Electronic Supplementary Information

Cu^{2+}-binding to S100B triggers polymerization of disulfide cross-linked tetramers with enhanced chaperone activity against amyloid-β aggregation

Joana S. Cristóvão[1,2], Guilherme G. Moreira[1,2], Filipe E. P. Rodrigues[1,2], Ana P. Carapeto[1,3], Mário S. Rodrigues[1,3], Isabel Cardoso[4,5], António Ferreira[2], Miguel Machuqueiro[1,2], Güenter Fritz[6], Cláudio M. Gomes[1,2*]

[1] Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
[2] Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade Lisboa, Lisbon, Portugal
[3] Departamento de Física, Faculdade de Ciências, Universidade Lisboa, Lisbon, Portugal
[4] IBMC - Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal.
[5] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[6] Institute of Microbiology, University of Hohenheim, 70599 Stuttgart, Germany

* Corresponding author: C.M. Gomes (cmgomes@fc.ul.pt)

1. Supplementary methods and analysis

Materials. All reagents were of the highest grade commercially available. Thioflavin T (ThT) was obtained from Sigma. A Chelex resin (Bio-Rad) was used to remove contaminant trace metals from all solutions. Human S100B and mutant ΔCys (cysteine deficient variant (C68S-C84S) were expressed using the pGEMEX vector in *E. coli* and purified to homogeneity using a previously established protocol [1]. ApoS100B was prepared with incubation at 37ºC for 2

hours with a 300-fold excess of dithiothreitol (DTT) and 0.5 mM ethylenediamine tetraacetic acid (EDTA) and eluted in a Superdex S75 Tricorn column (GE Healthcare). Aβ42 was recombinantly expressed in *E. coli* and purified as in [2]. To obtain the monomeric form, 1 mg of Aβ42 was dissolved in 7 M guanidine hydrochloride and eluted in a Superdex S75 Tricorn Column (GE Healthcare) with 50 mM HEPES pH 7.4. Low-binding tubes (ref: MCT-150-L-C, Axygen Scientific, Corning) were used for all manipulations.

Electron Paramagnetic Resonance. S100B dimer was 0.58 mM in 25 mM NEM, 100 mM NaCl, 20% glycerol, in D_2O, pD 7.6. The samples were prepared under exclusion of dioxygen in a glove box. Oxygen was removed from all solutions applying ten cycles of vacuum and saturation with argon. The isotope ^{63}Cu was used to reconstitute S100B and 1.16 mM $CuCl_2$ were added per S100B dimer. For the sample containing Ca^{2+} 2 mM of $CaCl_2$ was added. The EPR samples (255 µl, 4-mm standard quartz tubes) were frozen in an *iso*-pentane bath cooled with liquid nitrogen to 173 K and stored in liquid nitrogen prior to the measurement. EPR spectra were recorded at 70 K with a Bruker Elexsys 500 instrument, with the following settings: microwave frequency 9.65 GHz (X-band), microwave power 0.2 mW, modulation frequency 100 kHz, modulation amplitude 10 Gauss; 4 spectra between 2500 and 3700 Gauss were averaged.

Circular Dichroism (CD) Spectroscopy. CD measurements were performed on a Jasco J-1500 spectropolarimeter equipped with a Peltier-controlled thermostated cell support. Far-UV CD spectra (200 to 260 nm) were recorded at 5 µM S100B with or without molar equivalents of $CuCl_2$ in 50 mM TRIS pH 7.4, using a 1 mm pathlength cell. Eight scans were recorded. Visible UV-CD (260 to 800 nm) were recorded at 100 µM S100B with and without molar equivalents of $CuCl_2$ in 50 mM TRIS pH 7.4, using a 1 cm pathlength cell. Five scans were recorded, baseline spectra were subtracted from each spectrum followed by smoothing.

Fluorescence Spectroscopy. Fluorescence measurements were performed on a Jasco FP-8200 spectrofluorometer. Temperature was kept at 25 ºC by a peltier-controlled cell support. For ANS analysis, S100B (40 µM) was incubated with up to 4 molar equivalents of $CuCl_2$ for 30 minutes. ANS (80 µM) was added and incubated for more 30 minutes. ANS emission spectra were recorded using 10 nm excitation and emission slits upon 370 nm excitation. For ThT binding analysis 20 µM S100B were incubated with up to 4 molar equivalents of Cu^{2+} for

2 hours. ThT (10 µM) was added and the emission spectrum recorded upon excitation at 440 nm using 10 nm emission and excitation slits.

Visible absorption Spectroscopy. Visible absorption measurements were performed on a Shimadzu spectrophotometer. The absorption spectra were acquired from 350 nm to 850 nm using S100B (100 µM, wild-type and S100BΔCys) with and $CuCl_2$ in 50 mM TRIS pH 7.4, using a 1 cm pathlength cell.

Electrophoretic Assays. SDS-PAGE analysis of AMS binding to S100B under reducing/non-reducing conditions was carried out using 12% acrylamide gel electrophoresis. Samples were incubated with and without 10 mM DTT for 10 minutes and then incubated 1 hour with 10 mM AMS in 50 mM TRIS pH 7.4. For the analysis at different Cu^{2+}:S100B ratios, S100B was incubated for 1h at room temperature in the presence of copper. After incubation, loading buffer was added and samples were subsequently analysed by SDS-PAGE gel using Low Molecular Weight Protein Markers as standards.

Analytical size exclusion chromatography. Analytical size exclusion chromatography was performed at room temperature on a Superdex 75 Tricorn high performance column (GE Healthcare, v=24 mL) typically eluted at 1 ml/min with 50 mM TRIS pH 7.4 0.2M NaCl. For the separation of S-S S100B oligomers, S100B (20-150µM) was incubated 2, 4, 17 and 28h at 37ºC at different Cu^{2+}:S100B ratios, centrifuged 5 min at 15000 rpm in a bench centrifuge at 4ºC, and applied in the column for elution.

Transmission electron microscopy. For the structural and morphological analysis, samples (5 µl) were adsorbed to carbon-coated collodion film supported on 400-mesh copper grids, and negatively stained with 1% uranyl acetate. The grids were visualized with a JEOL JEM-1400 transmission electron microscope equipped with an Orious Sc1000 digital camera, and exhaustively observed.

Atomic Force Microscopy. Samples were imaged with a PicoSPM LE system of Molecular Imaging in air at room temperature in dynamic mode. About 10 µl of solution was deposited in

freshly cleaved mica from Agar Scientific. The sample rested for 20 min before rinsing the mica with water which was then dried in air prior to analysis.

Molecular Dynamics Simulations

Interestingly, the crystal symmetry units suggest contacts between S100B dimers in which Cys84 residues from different subunits are near each other (Fig. S2a). However, since this tetrameric arrangement is constrained by the crystallographic packing, we have performed molecular dynamics simulations to probe for the dynamics of tetrameric assemblies. Initially we performed simulations of the S100B dimer, that showed a very stable protein complex with all structural properties equilibrating within the first 50 ns (Fig. S2b-c). The solvent accessible surface (SAS) area results showed that Cys84 remains significantly more exposed, confirming that it should be the residue more susceptible to oxidation (Fig. S2b). Furthermore, Cys68 is consistently buried in the simulations, which clearly excludes its involvement. We then carried out simulations on tetramer configurations built from two dimer structures assembled to allow for a S-S bond between two Cys84 residues (Fig. S2d). The two configurations were selected to maximize the interfacial area and differ in that the second dimer can be rotated by 180° along the S-S bond. Even though the interfacial SAS area in all tetramer simulations converged to the same value (Fig. S2e), the first assembly (Tetramer 1) establishes wider stabilizing interactions across the interface between the two S100B dimers, hence promoting a higher shape complementarity (Fig. 2a-d).

All simulations were performed using GROMACS v.2018.6 [3, 4] and the GROMOS 54a7 force field [5, 6]. Due to the presence of metal-binding sites (Ca^{2+} and Zn^{2+}) in the protein, these were parameterized using the Gaussian 09 software [7] at the B3LYP level and using the 6-31G* basis sets [8]. The bonded parameters were obtained from the structural optimizations of the metal coordination spheres. The partial charges were obtained from the electrostatic potential (ESP) calculations using the Merz-Singh-Kollman scheme [9] and manually curated to adhere to the GROMOS charge group approach.

The S100B dimer was obtained from the human S100B crystal structure available in the Protein Data Bank (PDB code 3D0Y [10]), which was stripped of waters and the initial residue of the first chain, an N-formylmethionine, was mutated to a methionine. The two final residues of the sequence, which were missing, were also added. The two tetramer initial structures were built from two dimer structures assembled to allow for a disulfide bond. There are 4 cysteine residues in each dimer, but only two of them are solvent-exposed enough to allow bridging over without a significant conformational rearrangement. Upon disulfide formation, the two dimer structures can still rotate along the bond axis and there are two configurations

that maximize the interfacial area, differing by a 180º rotation along this disulfide bond axis. (**Figure S2d**). To form the disulfide bond, we brought together two aligned dimer conformations (in each tetramer assembly) and performed a short steered MD segment (~10 ns) to smoothly bring together the two sulfur atoms (0.1 $nm.ns^{-1}$; 1000 $kJ.mol^{-1}.nm^{-2}$) to a bond forming distance.

All systems were solvated in SPC water molecules [11] in a dodecahedral box with a minimum distance of 0.9 nm between protein and the end of the box. The final systems were composed of ~10000 water molecules for the dimer and ~20000 water molecules for the tetramers. The protonation states of all residues were chosen according to their abundance at physiological pH, while for histidines, the electrostatic environment was considered to choose the appropriate tautomeric form. In all cases, overall negative charges were obtained for the protein and Na^+ ions were added to achieve charge neutrality.

In all simulations, a v-rescale heat bath [12] at 310 K was used with separate couplings for the solute (which includes the protein content) and the solvent (includes water and ions) with a relaxation time of 0.1 ps. The Parrinello-Rahman isotropic pressure coupling [12-14] was used to keep the pressure at 1 bar, with isothermal compressibility of 4.5×10^{-5} bar^{-1}, and a relaxation time of 1.0 ps. All protein bonds were constrained using the LINCS algorithm [15], while all water molecules were constrained using the SETTLE algorithm [16]. The equations of motions were integrated using a 2 fs time step, with the lists of neighbors being updated every 20 fs. The Particle-Mesh Ewald (PME) electrostatics [17] was applied using 0.12 nm for the maximum grid spacing of the Fast Fourier Transform and a cutoff distance of 1.4 nm for Lennard-Jones and Coulomb interactions. The interpolation order for PME was 4.

All systems were energy minimized using two steps (~30000 steps each) using the steepest descent algorithm, one unconstrained followed by a second step with the p-LINCS algorithm applied on all bonds. The initialization process consisted of three steps: 100 ps NVT MD simulation with initial velocities being generated from a Maxwell velocity distribution at 310 K and position restraints (1000 $kJ.mol^{-1}.nm^{-2}$) on the C_α atoms of the protein; 200 ps of NPT MD simulation using the Parrinello-Rahman isotropic pressure couple and weaker position restraints (100 $kJ.mol^{-1}.nm^{-2}$); while in the last step, only the position restraints were weakened (10 $kJ.mol^{-1}.nm^{-2}$). For the dimer system, 3 replicates of 200 ns were produced while for the two tetramer systems, single 500 ns simulations were performed.

All analyses were performed using the GROMACS tools and/or in-house scripts. Equilibrium properties were performed on the equilibrated segments where the initial 50 and 200 ns were discarded for the dimer and tetramer simulations, respectively. The hydrophobic/hydrophilic solvent accessible surface (SAS^{hydro}) index was calculated as the sum of the product between

the solvent-accessible surface (SAS) area and the hydrophobicity (Wimley--White scale) [18] of each residue, which is a variant of the spatial aggregation propensity (SAP) method developed by Trout and coworkers [19, 20]. Negative SAS^{hydro} values are obtained for hydrophobic residues, while positive values are obtained for hydrophilic ones. The data can be averaged out over time and plotted per residue or can be averaged out over residue and plotted over time. In this case, positive (hydrophilic) and negative (hydrophobic) values are separated in two distinct time series to avoid their canceling. Images were rendered using PyMOL [21].

Kinetic modelling and fitting

Under the outlined assumptions for model building, only the following oligomer types which might occur during S100B polymerization were considered: D_n, *n*-dimer oligomers with both reactive cysteines in reduced state; D_na, *n*-dimer oligomers with one oxidized (activated) cysteine; $D_n$aa, *n*-dimer oligomers with both reactive cysteines oxidized (activated). As the model intends to describe the early stages of S100B polymerization, $1 \leq n \leq 12$. The reactions and reaction rates postulated in the model are defined in Table S1:

Table S1 - Reactions and rate equations considered in the kinetic model		
Reaction	Stoichiometry	Rate
Cysteine oxidations	$D_n + Cu^{2+} \rightarrow D_n a + Cu^{+}$	$v_{ox} = k_{ox} \cdot D_n \cdot Cu^{2+}$
Cysteine oxidations	$D_n a + Cu^{2+} \rightarrow D_n aa + Cu^{+}$	$v_{ox} = k_{ox} \cdot D_n a \cdot Cu^{2+}$
Polymerization	$D_n a + D_m a \rightarrow D_{(n+m)}$	$v_{poly2} = k_{poly2} \cdot D_n a \cdot D_m a$
Polymerization	$D_n a + D_m aa \rightarrow D_{(n+m)} a$	$v_{poly3} = 2\, k_{poly3} \cdot D_n a \cdot D_m aa$
Polymerization	$D_n aa + D_m aa \rightarrow D_{(n+m)} aa$	$v_{poly4} = 4\, k_{poly4} \cdot D_n aa \cdot D_m aa$

The parameters k_{ox}, k_{poly2}, k_{poly3} and k_{poly4} were then estimated by fitting the model to four experimental time courses of ThT-monitored S100B-dimer polymerization experiments, as described in section 'S100B polymerization Kinetics' under Materials and Methods section. These time courses were obtained from polymerization of dimeric S100B (40 μM) resulted from adding Cu^{2+} at increasing in Cu^{2+}:S100B ratios (1:1, 1:2, 1:3 and 1:4). Model predictions were fit to data of ThT signal by including in the model a variable T representing that signal with expression

$$T = K_T \sum_{i=2}^{12} (i-1) \cdot (D_i + D_i a + D_i aa)$$

where K_T is a scaling constant also subject to fitting. The fitting objective function to minimize was

$$SS = \sum_{time\ courses} \sum_{time\ points} \left(T - T_{exp}\right)^2$$

where T_{exp} is the experimental ThT signal and T is defined as above. In fitting and simulations, model ODE solutions were computed using the LSOLDA adaptive size method as implemented in function `odeint` of the scipy [22] computing platform.

The parameters corresponding to the Cys oxidation and polymerization rates were determined by fitting the model to the experimental time-courses of S100B-dimer polymerization, which provided accurate descriptions of the experimental kinetic traces. Parameter fitting was performed by minimization of SS using differential evolution [23] with simplex local refinement optimizer. The best parameters from model fitting are the following: k_{ox} = 0.0037 μM^{-1} hr^{-1}; k_{poly2} = 0.0012 μM^{-1} hr^{-1}; k_{poly3}= 0.0017 μM^{-1} hr^{-1}; $k_{poly4=}$0.000044 μM^{-1} hr^{-1}; K_T = 0.14. We then used the fitted model to simulate the evolution of the several oligomeric forms of S100B. At the highest tested ratio of Cu^{2+}:S100B=4, the fully oxidized Cys forms are the most abundant (Fig. 3b), and the species in which half of the Cys residues is oxidized peaks at 2.5h, in agreement with the rapid formation of a cross-linked S100B tetramer, as noted above. In agreement, relatively early at the simulation, nearly 25% of the species is composed by the S-S cross-linked tetramer (4-mer), which gradually evolves to longer polymers (n-mers) (Fig. 3c). Fitting also evidenced that oxidation at both Cys decreases the reactivity towards polymerization. In agreement, most of the S100B polymers correspond to species composed by up to 8 dimers.

Bioimaging

TEM and AFM analysis of apo S100B evidences a uniform distribution of small oligomers (5-10 nm and some 50 nm species, Fig. 3h). TEM analysis of the S100B species formed after 70 h incubation at 37 °C at Cu2+:S100B=4 depicts spherical species with a diameter around 5 nm, as well as what seems to be an association of two or more of these species, producing longer structures with up to 10 nm. AFM investigations evidence a relatively homogeneous distribution of oligomers with comparable dimensions. There is vast evidence showing that oligomerization of chaperone proteins constitutes an effective enhancer mechanism for improved function [24-26].

2. Supplementary reaction scheme

The native S100B dimer contains two Cys per monomer, one of which (Cys84) locates next to one of the Cu^{2+} coordinating residues (His85) and is thus a target for a temporary interaction with Cu^{2+}. This involves Cu^{2+} reduction to Cu^{+} by sulfur, followed by redox cycling with O_2:

$Cu^{2+} + S \leftrightarrow Cu^{+} + S^{+}$	[Cu^{2+} is reduced to Cu^{+} by sulfur]
$Cu^{+} + O_2 \leftrightarrow O_2^{-} + Cu^{2+}$	[Copper is reoxidised]
$Cu^{2+} + S \leftrightarrow Cu^{+} + S^{+}$	[A second reduction of Cu^{2+} to Cu^{+} by sulfur]
$Cu^{+} + O_2^{-} + 2H^{+} \leftrightarrow H_2O_2 + Cu^{2+}$	[Copper is again reoxidised]
$S^{+} + S^{+} \leftrightarrow S\text{-}S$	[disulfide is formed]

Which translates to the global reaction scheme:

$2Cu^{2+} + 2S + O_2 + 2H^{+} \leftrightarrow 2Cu^{2+} + S\text{-}S + H_2O_2$

3. Supplementary Figures

Supplementary Figure S1

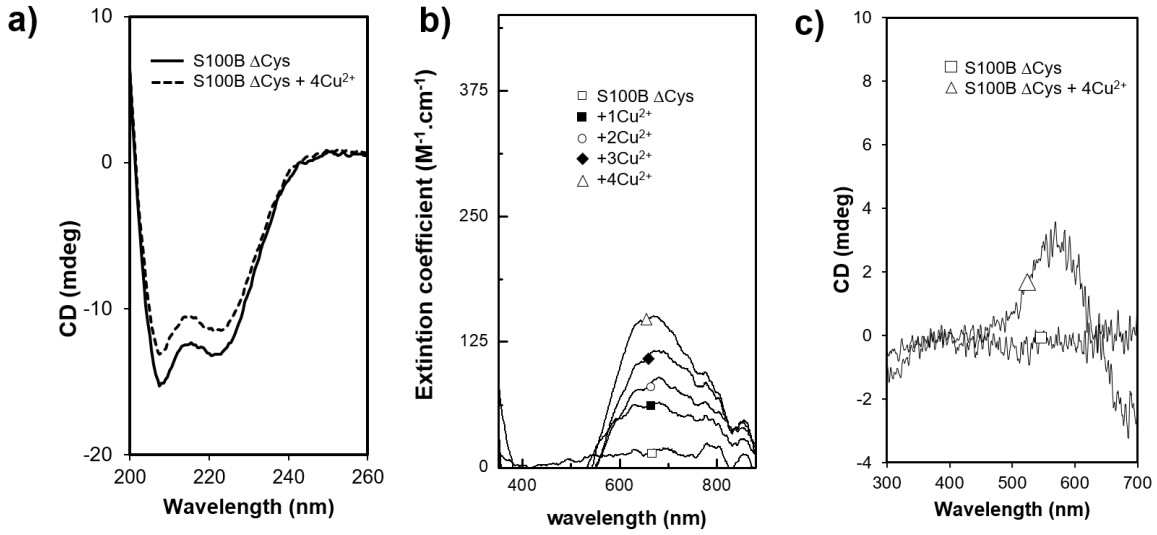

Figure S1 – Spectroscopic analysis of S100BΔCys in the presence of Cu^{2+}. a) far-UV CD spectroscopy; **b)** visible-absorption spectroscopy; **c)** visible circular dichroism. See Materials and methods for details.

Supplementary Figure S2

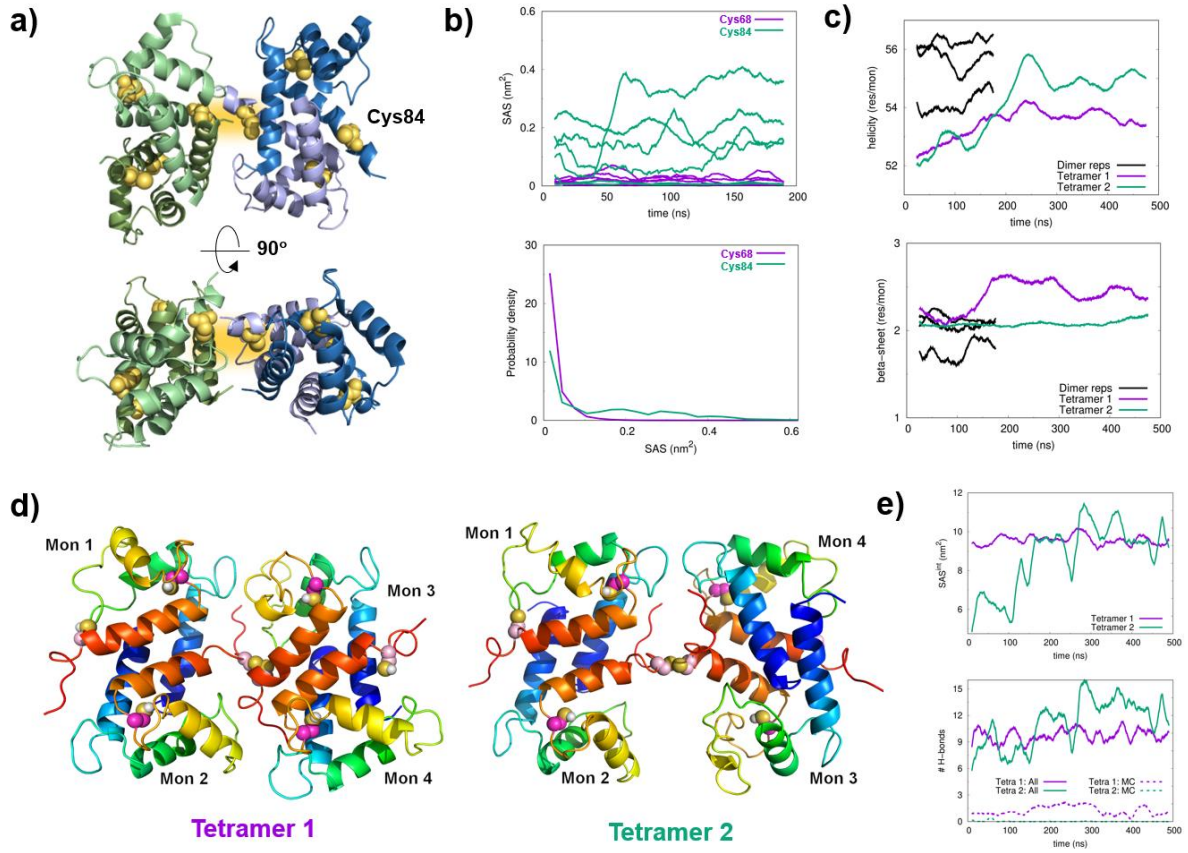

Figure S2 – Structural analysis from the MD simulations.
a) Contacts between S100B dimers symmetrically packed in a crystal structure (PDBID: 2h61). Cys84 and Cys69 are indicated as yellow spheres. The yellow shadow highlights proximity between Cys84, which might lead to disulfide formation; **b)** Solvent accessible surface (SAS) area over time (top) and histogram distributions (bottom) for the two different cysteine residues in S100B dimer simulations. The equivalent cysteine residues in each monomer and dimer simulation replicates were combined in the histograms; **c)** Number of residues (normalized per monomer) in helical (left) and β-sheet (right) conformations over time for simulations of S100B dimers (triplicates) and tetramers (the two configurations). A floating window average smoothing function with 200 points was used for clarity; **d)** The two tetramer assemblies used in this work. The secondary structure of each monomer was coloured like a rainbow (from blue to red) according to the residue number. The second dimer (monomers 3 and 4) of tetramer 1 (left) was rotated 180º along the disulfide bond to obtain tetramer 2 (right). All cysteine residues are shown as light (more solvent-exposed) or darker (more buried) pink spheres; **e)** The interfacial surface area (top) and the hydrogen bonds (bottom) between S100B dimers in the two tetramer configurations. The hydrogen bonds across dimers' main chain (MC) (cross-β structure) are also shown. A floating window average smoothing function with 200 points was used for clarity.

Supplementary Figure S3

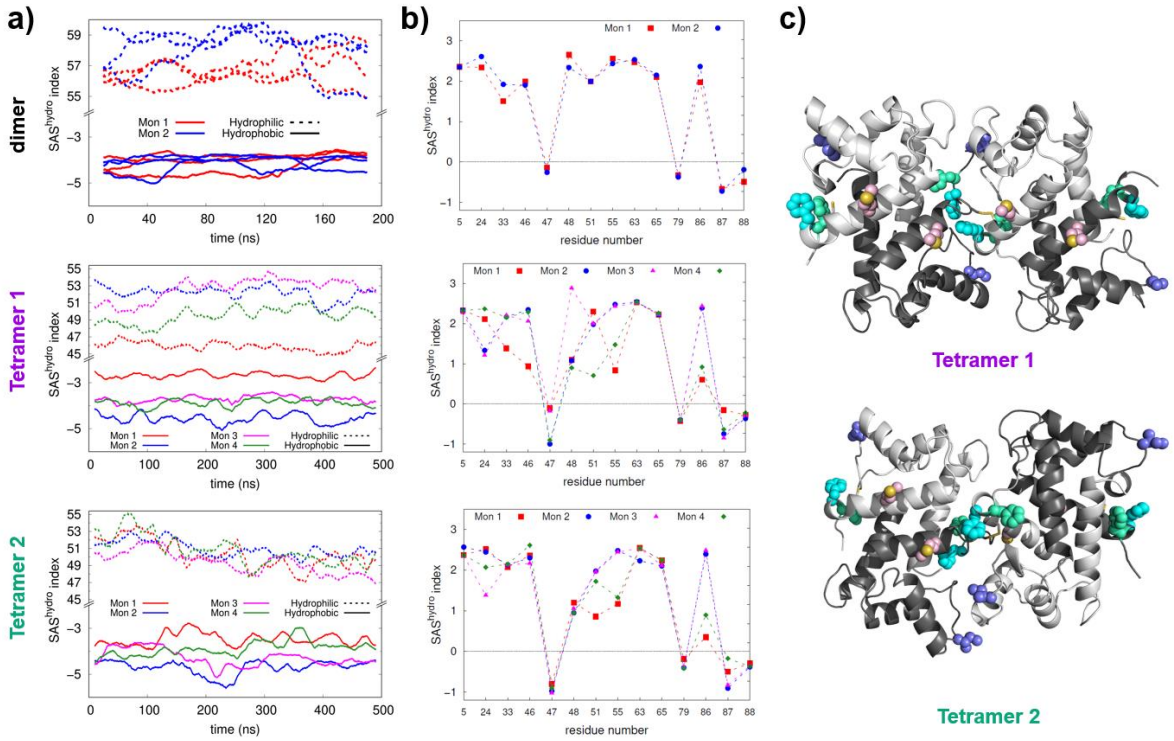

Figure S3 – Surface analysis from the SAS^{hydro} index values.
a) Hydrophobic/hydrophilic solvent accessible surface (SAS^{hydro}) index values over time for each monomer in S100B dimer (top), tetramer 1 (middle) and tetramer 2 (bottom) simulations. 3 replicates of the dimer simulations are shown. A floating window average smoothing function with 200 points was used for clarity; **b)** SAS^{hydro} index values over residue for the average of the three replicates of S100B dimer (top) and tetramer 1 (middle) and tetramer 2 (bottom) simulations. To highlight only the most relevant residues, a selection criteria was applied (SAS^{hydro} < -0.4 for hydrophobic and SAS^{hydro} > +2.2 for hydrophilic residues); **c)** The most solvent exposed hydrophobic residues are shown as spheres in the final conformations of the two tetramer configurations of S100B (tetramer 1, top; tetramer 2, bottom). Monomers 1,3 and 2,4 are shown as light and dark grey, respectively. Ile47, Met79, Phe87 and Phe88 are shown in sky blue, light pink, cyan, and pale green, respectively. The four Cys84 residues are shown as sticks to help identify the covalent bond between dimers and the dangling thiol groups available for further polymeric growth.

Supplementary Figure S4

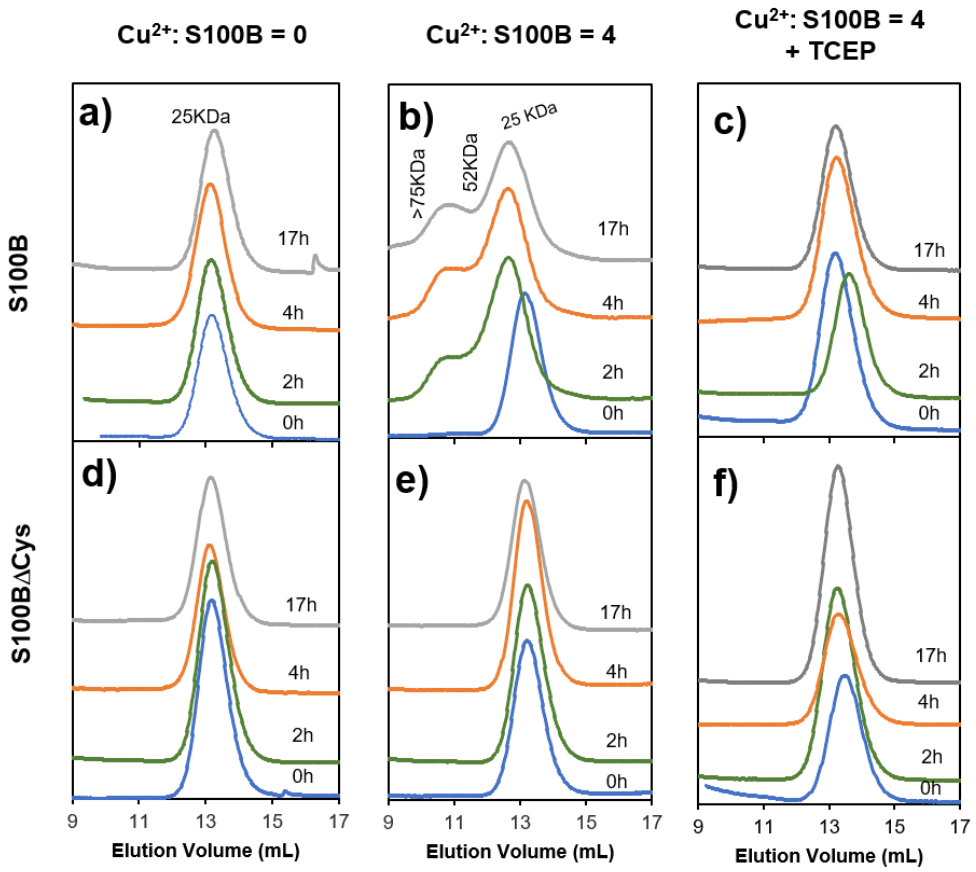

Figure S4 – SEC analysis of S100B polymerization.

The time course elution profiles of S100B species were determined for S100B wild type (**a-c**) and S100BΔCys (**d-f**) upon incubation with Cu^{2+} at a Cu^{2+}:S100B=4 with (**c & f**) and without (**b & e**) added TCEP (10 mM), in comparison to the control S100B with no added Cu^{2+} (**a & d**). Samples were prepared as described in materials and methods.

Supplementary References

1. H. M. Botelho, G. Fritz and C. M. Gomes, *Methods Mol Biol*, 2012, **849**, 373-386.
2. D. M. Walsh, E. Thulin, A. M. Minogue, N. Gustavsson, E. Pang, D. B. Teplow and S. Linse, *Febs J*, 2009, **276**, 1266-1281.
3. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *J Comput Chem*, 2005, **26**, 1701-1718.
4. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J Chem Theory Comput*, 2008, **4**, 435-447.
5. N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark and W. F. van Gunsteren, *Eur Biophys J*, 2011, **40**, 843-856.
6. W. Huang, Z. Lin and W. F. van Gunsteren, *J Chem Theory Comput*, 2011, **7**, 1237-1243.
7. G. W. T. M. J. Frisch, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Journal*, 2009.
8. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *The Journal of Physical Chemistry*, 1994, **98**, 11623-11627.
9. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner, *Journal of the American Chemical Society*, 1984, **106**, 765-784.
10. T. Ostendorp, J. Diez, C. W. Heizmann and G. Fritz, *Biochim Biophys Acta*, 2011, **1813**, 1083-1091.
11. J. Hermans, H. J. C. Berendsen, W. F. van Gunsteren and J. P. M. Postma, *Biopolymers*, 1984, **23**, 1513-1518.
12. G. Bussi, D. Donadio and M. Parrinello, *J Chem Phys*, 2007, **126**, 014101.
13. S. Nosé and M. L. Klein, *Molecular Physics*, 1983, **50**, 1055-1076.
14. M. Parrinello and A. Rahman, *Journal of Applied Physics*, 1981, **52**, 7182-7190.
15. B. Hess, *J Chem Theory Comput*, 2008, **4**, 116-122.
16. S. Miyamoto and P. A. Kollman, *Journal of Computational Chemistry*, 1992, **13**, 952-962.
17. T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1993, **98**, 10089-10092.
18. W. C. Wimley and S. H. White, *Nat Struct Biol*, 1996, **3**, 842-848.
19. N. Chennamsetty, V. Voynov, V. Kayser, B. Helk and B. L. Trout, *Proc Natl Acad Sci U S A*, 2009, **106**, 11937-11942.
20. N. Chennamsetty, V. Voynov, V. Kayser, B. Helk and B. L. Trout, *J Phys Chem B*, 2010, **114**, 6614-6624.
21. L. Schrödinger, *Journal*, 2020.
22. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C.

Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G.-L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, Y. Vázquez-Baeza and C. SciPy, *Nature Methods*, 2020, **17**, 261-272.

23. R. Storn and K. Price, *Journal of Global Optimization*, 1997, **11**, 341-359.
24. M. Haslbeck and E. Vierling, *Journal of molecular biology*, 2015, **427**, 1537-1548.
25. G. Chen, A. Abelein, H. E. Nilsson, A. Leppert, Y. Andrade-Talavera, S. Tambaro, L. Hemmingsson, F. Roshan, M. Landreh, H. Biverstal, P. J. B. Koeck, J. Presto, H. Hebert, A. Fisahn and J. Johansson, *Nature communications*, 2017, **8**, 2081.
26. A. Mainz, J. Peschek, M. Stavropoulou, K. C. Back, B. Bardiaux, S. Asami, E. Prade, C. Peters, S. Weinkauf, J. Buchner and B. Reif, *Nature structural & molecular biology*, 2015, **22**, 898-905.