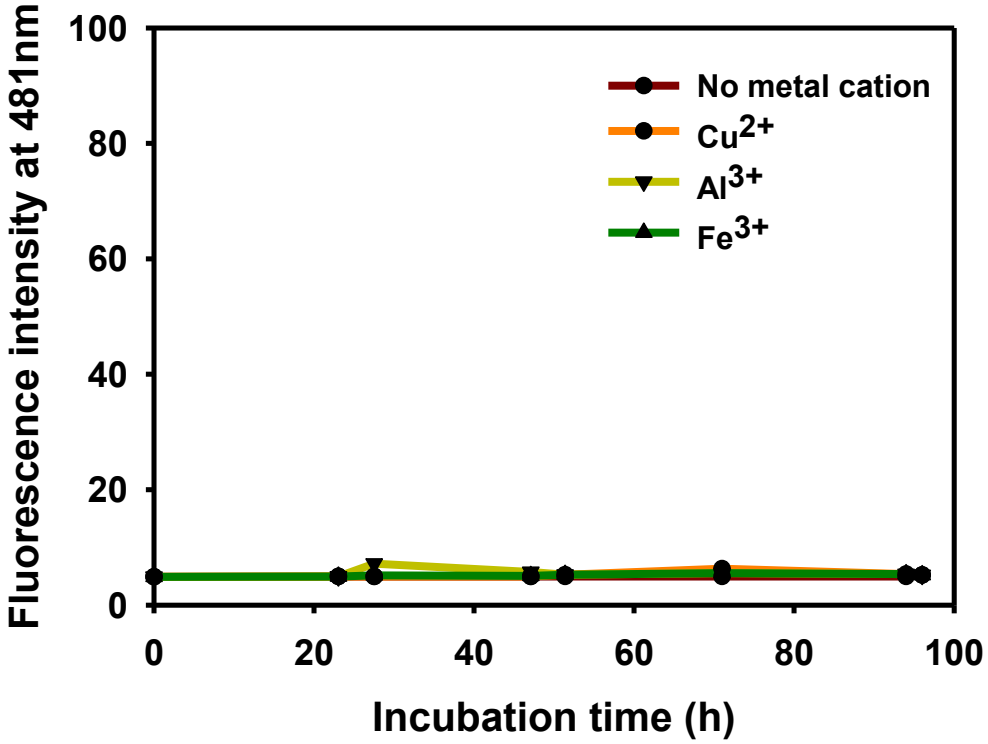

Figure S27. Changes in the ThT fluorescence intensity (λ_{exc} 440 nm) as a function of the incubation time (at 37°C while shaking at 1000rpm) of solutions containing αS-CEL (10μM) either alone or in the presence of $FeCl_3$, $AlCl_3$ or $CuCl_2$ (150μM). All the reaction mixtures were prepared in 20mM phosphate buffer (pH 7.4) in the presence of NaCl (150mM). ThT (50μM) was added before measurements.

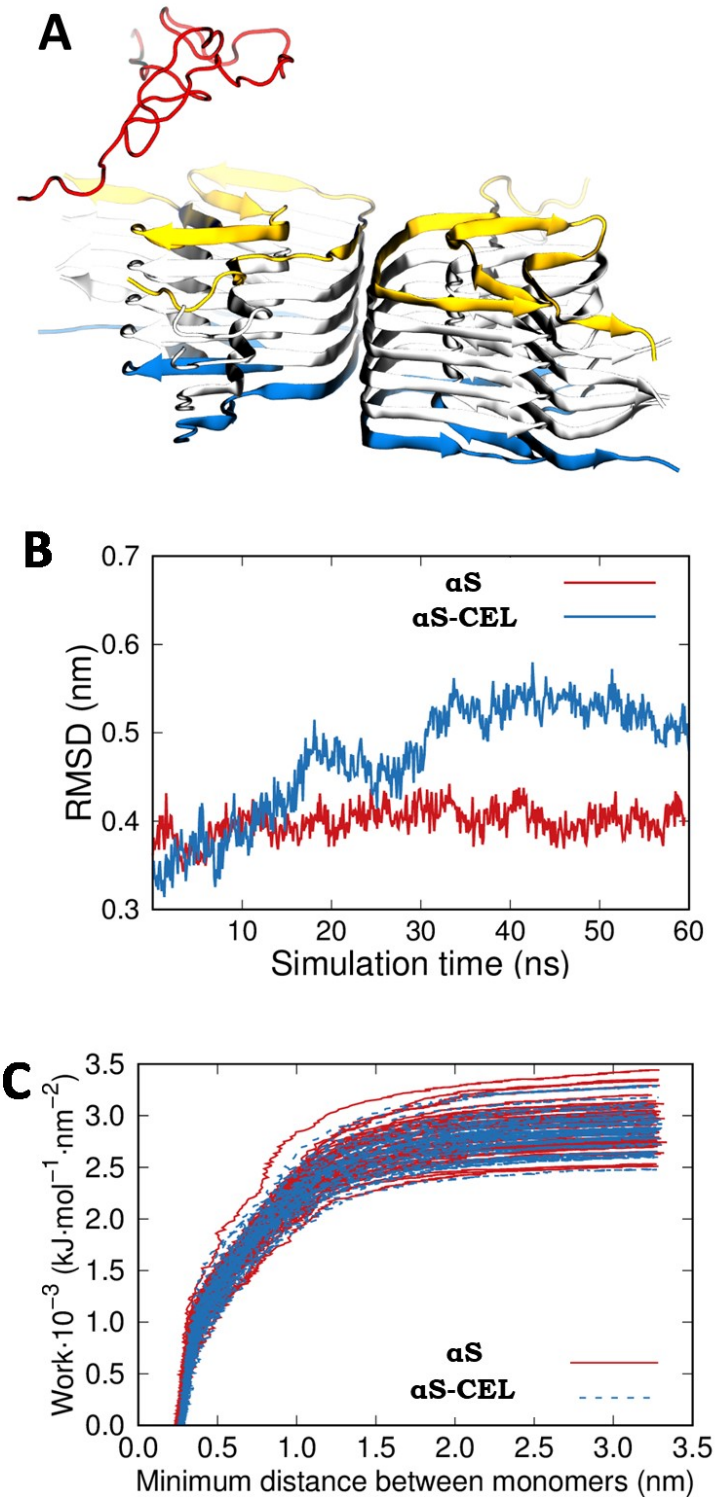

Figure S28. All-atom MD and SMD simulations carried out on the native and CEL-modified cryo-EM structure of αS and αS-CEL fibrils. **(A)** Detachment of a αS or αS-CEL monomer (red) from the ending extreme of the amyloid fibril. The CV for the detachment in the SMD simulations was defined as the minimum distance with respect to selected atoms of the closest monomers (yellow). Position restraints were applied to the monomers at the beginning extreme of the fibril (blue). **(B)** Root mean square deviation (RMSD) of the backbone atoms of the αS or αS-CEL monomers at the ending extreme of the fibrils. **(C)** Work performed for the detachment of the ending monomers from the respective native and CEL-modified fibrils. Each line corresponds to one replica.

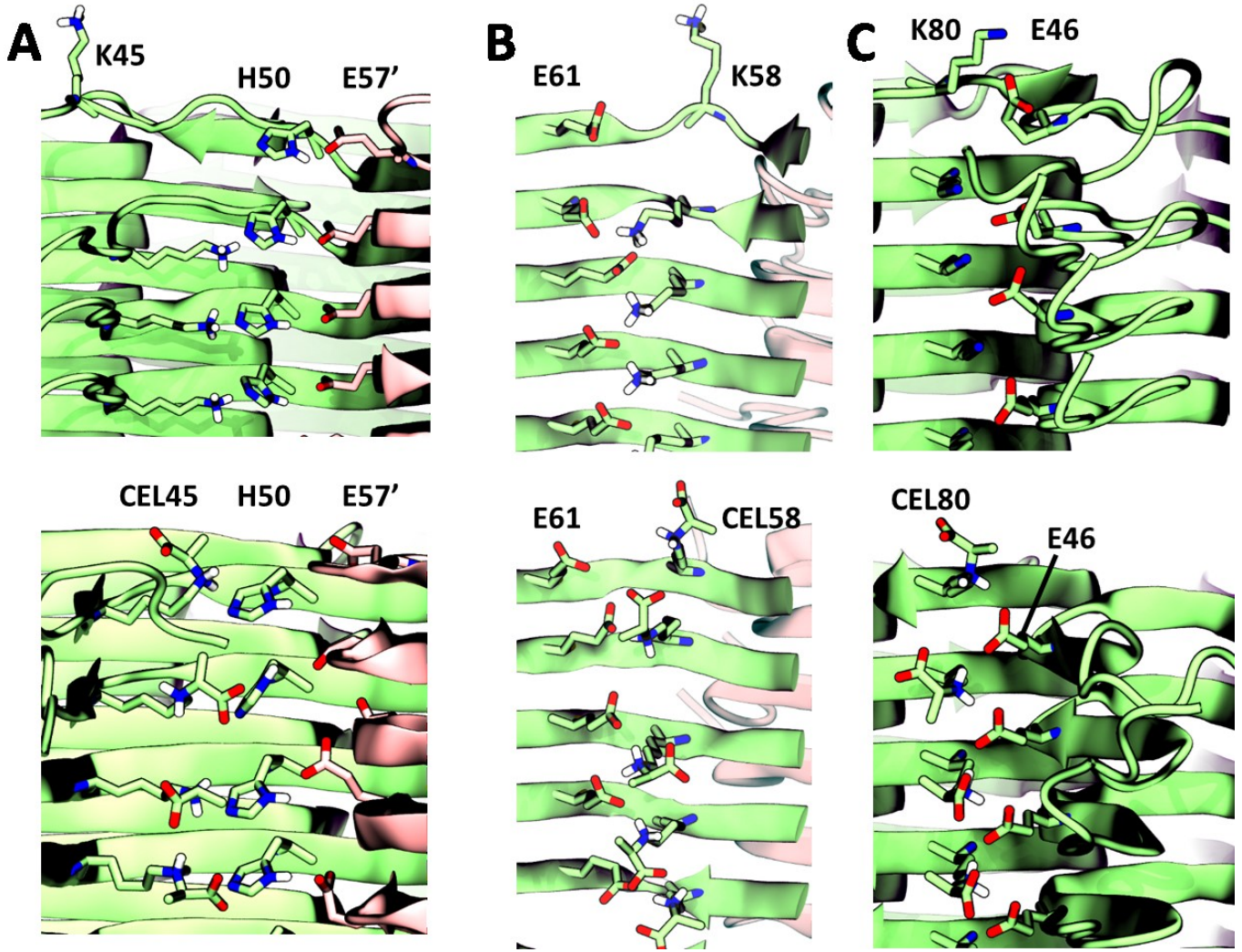

Figure S29. Comparison of key Lys-involving interactions in the αS fibril assembly between αS fibrils (*top*) and CEL-modified αS fibrils (*bottom*). **(A)** Electrostatic interactions between K45/CEL45, H50 and E57'. **(B-C)** Salt bridges formed between E61 and K58 **(B)**, and between K80 and E46 **(C)**. In the representations, the αS molecules corresponding to each intertwining protofilament are colored in green and red, respectively.

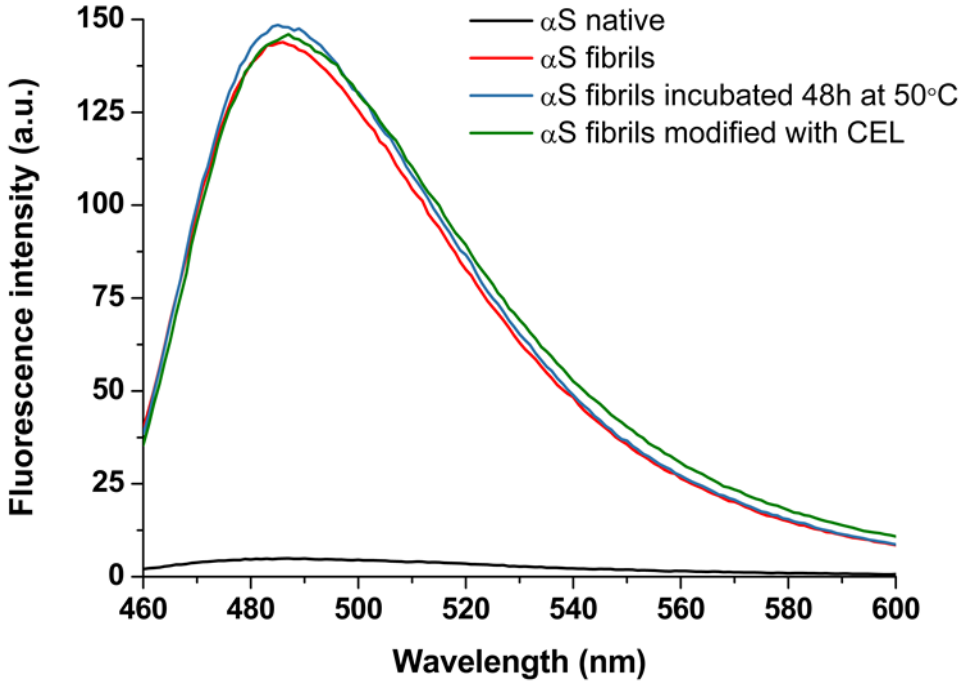

Figure S30. ThT fluorescence spectra of solutions containing monomeric native αS (black), amyloid fibrils obtained from native αS (red), amyloid fibrils obtained from native αS, which were incubated 48h at 50°C (blue), and αS amyloid fibrils modified with CEL (green). To carry out this experiment, the sample containing native αS amyloid fibrils (red), was split in two different samples and one of them was incubated in the presence of pyruvic acid and $NaBH_3CN$ (green), whereas the other was incubated in their absence (blue).

Table S1. *Secondary structure content of αS and αS-CEL derived from far-UV CD spectra using the BeStSel on-line platform (http://bestsel.elte.hu)[4] and from CG-MD simulations.*

Secondary structure (%)		αS		αS-CEL	
		CD	**CG-MD**	**CD**	**CG-MD**
	Helix	1.7	0.2±0.4	0.0	0.1±0.3
	Antiparallel	24.9	26.8±6.8	23.1	22.8±6.8
	Parallel	0.0	-	0.0	-
	Turns	18.6	-	19.5	-
	Others	54.8	73.0±6.8	57.4	77.1±5.8

Table S2. ^{1}H and ^{13}C chemical shifts obtained for non-modified and CEL-modified N^{α}-Ac-Lys.

	^{1}H (ppm)		**^{13}C (ppm)**	
	N^{α}-Ac-Lys	**CEL-modified N^{α}-Ac-Lys**	**N^{α}-Ac-Lys**	**CEL-modified N^{α}-Ac-Lys**
-NH	7.93	7.92	--	--
-CH_{α}	4.12	4.12	57.7	57.7
-CH_{β}	1.79/1.69	1.79/1.69	33.7	33.7
-CH_{γ}	1.40	1.40	24.8	25.0
-CH_{δ}	1.67	1.69	29.0	28.1
-CH_{ε}	2.99	3.01	42.2	48.7
-CH_3 (methyl)	2.04	2.01	24.6	24.6

Table S3. *Fractions (%) of cis-Pro bonds in αS and in αS-CEL.*

	Residue	***αS***	***αS-CEL***
[a]Promega (% cis-Pro)	P108	7.5	7.5
	P117	8.3	7.7
	P120	2.1	2.8
	P128	8.2	9
	P138	8.8	7.1
[b]$I_{cis}/(I_{cis}+I_{trans})$ (% cis-Pro)	P108	[c]-	5.5
	P117	5.5	7.2
	P120	5.7±0.2	6.0±1.2
	P128	4.0±1.1	4.1±0.5
	P138	2.6±0.5	3.5±0.9

[a]The fractions of cis conformation for each Pro were determined using the Promega server facility (https://spin.niddk.nih.gov/bax/nmrserver/promega/) using the N, H_N, H_α, CO, C_α and C_β chemical shifts[11c].

[b]The fractions of cis conformation for each Pro were determined using the ^{15}N-HSQC peak intensities from resonances affected by the cis and trans states accordingly to the data and the approach recently reported by Alderson et al.[45]. Hence, the ratios between cis and trans-Pro were obtained from the intensity ratios of the double ^{15}N-HSQC peaks corresponding to: A107, sensitive to the isomerization of P108; D119, sensitive to the isomerization of P117; D119 and A124, both sensitive to the isomerization of P120; A124 and S129, both sensitive to the isomerization of P128; E137 and A140, both sensitive to the isomerization of P138.

[c]The fraction could not be determined because the peak corresponding to the cis-Pro was not observed.

Table S4. *Number of resulting clusters generated for ensemble analysis as a function of the cut off.*

Cut off (nm)	N
2.0	69
1.9	114
1.8	216
1.7	453
1.6	1098
1.5	2719

The random structures were generated by using the TraDES-2 software, and they were clustered with the gromos method, implemented in Gromacs 2016.4 (*see materials and methods for description).*

Table S5. *RMS residuals and the determination coefficient R^2 of the linear regression between the predicted ensemble-weighted average δH_a with respect to the experimental values*

Cutoff (nm)	RMS αS	RMS αS-CEL	R^2 αS	R^2 αS-CEL
2.0	1.5	1.5	0.69	0.45
1.9	1.5	1.5	0.69	0.71
1.8	1.5	1.4	0.67	0.69
1.7	1.5	1.5	0.6	0.49

Table S6. *RMS residuals and the determination coefficient R^2 of the linear regression between the predicted ensemble-weighted average δC_a with respect to the experimental values*

Cutoff (nm)	RMS αS	RMS αS-CEL	R^2 αS	R^2 αS-CEL
2.0	0.5	0.7	0.99	0.98
1.9	0.6	0.6	0.99	0.99
1.8	0.6	0.6	0.99	0.99
1.7	0.6	0.8	0.99	0.98

Table S7. *RMS residuals and the determination coefficient R^2 of the linear regression between the predicted ensemble-weighted average δCO with respect to the experimental values*

Cutoff (nm)	RMS αS	RMS αS-CEL	R^2 αS	R^2 αS-CEL
2.0	0.7	0.7	0.87	0.79
1.9	0.8	0.7	0.85	0.85
1.8	1.0	0.7	0.81	0.82
1.7	0.9	0.7	0.80	0.78

REFERENCES

1 C. Roodveldt, A. Andersson, E. J. De Genst, A. Labrador-Garrido, A. K. Buell, C. M. Dobson, G. G. Tartaglia and M. Vendruscolo, *Biochemistry*, 2012, **51**, 8771–8778.

2 R. F. Borch, M. D. Bernstein and H. D. Durst, *J. Am. Chem. Soc.*, 1971, **93**, 2897–2904.

3 H. Martínez-Orozco, L. Mariño, A. B. Uceda, J. Ortega-Castro, B. Vilanova, J. Frau and M. Adrover, *ACS Chem. Neurosci.*, 2019, **10**, 2919–2930.

4 A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers and J. Kardos, *Proc. Natl. Acad. Sci.*, 2015, **112**, E3095–E3103.

5 M. Piotto, V. Saudek and V. Sklenár, *J. Biomol. NMR*, 1992, **2**, 661–666.

6 D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley and B. D. Sykes, *J. Biomol. NMR*, 1995, **6**, 135–40.

7 F. Delaglio, S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277–293.

8 W. Bermel, I. Bertini, I. C. Felli, Y. M. Lee, C. Luchinat and R. Pierattelli, *J. Am. Chem. Soc.*, 2006, **128**, 3918–3919.

9 J. A. Marsh, V. K. Singh, Z. Jia and J. D. Forman-Kay, *Protein Sci.*, 2006, **15**, 2795–804.

10 K. Tamiola and F. A. A. Mulder, *Biochem. Soc. Trans.*, 2012, **40**, 1014–20.

11 *(a)* Y. Shen, F. Delaglio, G. Cornilescu and A. Bax, *J. Biomol. NMR*, 2009, **44**, 213.; *(b)* Y. Shen and A. Bax, *J. Biomol. NMR*, 2010, **48**, 13–22; *(c)* Y. Shen and A. Bax, *J. Biomol. NMR*, 2010, **46**, 199–204

12 G. W. Vuister and A. Bax, *J. Am. Chem. Soc.*, 1993, **115**, 7772–7777.

13 G. Cornilescu, A. Bax and D. A. Case, *J. Am. Chem. Soc.*, 2000, **122**, 2168–2171.

14 *(a)* A. B. Mantsyzov, A. S. Maltsev, J. Ying, Y. Shen, G. Hummer and A. Bax, *Protein Sci.*, 2014, **23**, 1275–1290; *(b)* A. B. Mantsyzov, Y. Shen, J. H. Lee, G. Hummer and A. Bax, *J. Biomol. NMR*, 2015, **63**, 85–95.

15 J. H. Lee, F. Li, A. Grishaev and A. Bax, *2J. Am. Chem. Soc.*, 2015, **137**, 1432–1435.

16 J. M. Schmidt, M. J. Howard, M. Maestre-Martínez, C. S. Pérez and F. Löhr, *Magn. Reson. Chem.*, 2009, **47**, 16–30.

17 *(a)* D. Wu, A. Chen and C. S. Johnson Jr, *J. Magn. Reson., Ser. A,* 1995, **115** , 260-264; *(b)* E. Martínez-Viviente and P. S. Pregosin, *Helv. Chim. Acta*, 2003, **86**, 2364-2378.

18 *(a)* M. Caffrey, J. Kaufman, S. J. Stahl, P. T. Wingfield, A. M. Gronenborn and G. M. Clore, *J. Magn. Reson.*, 1998, **135**, 368–372; *(b)* K. Chen and N. Tjandra, *J. Magn. Reson.*, 2009, **197**, 71–76.

19 T. Yuwen and N. R. Skrynnikov, *J. Magn. Reson.*, 2014, **241**, 155–169.

20 N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay and L. E. Kay, *Biochemistry*, 1994, **33**, 5984–6003.

21 A. F. Hansen, P. Vallurupalli, L. E. Kay. *J. Phys. Chem. B.*, 2008, **112**, 5898–5904.

22 H. J. Feldman and C. W. V. Hogue, *Proteins Struct. Funct. Genet.*, 2000, **39**, 112–131.

23 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chemie Int. Ed.*, 1999, **38**, 236–240.

24 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

25 Z. Ziegler, M. Schmidt, T. Gurry, V. Burger and C. M. Stultz, *Bioinformatics*, 2016, **32**, 2545–2547.

26 S. S. Nielsen, K. Noergaard Toft, D. Snakenborg, M. G. Jeppesen, J. K. Jacobsen, B. Vestergaard, J. P. Kutteraand L. Arleth. *J. Appl. Cryst.,* 2009, **42**, 959-964.

27 R. Ramis, J. Ortega-Castro, R. Casasnovas, L. Mariño, B. Vilanova, M. Adrover and J. Frau, *J. Chem. Inf. Model.*, 2019, **59**, 1458–1471.

28 J. R. Allison, P. Varnai, C. M. Dobson and M. Vendruscolo, *J. Am. Chem. Soc.*, 2009, **131**, 18314–18326.

29 Y. Li, C. Zhao, F. Luo, Z. Liu, X. Gui, Z. Luo, X. Zhang, D. Li, C. Liu and X. Li, *Cell Res.*, 2018, **28**, 897–903.

30 Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *J. Comput. Chem.*, 2003, **24**, 1999–2012.

31 W. L. Jorgensen, *J. Am. Chem. Soc.*, 1981, **103**, 335–340.

32 I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B*, 2008, **112**, 9020–9041.

33 L. Mariño, R. Casasnovas, R. Ramis, B. Vilanova, J. Ortega-Castro, J. Frau and M. Adrover, *Int. J. Biol. Macromol.*, 2019, **129**, 254–266.

34 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

35 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

36 H. Grubmüller, B. Heymann and P. Tavan, *Science*, 1996, **271**, 997–9.

37 C. Jarzynski, *Phys. Rev. Lett.*, 1997, **78**, 2690–2693.

38 P. Ulrich and A. Cerami, *Recent Prog. Horm. Res.*, 2001, **56**, 1–22.

39 I. Allaman, M. Bélanger and P. J. Magistretti, *Front. Neurosci.*, 2015, **9**, 23.

40 H. Vicente Miranda, É. M. Szegő, L. M. A. Oliveira, C. Breda, E. Darendelioglu, R. M. de Oliveira, D. G. Ferreira, M. A. Gomes, R. Rott, M. Oliveira, F. Munari, F. J. Enguita, T. Simões, E. F. Rodrigues, M. Heinrich, I. C. Martins, I. Zamolo, O. Riess, C. Cordeiro, A. Ponces-Freire, H. A. Lashuel, N. C. Santos, L. V. Lopes, W. Xiang, T. M. Jovin, D. Penque, S. Engelender, M. Zweckstetter, J. Klucken, F. Giorgini, A. Quintas and T. F. Outeiro, *Brain*, 2017, **140**, 1399–1419.

41 T. E. Friedemann, *J. Biol. Chem.*, 1927, **73**, 331–334.

42 *(a)* Z. Dosztanyi, V. Csizmok, P. Tompa and I. Simon, *Bioinformatics*, 2005, **21**, 3433–3434; *(b)* Z. R. Yang, R. Thomson, P. McNeil and R. M. Esnouf, *Bioinformatics*, 2005, **21**, 3369–3376; *(c)* L. P. Kozlowski and J. M. Bujnicki, *BMC Bioinformatics*, 2012, **13**, 111; *(d)* X. Fan and L. Kurgan, *J. Biomol. Struct. Dyn.*, 2014, **32**, 448–464; *(e)* T. Ishida and K. Kinoshita, *Nucleic Acids Res.*, 2007, **35**, W460; *(f)* R. Linding, L. J. Jensen, F. Diella, P. Bork, T. J. Gibson and R. B. Russell, *Structure*, 2003, **11**, 1453–9; *(g)* M. J. Mizianty, W. Stach, K. Chen, K. D. Kedarisetti, F. M. Disfani and L. Kurgan, *Bioinformatics*, 2010, **26**, i489-96; *(h)* L. J. McGuffin, *Bioinformatics*, 2008, **24**, 1798–1804; *(i)* D. T. Jones and D. Cozzetto, *Bioinformatics*, 2015, **31**, 857–63.

43 *(a)* D. W. A. Buchan, F. Minneci, T. C. O. Nugent, K. Bryson and D. T. Jones, *Nucleic Acids Res.*, 2013, **41**, W349-57; *(b)* I. Walsh, A. J. M. Martin, T. Di Domenico, A. Vullo, G. Pollastri and S. C. E. Tosatto, *Nucleic Acids Res.*, 2011, **39**, W190-6.

44 J. A. Marsh, V. K. Singh, Z. Jia and J. D. Forman-Kay, *Protein Sci.*, 2006, **15**, 2795–804.

45 T. R. Alderson, J. H. Lee, C. Charlier, J. Ying and A. Bax, *ChemBioChem*, 2018, **19**, 37–42.