RESEARCH ARTICLE

Conformational ensemble of native α-synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations

Nicholas I. Brodie[1ʘ], Konstantin I. Popov[2ʘ], Evgeniy V. Petrotchenko[1,3¤], Nikolay V. Dokholyan[2,4]*, Christoph H. Borchers[1,3,5,6]*

1 University of Victoria -Genome British Columbia Proteomics Centre, Vancouver Island Technology Park, Victoria, British Columbia, Canada, **2** Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina, United States of America, **3** Segal Cancer Proteomics Centre, Lady Davis Institute, Jewish General Hospital, McGill University, Quebec, Canada, **4** Departments of Pharmacology, and Biochemistry and Molecular Biology, Pennsylvania State College of Medicine, Hershey, Pennsylvania, United States of America, **5** Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada, **6** Gerald Bronfman Department of Oncology, Jewish General Hospital, McGill University, Montreal, Quebec, Canada

ʘ These authors contributed equally to this work.
¤ Current address: Segal Cancer Proteomics Centre, Lady Davis Institute, Jewish General Hospital, McGill University, Quebec, Canada
* dokh@psu.edu (NVD); christoph@proteincentre.com (CHB)

Abstract

Combining structural proteomics experimental data with computational methods is a powerful tool for protein structure prediction. Here, we apply a recently-developed approach for *de novo* protein structure determination based on the incorporation of short-distance crosslinking data as constraints in discrete molecular dynamics simulations (CL-DMD) for the determination of conformational ensemble of the intrinsically disordered protein α-synuclein in the solution. The predicted structures were in agreement with hydrogen-deuterium exchange, circular dichroism, surface modification, and long-distance crosslinking data. We found that α-synuclein is present in solution as an ensemble of rather compact globular conformations with distinct topology and inter-residue contacts, which is well-represented by movements of the large loops and formation of few transient secondary structure elements. Non-amyloid component and C-terminal regions were consistently found to contain β-structure elements and hairpins.

Author summary

As the population ages, neurodegenerative diseases such as Parkinson's disease will become an increasing problem in many countries. Aggregation of the protein α-synuclein is the primary cause of Parkinson's disease, but there is still a dearth of structural information pertaining to the native, non-aggregating form of this protein. A better understanding the structural state of the native protein may prove useful for the design of new

therapeutics to combat this disease. In order to obtain more structural information on this protein, we have recently modelled the native α-synuclein protein. These models were generated using a novel approach which combines protein crosslinking and discrete molecular dynamics simulations. We have found that the α-synuclein protein can adopt several shapes, all with a similar topology, resembling a three fingered closed claw. A region of the protein important for aggregation was found to be protected from the surrounding biological environment in these conformations, and the stabilization of these structures may be a fruitful avenue for future drug research into mitigating the cause and effect of Parkinson's disease.

Introduction

α-Synuclein is involved in the pathogenesis of misfolding-related neurodegenerative diseases, in particular Parkinson's disease [1, 2]. A misfolding event leads to the formation of oligomers which are believed to result in cell toxicity and which eventually lead to the death of neuronal cells [3]. α-Synuclein is thought to interact with lipid vesicles *in vivo* [4] and the toxicity is thought to be mediated via membrane disruption by misfolded oligomers [5]. Moreover, a prion-like spread of the pathology *via* the conversion of native α-synuclein molecules by toxic oligomers has been suggested [6]. Native α-synuclein is considered to be an intrinsically disordered protein, although there is evidence that some globular structure exists in solution, which may serve as a basis for understanding the mis-folding and oligomerization pathways. A number of biophysical methods, such as NMR, EPR, FRET, and SAXS—in combination with computational methods—have been applied to the study of intrinsically disordered proteins, including the structure of α-synuclein in solution [7–10]. In all of these cases, even a limited amount of experimental structural data was helpful in the characterization of the conformational ensemble of α-synuclein in solution.

Recently, we developed a method for determination of protein structures, termed short-distance crosslinking constraint-guided discrete molecular dynamics simulations (CL-DMD), where the folding process is influenced by short-distance experimental constraints which are incorporated into the DMD force field [11]. Adding constraints to DMD simulations results in a reduction of the possible conformational space and allows the software to achieve protein folding on a practical time scale. We have tested this approach on well-structured proteins including myoglobin and FKBP and have observed clear separation of low-energy clusters and a narrow distribution of structures within the clusters. The conformational flexibility of intrinsically disordered proteins, such as α-synuclein, brings additional challenges to the computational process [12]. In cases like this, proteins exist as a collection of inter-converting conformational states, and crosslinking data represents multiple conformations of a protein rather than a single structure. In addition, recent research indicates that traditional force fields with their parametrization are not ideal for providing an accurate description of disordered proteins, and tend to produce more compact structures [13]. Recently research has been focused on improving traditional state-of-the-art force fields and their ability to predict structures of disordered proteins without losing their accuracy for structured proteins [14]. In this work we use a Medusa force field [15–17] that is utilized in DMD simulations is discretized to mimic continuous potentials. DMD uses a united atom representation for the protein where all heavy atoms and polar hydrogens are explicitly accounted. The solvation energy is described in terms of the discretized Lazaridis-Karplus implicit solvation model [18] and

inter-atomic interactions, such van der Waals and electrostatics, are approximated by a series of multistep square-well potentials.

Other additional potentials, such as pair-wise distance constraints [19, 20] and solvent accessibility information [21, 22] can also be readily integrated. During CL-DMD simulations there are no continuous forces that would drive the atoms to satisfy all constraints, rather generating conformational ensembles, which satisfy an optimal number of the constraints are generated. This, to some degree, naturally resolves conflicting experimental constraints. Thus, CL-DMD simulations are a viable computational platform for the structural analysis of intrinsically disordered proteins [23] in general, and α-synuclein in particular.

Here, we used the CL-DMD approach [11] to determine conformational ensembles of the α-synuclein protein in solution. During this process, α-synuclein was crosslinked with a panel of short-range crosslinkers, crosslinked proteins were enzymatically digested, crosslinked residues were determined by LC-MS/MS analysis, and the resulting data on inter-residues distances were introduced into DMD force field as external constraints. To experimentally validate the predicted structures, we analyzed α-synuclein using surface modification (SM), circular dichroism (CD), hydrogen-deuterium exchange (HDX), and long-distance crosslinking (LD-CL).

Methods

All materials were from Sigma-Aldrich, unless noted otherwise.

Structural proteomics

α-Synuclein was crosslinked with a panel of short-range reagents azido-benzoic acid succinimide (ABAS-$^{12}C_6$/$^{13}C_6$), succinimidyl 4,4'-azipentanoate (SDA), [24] triazidotriazirine (TATA-$^{12}C_3$/$^{13}C_3$), [25] and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) [26]. ABAS and SDA are hetero-bifunctional amino group-reactive and photo-reactive reagents, TATA is a homo-bifunctional photo-reactive reagent, and EDC is a zero-length carboxyl-to-amino group crosslinker. Crosslinked proteins were digested with proteinase K or trypsin proteolytic enzymes, and the digest was analyzed by LC-MS/MS to identify crosslinked peptides (S1 Table). We used an equimolar mixture of ^{14}N- and ^{15}N-metabolically labeled α-synuclein to exclude potential inter-protein crosslinks from the analysis and to facilitate the assignment of crosslinked residues based on the number of nitrogen atoms in the crosslinked peptides and the MS/MS fragments [27]. The distances between crosslinked residues are based on the length of the crosslinker reagents, and were introduced as constraints into the DMD potentials (see section below and [11] for additional details). A total of 30 crosslinking constraints were used in these DMD simulations (S1 Table). In addition, α-synuclein was characterized by top-down ECD- and UVPD-FTMS HDX and CD to determine the secondary-structure content (Fig 1 and S1 Fig). Quantitative differential surface modification experiments were performed with and without 8 M urea to determine the characteristics of the residues as exposed or buried (S2 Table). LD-CL was used to estimate the overall protein topology (S3 Table).

Expression and purification of α-synuclein

α-Synuclein was expressed using a pET21a vector provided by Dr. Carol Ladner of the University of Alberta. The protein was expressed in E. coli BL21(DE3) bacteria and was purified as in [25]. Briefly, the protein was overexpressed with 1 mM IPTG in 1L LB cultures of BL21DE3 E. coli for 4 hours at 30˚C. Cells were lysed with a French press and the lysate was heated at 70˚C for 10 minutes and then centrifuged at 14000 g for 30 minutes. The soluble fraction was precipitated for 1 hour in 2.1 M (NH$_4$)$_2$SO$_4$. α-Synuclein was then purified by fast protein liquid
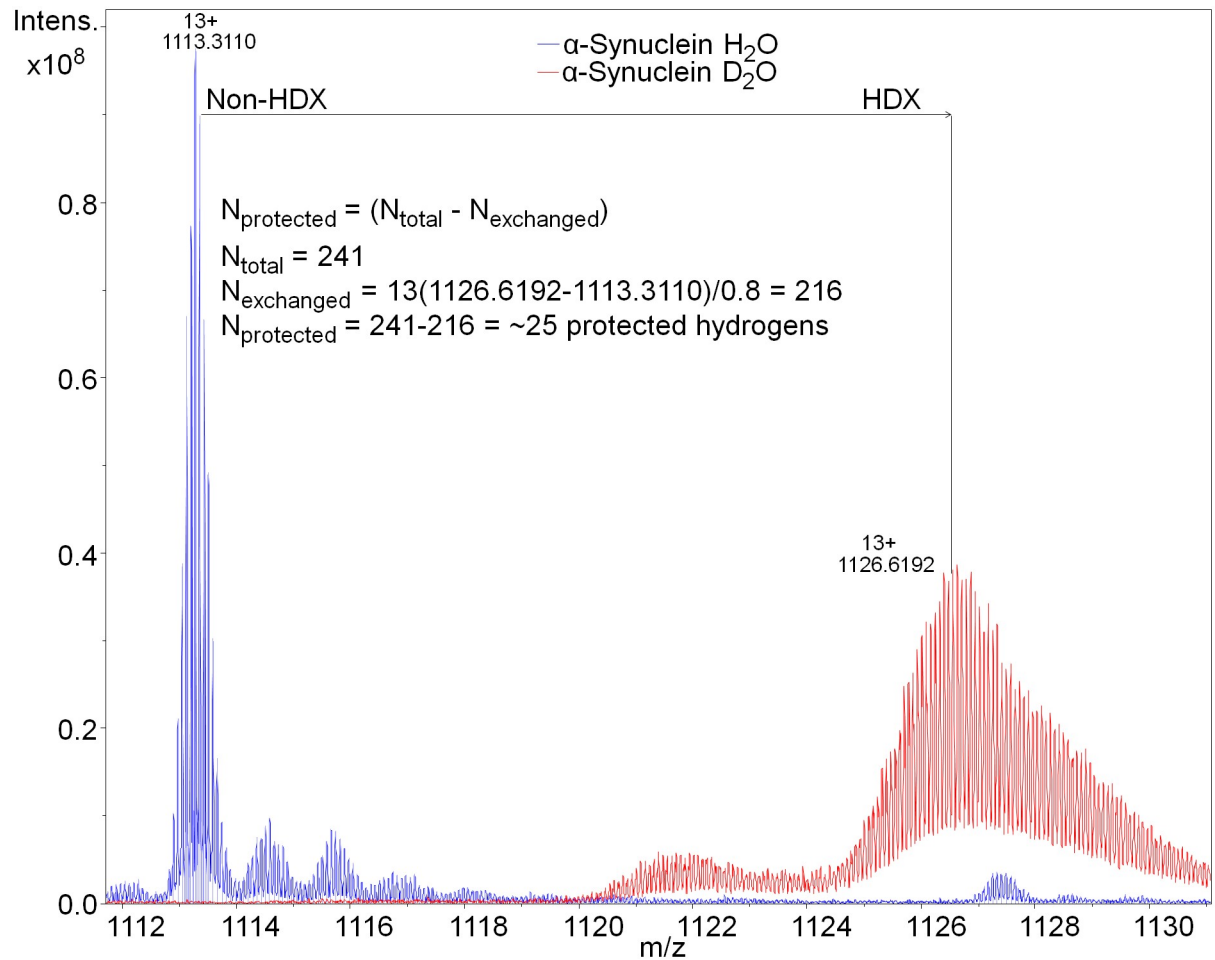
Fig 1. Hydrogen-deuterium exchange of α-synuclein. The total number of protected hydrogens was measured to be approximately 25 at 2s of exchange, indicating the existence of transient secondary structure.

chromatography on a Mono Q 4.6/100 SAX column (GE Life Science), using a gradient from 50–500 mM NaCl, in 50 mM Tris at pH 8.0. Elution fractions containing α-synuclein were further purified by size exclusion on a Superdex 200 30/100 GL column (GE Life Science). For the expression of metabolically labeled ^{15}N α-synuclein, 1L of M9 Minimal media was prepared with 1 g/L $^{15}NH_4Cl$ (Cambridge Isotopes) as the sole source of nitrogen. BL21(DE3) cells were grown overnight in 50mL of this media, then seeded into 1 L, grown to an A_{600} of approximately 0.8, and induced using 1 mM IPTG. After expression overnight at 30°C, ^{15}N α-synuclein was purified as described above.

Crosslinking

Unlabeled and ^{15}N metabolically-labeled α-synuclein were mixed in a 1:1 ratio at a concentration of 20 μM in 50 mM Na_2HPO_4 and incubated overnight at room temperature prior to crosslinking. α-synuclein aliquots of 38 μL were then crosslinked using either 1 mM of the ABAS-$^{12}C_6/^{13}C_6$ crosslinker (Creative Molecules) or 30 mM of the EDC crosslinker. ABAS crosslinking reaction mixtures were incubated for 10 minutes in the dark to allow the NHS-ester reaction to take place, followed by 10 minutes of UV irradiation under a 25 W UV lamp

(Model UVGL-58 Mineralight lamp, UVG) with a 254 nm wavelength filter. ABAS reaction mixtures were quenched with 10 mM ammonia bicarbonate. EDC reaction mixtures were incubated for 20 minutes. A portion of each crosslinking reaction mixture was checked by SDS-PAGE gel to see the extent of potential intermolecular crosslinked products. Aliquots were subsequently split and digested with either trypsin or proteinase K at an enzyme: protein ratio of 1:10. Digestion was quenched using a final concentration of 10 mM AEBSF (ApexBio), and samples were then acidified with formic acid for analysis by mass spectrometry. For TATA, 100 μM synuclein in 50 mM sodium phosphate buffer was reacted with 0.5 mM TATA-$^{12}C_3$/$^{13}C_3$ (Creative Molecules). Samples were incubated for 5 minutes with 254 nm UV light from the same lamp as was used for the ABAS reactions. Samples were then split and digested with either proteinase K or trypsin at an enzyme: protein ratio of 1:20. For SDA reactions, 20 μL of 1mg/mL α-synuclein was crosslinked using 1 mM SDA (Creative Molecules, Inc.). Aliquots were incubated for 15 minutes in the dark prior to incubation under the same UV lamp as used previously for ABAS reactions but changing the wavelength to 366 nm. Samples were then run on an SDS-PAGE gel, and bands representing the α-synuclein monomer were excised and subjected to in-gel trypsin digestion. After in-gel digestion, samples were acidified using formic acid prior to mass spectrometric analysis. The CBDPS crosslinking reaction mixture consisted of 238 μL of 50 μM α-synuclein, with 0.12 mM CBDPS. Samples were split and digested with either proteinase K or trypsin at an enzyme: protein ratio of 1:10. Digests were quenched with 10 mM AEBSF and samples were enriched using monomeric avidin beads (Thermo Scientific). Enriched samples were acidified for mass spectrometric analysis using formic acid.

LC-MS/MS analysis

Mass spectrometric analysis was then performed using a nano-HPLC system (Easy-nLC II, ThermoFisher Scientific), coupled to the ESI-source of an LTQ Orbitrap Velos or Fusion (ThermoFisher Scientific), using conditions described previously [11]. Briefly, samples were injected onto a 100 μm ID, 360 μm OD trap column packed with Magic C18AQ (Bruker-Michrom, Auburn, CA), 100 Å, 5 μm pore size (prepared in-house) and desalted by washing with Solvent A (2% acetonitrile:98% water, both 0.1% formic acid (FA)). Peptides were separated with a 60-min gradient (0–60 min: 4–40% solvent B (90% acetonitrile, 10% water, 0.1% FA), 60–62 min: 40–80% B, 62–70 min: 80% B), on a 75 μm ID, 360 μm OD analytical column packed with Magic C18AQ 100 Å, 5 μm pore size (prepared in-house), with IntegraFrit (New Objective Inc., Woburn, MA) and equilibrated with solvent A. MS data were acquired using a data-dependent method. The data dependent acquisition also utilized dynamic exclusion, with an exclusion window of 10 ppm and exclusion duration of 60 seconds. MS and MS/MS events used 60000- and 30000-resolution FTMS scans, respectively, with a scan range of m/z 400–2000 in the MS scan. For MS/MS, the CID collision energy was set to 35%. Data were analyzed using the $^{14}N^{15}N$ DXMSMS Match program from the ICC-CLASS software package [27]. SDA crosslinking data was analyzed using Kojak [28] and DXMSMS Match. For scoring and assignment of the MS/MS spectra, b- and y-ions were primarily used, with additional confirmation from CID-cleavage of the crosslinker where this was available.

Differential surface modification

Chemical surface modification with pyridine carboxylic acid N-hydroxysuccinimide ester (PCAS) (Creative Molecules) was performed as described previously [29]. Briefly, α-synuclein was prepared at 50 μM in 8 M urea in PBS, pH 7.4 (unfolded state), or in only PBS (folded state). Either the light or the heavy form of the ^{13}C-isotopically-coded reagent (PCAS-$^{12}C_6$ or

PCAS-$^{13}C_6$) was then added to give a final concentration of 10 mM. Reaction mixtures were incubated for 30 minutes and quenched with 50 mM ammonium bicarbonate. Samples were then mixed at a 1:1 ratio, combining folded (PCAS-^{12}C) with unfolded (PCAS-^{13}C) samples, as well as in reverse as a control. Samples were acidified with 150 mM acetic acid and digested with pepsin at a 20:1 protein: enzyme ratio overnight at 37°C. After digestion samples were prepared for mass spectrometry analysis using C_{18} zip-tips (Millipore). Zip-tips were equilibrated with 30 μL 0.1% TFA, sample was introduced, then washed with 30 μL 0.1% TFA and eluted with 2 μL of 0.1% formic acid/50% acetonitrile. Samples were analyzed by LC-MS/MS as described above.

Hydrogen/deuterium exchange

Top-down ECD-FTMS hydrogen/deuterium exchange was performed as described previously [30]. Briefly, protein solution and D_2O from separate syringes were continuously mixed in a 1:4 ratio (80% D_2O final) via a three-way tee which was connected to a 100 μm x 5 cm capillary, providing a labeling time of 2 s. The outflow from this capillary was mixed with a quenching solution containing 0.4% formic acid in 80% D_2O from the third syringe via a second three-way tee and injected into a Bruker 12 T Apex-Qe hybrid Fourier Transform mass spectrometer, equipped with an Apollo II electrospray source. In-cell ECD fragmentation experiments were performed using a cathode filament current of 1.2 A and a grid potential of 12 V. Approximately 800 scans were accumulated over the m/z range 200–2000, corresponding to an acquisition time of approximately 20 minutes for each ECD spectrum. Deuteration levels of the amino acid residues were determined using the HDX Match program [31] (S1 Fig).

Synuclein UVPD spectra were collected on a Thermo Scientific Orbitrap Fusion Lumos Tribrid mass spectrometer equipped with a 2.5-kHz repetition rate (0.4 ms/pulse) 213 nm Nd:YAG (neodymium-doped yttrium aluminum garnet) laser (CryLas GmbH) with pulse energy of 1.5 ± 0.2 μJ/pulse and output power of 3.75 ± 0.5 mW for UVPD. The solution was exchanged with deuterium using the same three-way tee setup, although in this case a 50 μm x 7cm capillary provided a labeling time of ~1s. Spectra were acquired for 8 or 12 ms, and resultant spectra were averaged and used for the data analysis with the HDX Match program as above.

Circular dichroism

CD spectra were recorded on Jasco J-715 spectrometer under a stream of nitrogen. The content of α-helical and β-sheet structures was calculated using BeStSel web server [32].

Discrete molecular dynamics modelling

Crosslink guided discrete molecular dynamics (CL-DMD) simulations were performed according to the protocol described in our previous work [11]. Briefly, discrete molecular dynamics (DMD) is a physically based and computationally efficient approach for molecular dynamics simulations of biological systems [16, 17]. In DMD, continuous inter-atom interaction potentials are replaced with their discretized analogs, allowing the representation of interactions in the system as a series of collision events where atoms instantaneously exchange their momenta according to conservation laws. This approach significantly optimizes computations by replacing integration of the motion equations at fixed time steps with the solution of conservation-law equations at event-based time points [33]. In order to incorporate experimental data for inter-residue distances between corresponding atoms into DMD simulations, we introduced a series of well-shape potentials that energetically penalize atoms whose interatomic distance do not satisfy experimentally determined inter-atom proximity constraints.

The widths of these potentials are determined by the cross-linker spacer length and side chain flexibility [11]. Starting from the completely unfolded structure of α-synuclein molecule, we performed an all-atom Replica Exchange (REX) [34] simulations of the protein where 24 replicas with temperatures equally distributed in the range from 0.375 to 0.605 kcal/(mol k_B), are run for 6 x 10^6 DMD time steps (S2 Fig). The simulation temperature of each of the replicates periodically exchanged according to the Metropolis algorithm allowing the protein to overcome local energetic barriers and increase conformational sampling. During the simulations we monitored the convergence of the system energy distribution specific heat curve, calculated by Weighted Histogram Analysis Method (WHAM) [35] which was used as the indicator of system equilibration. We discarded the first 2 x 10^{6-time} steps of system equilibration during the analysis. Next, we ranked all of the structures among all of the trajectories, and selected the ones with lowest 10% of the energies, as determined by the DMD Medusa force field [36]. These structures were then clustered using the GROMACS [37] distance- based algorithm described by Daura et al. [38]. It uses root-mean-square deviation (RMSD) between backbone Cα atoms as a measure of structural similarities between the cluster representatives. A RMSD cut-off was chosen to correspond to the peak of the distribution of pair-wise RMSDs for all of the low-energy structures. Because the energies of the resulting centroids representative of the clusters are very close to each other (S3 Fig) and picking one of them would potentially introduce a bias related to our scoring energy function, we presented them all as our predicted models of the α-synuclein globular structure. We then calculated the root-mean-square deviation of atomic positions within each cluster and used this as a measure of fluctuations of the structures of corresponding centroids (Figs 2 and 3). In order to obtain information on the global folding of α-synuclein, we performed clustering analysis on the lowest-energy structures obtained during CL-DMD simulations.

Results and discussion

In contrast to ordered proteins, where the lowest energy structures are usually represented by a few centroids with distinct conformations and folds close to the corresponding native structure [11], disordered proteins are usually represented by a broader variety of structures, which are distinct from each other, reflecting the conformational freedom of the intrinsically disordered protein [39]. In our case, the lowest energy structures forming the conformational ensemble for α-synuclein are shown as overlays in Fig 4A. They can be described as rather compact globular structures with a common general topology, containing a few distinct secondary structure fragments. Overall, a distinct topology can be observed for all these conformers (Fig 4B). The protein forms a three-prong closed claw-like shape with the N-terminal (blue), intermediate (yellow), and C-terminal (red) subdomains having converged at the top, with the connecting subdomain (green) being located at the bottom of the structure (Fig 4). The C-terminal portion of the molecule was found in close proximity to the N-terminal portion, possibly reflecting long-range electrostatic interactions between these subdomains (Fig 4) [8]. The C-terminal portion was also in contact with the intermediate subdomain—in some conformers it protruded deeper into the structure and was positioned between the N-terminal and the intermediate subdomains. It should be noted that non-N-acetylated α-synuclein protein was used in this study, although in the cell, α-synuclein is predominantly N-acetylated [40], and this post-translational modification may have some effect on the protein structure by increasing the propensity of the N-terminus to form an α-helix [41], and that this may affect the aggregation of synuclein under some conditions [42, 43]. N-terminal acetylation, however, does not appear to have a significant effect on the propensity of synuclein to aggregate [44] under lipid-free conditions, and it therefore seems that the structural features important for
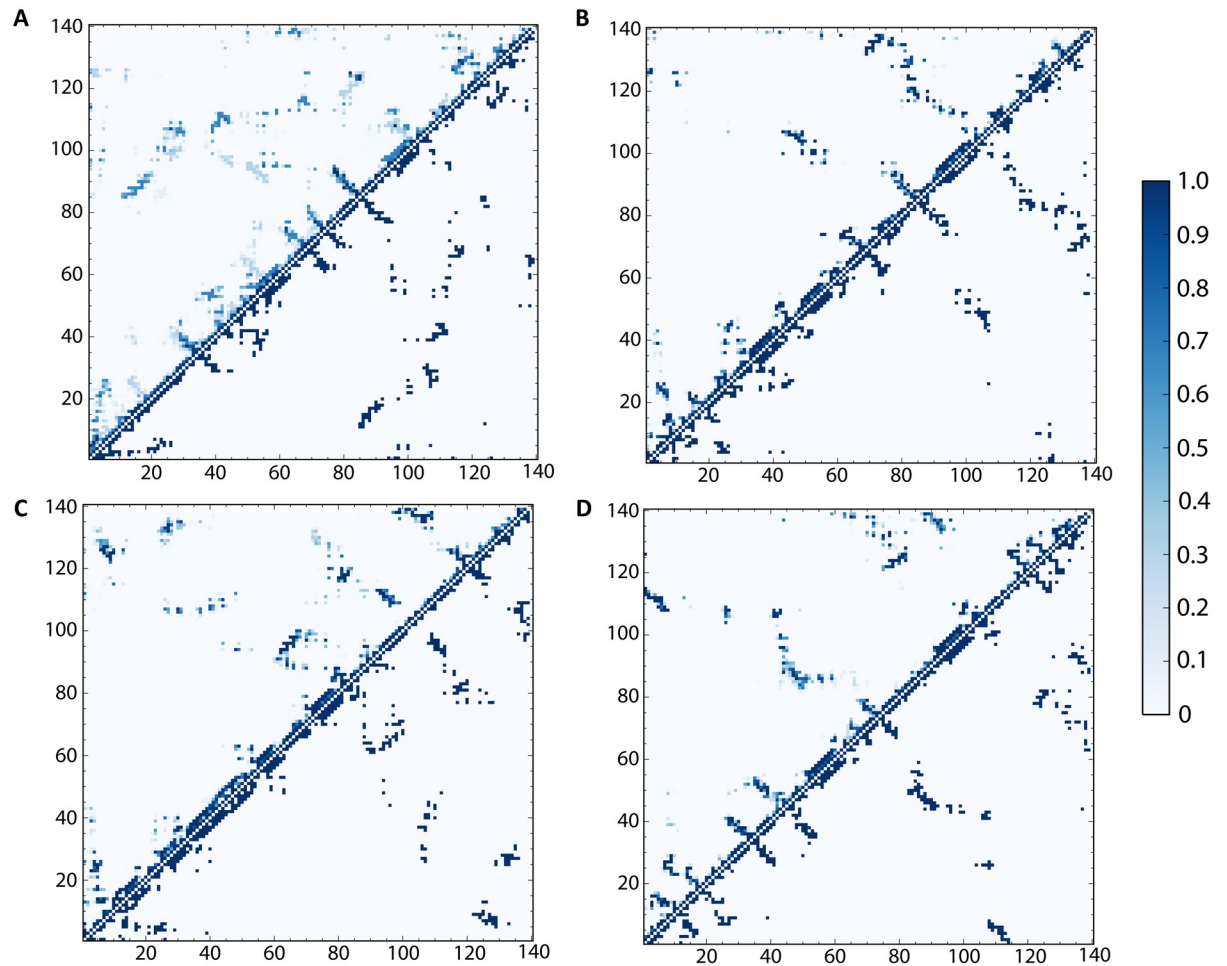
Fig 2. Contact frequency maps for representative clusters of α-synuclein model. Plots (A-D) represent static and frequency contact maps for the major representative clusters of protein structures during the simulations (populations: cluster 1 ~37%, cluster 2 ~28%, cluster 3 ~20%, cluster 4 ~8%). Points below the diagonals correspond to contacts between residues. Note that 0 means that the atoms are not in contact; 1 means that the two atoms are in contact. Two residues form a contact if their Cα atoms are within 8 Å of each other. Data points above the diagonal in plots (A-D) indicate how often each contact between two residues can be found within the clusters for which corresponding structure from bottom diagonals are centroids. The colour map quantifies how often these contact form, where white/zero means that contacts never form, and dark blue/one means that contacts always form).

predicting aggregation and designing small molecules to affect misfolding are also present in the absence of acetylation.

Some of the experimental constraints may conflict with each other due to the fact that proteins exist in multiple conformations. During DMD simulations, the total potential is minimized, satisfying an optimal number of constraints and generating an ensemble of structures satisfying subsets of consistent experimental constraints. This ensemble is further clustered into substates that represent subsets of non-conflicting constraints. Depending on the conformation and the number of constraints that are satisfied, these structures will have different energy scores assigned by the force field [16, 36]. This allows us to account for both the nature of conformational changes of the protein in solution (i.e., while the structure of the protein is "breathing" and transitioning between conformations), and possible errors during the determination of the experimental constraints [12]. Nevertheless, we found that the lowest-energy conformers satisfy most of the experimental distance constraints (S4 Fig), while the models
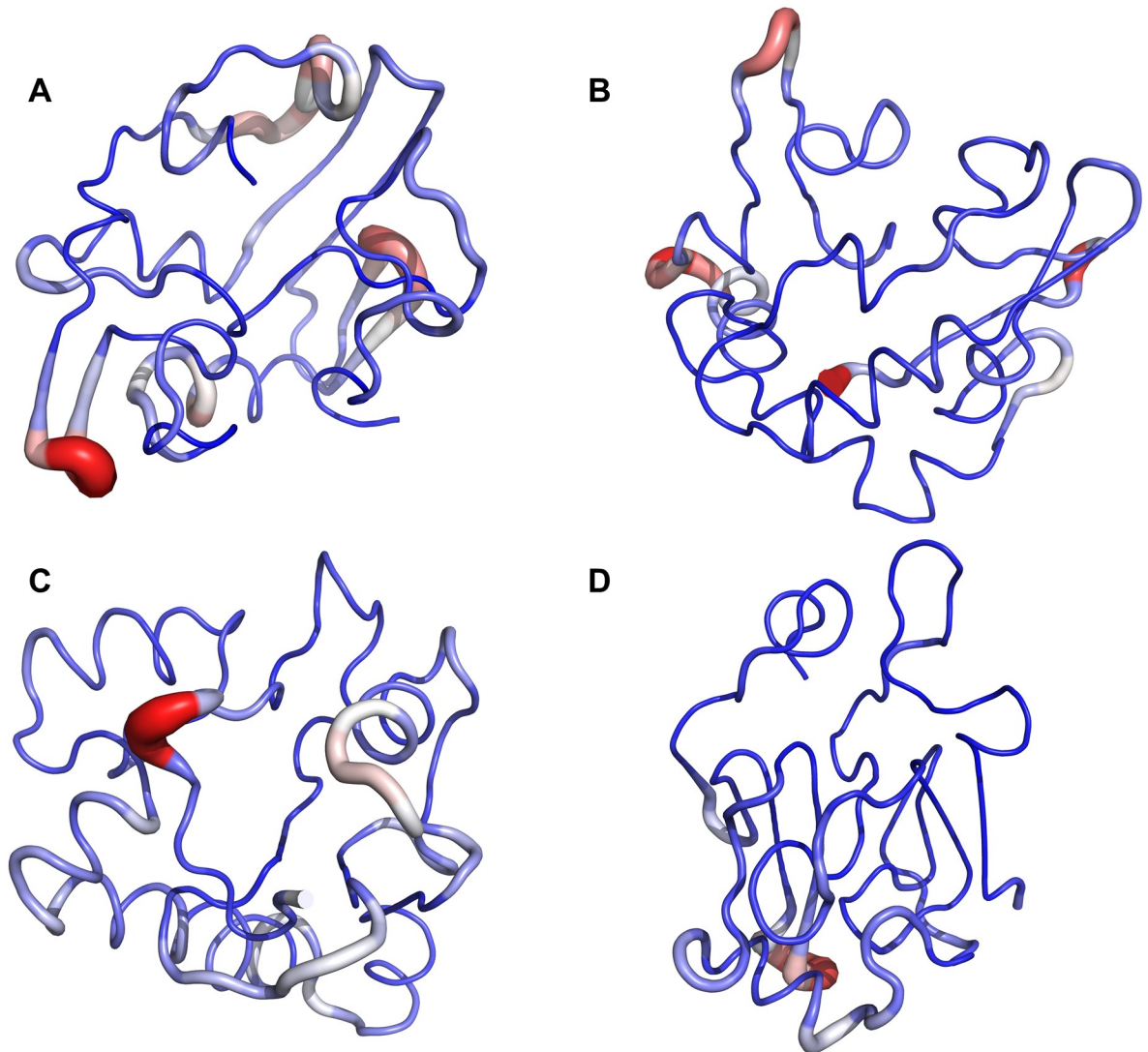
Fig 3. Tube representation of the fluctuations of the clusters. Tube models A-D represent the 37%, the 28%, the 20%, and the 8% clusters, respectively. The thickness and colour of the tubes indicate the dynamics of the corresponding regions within each cluster during the simulations. The tubes are coloured from blue (low flexibility) to red (high flexibility).

with higher energy satisfy a lower number of constraints. No quantitative crosslinking data or weighting of observed crosslinks were used in the simulations, but more-populated conformations have a higher chance of being represented in the resulting crosslinking dataset, given that crosslinks are not randomly distributed and are most likely to form between residues which spend a significant amount of time within close contact. A comparison of predicted structures within the ensemble revealed that the conformational changes to the protein structure were mainly due to movement of the large loops and hairpins constituting the N-, C-terminal, and intermediate subdomains.

The validity of this ensemble was tested by removing all of the crosslinking constraints and allowing the structure to relax over 2×10^6 steps, to see if the structure was stabilized only by the potentials enforcing satisfaction of the experimental crosslinks. The predicted representatives of the clusters were allowed to fluctuate in this way, and the result was a relatively
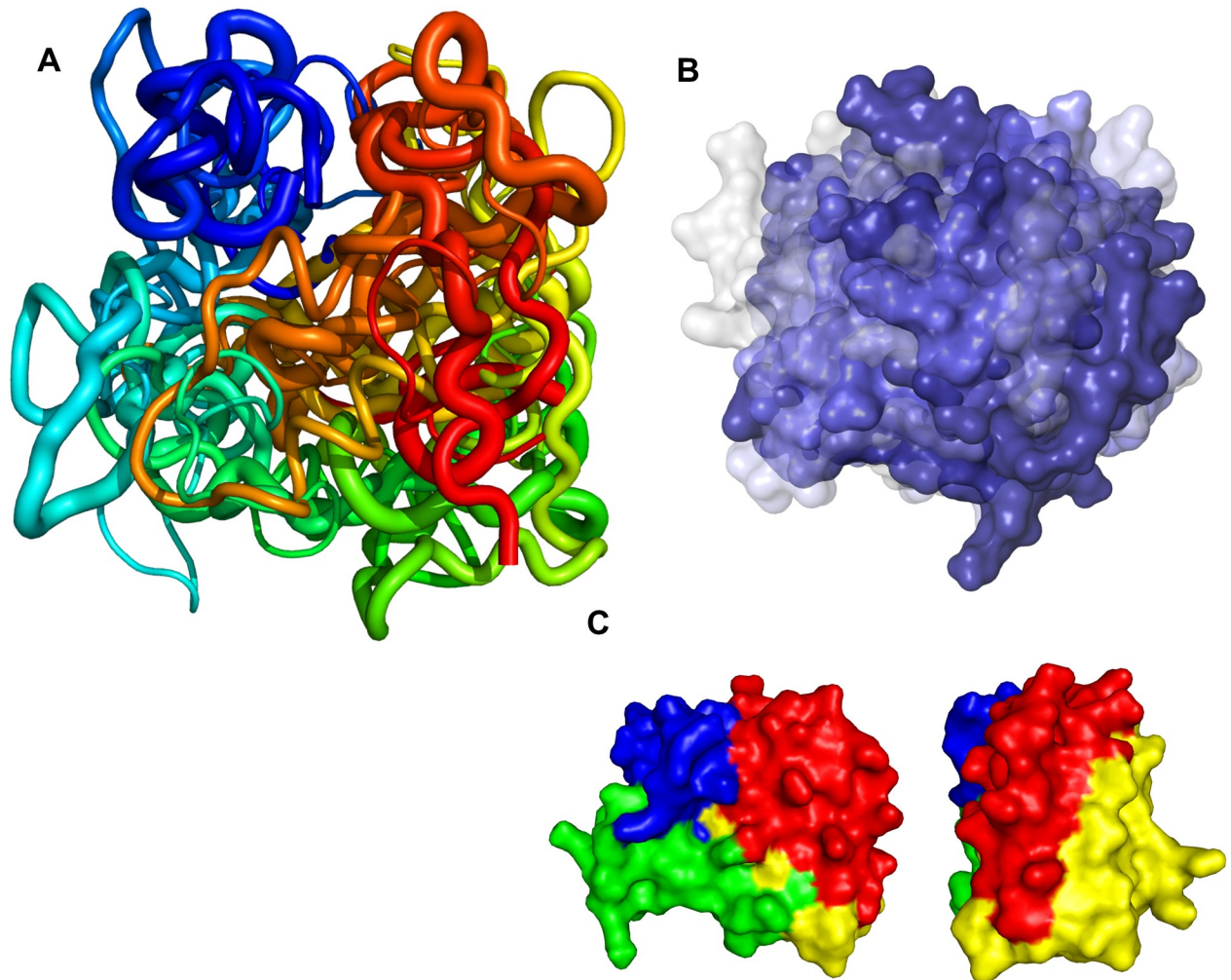
Fig 4. Structure of native α-synuclein in solution as determined by CL-DMD. A. Conformational ensemble of α-synuclein. Representatives of the four major clusters are aligned and coloured N-to-C from blue-to-red. The thickness of the cartoon representations corresponds to the content of approximately 37%, 28%, 20% and 8% of the overall population of the structures in each cluster. **B.** Surface representation of the ensemble. Representatives of the four major clusters as in panel A are aligned and coloured blue-to-white according to percent of all structures in each cluster. **C.** Subdomain topology of the α-synuclein conformational ensemble. Residues are coloured 1–27 blue, 28–52 green, 53–108 yellow, 109–140 red.

minor expansion of the protein (results for the centroid of the most representative lowest energy cluster, Fig 3A, are shown for clarity), from a radius of gyration of 14.1 Å initially, to an average of 15 Å (S5 Fig). This indicates that the structure has not been overfit to the experimental data during the simulation, which can be a problem when modelling intrinsically disordered proteins with experimental distance constraints [14].

The crosslinking analysis applied here is conceptually similar to the PRE-NMR [8, 40, 45, 46] or FRET [47, 48] techniques used for the prediction of α-synuclein conformational ensembles, with pairwise distances between crosslinked residues being analogous to the pairwise distances between the spin-label and the atomic nuclei of the protein in case of PRE-NMR (S6 Fig) or between fluorophores, in the case of FRET. The lowest energy-cluster centroids in our ensemble tend to be more compact than those based on PRE-NMR (S6 and S7 Figs), FRET ensembles, or SAXS data. It has been suggested that the higher values of the radius of gyration determined by these techniques may be caused by an existing equilibrium between the

monomeric and multimeric states of the synuclein protein under the experimental conditions used [49, 50]. SAXS data can be uniquely sensitive to the more-extended conformations of the ensemble [49]. Unconstrained all-atom molecular dynamics simulations produced multimodal distributions which included compact states [49], although it was noted that modelling disordered proteins may require the development of specialized force fields [43, 49]. It was also shown that the addition of experimental distance constraints to the all-atom Monte Carlo conformational search leads to more compact structures [49]. Thus, representation of the slightly more compact states may still be caused by bias due to the short-range crosslinking reagents used in this study, as well as by the traditional parameterization of the Medusa force field [15–17].

In addition, we selected the lowest-energy structures for our clustering and model selection, which eliminated some of the unfolded less-compact states, while in the PRE-NMR based approach, the authors used all of the generated structures for the analysis. We are confident in this approach as these lower energy conformations were also compatible with the long-distance crosslinks. The other difference between these two techniques is the range of distance constraints produced. The PRE-NMR technique uses distances up to ~2 nm. Local residue-specific conformational preferences are not well reproduced in the PRE-RMD ensembles [46]. When applying CL-DMD for well-structured model proteins, i.e., proteins with known crystal structure, we also found that long-distance constraints (>15 A) were not particularly helpful in finding the true protein structure solutions [11]. Short-distance constraints were critical in finding the correct local conformations in well-structured proteins and might possibly be helpful in producing more detailed structures in the case of the CL-DMD ensembles described here.

Some transient secondary-structure elements were observed in the conformers of the ensemble (Fig 5). The extent and the location of the predicted secondary-structure motifs in the lowest-energy conformer were in good agreement with the experimental data obtained by HDX and CD. Thus, ~25 of the total number of protected protons were detected by HDX (Fig 1) of the whole protein, and ~2.4% of α- and 29.1% of β-structure was detected by CD. We were able to detect amide HDX protection only when the exchange time was reduced from ten seconds to two seconds, which indicates the transient nature of the secondary structure observed. The location of the predicted secondary structure within the N-terminal portion of the protein was confirmed by determining the protection status of individual residues using the c-ion fragment series (i.e., the fragments starting from N-terminus) as obtained by ECD MS/MS (Fig 1 and S1 Fig). Here, we observed some propensity for the protein to form α-helical secondary structure near the N-terminus, particularly between residues 25–55. However, it is not nearly as extensive as the secondary structure formed in the presence of detergent micelles. In particular, two centroids (Fig 5B and 5C) show a greater tendency towards α-helicity, and together represent some 48% of the ensemble.

Unfortunately, for this protein, using ECD or ETD fragmentation we were unable to obtain reliable z-ion fragment series (i.e., fragments starting from the C-terminus) which would have spanned the predicted β-structural elements, probably due to the high net negative charge on the C-terminal portion of the molecule. However, the amount of protection predicted in the structure of the C-terminal part of the model, added to observed protection of the N-terminal part from ECD-HDX data, is in agreement with the total protection value of the protein (Figs 1 and 3, S8 Fig).

We then used our recently developed top-down UVPD-HDX method [51] in an attempt to determine the possible locations of secondary structure in the C-terminal portion of the α-synuclein molecule. We were able to observe a number of UVPD-specific fragment ions spanning the C-terminal portion of the protein sequence—a region which had previously not been covered. Deuteration analysis of these fragments allowed us to confirm the presence of
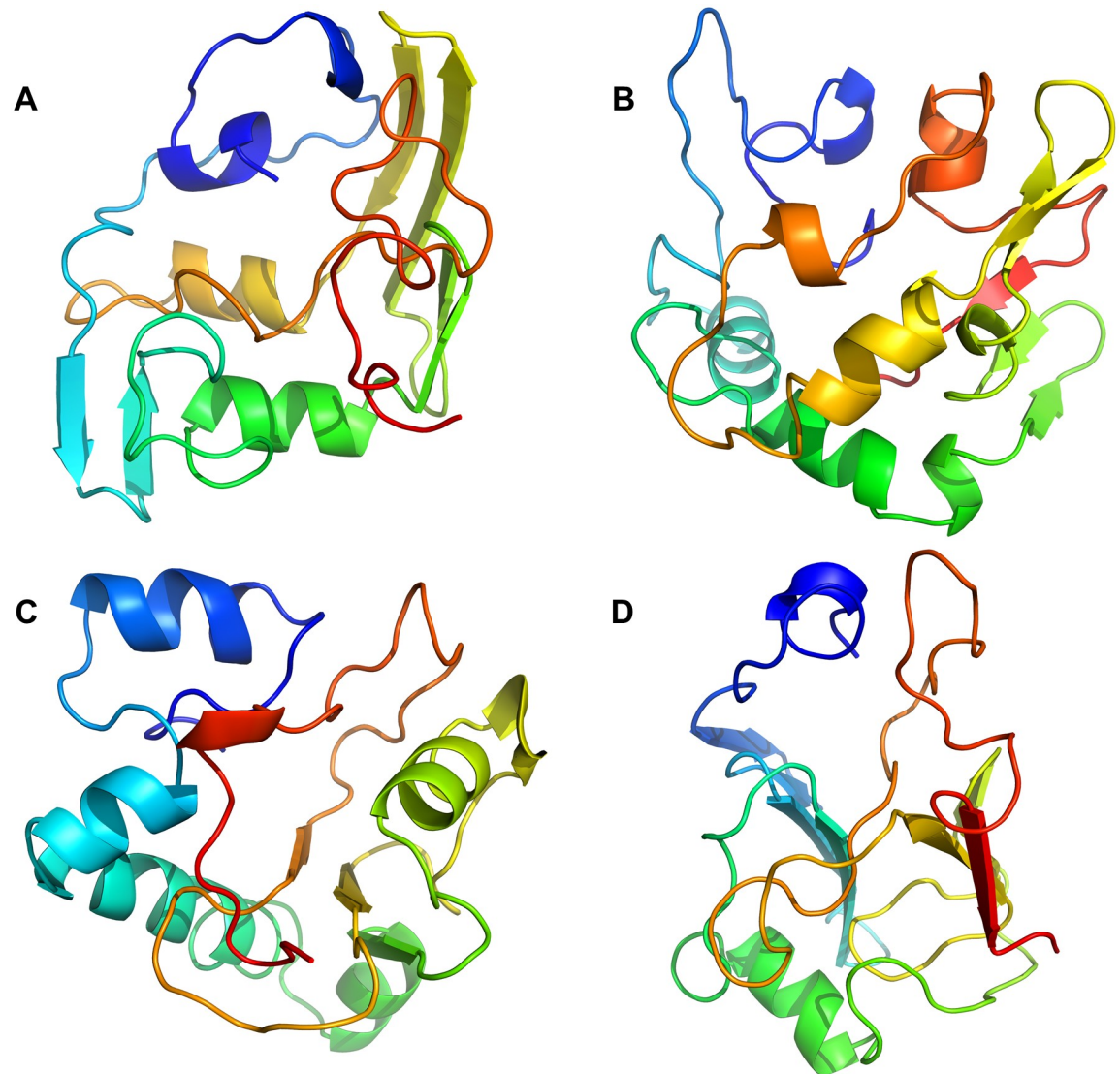
Fig 5. Comparison of the transient secondary structure in the α-synuclein conformational ensemble. Shown here are representative conformers of the α-synuclein ensemble shown in Fig 4. Structures A, B, C, and D represent clusters containing approximately 37%, 28%, 20%, and 8% of the overall population of the structures, respectively. Structures are coloured from blue to red, from the N- to C-terminus; the NAC region 61–95 corresponds to the light green-to-yellow section.

https://doi.org/10.1371/journal.pcbi.1006859.g005

secondary structure in this region and to further narrow the location of the secondary-structure elements to residues 116–119 and 128–138, and ~10 protected residues in the residue 71–116 segment of the sequence (Figs 1 and 3, S1 and S8 Figs). Surface modification results were in agreement with the final α-synuclein structures. As expected, lysine residues which ended up being exposed to the solvent in the model were equally modified with PCAS in the unfolded and folded states (i.e., with and without 8 M urea), and residues which were partially buried in the model exhibited a larger extent of modification in the fully unfolded state with 8 M urea (Fig 6). The long-distance intra-protein CBDPS crosslinks were also in good agreement with the final models of the protein.

Interestingly, the aggregation-prone non-amyloid-β component (NAC) region, which is reportedly involved in the nucleation of the aggregation [52], was consistently predicted to
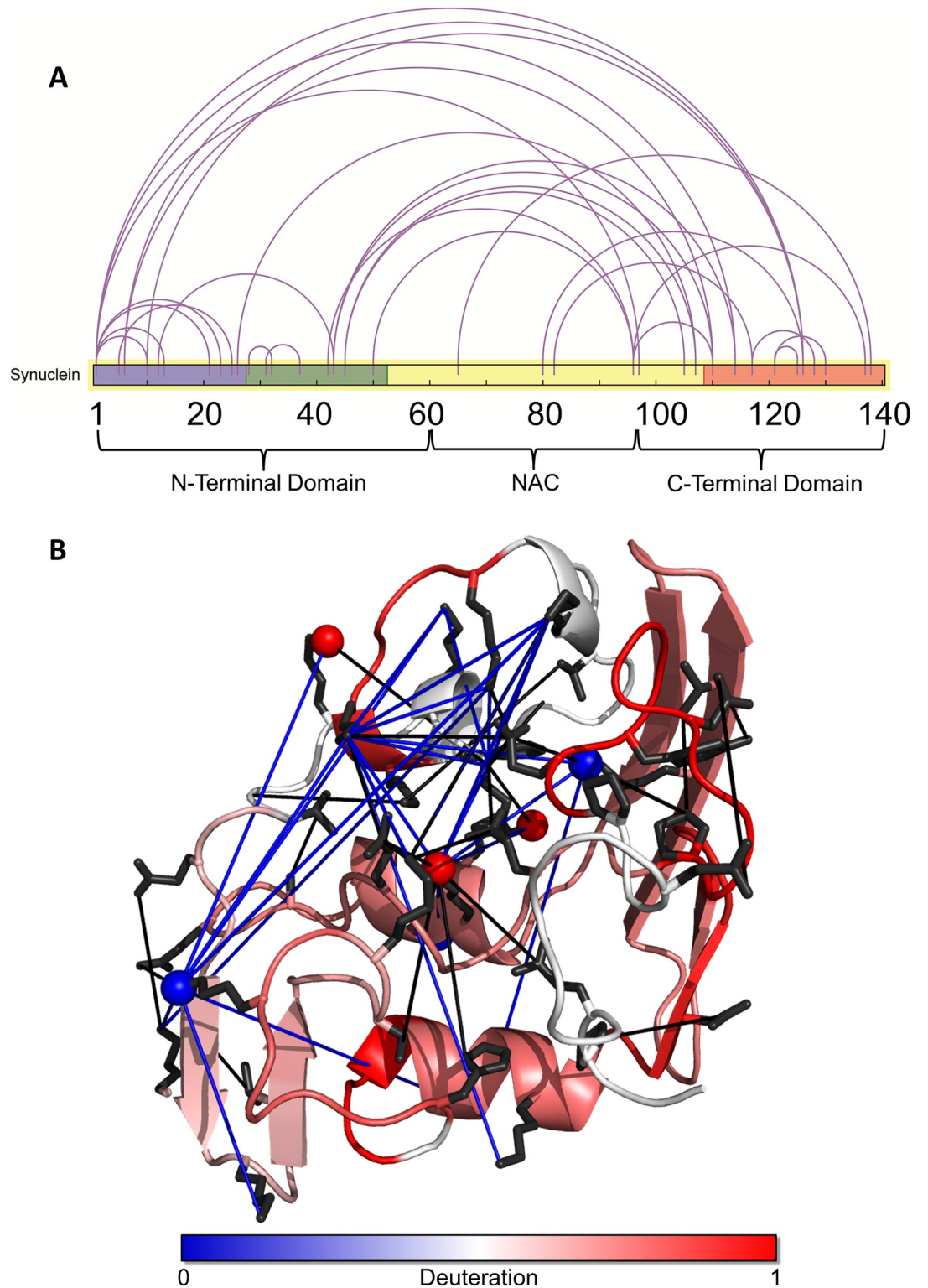
Fig 6. Experimental validation of the α-synuclein structure with SM, HDX, and LD-CL. A. A schematic diagram of the synuclein protein which each crosslink used for the CL-DMD simulation shown as a line. The diagram is coloured in accordance with the subdomain topology described in Fig 4. **B.** The lowest energy conformer from the ensemble as in Fig 1 is shown, with crosslinks and other validating data indicated on the structure. Short-distance crosslinks used in the CL-DMD simulations are

shown as black lines. The main chain is coloured from blue to red according to the experimentally observed backbone amide protection values from the HDX experiments. Atoms that are equally modified with PCAS in the native and denatured states are shown in blue; atoms that are preferentially modified after denaturation with 8M urea atoms are shown in red. LD-CL CBDPS crosslinks are shown as blue lines.

contain hairpins in extended conformation, which were in contact with the C-terminal loop, which, in turn, interacted with the N-terminal sub-domain. It would be intriguing to see if a distortion of any of these interactions could lead to an "opening up" of this region which might facilitate stacking interactions with similar portions of other monomers during the aggregation initiation. We are currently performing CL-DMD-based differential structural characterization of α-synuclein oligomers, in comparison to the native structures reported here, in order to localize critical mis-folding events leading to the pathological aggregation. Two recent studies, by Lautenschlager *et al.* [4] and Fusco et al [5] have indicated that cellular toxicity of synuclein oligomers may be dependant on both the N- and C-terminal domains and their interactions with lipid bilayers. In toxic oligomeric species, the synuclein N-terminus has less interaction with the membrane surface, and this interaction is replaced by an interaction with the NAC region, leading to toxicity. In the case of the C-terminus, increased concentrations of calcium may create a charge-neutralizing effect, which leads to increased toxicity and oligomer formation, possibly by disrupting its interaction with the NAC region.

The β-structure in the NAC-containing intermediate domain and in the C-terminal subdomain of the predicted ensemble is in general agreement with the C-terminal half of the protein's propensity towards formation of β-structure, as determined by NMR [7]. The fact that, in the ensemble, the NAC region was repeatedly found to contain β-structure-like hairpins may explain the propensity of this region to form inter-molecular β-structures which appear early in the aggregation process. These hairpins also resemble those observed in the aggregating form of Aβ1–40 found in fibrils [53], and in the mature fibrillar form of the synuclein protein [54]. This increases our confidence in the significance of this particular transient structure in the native protein. In agreement with previous studies [8, 45], we also found that the NAC region was in contact with C- terminal portion of the molecule, but in our ensemble it is not quite shielded by it, as was suggested. Rather, it is stabilized by β-sheet-like contacts. Additionally, the residues V74-V82 within the NAC region, which has been shown to be critical for incorporation into newly forming fibrils (Ricardo *et al.*, 2018), remains somewhat exposed in the most favorable conformations of the ensemble, and with this region already incorporating important β-sheet contacts, it seems primed for incorporation into newly-forming fibrils, which may help explain the tendency for native synuclein to spontaneously form fibrils with mere agitation.

The structures of the major α-synuclein conformational clusters determined here and the presence of the secondary structure detected in the NAC region and C-terminal portions of the molecule allowed us to hypothesize a possible mechanism for the misfolding conformational change and the early aggregation events. The existence of the transient hairpin/beta-structure in the NAC and C-terminal subdomains predisposes the protein to form the beta-nucleation sites of the early oligomers, which may further mature to the cross-beta-structure of the fibrils. The mis-folding conformational change that is necessary for the formation of inter-molecular contacts would then be the detachment of the NAC-containing hairpin (possibly in concert with stabilizing C-terminal beta-strands) from the core of the molecule. This would provide a stacking template for the similarly detached NAC hairpin of the other α-
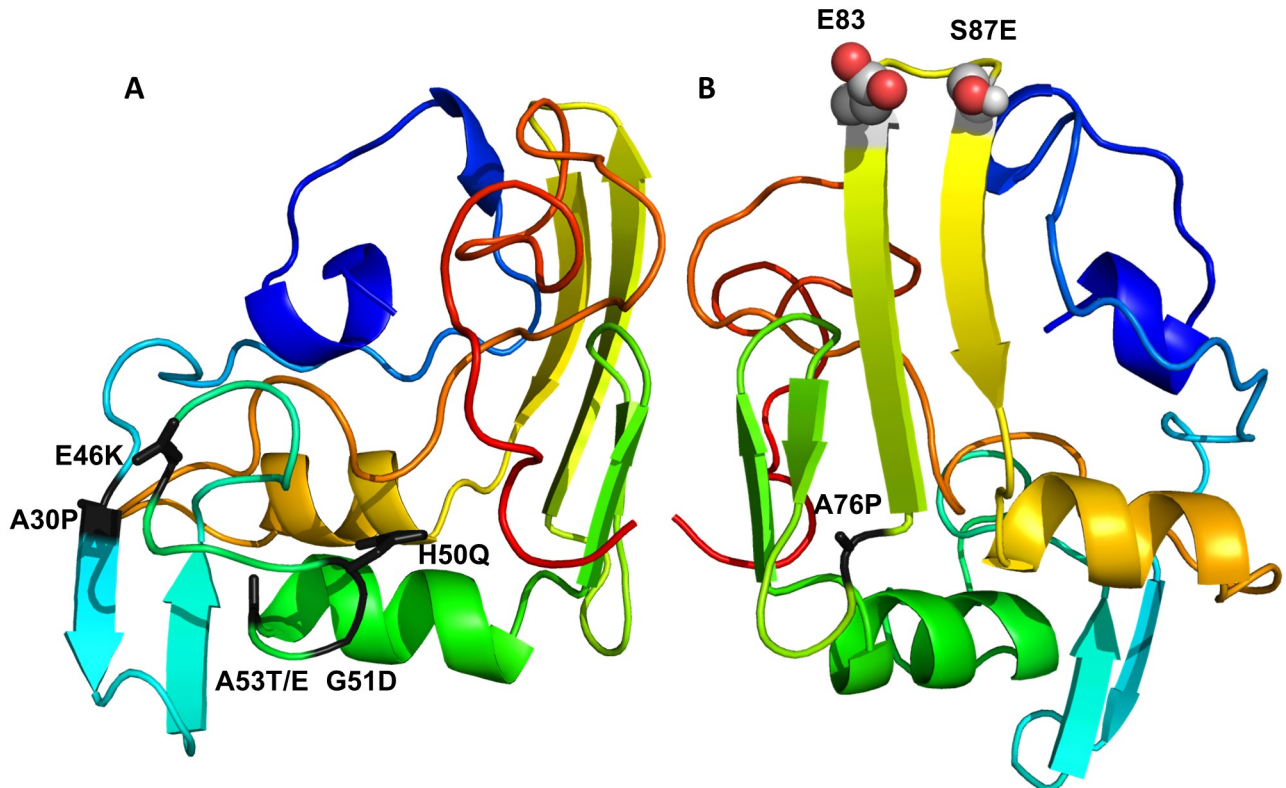
Fig 7. Location of the mutations influencing aggregation of α-synuclein. A. Familial mutations located in the connecting subdomain are highlighted in black in the centroid structure of the cluster 1. B. Mutations potentially influencing the NAC β-hairpin.

synuclein molecule. This hypothesis can be tested by determining the structure of the early oligomers, which is currently underway in our laboratories.

The locations of the mutations shown to be influencing the aggregation of the α-synuclein within the proposed structure is interesting (Fig 7) [55]. The familial mutations A30P, E46K, H50Q, G51D and A53 T/E [56] were found to located on the same surface of the connecting subdomain. Although the effects of the these mutations could possibly be produced via different mechanisms (such as modified membrane binding or fibril stability), at least the effect of the oligomer-promoting mutations A30P and A56P can be explained in terms of the structure described here, where they would increase the segmental flexibility and possibly relax the location and interaction of the N-terminal subdomain with the core of the molecule. A76P oligomer-promoting mutation similarly would facilitate segmental flexibility and the detachment of the NAC-containing β-hairpin postulated above. The S87E mutation, which blocks α-synuclein oligomerization and fibrilogenesis [57], would possibly disrupt NAC β-hairpin via electrostatic repulsion with the E83 residue.

Conclusions

In summary, we have determined *de novo* the conformational ensemble of native α-synuclein in solution by short-distance crosslinking constraint-guided DMD simulations, and validated this structure with experimental data from CD, HDX, SM, and LD-CL experiments. The predicted conformational ensemble is represented by rather compact globular conformations

with transient secondary structure elements. The obtained structure can serve as a starting point for understanding the mis-folding and oligomerization of α-synuclein.

Supporting information

S1 Fig. Hydrogen-deuterium exchange of α-synuclein. A. The number of protected hydrogen atoms in c-ions detected by ECD-FTICR mass spectrometry. The upper blue line represents the theoretical situation of perfect protection of all backbone hydrogen atoms. The lower line represents the number of actual protected hydrogen atoms detected at that fragment. **B.** The deuteration status of backbone amides from ECD data. A value of 1 indicates that the residue is unprotected; a value of zero represents complete protection from exchange with deuterium. **C.** The number of protected hydrogen atoms in fragment ions detected by UVPD-FT mass spectrometry. **D.** The deuteration status of backbone amides from UVPD data. N-terminal and C-terminal fragment data are shown in blue and red, respectively.
(TIF)

S2 Fig. Replica exchange (REX) rate illustration. The plot shows the exchange rates between 24 replicas during DMD REX simulation. Each replicate is represented by line of its own colour (from red to blue). Each 24 of the replicates starts with temperatures equally distributed in the range from 0.375 to 0.605 kcal/(mol k_B) and is run for 6 x 10^6 DMD time steps. The first 2 x 10^6 DMD time steps were disregarded during the analysis to account for equilibration of the simulation. During the simulation, replicates exchange their temperatures allowing "colder" replicates to "heat up" and "hot" replicates to "cool down" (lines moving up and down in the plot), which assists in overcoming small local energy barriers. As it can be seen, each replicate (line of each colour) passes through different temperatures from low to high, which indicates that all replicates are well mixed (the plot is highly mosaic) during the simulation. For each replicate, states with different temperatures are visited during REX simulation, which suggests that the REX analysis is adequate.
(TIF)

S3 Fig. Heat capacity curve of native α-synuclein. The heat capacity (Cv) curve was computed using WHAM [58] on the REX/DMD trajectories for α-synuclein in the range of 0.375 to 0.605 kcal/(mol kB) DMD temperature units. The large peak corresponds to the unfolding temperature of the protein. States corresponding to all representative structures are located on the wide shoulder of the curve on the left. The absence of major peaks in this area indicates that the protein can coexist in multiple compact states.
(TIF)

S4 Fig. Comparison of crosslinking constraints satisfied by each cluster. The distances at which a crosslink is considered satisfied for ABAS, EDC, SDA, and TATA crosslinks are < 17 Å, < 10 Å, < 15 Å, and < 14 Å, respectively. Unsatisfied crosslinks were counted in the simulation as a penalty.
(TIF)

S5 Fig. Synuclein structure fluctuations in the absence of crosslinking constraints. The centroid structure from the lowest energy cluster in the ensemble (Fig 3A) was allowed to relax during simulation without any crosslinking constraints at the temperature of 0.45 (kcal/mol/kb DMD units) for 2 x 10^6 DMD time steps. The first 500k steps were discarded as the system equilibration.
(TIF)

S6 Fig. Comparison of the CL-DMD and PRE-NMR ensembles. PRE-NMR α-synuclein ensemble corresponding to the publication by Allison et al., JACS 2009 [59] was obtained from Protein Ensemble Database http://pedb.vib.be. Ensemble average inter-residue distances from the residues, which were spin labeled in the PRE-NMR study. Each graph shows the average inter-residue differences between all residues in the ensemble and the spin-labelled residues in the PRE-NMR experiment. Panels **A-D** correspond to the four major CL-DMD clusters presented in this work, the **E** panels correspond to the PRE-NMR ensemble. (TIF)

S7 Fig. Comparison of the CL-DMD and PRE-NMR ensembles' radii of gyration probabilities. Radius of gyration (Rg) probability distributions for the ensemble of α-synuclein structures obtained by CL-DMD (combination of four major clusters) and for PRE-NMR structures downloaded from Protein Ensemble Database http://pedb.vib.be. (TIF)

S8 Fig. Comparison of top-down HDX data from UVPD fragmentation of hydrogen/deuterium exchanged synuclein with the expected protection for each cluster. The expected protection is represented by blue boxes. The deuteration levels determined from the HDX experiments are shown using a blue-white-red gradient, with blue at 0 (protected from exchange) and red at 1 (exchanged). (TIF)

S1 Table. Crosslinks used for molecular modelling. Crosslinks found which were used during CL-DMD simulations. *These crosslinks were previously reported in [25]. (XLSX)

S2 Table. Surface modification of α-synuclein. Residues modified by PCAS-$^{12}C_6{}^{13}C_6$ in a urea-PCAS surface modification experiment. H/L ratios measure the surface exposure of a residue by comparing the ratio of native to urea-treated proteins modified by PCAS-$^{12}C_6{}^{13}C_6$. A ratio of over 1.5 indicates that the residue in question is somewhat protected in the native structure. (XLSX)

S3 Table. Long distance CBDPS crosslinking of α-synuclein. Long-distance crosslinks detected using CBDPS crosslinker. All crosslinks were within the 25Å distance maximum for CBDPS. (XLSX)

Acknowledgments

Author Contributions

Conceptualization: Evgeniy V. Petrotchenko, Nikolay V. Dokholyan, Christoph H. Borchers.

Funding acquisition: Nikolay V. Dokholyan, Christoph H. Borchers.

Investigation: Nicholas I. Brodie, Konstantin I. Popov.

Methodology: Konstantin I. Popov, Evgeniy V. Petrotchenko, Nikolay V. Dokholyan, Christoph H. Borchers.

Software: Nikolay V. Dokholyan.

Supervision: Evgeniy V. Petrotchenko, Nikolay V. Dokholyan, Christoph H. Borchers.

Writing – original draft: Nicholas I. Brodie.

Writing – review & editing: Konstantin I. Popov, Evgeniy V. Petrotchenko, Nikolay V. Dokholyan, Christoph H. Borchers.

References

1. Lashuel HA, Overk CR, Oueslati A, Masliah E. The many faces of alpha-synuclein: from structure and toxicity to therapeutic target. Nat Rev Neurosci. 2013; 14(1):38–48. https://doi.org/10.1038/nrn3406 PMID: 23254192

2. Bendor JT, Logan TP, Edwards RH. The Function of alpha-Synuclein. Neuron. 2013; 79(6):1044–66. https://doi.org/10.1016/j.neuron.2013.09.004 PMID: 24050397

3. Cremades N, Chen SW, Dobson CM. Structural Characteristics of alpha-Synuclein Oligomers. Int Rev Cell Mol Biol. 2017; 329:79–143. Epub 2017/01/23. https://doi.org/10.1016/bs.ircmb.2016.08.010 PMID: 28109332.

4. Lautenschläger J, Stephens AD, Fusco G, Ströhl F, Curry N, Zacharopoulou M, et al. C-terminal calcium binding of α-synuclein modulates synaptic vesicle interaction. Nat Commun. 2018; 9(1):712. https://doi.org/10.1038/s41467-018-03111-4 PMID: 29459792.

5. Fusco G, Chen SW, Williamson PTF, Cascella R, Perni M, Jarvis JA, et al. Structural basis of membrane disruption and cellular toxicity by α-synuclein oligomers. Science. 2017; 358(6369):1440–3. https://doi.org/10.1126/science.aan6160 PMID: 29242346.

6. Brundin P, Ma JY, Kordower JH. How strong is the evidence that Parkinson's disease is a prion disorder? Curr Opin Neurol. 2016; 29(4):459–66. https://doi.org/10.1097/WCO.0000000000000349 PMID: 27257944

7. Eliezer D. Biophysical characterization of intrinsically disordered proteins. Curr Opin Struct Biol. 2009; 19(1):23–30. https://doi.org/10.1016/j.sbi.2008.12.004 PMID: 19162471

8. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM. Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J Am Chem Soc. 2005; 127(2):476–7. https://doi.org/10.1021/ja044834j PMID: 15643843

9. Nath A, Sammalkorpi M, DeWitt DC, Trexler AJ, Elbaum-Garfinkle S, O'Hern CS, et al. The Conformational Ensembles of alpha-Synuclein and Tau: Combining Single-Molecule FRET and Simulations. Biophys J. 2012; 103(9):1940–9. https://doi.org/10.1016/j.bpj.2012.09.032 PMID: 23199922

10. Li J, Uversky VN, Fink AL. Conformational behavior of human alpha-synuclein is modulated by familial Parkinson's disease point mutations A30P and A53T. Neurotoxicology. 2002; 23(4–5):553–67. https://doi.org/10.1016/S0161-813x(02)00066-9 PMID: 12428728

11. Brodie NI, Popov KI, Petrotchenko EV, Dokholyan NV, Borchers CH. Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. Sci Adv. 2017; 3(7):e1700479. Epub 2017/07/12. https://doi.org/10.1126/sciadv.1700479 PMID: 28695211.

12. Bonomi M, Heller GT, Camilloni C, Vendruscolo M. Principles of protein structural ensemble determination. Curr Opin Struc Biol. 2017; 42:106–16. Epub 2017/01/08. https://doi.org/10.1016/j.sbi.2016.12.004 PMID: 28063280.

13. Piana S, Donchev AG, Robustelli P, Shaw DE. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. J Phys Chem B. 2015; 119(16):5113–23. https://doi.org/10.1021/jp508971m PMID: 25764013.

14. Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. Proc Natl Acad Sci U S A. 2018; 115(21):E4758–E66. https://doi.org/10.1073/pnas.1800690115 PMID: 29735687

15. Ding F, Dokholyan NV. Emergence of protein fold families through rational design. PLoS Comput Biol. 2006; 2:e85 https://doi.org/10.1371/journal.pcbi.0020085 PMID: 16839198

16. Ding F, Tsao D, Nie HF, Dokholyan NV. Ab initio folding of proteins with all-atom discrete molecular dynamics. Structure. 2008; 16(7):1010–8. https://doi.org/10.1016/j.str.2008.03.013 PMID: 18611374

17. Shirvanyants D, Ding F, Tsao D, Ramachandran S, Dokholyan NV. Discrete molecular dynamics: an efficient and versatile simulation method for fine protein characterization. J Phys Chem B. 2012; 116 (29):8375–82. Epub 2012/01/28. https://doi.org/10.1021/jp2114576 PMID: 22280505.

18. Lazaridis T, Karplus M. Effective energy functions for protein structure prediction. Curr Opin Struct Biol. 2000; 10(2):139–45. https://doi.org/10.1016/S0959-440X(00)00063-4 PMID: 10753811.

19. Chen Y, Ding F, Dokholyan NV. Fidelity of the protein structure reconstruction from inter-residue proximity constraints. J Phys Chem B. 2007; 111(25):7432–8. Epub 2007/06/05. https://doi.org/10.1021/jp068963t PMID: 17542631.

20. Proctor EA, Ding F, Dokholyan NV. Discrete molecular dynamics. Wiley Interdisciplinary Reviews-Computational Molecular Science. 2011; 1(1):80–92. https://doi.org/10.1002/wcms.4

21. Dixon RD, Chen Y, Ding F, Khare SD, Prutzman KC, Schaller MD, et al. New insights into FAK signaling and localization based on detection of a FAT domain folding intermediate. Structure. 2004; 12 (12):2161–71. https://doi.org/10.1016/j.str.2004.09.011 PMID: 15576030.

22. Ding F, Lavender CA, Weeks KM, Dokholyan NV. Three-dimensional RNA structure refinement by hydroxyl radical probing. Nat Methods. 2012; 9(6):603–8. https://doi.org/10.1038/nmeth.1976 PMID: 22504587.

23. Szöllősi D, Horváth T, Han KH, Dokholyan NV, Tompa P, Kalmár L, et al. Discrete molecular dynamics can predict helical prestructured motifs in disordered proteins. PLoS One. 2014; 9(4):e95795. https://doi.org/10.1371/journal.pone.0095795 PMID: 24763499.

24. Brodie NI, Makepeace KA, Petrotchenko EV, Borchers C. Isotopically-coded short-range hetero-bifunctional photo-reactive crosslinkers for studying protein structure. J Prot. 2015; 118:12–20. https://doi.org/10.1016/j.jprot.2014.08.012 PMID: 25192908.

25. Brodie NI, Petrotchenko EV, Borchers CH. The novel isotopically coded short-range photo-reactive crosslinker 2,4,6-triazido-1,3,5-triazine (TATA) for studying protein structures. J Prot. 2016:pii: S1874-3919(16)30046-X. https://doi.org/10.1016/j.jprot.2016.02.024 PMID: 26931439.

26. Boulikas T, Wiseman JM, Garrard WT. Points of Contact between Histone H-1 and the Histone Octamer. Proc Natl Acad Sci U S A. 1980; 77(1):127–31. https://doi.org/10.1073/pnas.77.1.127 PMID: 6928607.

27. Petrotchenko EV, Serpa JJ, Makepeace KAT, Brodie NI, Borchers CH. $^{14}N^{15}N$ DXMSMS Match program for the automated analysis of LC/ESI-MS/MS crosslinking data from experiments using ^{15}N metabolically labeled proteins. J Prot. 2014; 109:104–10. https://doi.org/10.1016/j.jprot.2014.06.014 PMID: 24972318

28. Hoopmann MR, Zelter A, Johnson RS, Riffle M, MacCoss MJ, Davis TN, et al. Kojak: Efficient Analysis of Chemically Cross-Linked Protein Complexes. J Proteome Res. 2015; 14(5):2190–8. https://doi.org/10.1021/pr501321h PMID: 25812159

29. Serpa JJ, Patterson AP, Pan J, Han J, Wishart DS, Petrotchenko EV, et al. Using multiple structural proteomics approaches for the characterization of prion proteins. J Prot. 2013; 81:31–42. https://doi.org/10.1016/j.jprot.2012.10.008 PMID: 23085224.

30. Pan J, Han J, Borchers CH, Konermann L. Hydrogen/deuterium exchange mass spectrometry with top-down electron capture dissociation for characterizing structural transitions of a 17 kDa protein. J Am Chem Soc. 2009; 131(35):12801–8. Epub 2009/08/13. https://doi.org/10.1021/ja904379w PMID: 19670873.

31. Petrotchenko EV, Borchers CH. HDX Match software for the data analysis of top-down ECD-FTMS hydrogen/deuterium exchange experiments. J Am Soc Mass Spectrom. 2015; 26(11):1895–8. https://doi.org/10.1007/s13361-015-1213-z PMID: 26162650

32. Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc Natl Acad Sci U S A. 2015; 112 (24):E3095–103. Epub 2015/06/04. https://doi.org/10.1073/pnas.1500851112 PMID: 26038575.

33. Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Discrete molecular dynamics studies of the folding of a protein-like model. Fold Des. 1998; 3(6):577–87. https://doi.org/10.1016/S1359-0278(98)00072-8 PMID: 9889167.

34. Okamoto Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. Journal of Molecular Graphics & Modelling. 2004; 22(5):425–39. https://doi.org/10.1016/j.jmgm.2003.12.009 PMID: 15099838.

35. Chodera JD, Swope WC, Pitera JW, Seok C, Dill KA. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. J Chem Theory Comput. 2007; 3 (1):26–41. https://doi.org/10.1021/ct0502864 PMID: 26627148.

36. Yin S, Biedermannova L, Vondrasek J, Dokholyan NV. MedusaScore: An accurate force field-based scoring function for virtual drug screening. J Chem Inf Model. 2008; 48(8):1656–62. https://doi.org/10.1021/ci8001167 PMID: 18672869

37. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013; 29(7):845–54. https://doi.org/10.1093/bioinformatics/btt055 PMID: 23407358.

38. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide Folding: When Simulation Meets Experiment. Angew Chem. 1999 38(1–2):236–40.

39. Pauwels K, Lebrun P, Tompa P. To be disordered or not to be disordered: is that still a question for proteins in the cell? Cell Mol Life Sci. 2017. Epub 2017/06/15. https://doi.org/10.1007/s00018-017-2561-6 PMID: 28612216.

40. Theillet FX, Binolfi A, Bekei B, Martorana A, Rose HM, Stuiver M, et al. Structural disorder of monomeric alpha-synuclein persists in mammalian cells. Nature. 2016; 530(7588):45–50. Epub 2016/01/26. https://doi.org/10.1038/nature16531 PMID: 26808899.

41. Rossetti G, Musiani F, Abad E, Dibenedetto D, Mouhib H, Fernandez CO, et al. Conformational ensemble of human α-synuclein physiological form predicted by molecular simulations. Phys Chem Chem Phys. 2016; 18(8):5702–6. https://doi.org/10.1039/c5cp04549e PMID: 26553504

42. Bartels T, Kim NC, Luth ES, Selkoe DJ. N-alpha-acetylation of α-synuclein increases its helical folding propensity, GM1 binding specificity and resistance to aggregation. PLoS ONE. 2014; 9(7):e103727. https://doi.org/10.1371/journal.pone.0103727 PMID: 25075858.

43. Ruzafa D, Hernandez-Gomez YS, Bisello G, Broersen K, Morel B, Conejero-Lara F. The influence of N-terminal acetylation on micelle-induced conformational changes and aggregation of α-Synuclein. PLoS ONE. 2017; 12(5):e0178576. https://doi.org/10.1371/journal.pone.0178576 PMID: 28562630.

44. Maltsev AS, Ying J, Bax A. Impact of N-terminal acetylation of α-synuclein on its random coil and lipid binding properties. Biochemistry. 2012; 51(25):5004–13. https://doi.org/10.1021/bi300642h PMID: 22694188.

45. Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, Jovin TM, et al. Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. Proc Natl Acad Sci U S A. 2005; 102(5):1430–5. https://doi.org/10.1073/pnas.0407146102 PMID: 15671169.

46. Allison JR, Rivers RC, Christodoulou JC, Vendruscolo M, Dobson CM. A relationship between the transient structure in the monomeric state and the aggregation propensities of α-synuclein and β-synuclein. Biochemistry. 2014; 53(46):7170–83. https://doi.org/10.1021/bi5009326 PMID: 25389903.

47. Grupi A, Haas E. Segmental conformational disorder and dynamics in the intrinsically disordered protein α-synuclein and its chain length dependence. Journal of Molecular Biology. 2011; 405(5):1267–83. https://doi.org/10.1016/j.jmb.2010.11.011 PMID: 21108951.

48. Trexler AJ, Rhoades E. Single molecule characterization of α-synuclein in aggregation-prone states. Biophys J. 2010; 99(9):3048–55. https://doi.org/10.1016/j.bpj.2010.08.056 PMID: 21044603.

49. Schwalbe M, Ozenne V, Bibow S, Jaremko M, Jaremko L, Gajda M, et al. Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α-synuclein in solution from NMR and small angle scattering. Structure. 2014; 22(2):238–49. https://doi.org/10.1016/j.str.2013.10.020 PMID: 24361273.

50. Gurry T, Ullman O, Fisher CK, Perovic I, Pochapsky T, Stultz CM. The dynamic structure of α-synuclein multimers. J Am Chem Soc. 2013; 135(10):3865–72. https://doi.org/10.1021/ja310518p PMID: 23398399.

51. Brodie NI, Huguet R, Zhang T, Viner R, Zabrouskov V, Pan J, et al. Top-down hydrogen-deuterium exchange analysis of protein structures using ultraviolet photodissociation (UVPD). Anal Chem. 2018; 90(5):3079–82. https://doi.org/10.1021/acs.analchem.7b03655 PMID: 29336549.

52. Giasson BI, Murray IVJ, Trojanowski JQ, Lee VMY. A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. J Biol Chem. 2001; 276(4):2380–6. https://doi.org/10.1074/jbc.M008919200 PMID: 11060312

53. Nasica-Labouze J, Nguyen PH, Sterpone F, Berthoumieu O, Buchete NV, Coté S, et al. Amyloid β Protein and Alzheimer's Disease: When Computer Simulations Complement Experimental Studies. Chem Rev. 2015; 115(9):3518–63. https://doi.org/10.1021/cr500638n PMID: 25789869.

54. Tuttle MD, Comellas G, Nieuwkoop AJ, Covell DJ, Berthold DA, Kloepper KD, et al. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. Nature Structural and Molecular Biology. 2016; 23(5):409–15. https://doi.org/10.1038/nsmb.3194 PMID: 27018801.

55. Breydo L, Wu JW, Uversky VN. A-synuclein misfolding and Parkinson's disease. Biochim Biophys Acta. 2012; 1822(2):261–85. https://doi.org/10.1016/j.bbadis.2011.10.002 PMID: 22024360.

56. Sahay S, Ghosh D, Singh PK, Maji SK. Alteration of Structure and Aggregation of α-Synuclein by Familial Parkinson's Disease Associated Mutations. Curr Protein Pept Sci. 2017; 18(7):656–76. https://doi.org/10.2174/1389203717666160314151706 PMID: 26972727.

57. Oueslati A, Paleologou KE, Schneider BL, Aebischer P, Lashuel HA. Mimicking phosphorylation at serine 87 inhibits the aggregation of human α-synuclein and protects against its toxicity in a rat model of Parkinson's disease. J Neurosci. 2012; 32(5):1536–44. https://doi.org/10.1523/JNEUROSCI.3784-11.2012 PMID: 22302797.

58. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. Journal of Computational Chemistry. 1992; 13:1011–21. https://doi.org/10.1002/jcc.540130812

59. Allison JR, Varnai P, Dobson CM, Vendruscolo M. Determination of the Free Energy Landscape of alpha-Synuclein Using Spin Label Nuclear Magnetic Resonance Measurements. J Am Chem Soc. 2009; 131(51):18314–26. https://doi.org/10.1021/ja904716h PMID: 20028147