Revealing Well-Defined Soluble States during Amyloid Fibril Formation by Multilinear Analysis of NMR Diffusion Data

Kristine Steen Jensen,*† Sara Linse,*§ Mathias Nilsson,*§ Mikael Akke,*§ and Anders Malmendal†§

†Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, P.O. Box 124, SE-22100 Lund, Sweden
‡Biochemistry and Structural Biology, Center for Molecular Protein Science, Department of Chemistry, Lund University, P.O. Box 124, SE-22100 Lund, Sweden
§Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, U.K.
†Department of Science and Environment, Roskilde University, P.O. Box 260, DK-4000 Roskilde, Denmark

Supporting Information

ABSTRACT: Amyloid fibril formation is a hallmark of neurodegenerative disease caused by protein aggregation. Oligomeric protein states that arise during the process of fibril formation often coexist with mature fibrils and are known to cause cell death in disease model systems. Progress in this field depends critically on development of analytical methods that can provide information about the mechanisms and species involved in oligomerization and fibril formation. Here, we demonstrate how the powerful combination of diffusion NMR and multilinear data analysis can efficiently disentangle the number of involved species, their kinetic rates of formation or disappearance, spectral contributions, and diffusion coefficients, even without prior knowledge of the time evolution of the process or chemical shift assignments of the various species. Using this method we identify oligomeric species that form transiently during aggregation of human superoxide dismutase 1 (SOD1), which is known to form misfolded aggregates in patients with amyotrophic lateral sclerosis. Specifically, over a time course of 42 days, during which SOD1 fibrils form, we detect the disappearance of the native monomeric species, formation of a partially unfolded intermediate in the dimer to tetramer size range, subsequent formation of a distinct similarly sized species that dominates the final spectrum detected by solution NMR, and concomitant appearance of small peptide fragments.

Nuclear magnetic resonance (NMR) spectroscopy can provide detailed chemical information from individual signals for virtually every atom in a molecule and, hence, makes it possible to follow the evolution of molecular species. For simple mixtures, the use of DOSY (diffusion-ordered spectroscopy) can resolve the individual NMR signals by distinguishing different species through their different rates of translational diffusion. By contrast, complex samples containing interconverting species pose a formidable challenge to the interpretation of NMR data by conventional means. However, the additional complication posed by an evolving mixture can be an advantage, because the time and diffusion dependences can be encoded as independent dimensions enabling the use of multilinear data analysis methods, such as parallel factor analysis, PARAFAC.

PARAFAC is an extension of the two-dimensional principal component analysis (PCA) to higher order arrays. While PCA suffers from rotational ambiguity, resulting in fitted components that often are linear combinations of different processes and describe deviations from the mean rather than actual amplitudes, a successful PARAFAC decomposition directly provides the user with the variation of each species in a form representing meaningful amplitudes, resulting in a tremendous improvement in interpretability.

Here, we show that the combination of DOSY and PARAFAC provides powerful information on an elusive oligomerization process underlying formation of amyloid fibrils. We employed three-dimensional $^1$H–$^{15}$N-DOSY-HSQC to acquire a series of spectra during fibril formation by SOD1 (153 amino acid residues), a protein associated with the neurodegenerative disease amyotrophic lateral sclerosis. The present investigation involves three independent dimensions: the NMR resonance frequencies, the signal decays caused by diffusion, and time. PARAFAC translates these dimensions into the sought information: the number of (NMR detectable) species and their time evolution profiles, sizes, and spectra. This is achieved without any constraints or

Received: July 31, 2019
Published: November 8, 2019
presumptions of kinetic models, hydrodynamic properties, or spectral characteristics of the species.

We used the pseudo wild type SOD1 variant C6A/F50E/G51E/C57S/C111A/C146S (pwtSOD1^{-AC}) that mimics the monomeric metal-free and disulphide-reduced form of SOD1\(^{10,12,13}\) which is known to be a key precursor in the formation of aberrant oligomers and fibrils.

We performed \(^1\)H–\(^15\)N-DOSY-HSQC experiments based on the X-STE pulse sequence\(^3\) with a diffusion time of 500 ms and 10 different gradient strengths, yielding a total experimental time of 18.6 h. This experiment was repeated 10 times on the same sample over a period of 6 weeks, during which time the sample temperature was maintained at 37 °C, to cover the full fibril formation of pwtSOD1^{-AC} (Supporting Information (SI) methods). Over this time period the \(^1\)H–\(^15\)N-HSQC spectrum of pwtSOD1^{-AC} becomes crowded as peaks from partially unfolded species build up (Figure S1). For the present analysis we selected a total of 223 peaks, which may well have contributions from overlapping resonances.

The data were analyzed using PARAFAC models to identify the minimum number of species (factors) needed to describe the variance of the data set, see SI methods for details. We computed PARAFAC models including up to 5 factors. Anomalous behavior was identified for three weak signals from histidines affected by exchange processes at this pH, and these were removed from further analysis. The analysis identified four factors based on various quality assessment tools, such as residuals and core consistency diagnostics\(^{19}\) (Table S1). Factor 4 comprises only three signals that we excluded to generate a separate set of PARAFAC models with improved descriptors for factors 1–3.

The resulting factors are characterized by unique combinations of time evolution, size, and spectral contribution (Figure 1) that make physical and chemical sense: the amplitudes of individual factors change smoothly with time (Figure 1A–D) and decrease exponentially with the square of the gradient strength, indicating that each factor represents a unique, monodisperse species (Figure 1E–H). Due to extensive overlap, individual peaks may represent more than one molecular species, but this is resolved by the PARAFAC decomposition. Most peaks include contributions from more than one factor (Figure 1I–L), indicating that species I, II, and III contain different levels of similarly folded and unfolded segments.

At the starting point of the reaction the dominant species (state I, factor 1) is the folded, monomeric apo form of pwtSOD1^{-AC}, as identified by the chemical shifts.\(^3\) The diffusion coefficient, \(D = (1.40 \pm 0.02) \times 10^{-10} \text{ m}^2\text{s}^{-1}\), for this species corresponds to a Stokes’ radius, \(r_s = 23 \text{ Å} \text{(eq S2)}\), which corresponds very well with the previously determined value of \(r_s = 22.5 \text{ Å}\) for monomeric apo disulphide-reduced SOD1.\(^{10}\)

Factor 2 describes an intermediate (state II) that builds up rapidly from the start, displays the same short lag phase as the monomer curve, peaks around day 8–15, and declines over days 15–42 (Figure 1B). The time evolution of factor 2 is similar to that of prefibrillar intermediates of Aβ42 and transthyretin, associated with Alzheimer’s disease and transthyretin amyloidosis, respectively.\(^{21,22}\) State II has a lower diffusion coefficient than state I, \(D = (0.86 \pm 0.01) \times 10^{-10} \text{ m}^2\text{s}^{-1}\), corresponding to \(r_s = 38 \text{ Å}\). If states I and II behave like globular species (despite the presence of disordered segments), the ratio of their \(r_s\) values suggests that state II is a tetramer.

For comparison, the native Cu,Zn-SOD1 dimer has \(r_s = 30.3 \text{ Å}\), further supporting the conclusion that state II is tri- or tetrameric (SI). A trimeric SOD1 species with partially unfolded regions has previously been implicated.\(^{23}\)

Factor 3 describes the final visible state III, which has essentially the same diffusion coefficient as state II, \(D = (0.89 \pm 0.01) \times 10^{-10} \text{ m}^2\text{s}^{-1}\), corresponding to \(r_s = 37 \text{ Å}\). The similar diffusion coefficients and the large number of peaks with similar contributions from factors 2 and 3 (green in Figure 1J,K) imply that states II and III have similar structures.

Factor 4 describes state IV, which has a much higher diffusion coefficient than the other three states, \(D = (5.90 \pm 0.2) \times 10^{-10} \text{ m}^2\text{s}^{-1}\), yielding \(r_s \approx 5 \text{ Å}\). State IV involves only three signals and appears to build up simultaneously with state III (Figure 1C,D), suggesting that the two events are linked; most likely, state IV represents a short peptide that is cleaved off from the protein.

Variations in intensity and diffusion rate for the different types of signals are shown in Figure 2. Note that the data in Figure 2A describe individual peaks, many of which contain contributions from more than one PARAFAC factor. The evolution profiles and diffusion coefficients of the three main states suggest the following linear reaction model: state I (monomer) \(\rightarrow\) state II (partly unfolded oligomer) \(\rightarrow\) state III (partly unfolded and cleaved oligomer) \(\rightarrow\) state IV (peptide). The observed buildup of state III to a plateau value (Figure
The results by small-angle X-ray scattering, which unequivocally shows that pwtSOD$^{1\Delta C}$ is monomeric at day 0, but polydisperse with larger oligomeric species by day 19 (SI Figure S2). Both samples are dominated by globular species with disordered regions, but the disordered contribution is only slightly increased at the later time point (SI Figure S2B), implying that states II and III are far from fully unfolded. The buildup of oligomers was further confirmed by chemical cross-linking of the oligomeric state II, implying that states II and III are far from fully unfolded. The buildup of oligomers was further confirmed by chemical cross-linking of cross-linked, intact pwtSOD$^{1\Delta C}$ samples obtained at day 0 and day 19 (SI Figure S3). We also used cryo-transmission electron microscopy to verify that fibrils formed during the course of the 6-week experiment (SI Figure S4). It should be noted that our results do not exclude the possibility that the formation of the oligomeric state II takes place via the unfolded monomer, even though this state is weakly populated and the equilibrium unfolding dynamics is not rate limiting. Future experiments will reveal whether the identified states are on-path precursors to fibril formation.

We have shown that $^{1}\text{H} - ^{15}\text{N}$-DOSY-HSQC spectroscopy in combination with PARAFAC analysis allows straightforward identification of four NMR-detectable species of pwtSOD$^{1\Delta C}$ present during fibril formation. PARAFAC yields the time evolution profiles and diffusion coefficients of these species, and identifies their chemical shifts. Thus, the approach enables powerful characterization of molecular processes that occur alongside fibril formation, or other macromolecular transitions occurring on similar time scales.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b07952.

Experimental procedures and supporting figures with additional data (PDF)

### AUTHOR INFORMATION

#### Corresponding Authors

*E-mail: kristine.steen_jensen@bpc.lu.se (K.S.J.).
*E-mail: mikael akka@bpc.lu.se (M.A.).
*E-mail: amalm@ruc.dk (A.M.).

#### ORCID

Sara Linse: 0000-0001-9629-7109
Mathias Nilsson: 0000-0003-3301-7899
Mikael Akke: 0000-0002-2395-825X
Anders Malmendal: 0000-0002-8413-9717

#### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

We thank Ulrich Weininger for help with HSQC-based DOSY experiments; Annette Langkilde, Saskia Bucciarelli, Bente Vestergaard, and the P12 EMBL beamline staff (DESY, Hamburg) for assistance with SAXS experiments; Katja Bernfur for assistance with mass-spectrometry; and D. Flemming Hansen for inspiring discussions. This work was supported by the Swedish Research Council (2014-5815 to M.A., 2015-00143 to S.L.), Danish Council for Independent Research, Sapere Aude: DFF Research Talent (DFF – 4002-00258 to K.S.J.), the European Research Council (340890 to M.A., 2015-00143 to S.L.), Danish Council for Independent Research, Sapere Aude: DFF Research Talent (DFF – 4002-00258 to K.S.J.), the European Research Council (340890 to M.A., 2015-00143 to S.L.).

#### REFERENCES

1. H. N. R. Thiery, A. Malmendal, S. Linse, J. Am. Chem. Soc. 2019, 141, 18649–18652.
REFERENCES

(1) Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. Naturally Secreted Oligomers of Amyloid β Protein Potently Inhibit Hippocampal Long-Term Potentiation in Vivo. Nature 2002, 416, 535–539.

(2) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. Inherent Toxicity of Aggregates Implies a Common Mechanism for Protein Misfolding Diseases. Nature 2002, 416, 507–511.

(3) Teilum, K.; Smith, M. H.; Schulz, E.; Christensen, L. C.; Solomentsev, G.; Oliveberg, M.; Akke, M. Transient Structural Distortion of Metal-Free Cu/Zn Superoxide Dismutase Triggers Aberrant Oligomerization. Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 18273–18278.

(4) Sekhar, A.; Rumfaldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Bouvignies, G.; Meiering, E. M.; Kay, L. E. Thermal Fluctuations of Immature SOD1 Lead to Separate Folding and Misfolding Pathways. eLife 2015, 4, No. e07296.

(5) Khajeh, M.; Botana, A.; Bernstein, M. A.; Nilsson, M.; Morris, G. A. Reaction Kinetics Studied Using Diffusion-Ordered Spectroscopy and Multiway Chemometrics. Anal. Chem. 2010, 82 (5), 2102–2108.

(6) Nilsson, M.; Khajeh, M.; Botana, A.; Bernstein, M. A.; Morris, G. A. Diffusion NMR and Trilinear Analysis in the Study of Reaction Kinetics. Chemom. Commun. 2009, No. 10, 1252–1254.

(7) Bro, R. PARAFAC. Tutorial and Applications. Chemom. Intell. Lab. Syst. 1997, 38, 149–171.

(8) Dal Poggetto, G.; Castaillar, L.; Adams, R. W.; Morris, G. A.; Nilsson, M. Relaxation-Encoded NMR Experiments for Mixture Analysis: REST and Beer. Chem. Commun. 2017, 53, 7461–7464.

(9) Colbourne, A. A.; Meier, S.; Morris, G. A.; Nilsson, M. Unmixing the NMR Spectra of Similar Species – Vive La Difference. Chem. Commun. 2013, 49, 10510–10512.

(10) Colbourne, A. A.; Morris, G. A.; Nilsson, M. Local Covariance Order Diffusion-Ordered Spectroscopy: A Powerful Tool for Mixture Analysis. J. Am. Chem. Soc. 2011, 133, 7640–7643.

(11) Björnersjö, J.; Botana, A.; Morris, G. A.; Nilsson, M. Resolving Complex Mixtures: Trilinear Diffusion Data. J. Biol. Chem. 2014, 251, 251–257.

(12) Bertini, I.; Piccioli, M.; Viezzi, M. S.; Chiu, C. Y.; Mullenbach, G. T. A Spectroscopic Characterization of a Monomeric Analog of Copper, Zinc Superoxide-Dismutase. Eur. Biophys. J. 1994, 23, 167–176.

(13) Khan, M. A. I.; Respondek, M.; Kjellström, S.; Deep, S.; Linse, S.; Akke, M. Cu/Zn Superoxide Dismutase Forms Amyloid Fibris Under-Physiological Quiescent Conditions: The Roles of Disulfide Bonds and Effects of Denaturant. ACS Chem. Neurosci. 2017, 8, 2019–2026.

(14) Lindberg, M. J.; Byström, R.; Boknås, N.; Andersen, P. M.; Oliveberg, M. Systematically Perturbed Folding Patterns of Amyotrophic Lateral Sclerosis (ALS)-Associated SOD1 Mutations. Proc. Natl. Acad. Sci. U. S. A. 2005, 102, 9754–9759.

(15) Furukawa, Y.; Kaneko, K.; Yamanaka, K.; O’Halloran, T. V.; Nukina, N. Complete Loss of Post-Translational Modifications Triggers Fibrillar Aggregation of SOD1 in the Familial Form of Amyotrophic Lateral Sclerosis. J. Biol. Chem. 2008, 283, 24167–24176.

(16) Rakhit, R.; Crow, J. P.; Lepock, J. R.; Kondejewski, L. H.; Cashman, N. R.; Chakrabarty, A. Monomeric Cu/Zn-Superoxide Dismutase Is a Common Misfolding Intermediate in the Oxidation Models of Sporadic and Familial Amyotrophic Lateral Sclerosis. J. Biol. Chem. 2004, 279, 15499–15504.

(17) Khare, S. D.; Caplow, M.; Dokholyan, N. V. The Rate and Equilibrium Constants for a Multistep Reaction Sequence for the Aggregation of Superoxide Dismutase in Amyotrophic Lateral Sclerosis. Proc. Natl. Acad. Sci. U. S. A. 2004, 101, 15094–15099.