

Three long-range distance constraints and an approach towards a model for the αSynuclein-fibril fold

Maryam Hashemi Shabestari^[a], Pravin Kumar^[a], Ine M.J. Segers-Nolten^[b], Mireille M.A.E. Claessens^[b], Bart D. van Rooijen^[b, #], Vinod Subramaniam*^[b, c], and Martina Huber*^[a]

^[a] Department of Physics, Huygens-Kamerlingh-Onnes Laboratory, Leiden University, PO Box 9504, 2300 RA Leiden, The Netherlands

^[b] Nanobiophysics, MESA+ Institute for Nanotechnology & ^[c] MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, PO Box 217, 7500AE Enschede, The Netherlands

#present address: Department of Radiology, Maastricht University Medical Center, Maastricht, the Netherlands.

Dedicated to Giovanni Giacometti on the occasion of his 85th birthday.

Abstract

Amyloid fibrils and plaques are hallmarks of neurodegenerative diseases. In Parkinson's disease, plaques (Lewy bodies) consist predominantly of the α-synuclein (αS) protein. To understand aggregation, the molecular architecture of αS fibrils needs to be known. Here we determine nm-distance constraints for the protein in the fibril by double electron-electron paramagnetic resonance (DEER) on doubly spin labeled αS-variants, diamagnetically diluted with wild-type αS to suppress intermolecular interactions. Intramolecular distances in three pairs (56/69, 56/90 and 69/90) are reported. An approach to derive a model for the fibril-fold from sparse distance data assuming only parallel β-sheets is described. Using the distances obtained in this study as input, a model is obtained with three strands, comprising residues 56 to 90, in which the strands consist of eight to twelve residues each. Limitations of the approach are discussed in detail, showing that the interpretation of the data does not yet yield an unambiguous structure model. Possible avenues to improve this situation are described.

Keywords: amyloid fibrils; alpha-synuclein; amyloid fibril fold; double electron-electron resonance (DEER or PELDOR)

1. Introduction

Amyloid fibrils are the hallmark of many neurodegenerative diseases. Detailed structural information about the fold of the protein monomers comprising the fibril is required to understand the architecture of the fibril, to identify residues that are crucial for fibrillization, and to understand the effect of mutations related to disease. Whereas atomic-resolution information is available for several short fibril-forming peptides,[1-5] the fold of longer peptides/proteins is more complicated and much less understood[6], as reviewed by Margittai and Langen [7].

We focus on full-length human α -synuclein (α S), a protein containing 140 amino acid residues. This protein, which is implicated in Parkinson's disease (PD), is the major component of the Lewy Bodies characteristic of PD. Results from solid state NMR (ssNMR) and EPR studies have identified residues involved in the formation of β -sheets in the fibrils [8;9] and proposals for the fold of stretches of residues in close contact were made[10-19]. However, an unambiguous model could not be derived, largely because of the lack of long-range (> 2 nm) distance constraints. Here we discuss two-frequency, pulsed EPR (DEER or PELDOR) to obtain such constraints in α S. While the DEER method is becoming increasingly used for biological structure determination, reports on fibril structures are scarce. The main challenge is the stacking of proteins along the fibrils axis (Fig. 1a), which gives rise to intermolecular distances and interfering DEER background contributions. Recently, the fold of the 38-aminoacid residue hIAPP protein has been determined with DEER constraints as input [20]. As we have shown in an initial report, also for α S intramolecular distances within the fibrillized protein can be obtained [21].

Subsequently, two publications appeared, in which DEER experiments on α S fibrils are described [22;23].

Here DEER results from three double mutants, α S56/69, α S56/90 and α S69/90, in which the denoted residues were mutated to cysteines and labeled with the MTSL ([(1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate]) spin label. These proteins were co-fibrillized with unlabeled wt- α S protein (diamagnetic dilution) [8;9] and we show that by stepwise increasing the dilution we can identify intramolecular distances, by observing the diminishing contribution of intermolecular distances. In addition, a first-order model for the fibril fold based on a simplified schematic of the basic fibril architecture is described. It enables to combine sparse distance constraints to generate a visualization of a possible fibril fold. We discuss the challenges involved in deriving α S-fibril structure, and show that the distance constraints derived from our experiments are compatible with a five-strand model that positions residue 41 on the first strand (strand I in Fig. 1c) and residue 90 on the last strand (strand V).

2. Materials and Methods

Expression and purification of cysteine variants of the alpha-synuclein protein (αS)

Single and double cysteine mutations were introduced into the αS gene by site-directed mutagenesis. Mutants were expressed in *Escherichia coli* strain BL21(DE3) and subsequently purified in the presence of 1 mM DTT [24;25]. Prior to labeling, αS mutant proteins were reduced with a six-fold molar excess (per cysteine) of DTT for 30 minutes at room temperature. Subsequently, samples were desalted on Pierce Zeba 5mL desalting columns, followed by an immediate addition of a six-fold molar excess (per cysteine) of the MTSL spin label [(1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl))-methanethiosulfonate] and incubated for one hour in the dark at room temperature. Free label was removed using two additional desalting steps. Protein samples were applied onto Microcon YM-100 spin columns to remove any precipitated and/or oligomerized proteins and diluted into 10 mM Tris-HCl, pH 7.4 to typical protein concentrations of approximately 0.25 mM [26].

Preparation and harvesting of fibrillar αS

Fibrils of αS were formed by incubating monomer solutions at a total protein concentration of 100 μM. Initially, fibrils of all mutants were prepared using a diamagnetic dilution of 10 μM spin-labeled α-synuclein in the presence of 90 μM wild-type α-synuclein (1 in 10). For selected mutants, fibrillization of the double cysteine mutants was carried out in a 1 in 20 diamagnetic dilution, using 5 μM spin-labeled α-synuclein α-synuclein together with 95 μM wild-type protein. The corresponding single mutants were fibrillized in a 1 in 10 ratio, to keep the spin-label concentrations constant. For the diamagnetic dilution series with αS56/69, fibrils were prepared with 20 μM αS56/69 and 80 μM wild-type (1 in 5), 10 μM αS56/69 and 90 μM wild-type (1 in 10) and 5 μM αS56/69 and 95 μM wild-type (1 in 20).

All aggregations were performed in 10 mM Tris-HCl, 50 mM NaCl, pH 7.4 buffer at 37 °C in 2 ml LoBind Eppendorf tubes with constant shaking at 1000 rpm in a Thermo mixer (Eppendorf). The time evolution of α-synuclein aggregation was monitored by the standard Thioflavin T (ThioT) fluorescence assay. The fibrillization was generally completed in 3-4 days. The α-synuclein fibrils were harvested by centrifuging for 90 minutes at 18,000 x g in an Eppendorf microcentrifuge. The supernatant was carefully removed, leaving a fibril pellet typically with a volume of 80-100 μl. The fibril pellet was used in the CW EPR and DEER measurements.

Atomic Force Microscopy (AFM)

The formation of alpha-synuclein fibrils was confirmed by tapping mode AFM in air. Aggregation aliquots were diluted in 10 mM Tris, 50 mM NaCl, pH 7.4, adsorbed onto mica, washed twice with 100 μl MilliQ water and gently dried under nitrogen gas. Tapping mode AFM height images were made on a

custom built instrument [27;28]. SPIP software (Image Metrology A/S, Lyngby, Denmark) was used for visualization.

CW EPR at 80K and at room temperature

The X-band cw EPR measurements were performed using an Elexsys E680 spectrometer (Bruker, Rheinstetten, Germany) with a rectangular cavity, using a modulation frequency of 100 kHz. For measurements at 80 K a helium gas-flow cryostat (Oxford Instruments, United Kingdom) with an ITC502 temperature controller (Oxford Instruments, United Kingdom) was used. For the measurements in frozen solution, 3 mm outer diameter quartz sample tubes were used. To obtain a frozen glass 20% glycerol was added to the samples before freezing them in liquid nitrogen. The frozen samples were inserted in the pre-cooled helium gas-flow cryostat. The EPR spectra were recorded using a modulation amplitude of 2 G and a microwave attenuation of 0.159 mW. Typical accumulation times were 40 min. For room-temperature measurements, 10-15 μ l samples of α S fibrils were drawn into Blaubrand 50 μ l capillaries. The accumulation time for the spectra was 40 min. per spectrum. Measurements were done at 20°C, using 6.331 mW of microwave power and a modulation amplitude of 1.4 G.

DEER measurements

The DEER measurements were performed at X-band (9.5 GHz) on an Elexsys E680 spectrometer (Bruker, Rheinstetten, Germany) using a 3 mm split-ring resonator (ER 4118XMS-3-W1). The temperature was kept at 40 K with a helium gas stream using a CF935 (Oxford Instruments, United Kingdom) cryostat with an ITC502 temperature controller (Oxford Instruments, United Kingdom). Samples were the same ones as used for frozen solution cw EPR and were handled in the same way to insert them into the precooled cryostat (see above) The pump and observer frequencies were separated by 70 MHz and adjusted as reported before[29]. The pump-pulse power was adjusted to invert the echo maximally [30-33]. The lengths of the pulses at the observer frequencies were 16 and 32 ns for $\pi/2$ - and π -pulses, respectively. The pump pulse length was 12 ns. All the DEER measurements were performed as two-dimensional experiments, in order to suppress the proton modulation. To do so, DEER time traces were measured for ten different τ_1 -values spaced by 8 ns starting at $\tau_1 = 200$ ns. The typical accumulation times per sample were 16 hours. Distance distributions were extracted using Tikhonov regularization with DeerAnalysis2009 [30-33].

Structure parameters of fibrils

To relate the distances measured to a proposed structure, some generally accepted parameters of the parallel β -sheet architecture of amyloid fibrils were used (see also Fig. 1 b): the intrasheet distance ($d =$

1.09 nm) and the distance between two successive amino-acids in the β -strand ($l = 0.35$ nm) [7]. From the distance l and the number of intervening residues between a pair of spin-labeled residues, the distance expected for same-strand location of the residues was determined.

3. Results

Fibrils of three double mutants, α S56/69, α S56/90 and α S69/90, and their corresponding singly labeled counterparts were prepared and investigated. In the following we describe first the experiments to characterize the fibrils and then the outcome of the EPR investigation.

Reproducibility, morphology of fibrils

Thioflavin T fluorescence assays performed on all samples showed no significant difference in fibril-formation rate, irrespective of the spin-label position and the degree of diamagnetic dilution. The morphology of the fibrils, as seen by AFM was similar for wild-type and spin-labeled mutants. Images are shown in Fig. 2. No evidence for spin-spin interaction, a telltale sign in EPR of clustering of spin-labeled proteins [8;9] was found. Systematic changes in the DEER results between samples of different diamagnetic dilutions (see below) provide evidence that the spin-labeled protein is incorporated well into the fibrils. The variability of the outcome was checked by multiple fibrillization experiments, which resulted in similar fibril behavior and identical distances, with only minor variations in the widths of the distance peaks.

Characterization of fibrils by continuous wave EPR at room temperature and 80 K

In Fig. 3, frozen and liquid solution EPR spectra are shown. The absence of line broadening in the frozen solution spectra (Fig. 3, d) – f)) is evidence for the absence of short distances (< 1.5 nm). Room temperature EPR on the harvested fibrils (Fig. 3 a) – c)) shows the immobilization of the spin labels in the fibrils. Neither the doubly-labeled nor the respective mono-labeled α S samples show signs of spin-spin interactions as observed for fibrils of pure spin-labeled α S or of samples containing a higher fraction of spin-labeled protein than used in the present study [8;9]. A fraction of narrow-line signal, accounting in all cases for less than 10 % of the spins, is attributed to monomeric α S in solution, a contribution that is below the sensitivity threshold of the DEER experiments as performed here.

DEER time traces and data analysis

In Fig. 4, the experimental, raw DEER time traces of fibrils grown from the doubly spin-labeled protein mutants α S56/69, α S56/90, and α S69/90 and their respective singly labeled reference proteins are shown.

Fibrils derived from the singly labeled protein have DEER traces that can be fit to a background function with dimensionalities between one and two, showing that the traces can be described by randomly distributed distances of spins that are preferentially aligned along one dimension, rather than along three dimensions (see below).

The traces of fibrils of the doubly spin-labeled proteins clearly differ from their singly labeled counterparts, fibrillized under the same conditions with diamagnetic dilutions to provide the same density of spins (Fig. 4 a – c). In the data of α S56/69 (Fig. 4 a), a fast initial decay in the double-mutant trace that is absent in the reference traces, shows the presence of a short distance component. The trace of α S69/90 (Fig. 4 c), deviates predominantly at later times, i.e. $> 0.5 \mu\text{s}$, from the singly labeled references, suggesting a longer intermolecular distance than in α S56/69. Consequently, even at the level of raw data, differences in the distances between the different pairs are already evident.

In Fig. 5, the steps of data analysis are illustrated. In Fig. 5 a, the data and the baseline fit are shown. All traces were corrected with a 1:1 ratio of the background derived from the DEER curves of the respective singly-labeled α S variants (for details see Fig. caption). Fig. 5 b shows the background corrected data and the corresponding fits to the data, derived from Tikhonov regularization with DeerAnalysis2009 [30-33]. Distance distributions are shown in Fig. 5 c.

Diamagnetic-dilution experiments to identify intermolecular distances

To discriminate between intra- and intermolecular distances, different diamagnetic-dilution series were measured. Intermolecular distances are expected to be prominent in fibrils, because the interprotein separation within the fibrils, i.e. the distance between proteins in the direction parallel to the fibril axis (Fig. 1 a), is only 0.48 nm [7]. For fibrils with in-register, parallel β -sheets,[7;9;9;11;12;19] this is the distance expected for residues of two adjacent proteins in the fibril. Therefore, fibrils made up of purely spin-labeled protein (undiluted samples) would have a dominant interaction of spin labels on adjacent strands. Due to the short distance between strands in the direction along the fibril axis, mostly short range and intermolecular interactions would be observed. With increasing content of unlabeled, wild-type protein, this contribution should decrease. For the α S56/69 mutant, a set of diamagnetic dilutions was measured and analyzed. Diamagnetic dilution was changed from 1 in 5 to 1 in 20 in three steps (Fig. 6). Distance distributions derived from this dataset show that the intensity of distance peaks changes systematically as the ratio of wt- α S to spin-labeled α S is varied. The distance contribution around 2 nm, and the peaks at 2.3, 2.8 and 3.9 nm all decrease systematically with respect to the peak at ~ 2.1 nm upon increasing the diamagnetic dilution, i.e. in the series from 1 in 5 to 1 in 20. Therefore 2.1 nm is identified as an intramolecular distance in α S56/69.

Initially, all mutants were screened at a diamagnetic dilution of 1 in 10, because thereby, fibrillization could be performed at total protein concentrations and sample volumes similar to those used in earlier studies.[27] These samples have more pronounced modulations, but are less reliable due to possible contributions of intermolecular distances.

Intramolecular distances in fibrils

The most pronounced peak in the distance distribution of the α S56/69 mutant (Fig. 5 c) is at 2.1 nm and diamagnetic dilution experiments (Fig. 6) identify it as an intramolecular distance. In the distance distribution of the α S56/90 mutant (Fig. 5 c, middle trace) a component with distances smaller than 2.5 nm is observed, which derives from the fast initial decay in the time trace (first 100 ns, Fig. 5a). This component is even more pronounced at a diamagnetic dilution of 1 in 10 (data not shown), showing that it derives from inter-, rather than intramolecular interactions. Also, this feature is similar in shape to the short distance contribution in the α S56/69 mutant at lower diamagnetic dilution (Fig. 6). Therefore, we attribute this contribution to an inter-, rather than intramolecular interaction. Of the remaining peaks in the distance distribution of the α S56/90 mutant, the peak at 3.8 nm has the largest intensity, and the smaller peak at 4.3 nm was shown, by the suppression tool in DEER analysis, to be non-significant. Therefore, the peak at 3.8 nm is identified as the intramolecular distance in the α S56/90 mutant. In contrast to the α S56/69 mutant, the α S69/90 mutant time trace (Fig. 4 c) does not deviate significantly from the reference traces of α S69 and α S90 in the short-time regime, i.e. up to 0.5 μ s, a clear indication of the absence of short intramolecular distances in α S69/90. A pronounced deviation at longer times reveals the presence of longer distances, and gives rise to the peak in the distance distribution (Fig. 5 c) at 4 nm. Considering this distance, the evolution time is not long enough to accurately determine the shape of the peak, resulting in a larger uncertainty for the distance in α S69/90 compared to the other mutants. For all mutants, distances below 1.8 nm were excluded by continuous wave (cw) EPR, see Fig. 3 and Materials and Methods.

The intramolecular distances obtained (Table 1) range from 2.1 nm (α S56/69) to > 4 nm (α S69/90). These distances are significantly shorter than the distances expected for same-strand pairs of residues (Table 1), showing that there must be at least one turn between residues 56 and 69 and between 69 and 90, i.e. that there are at least two turns separating residue 56 and 90.

4. Discussion

In this study we explore the potential of DEER to obtain long-range distance constraints to determine the fold of the α S protein in the fibril. To do so we performed DEER experiments on fibrillized singly and doubly labeled α S and developed an approach to generate a model of fibril architecture.

The results of AFM, ThioT fluorescence and cw EPR did not depend on the protein construct or diamagnetic dilution, showing that the fibrils are reproducible and that the presence of the spin label does not influence fibrillization.

To assure reliable distance determination a careful experimental study of the properties of the fibrils of the singly labeled proteins was performed alongside the doubly labeled ones to derive background functions. Dimensions of the background functions between one and two show that the spin labels line up preferentially in one direction, evidently the long axis of the fibril. Considering the fibril architecture, intermolecular interactions should stem from spin labels aligned along the fibril axis, separated by multiples of 0.48 nm, the spacing between two proteins along the fibril axis. Therefore, also correlated spin pairs should be present, with separations that fall within the 2 to 5 nm sensitivity range of the DEER experiment as performed here. In spite of these suspected correlated-pair interactions, the entire DEER time traces of singly labeled proteins can be fit with an exponential function of lower dimensionality - for details of background dimensionality, see [34] - showing that the net result of the random distribution of labeled proteins within the fibril is a smooth background function.

The raw time traces of the fibrils show clear evidence of intramolecular distances, given that the trace of each doubly labeled protein-fibril sample differs significantly from the ones of its singly labeled counterparts (Fig. 4). Intramolecular distances are the most likely origin of this difference, since the intermolecular interactions contribute equally to the DEER traces of fibrils from the singly and doubly labeled proteins, and therefore should not lead to a difference between the time traces. In almost all aspects the intermolecular interactions should be identical in fibrils of singly and doubly labeled proteins: The fibrils were made to have the same spin concentration and therefore should have the same density of intermolecular interactions. The only class of intermolecular interactions present in the fibrils of doubly labeled protein, but not in the singly labeled ones, are interactions of residue A in protein 1 with residue B in protein 2¹. We cannot fully exclude that this has a subtle effect on the structure of distance distributions from intermolecular interactions, but we do not expect it to have a dominant influence on the time traces. This demonstrates that the raw DEER time traces already show proof of intramolecular interactions, a result that is further emphasized by the detailed shapes. These shapes show that the intramolecular distances differ per mutant, with overall shorter distances for α S56/69 and longer distances for α S69/90, results that are borne out by the subsequent analysis.

By diamagnetic-dilution experiments intramolecular distances are identified. We show by stepwise increasing dilution that intramolecular distances dominate at dilutions of 1 in 20, without, however, fully

¹ We thank one of the referees for pointing this out to us and suggesting that this could be circumvented by using mixed singly labeled fibrils as references.

suppressing contributions from intermolecular distances (see Results). Similar dilution factors were used by Karyagina et al. [22]. In Pornsuwan et al. [23], several cases are reported, where higher diamagnetic dilutions were used, in particular when longer distances, for example, between the two outside sheets of the fibrils, were targeted. For distances in the range of 4.5 nm and above [23] long DEER evolution times are required, making it more difficult to distinguish intra- from intermolecular distances, and, as a consequence, higher diamagnetic dilutions are needed. Measurements with longer evolution times profit also particularly from the higher sensitivity of Q-band DEER, also employed in [23]. Initial Q-band DEER experiments we performed on a α S56/69 fibril sample (unpublished results) were in excellent agreement with the results obtained at X-band, showing that higher sensitivity of Q-band DEER enables shorter accumulation times, but that Q-band DEER is not essential to obtain reliable data for shorter distances in fibrils. All intramolecular distances determined in the present study were obtained by Tikhonov regularization and are collected in Table 1.

Comments on fibril morphology

Several examples of polymorphism in α S fibrils have been reported in the literature, showing that fibrils with different properties can be produced in vitro. Gath et al. [35;36] describe fibrils, termed ribbons, that do not show ThioT fluorescence. Since all fibrils investigated here were ThioT active, we can exclude this type of fibril. More subtle differences in morphology are also possible, such as straight and twisted fibrils, where, in the latter case, twists with different pitches (see for example [37-39]) were reported. Also, peak doubling in ssNMR was observed [35;36;40;41]. By AFM, the outside shape of all fibrils investigated in the present study was similar, excluding grossly different fibril morphologies. The AFM experiments were not conducted under conditions that would reveal the periodicity of twisted fibrils so they were not sensitive to more subtle differences in morphology. Also, our experiments did not reveal other tell-tale signs of polymorphism, such as multiple intramolecular distances or differences in distance distributions, when multiple experiments were performed on the same mutant. The widths of the peaks in the DEER distance distributions, in the order of 1 nm full-width-at-half-height, is so large that structural differences that cause peak doubling in ssNMR could go unnoticed.

Thus, within the framework of the experiments performed, there is no evidence of fibril polymorphism. We do not, however, have sufficient information to match the fibril type observed to a particular of the morphologies observed in the literature [35;36;40;41], a situation that according to our view is common to all studies presented so far.

Fibril model

Next, we explore the possibility to derive a model for the fold of α S in the fibril from a limited number of long-range distance constraints. Such a model is described in the appendix. This model assumes in-register configuration of the fibrils and parallel β -sheets (Fig. 1a and b), as found in several studies for α S fibrils [7;9;9;11;12;19] and considers the protein to be confined to a plane perpendicular to the fibril axis. It neglects the possibility [6] that (pairs of) strands could be at an angle with respect to others, or that the protein extends over different planes.[20] These arguably simplistic, but pragmatic, assumptions generate a framework of plausible fibril architecture. Such a model is useful, if only a limited number of long distance constraints are available, as in the present study.

With the distances obtained by DEER here, a fold as shown in Fig. 1 c and several variants discussed in the appendix result. Incorporating additional constraints, such as the distances between the residues 40, 41 and 90, 91 reported by Karyagina et al.[22] and Pornsuwan et al. [23] would place residues 40 and 41 on strand I in Fig. 1 c. The resulting stack of five β -sheets (strands I to V) would have a thickness B (Fig 1a) of 4.5 – 5 nm and a width W of 3.5 to 4 nm, in good agreement with TEM and AFM data for α S protofibrils. [10-12;42] To validate this model, additional distances are needed. Here the model can serve as a guide to derive the most useful positions that should be spin labeled in the future. We note that the distance between residues 54 and 90 reported in Pornsuwan et al. [23], is not compatible with the model shown in Fig. 1 c. Whether this is because (i) the assumptions on which the model is based are too simplistic, (ii) the fibrils had a different morphology in the respective studies, or (iii) due to some other reason cannot be determined at present and future studies are needed to clarify this point. A different approach to modelling has been put forward in Pornsuwan et al. [23], one that makes explicit use of the known properties of β -sheet substituents at adjacent residues, allowing ‘to project residues in a Cartesian coordinate space’. Given the constraints of both studies, a direct comparison of the approaches is not possible at this moment.

Comparison with information on the fold of α S from literature

Table 2 summarizes the information obtained in the present study and compares it to previous investigations. The first column gives the minimal number of turns between the residues targeted in the present study. Since the distances measured exclude same-strand locations for the pairs 56/69, 56/90 and 69/90, there must be at least one turn between each residue probed. The second column contains the location of turns that derive from the model described in the appendix. Obviously, the latter information is a lot less certain than the first, however, it enables a comparison to literature data that is listed in the following columns. Most studies agree that the β -sheet regions in the fibril core encompass at least residues 35-98 [8;9]. Starting from the N-terminus, the earliest stretch of β -sheet is proposed [36] to contain residues 16 – 21, followed by 38 – 45.

Following residue 43, several studies predict a turn (see Table 2). The region from residue 57 to 60 is classified as a possible turn in Vilar et al.[10]. The location of the second turn (66-68) from Vilar et al. [10] suggests that there are two turns between residues 56 and 69, with an intervening strand between residues 56 and 69. If such a turn was present, it would align residues 56 and 69 in a direction perpendicular to the strand and therefore would not violate the distance constraint obtained for residues 56 and 69 in the present study. However, such an arrangement is only possible if the 56-90 distance is longer than the 69-90 distance, a scenario that is not compatible with our data. Comellas et al. [14] present a model in which long and short β -strands alternate, suggesting a fibril architecture that differs from the assumptions of our model. The turn positions (58-60 and 84-90) [14] nevertheless agree very well with our model.

The model of Heise et al. [11;12], with turns at 65-69 and 82-87 again is compatible with the 56-69 distance, but not with the other distances. To further illustrate the potential of the approach, a sketch of the model(Fig. 1 c) and additional information[10;22] is shown in Fig. 7. In addition to Fig. 1c, the results of H/D exchange data from Vilar et al. are included, showing that these models are compatible. None of the present information is sufficient to define a complete model of the fibril fold so far.

Conclusions and outlook

In spite of considerable effort of different groups, the fold of α S protein within the fibril is not yet resolved. Several studies provide overlapping and consistent data, without, however, giving sufficient detail to define a model. Evidently, long-range distance constraints from DEER, as those determined in the present study, would be the ideal complementation of short-range-distance constraints from ssNMR.[18;19] Such an approach was demonstrated for the membrane-bound form of α S [43].

At this point we cannot verify or to dismiss the basic assumption of a simple fold, e.g. that the protein is located in a plane perpendicular to the fibril axis and we stress, that the fold suggested in Fig. 1 c is just a representation that is consistent with the minimal data set presented here. Evidently, the rich fibril architecture found for amyloid proteins so far [10-12;14;15;44] raises the suspicion that we would have to look beyond the simple models of α S fibril folds.

Last but not least, the recent studies have emphasized that future experiments will have to address the problem of fibril polymorphism in a systematic way, irrespective of the experimental approach. Different fibril types are difficult to distinguish and identify, but may well possess different folds, further complicating the interpretation of data.

Acknowledgements

Financial support from the Netherlands Organisation for Scientific Research (NWO) and Nanoned is gratefully acknowledged. This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). We thank Malte Drescher, Marta Robotta and Patrick Braun for discussions, and Patrick Carl, Bruker BioSpin Rheinstetten for Q-band DEER measurements.

Table 1

Intramolecular distances from DEER experiments on α S fibrils and distances expected for same-strand locations of residue pairs. All distances are in nm.

mutants	intramolecular distances	distance expected for residues on same β -strand
α S56/69	2.1	4.5
α S56/90	3.4	11.9
α S69/90	≥ 4	7.35

Table 2

Comparing turn positions in α S fibrils derived from recent ssNMR studies [10;19] [36] and the present study

this work		literature		
sequence regions where turns must be located	turns suggested (parallel linker case, Fig. 1 c)	turns predicted by Lv et al. (in italics: mouse α S)[19]	non-beta sheet regions from ssNMR studies	
56	≤ 50	50-51	43-48	45-46, 50-52,57-62,
1 turn	58-62	58-59	59-63	
		67-68		66-68
69				73-75
≥ 1 turn	70-74	79 -80	78-88	80-83
	85-90	86-89		
90				

Appendix

As described in the main text, an approach to interpret scarce distance data in terms of a simplified fibril architecture is useful to guide further experiments. The approach assumes an in-register configuration of the fibrils, parallel β -sheets (Fig. 1a and b) and considers the protein to be confined to a plane perpendicular to the fibril axis. The β -sheets of the fibril are perpendicular to that plane, so when viewed along the fibril axis as in Fig. 1 b and c they appear as a set of parallels, separated by the intersheet distance (d). The distances measured between the residues 56, 69 and 90 define the relative positions of these residues as the corners of a triangle in the plane containing the protein. To perform an unbiased search for the location of the residues 56, 69 and 90 on such an idealized model fibril, an orientation is sought where the positions of each of the three residues is as close as possible to a particular one of these parallels by rotating the triangle with respect to the parallels.

Since, by DEER, the distance between the NO-groups of the pyrrolidine rings is measured, we need to relate the positions of the nitroxide groups to the positions of the C_β -atoms. The linker between the nitroxide and the C_β -atom comprises four single bonds, length approximately 0.5 nm, making the linker flexible. Programs to model the linker orientation [45-47] require atomic-level information on the structure, which for α S fibrils is not available. Therefore, we constructed a geometric model that considers sets of different linker orientations at the three positions 56, 69 and 90, thus providing an unbiased approach to eliminate the effect or at least derive worse-case scenarios for the influence of the orientation of the spin-label linker on the structure.

First, we describe the procedure to derive a fibril model for the case where all linkers point in the same direction (parallel linkers) and below, we will discuss the effect of other linker orientations. Residue 56 is taken to be at the origin of the axis system used, and the triangle is rotated about this origin. Several orientations are found. Arrangements in which the residues are not sequential on successive parallels are excluded on the grounds that such an arrangement would result in too many turns of the protein given the number of intervening residues. That leaves four possible orientations. Two of these solutions involve same-strand locations for the pairs 56 and 69 or 69 and 90 and are therefore excluded (see above).

In the remaining two solutions the residues are at the positions marked by black squares in Fig. 1c. To generate a model for the fold of the protein given the positions of the residues 56, 69 and 90, the protein sequence is threaded through these points. The protein follows the direction of the β -sheet and subsequent residues within a strand are considered to be separated by 0.35 nm (l). Turns are placed at sequence positions such that they optimally complement opposing strands. They are considered to be made up of three residues, where residue i is the last residue in the preceding β -strand, and residue $i + 4$ is the first

residue of the subsequent strand that is part of the opposite β -sheet. There are two ways of threading the protein through the three squares shown in Fig. 1c: Starting at residue 56, the protein sequence can be threaded on strand I in the direction away from the position of 69, i.e. in the direction of the red arrow in Fig. 1c. The other possibility is to thread the protein from position 56 towards 69, i.e. in the direction shown by the red arrow in Fig. A1 a. In the latter case, residues 70 to 74 would be needed to complement strand III. A turn involving residues 74 to 78 and a two-residue strand followed by a turn at 80 to 84 would allow to cover the distance to position 90, leaving a stretch of residues from 86 to 90 uncomplemented by an opposing β -sheet. This scenario, which is illustrated in Fig. A1 a, seems unlikely, and is therefore discarded. Consequently, the model shown in Fig. 1c is the most likely fold amongst the solutions found. It consists of four zippered β -sheets, (strands II - V, Fig. 1c) requiring that intermediate sheets (e.g. strand III) would form ‘dry interfaces’[1] at both sides of the sheet. The prediction of turns is somewhat arbitrary, but assuming similar lengths of β -strands, turns at residues 58-62 (residue i and residue $i + 4$), 70-74, and 85-89 are plausible. Next we discuss how different linker orientations could affect this model.

The influence of different spin-label-linker conformations

To assess the effect of the flexible linker joining the nitroxide group to the protein-backbone a geometric model was developed, which mimics different orientations of the nitroxide groups. In this model, the linker is considered as having a fixed length of 0.5 nm and it can have different orientations with respect to the triangle formed by the positions of residues 56, 69 and 90 (black squares in Fig. 1c). Three orientations are considered per residue, each 120° apart to account for the difference in position of the protein backbone and the nitroxide group. Since only relative orientations are of interest, there are nine possible combinations of linker orientations for the residues 56, 69 and 90. Besides the possibility that all linkers are parallel, five other arrangements are found. In Fig. A1 b) three combinations of linker orientations are shown: The black squares for parallel linker orientations, and the orange and pink dots, for which the linker direction at residue 69 has different orientations. The three-strand model prevails for residues 56 to 90, showing that the model shown in Fig. 1c is robust, but the strands are shifted parallel to each other, juxtaposing other residues than in Fig. 1c at opposite strands. So, even if the linkers at residue 69 and 90 are oriented nearly parallel to the strand (orange dots, Fig. A1 b), the corresponding model of the fold is similar to the one shown in Fig. 1c, with the exception that strand IV is shifted by two residue positions with respect to strand II, and turn positions shift by maximally three residues.

If, however, the linkers at residues 69 and 90 point in a direction opposite to that of residue 56, i.e. if the linker at residue 56 points to the left of the strand in Fig. 1c, linkers at residues 69 and 90 point to the right side of their respective strand, or vice versa (Fig. A1 c) and d)). This solution is possible for two sets

of orientations of linkers (green and purple dots, Fig. A1). The two ways of threading the protein in this case are shown in Fig. A1 c) and d). The main difference is that there is only one turn between residues 69 and 90. A single turn between residue 69 and 90 was suggested in [10-12;14], however, these models predict two turns between residues 56 and 69, which is not compatible with the model shown in Fig. 1 c) and A1. Also, some studies [9] position residue 90 outside the β -sheet core, indicating that it could be in a turn region. We consider the folds shown in Fig. A1 c and d less likely, because the overall extent of the fibrillar region seems smaller than expected from the outside dimensions of the protofibrils [10], but we note that we cannot fully exclude it.

Figure Captions

Figure 1.

Fold of the protein in the α S fibrils. a): Fibril side view (fibril long axis: black arrow), colored planes: β -sheets, dark blue arrows: proteins making up the fibril; b): View along fibril axis. Protein (light blue), turns connecting strand sections of the protein. Strand: protein section involved in β -sheet. Distance between residues (dots): l and distance between strands: d . c): Suggestion of a model for the fold that derives from the three distance constraints determined in the present study. Vertical grid lines: Strands I-V. Black squares: Positions determined from the DEER data (see text and appendix for details) for residues 56, 69 and 90. Coloured arrows: β -sheets viewed along fibril axis (colours see side view in a), dark blue: turns, italics: numbers of residues at the start and at the end of the turn (details, see text). Residue 41 would be located on strand I in that model.

Figure 2

AFM images (A-D) of fibrils grown for different diamagnetic dilutions. A: α S56C, at diamagnetic dilution 1 in 20, B: α S56/69C at 1 in 10, C: α S56/90C at 1 in 10, D: α S69/90C at 1 in 20. All images are $2.5 \times 2.5 \mu\text{m}^2$. Height is shown as color density with a range from 0 to 10 nm.

Figure 3

Continuous wave EPR spectra obtained at room temperature of the harvested fibrils of α S. Shown are the spectra of the doubly-labeled proteins superimposed on their corresponding singly-labeled proteins a)-c) at a 1 in 10 diamagnetic dilution. Narrow lines, marked by asterisk, stem from a small contribution of nonfibrillized proteins < 10 %, a) α S56/69C superimposed with α S56C and α S69C, b): α S56/90C superimposed with α S56C and α S90C, c): α S69/90C superimposed with α S69C and α S90C. d) – e) Continuous wave EPR spectra obtained at 80 K of the harvested fibrils of doubly-labeled proteins superimposed with the 1:1 added spectra of the respective singly-labeled counterparts. All proteins were at a 1 in 10 diamagnetic dilution d) α S56/69C (green solid line) superimposed with α S56C plus α S69C (black broken line), e) α S56/90C superimposed with α S56C plus α S90C, f) α S69/90C superimposed with α S69C plus α S90C.

Figure 4

Experimental DEER time traces of mutants a) α S56/69, b) α S56/90, c) α S69/90 and their respective singly labeled counterparts. Diamagnetic dilution of doubly labeled proteins 1 in 20, of singly labeled 1 in

10. In all cases clear differences are observed between traces of the doubly labeled variants compared to the singly labeled reference fibrils. Comparing time traces derived from fibrils of doubly and singly labeled protein, a) shows a fast initial decay of the time trace, indicative of a shorter distance component, which in b) is less pronounced and in c) no evidence for such a short distance component is seen, showing that the distances between the spin labels increase in this order.

Figure 5

Experimental DEER time traces (a) of mutants α S56/69, α S56/90, and α S69/90 at diamagnetic dilution 1 in 20. Shown are: a) Original data with the baseline derived from the sum of the respective singly-labeled protein traces (dotted line). All singly-labeled α S samples used as background were at diamagnetic dilutions 1 in 10. The maximum of the echo intensity in each DEER-time trace is normalized to 1 and the traces are shifted for better visibility. b) Baseline corrected data and the fit to the data corresponding to the distance distributions shown in c). c). Distance distributions derived from the data. Arrows mark the distances identified as intramolecular (see text for details). Distributions were individually scaled to improve visibility.

Figure 6 Effect of diamagnetic dilution on the DEER time traces (left) and the distance distributions (right) of the mutant α S56/69 fibrillized diamagnetic dilutions. The distance peaks at 4.5, 3.9 resp. 3.8 (1 in 5 resp. 1 in 10), 2.7 and 2.2 nm decrease in intensity relative to the broader peak at 2.1 nm in the dilution series from 1 in 5 to 1 in 20, revealing that the former peaks derive from intermolecular interactions.

Figure 7

Representation of all intramolecular distances measured for the fibrils of α S shown to scale with respect to the full length α S (grey thick line) Residues 56, 69 and 90 were spin labeled. Distances are given as arrows. Distance 56-69 red, 56-90 orange, 69-90 light green. Also shown: distance 41-90 (light blue) from Karyagina et al.[22]. Positions of residues 56, 69 and 90 derive from the distances measured (see also Fig. 1c, main text). Superimposed is one version of the fibril model presented by Vilar et al. focusing on the results of H/D exchange experiments[10] where amino acid residues are shown as squares with sizes fitted to scale. Filled grey squares: Amino acid in β -sheet (extremely slow H/D exchange), open squares: amino acids not in β -sheet (faster H-D exchange), revealing that the NMR H-D exchange data are consistent with the EPR results of the present study.

Figure A1

Models of the α S-fibril illustrating alternating threading (a) and the effect of different linker orientations (b –d). Colored arrows indicate successive β sheets, and dark blue, bent lines, turns. Italic numbers: Residues at the start and the end of the turn (see text) a) Alternative way of threading the protein for the positions shown in Fig. 1c. For details see text. b) Effect of linker orientation: Black squares, all linkers parallel. Pink and orange dots give the relative positions of residues 69 and 90 for the two other linker orientations that are compatible with the fold shown in Fig. 1c. The similarity of the solutions reveal that this fold is consistent with cases in which all linkers are parallel (black squares, shown also in Fig. 1c), and in which one or both linkers at 69 and 90 are almost parallel to the strand but on the same side of the respective strand, as the linker at residue 56 (for details see text). The threading model is shown for the pink positions. Hashed turns indicate turn positions for the model shown in Fig. 1c of the main text. With respect to the model shown in Fig. 1c, strand IV is shifted by 2 residues relative to strand II, and turn positions differ by one (turn between strand II and III and two residues (turn between strand IV and V). c) Fold for linker orientations with the linker at residue 56 opposite to those at residues 69 and 90 for two different linker orientations (green and purple dots). The purple dot and the square close to the green dot labeled 69 are the alternative positions of residue 69. The black square between strands IV and V shows the position of residue 90 in the model shown in Fig. 1c of the main text. d) Same positions as for the model in c) with alternative threading of the protein (see text).

References

- [1] M.R. Sawaya, S. Sambashivan, R. Nelson, M.I. Ivanova, S.A. Sievers, M.I. Apostol, M.J. Thompson, M. Balbirnie, J.J.W. Wiltzius, H.T. McFarlane, A.O. Madsen, C. Riek, D. Eisenberg, *Nature* 447, 453-457 (2007)
- [2] J.T. Nielsen, M. Bjerring, M.D. Jeppesen, R.O. Pedersen, J.M. Pedersen, K.L. Hein, T. Vosegaard, T. Skrydstrup, D.E. Otzen, N.C. Nielsen, *Angewandte Chemie-International Edition* 48, 2118-2121 (2009)
- [3] M.I. Apostol, M.R. Sawaya, D. Cascio, D. Eisenberg, *J.Biol.Chem.* 285, 29671-29675 (2010)
- [4] M.J. Bayro, G.T. Debelouchina, M.T. Eddy, N.R. Birkett, C.E. MacPhee, M. Rosay, W.E. Maas, C.M. Dobson, R.G. Griffin, *Journal of the American Chemical Society* null (2011)
- [5] T. Luhrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Doeli, D. Schubert, R. Riek, *Proceedings of the National Academy of Sciences of the United States of America* 102, 17342-17347 (2005)
- [6] H. Van Melckebeke, C. Wasmer, A. Lange, A.B. Eiso, A. Loquet, A. Böckmann, B.H. Meier, *Journal of the American Chemical Society* 132, 13765-13775 (2010)
- [7] M. Margittai, R. Langen, *Quarterly Reviews of Biophysics* 41, 265-297 (2008)
- [8] A. Der-Sarkissian, C.C. Jao, J. Chen, R. Langen, *J.Biol.Chem.* 278, 37530-37535 (2003)
- [9] M. Chen, M. Margittai, J. Chen, R. Langen, *Journal of Biological Chemistry* 282, 24970-24979 (2007)
- [10] M. Vilar, H.T. Chou, T. Luhrs, S.K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, R. Riek, *Proceedings of the National Academy of Sciences of the United States of America* 105, 8637-8642 (2008)
- [11] H. Heise, W. Hoyer, S. Becker, O.C. Andronesi, D. Riedel, M. Baldus, *Proceedings of the National Academy of Sciences of the United States of America* 102, 15871-15876 (2005)
- [12] H. Heise, M.S. Celej, S. Becker, D. Riedel, A. Pelah, A. Kumar, T.M. Jovin, M. Baldus, *Journal of Molecular Biology* 380, 444-450 (2008)
- [13] K.D. Kloepper, K.L. Hartman, D.T. Ladror, C.M. Rienstra, *Journal of Physical Chemistry B* 111, 13353-13356 (2007)
- [14] G. Comellas, L.R. Lemkau, A.J. Nieuwkoop, K.D. Kloepper, D.T. Ladror, R. Ebisu, W.S. Woods, A.S. Lipton, J.M. George, C.M. Rienstra, *Journal of Molecular Biology* 411, 881-895 (2011)
- [15] J. Gath, B. Habenstein, L. Bousset, R. Melki, B.H. Meier, A. Bockmann, *Biomolecular Nmr Assignments* 6, 51-55 (2012)
- [16] L.R. Lemkau, G. Comellas, K.D. Kloepper, W.S. Woods, J.M. George, C.M. Rienstra, *Journal of Biological Chemistry* 287, 11526-11532 (2012)

- [17] G. Comellas, L.R. Lemkau, D.H.H. Zhou, J.M. George, C.M. Rienstra, *Journal of the American Chemical Society* 134, 5090-5099 (2012)
- [18] A. Loquet, K. Giller, S. Becker, A. Lange, *Journal of the American Chemical Society* 132, 15164-15166 (2010)
- [19] G.H. Lv, A. Kumar, K. Giller, M.L. Orcellet, D. Riedel, C.O. Fernandez, S. Becker, A. Lange, *Journal of Molecular Biology* 420, 99-111 (2012)
- [20] S. Bedrood, Y.Y. Li, J.M. Isas, B.G. Hegde, U. Baxa, I.S. Haworth, R. Langen, *J.Biol.Chem.* 287, 5235-5241 (2012)
- [21] M.H. Shabestari, I.M.J. Segers-Nolten, M.M.A.E. Claessens, B.D. van Rooijen, V. Subramaniam, M. Huber, *Biophysical Journal* 102, 454A (2012)
- [22] I. Karyagina, S. Becker, K. Giller, D. Riedel, T.M. Jovin, C. Griesinger, M. Bennati, *Biophysical Journal* 101, L1-L3 (2011)
- [23] S. Pornsuwan, K. Giller, D. Riedel, S. Becker, C. Griesinger, M. Bennati, *Angewandte Chemie-International Edition* 52, 10290-10294 (2013)
- [24] M.E. van Raaij, I.M. Segers-Nolten, V. Subramaniam, *Biophys.J.* 91, L96-L98 (2006)
- [25] G. Veldhuis, I. Segers-Nolten, E. Ferlemann, V. Subramaniam, *Chembiochem* 10, 436-+ (2009)
- [26] M. Drescher, F. Godschalk, G. Veldhuis, B.D. van Rooijen, V. Subramaniam, M. Huber, *Chembiochem* 9, 2411-2416 (2008)
- [27] M.E. van Raaij, I.M. Segers-Nolten, V. Subramaniam, *Biophys.J.* 91, L96-L98 (2006)
- [28] K.O. Vanderwerf, C.A.J. Putman, B.G. Degrooth, F.B. Segerink, E.H. Schipper, N.F. Vanhulst, J. Greve, *Review of Scientific Instruments* 64, 2892-2897 (1993)
- [29] M. Drescher, G. Veldhuis, B.D. van Rooijen, S. Milikisyants, V. Subramaniam, M. Huber, *Journal of the American Chemical Society* 130, 7796-+ (2008)
- [30] G. Jeschke, *Chemphyschem* 3, 927-932 (2002)
- [31] G. Jeschke, A. Koch, U. Jonas, A. Godt, *Journal of Magnetic Resonance* 155, 72-82 (2002)
- [32] G. Jeschke, *Macromolecular Rapid Communications* 23, 227-246 (2002)
- [33] G. Jeschke, Y. Polyhach, *Physical Chemistry Chemical Physics* 9, 1895-1910 (2007)
- [34] G. Jeschke, A. Koch, U. Jonas, A. Godt, *Journal Of Magnetic Resonance* 155, 72-82 (1902)
- [35] L. Bousset, L. Pieri, G. Ruiz-Arlandis, J. Gath, P.H. Jensen, B. Habenstein, K. Madiona, V. Olieric, A. Bockmann, B.H. Meier, R. Melki, *Nature Communications* 4, (2013)
- [36] J. Gath, L. Bousset, B. Habenstein, R. Melki, A. Bockmann, B.H. Meier, *Plos One* 9, (2014)
- [37] K.K.M. Sweers, I.M.J. Segers-Nolten, M.L. Bennink, V. Subramaniam, *Soft Matter* 8, 7215-7222 (2012)

- [38] M.E. van Raaij, J. van Gestel, I.M.J. Segers-Nolten, S.W. de Leeuw, V. Subramaniam, *Biophysical Journal* 95, 4871-4878 (2008)
- [39] M.E. van Raaij, I.M.J. Segers-Nolten, V. Subramaniam, *Biophysical Journal* 91, L96-L98 (2006)
- [40] H. Heise, W. Hoyer, S. Becker, O.C. Andronesi, D. Riedel, M. Baldus, *Proceedings of the National Academy of Sciences of the United States of America* 102, 15871-15876 (2005)
- [41] H. Heise, M.S. Celej, S. Becker, D. Riedel, A. Pelah, A. Kumar, T.M. Jovin, M. Baldus, *Journal of Molecular Biology* 380, 444-450 (2008)
- [42] W.G. Hoyer, D. Cherny, V. Subramaniam, T.M. Jovin, *Journal of Molecular Biology* 340, 127-139 (2004)
- [43] J.N. Rao, C.C. Jao, B.G. Hegde, R. Langen, T.S. Ulmer, *Journal of the American Chemical Society* 132, 8657-8668 (2010)
- [44] F. Shewmaker, R.P. McGlinchey, R.B. Wickner, *J.Biol.Chem.* 286, 16533-16540 (2011)
- [45] Y. Polyhach, G. Jeschke, *Spectroscopy-An International Journal* 24, 651-659 (2010)
- [46] M.M. Hatmal, Y.Y. Li, B.G. Hegde, P.B. Hegde, C.C. Jao, R. Langen, I.S. Haworth, *Biopolymers* 97, 35-44 (2012)
- [47] G. Hagelueken, R. Ward, J.H. Naismith, O. Schiemann, *Appl.Magn.Reson.* 42, 377-391 (2012)

Figure 1

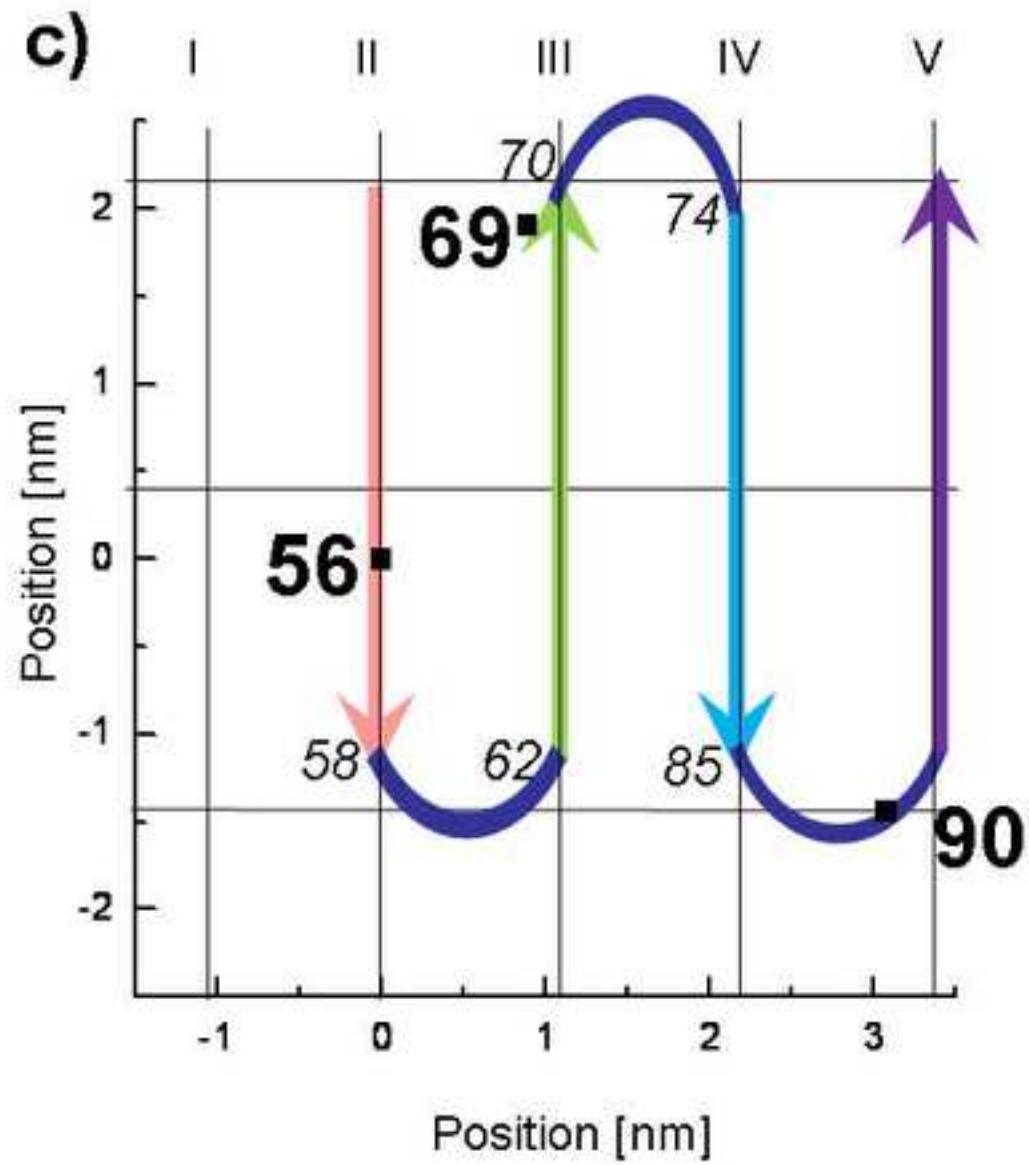
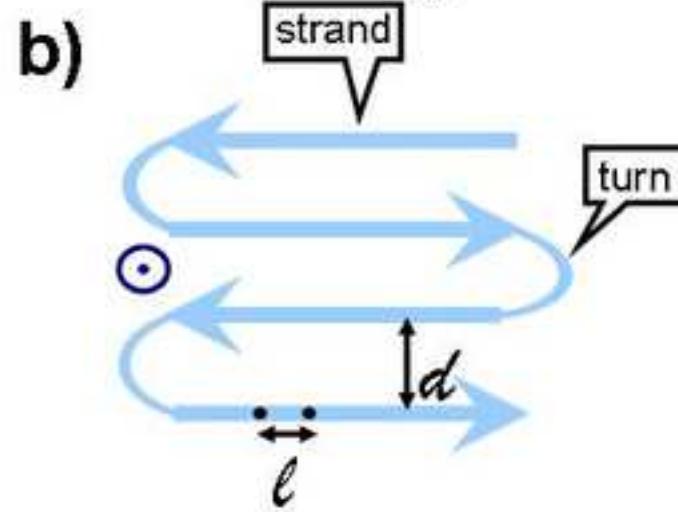
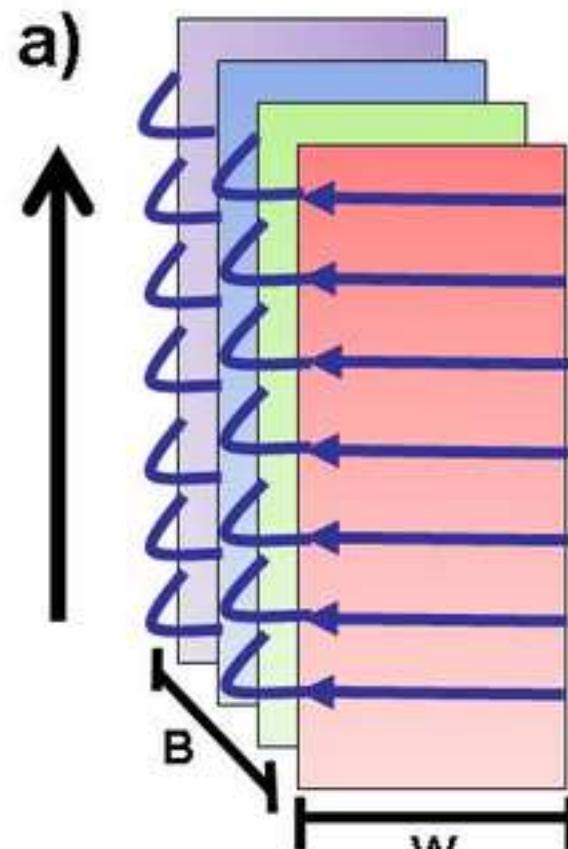


Figure 2

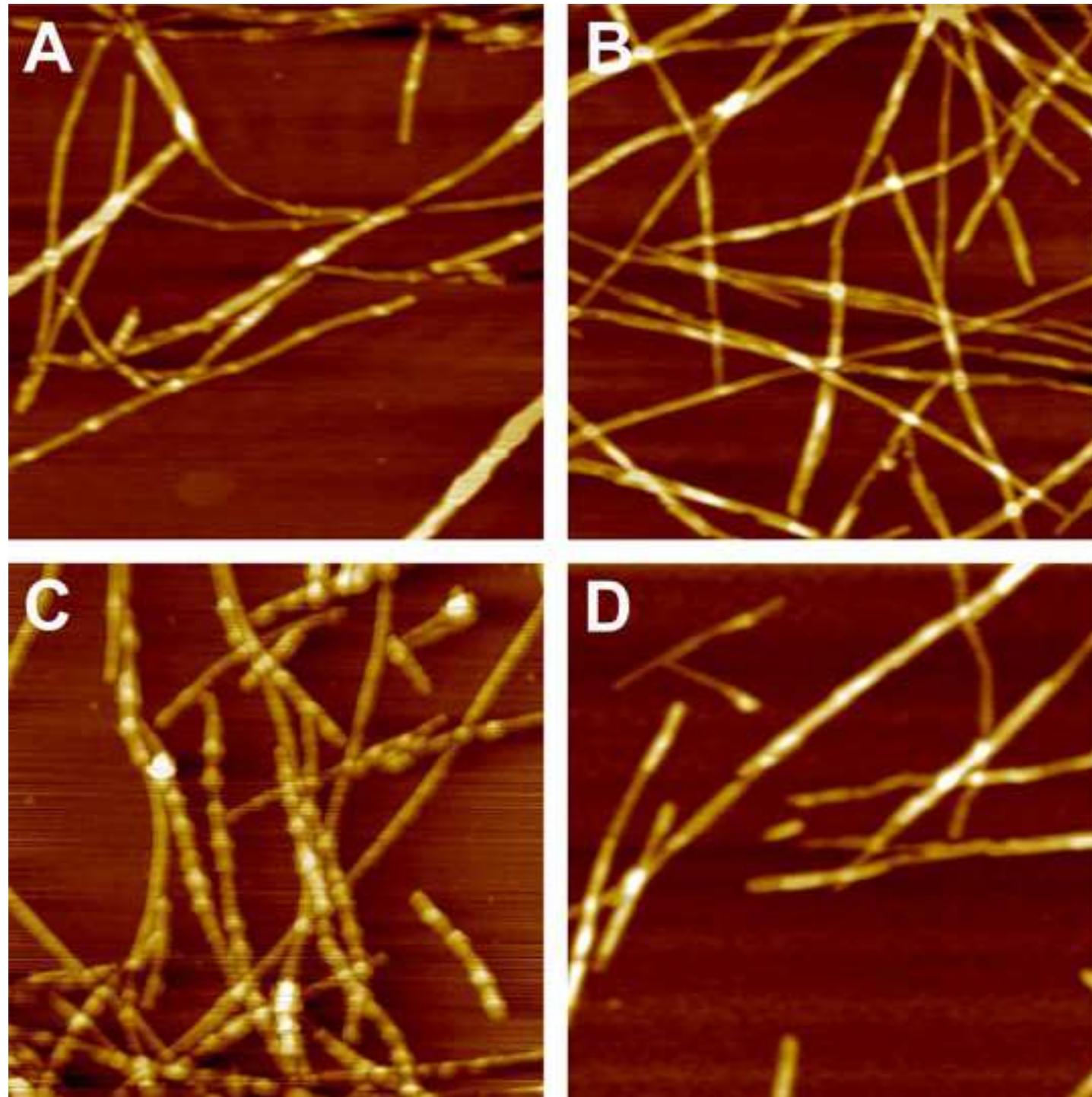


Figure 3

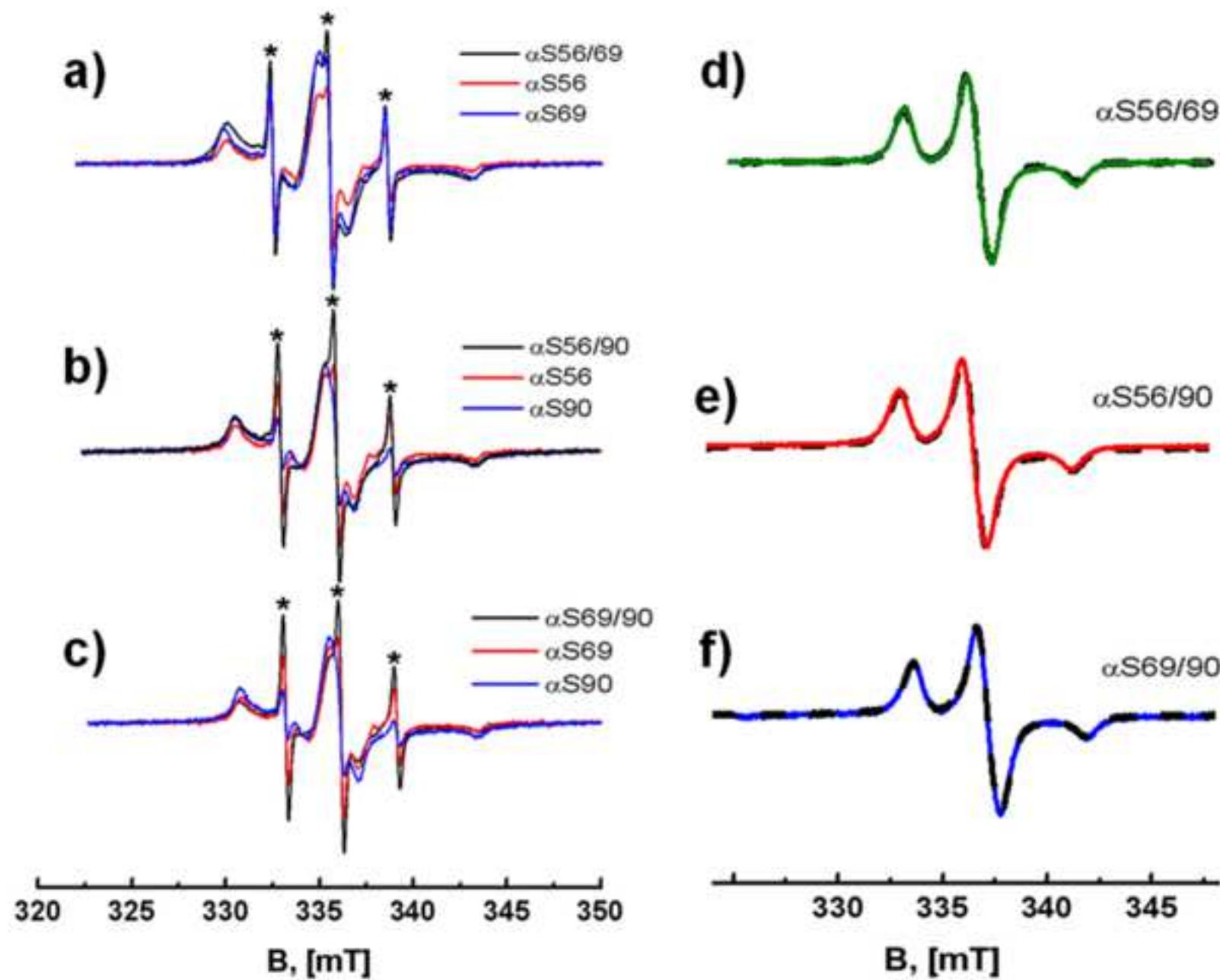


Figure 4

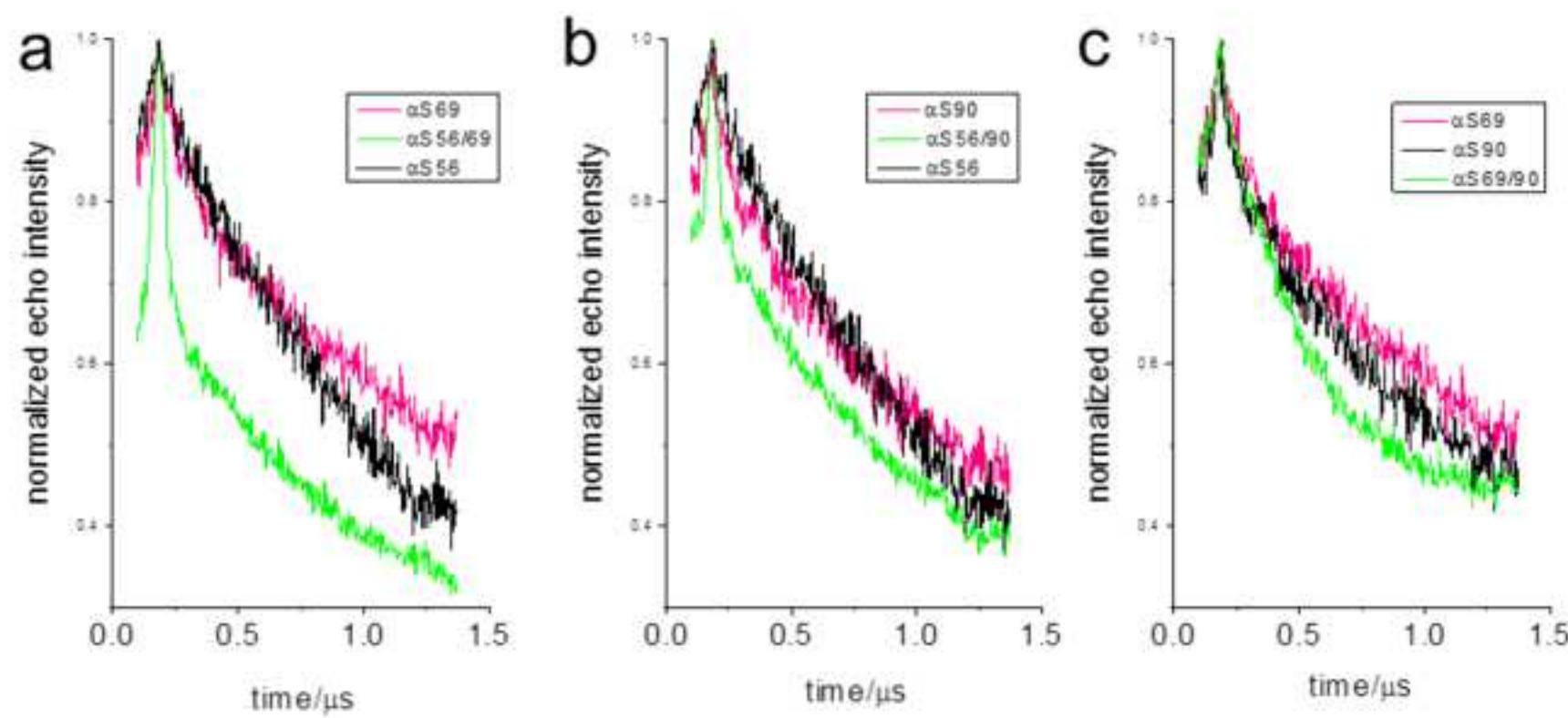


Figure 5

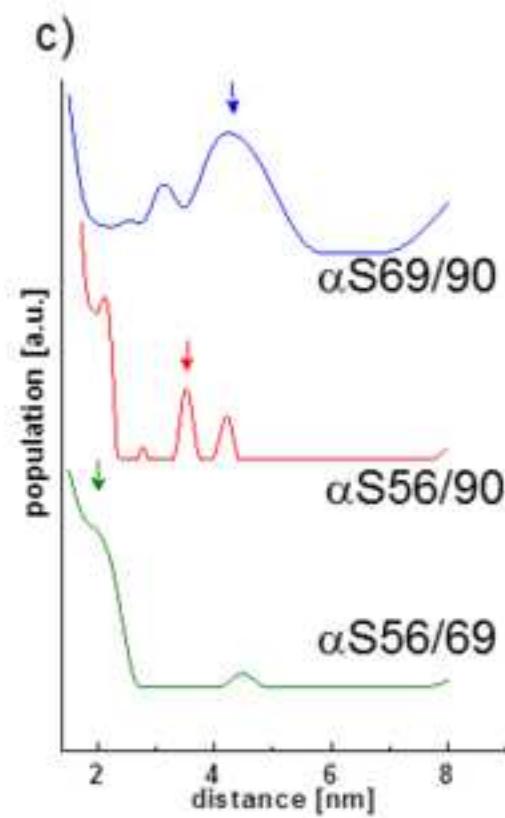
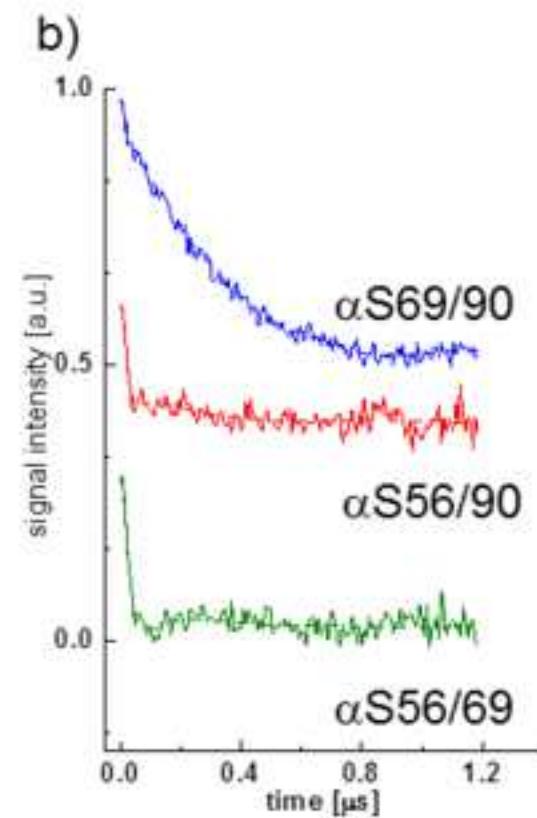
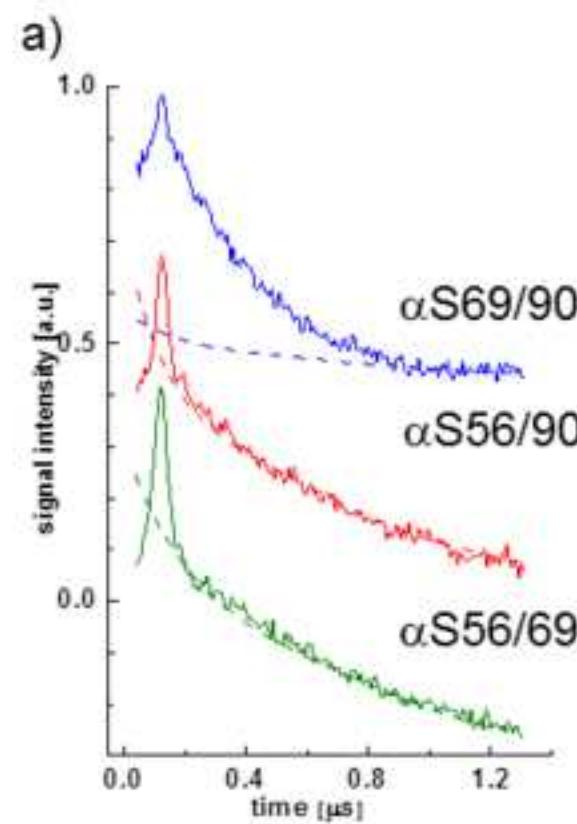


Figure 6

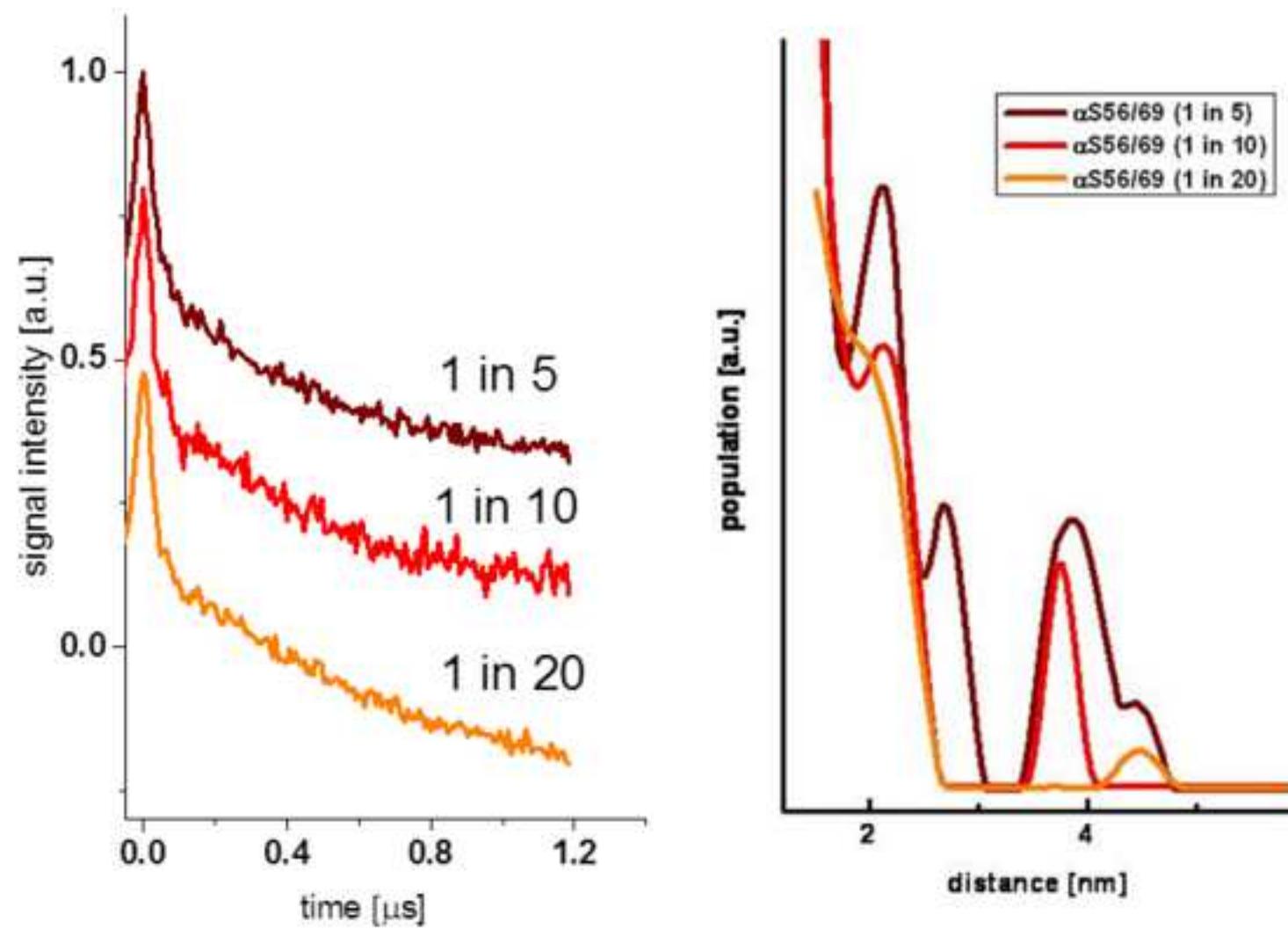


Figure 7

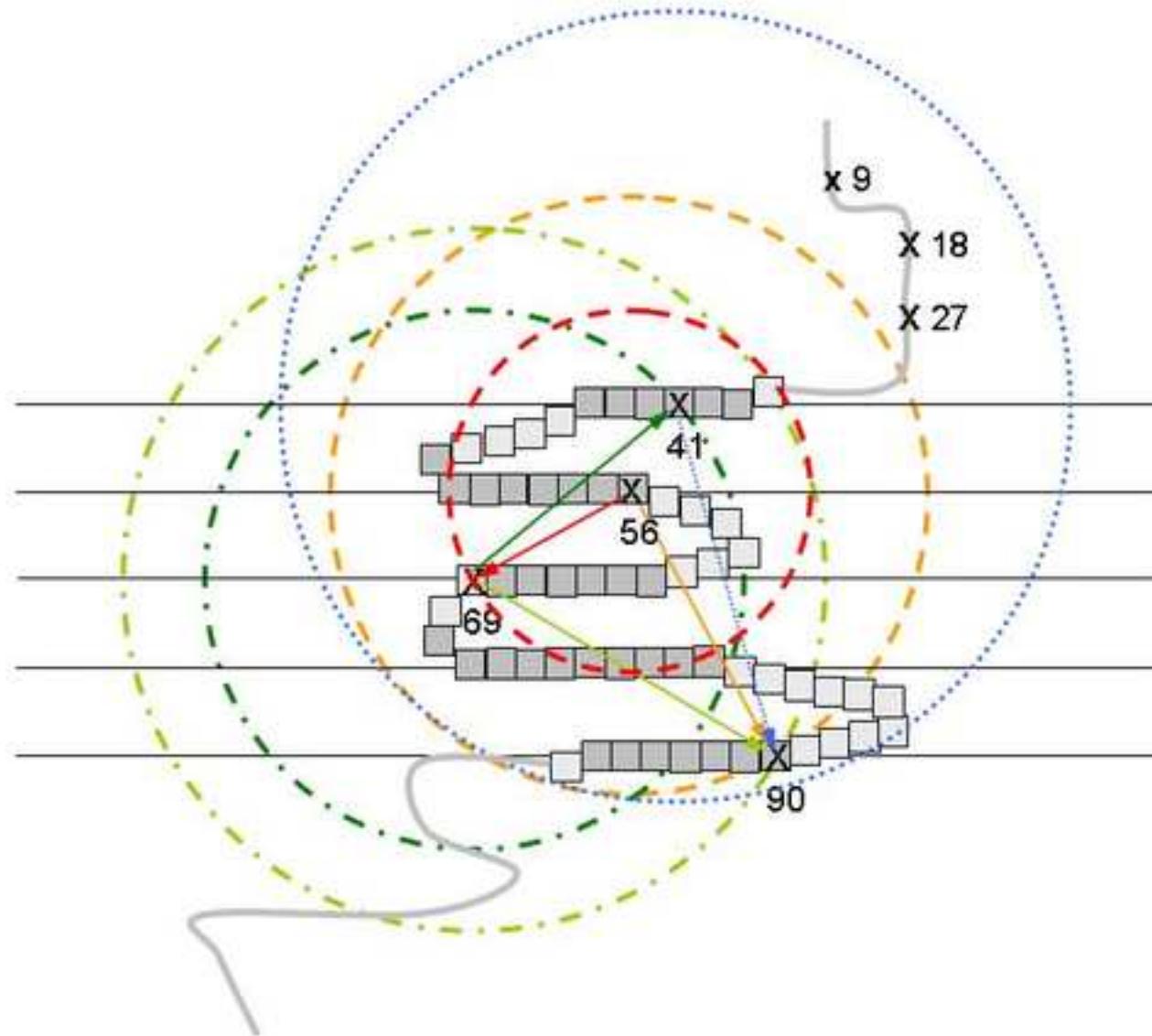


Figure A1

