Structural and Dynamic Insights Into α-Synuclein Dimer Conformations

Joanna Zamel[1,*], Jiaxing Chen[2,*], Sofia Zaer[1], Paul David Harris[1], Paz Drori[1], Mario Lebendiker[3], Nir Kalisman[1], Nikolay V. Dokholyan[2,4,5,#], Eitan Lerner[1,6,#]

[1] Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Faculty of Mathematics & Science, The Edmond J. Safra Campus, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel

[2] Department of Pharmacology, Penn State College of Medicine, Hershey, PA 17033, USA

[3] Wolfson Centre for Applied Structural Biology, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, The Edmond J. Safra Campus, Givat Ram, Jerusalem, 9190401, Israel

[4] Department of Biochemistry & Molecular Biology, Penn State College of Medicine, Hershey, PA 17033, USA

[5] Departments of Chemistry and Biomedical Engineering, Pennsylvania State University, University Park, PA 16802, USA

[6] The Center for Nanoscience and Nanotechnology, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel

[*] Shared first authorship

[#] Corresponding authors

Abstract

Parkinson's disease is associated with the aggregation of the protein α-synuclein. While α-synuclein can exist in multiple oligomeric states, the dimer has been a subject of extensive debates. Here, using an array of biophysical approaches, we demonstrate α-synuclein *in vitro* exhibits primarily a monomer-dimer equilibrium in nanomolar concentrations and up to a few micromolars. We then use spatial information from hetero-isotopic cross-linking mass spectrometry experiments as restrains in discrete molecular dynamics simulations to obtain the ensemble structure of dimeric species. Out of eight structural sub-populations of dimers we identify one that is compact, stable, abundant, and exhibits partially exposed β-sheet structures. This compact dimer is the only one where the hydroxyls of tyrosine 39 are in a proximity that can promote dityrosine covalent linkage implicated in amyloidogenesis. We propose that this α-synuclein dimer features etiological relevance to Parkinson’s disease.

Introduction

α-Synuclein (αSyn) is a small soluble protein localized mainly in nuclei and in presynaptic terminals of dopaminergic neurons in the *substantia nigra pars compacta*[1]. In the brain, αSyn is involved in several stages of neurotransmitter trafficking, including the stabilization of dopamine vesicles and supporting fast and efficient formation of the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex[2]. However, αSyn is involved in multiple disorders termed synucleinopathies, such as Parkinson's disease (PD). One of the hallmarks of PD is the accumulation of αSyn amyloid-like fibrils in Lewy bodies[1].

The protein is 140 amino acids long and can be divided into three peptide segments: 1) the N-terminal segment or domain (NTD; residues1-60), an amphipathic and lysine-rich region that plays critical roles in phospholipid attachment[3], 2) the non-amyloid-β component (NAC) segment (residues 61-95), a hydrophobic region that is involved in aggregation and fibrilization[4,5], and 3) the C-terminal segment or domain (CTD; residues 96-140), an acidic tail that contains multiple prolines and acidic amino acids that disrupt secondary structure formation[6].

αSyn attaches to different types of negatively charged membrane surfaces, including the ones that form dopamine vesicles in neurons[7,8]. Additionally, αSyn can spontaneously self-associate into oligomers. αSyn amyloid-like fibrils and certain oligomeric species are biochemical features of PD. Other oligomeric species have been documented, such as the α-helical stable tetramer[9–11]. These different oligomeric species may support different biological functions.

In aqueous solutions, αSyn is intrinsically disordered. Therefore, αSyn exhibits features of misfolded proteins, such as high structural heterogeneity and conformational plasticity. Yet, oftentimes when αSyn binds to a biomolecular target, such as the outer leaflet of phospholipid membranes or undergoes self-association, it tends to co-fold in a given pathway[12].

αSyn aggregation kinetics include a long (tens of hours) delay, which is thought of as the time it takes for αSyn to form a fibrillization nucleus[13]. Since exposure of the hydrophobic residues of the NAC segment to the solvent is one factor that can help nucleate oligomerization in the fibril formation pathway, delay of the fibrillization nucleation can be attained via αSyn forms that sterically occlude the NAC from exposure to the solvent. In the free-form monomer, the CTD and NTD sterically interact to shield the NAC from being exposed to the solvent, and by that protect the NAC segment from aggregation-prone interactions with the NAC segment of other αSyn molecules[14]. However, there may be other αSyn forms, including small oligomers, which may act to sterically occlude the NAC.

αSyn self-associates and some of its self-associated species lead to pathophysiological-promoting conditions, such as amyloid-like fibrils or toxic oligomers. Therefore, it is of utmost importance to study the mechanism by which αSyn self-associates, and specifically the earliest stages of oligomerization. Logically, the earliest stage of self-association would be the result of dimerization. It is noteworthy that αSyn dimers and trimers have previously been implicated with neurotoxicity[15]. Additionally, the αSyn dimer would serve as the smallest and most elementary self-associated species. Therefore, studying the structural features of an αSyn dimer is important for understanding how self-association starts in αSyn at the first place. Additionally, it is hypothesized that specific inter-molecular interactions within αSyn oligomers can act as seeds of aggregation, especially irreversible interactions, such as covalent bonds between pairs of amino acid side chains of two αSyn subunits[16]. One such covalent linkage is related to the stabilization of the fibrillization pathway is dityrosine covalent linkages[16–19]. Dityrosine linkages can form only after hydroxyl radicalization of tyrosines, which can occur in specific stress conditions, implicated in PD[20]. Therefore, upon elucidation of structure models of the αSyn dimer, inter-molecular proximal tyrosines may serve as a signature of the potential for this covalent linkage to occur under such stress conditions.

Several experimental and computational studies have previously focused on the αSyn dimer[16,21–27]. Using high-speed atomic force imaging, Zhang *et al.*[25] have revealed the αSyn dimer exists in several different configurations, at low spatial resolution: (i) with one compact globular subunit and another expanded and dynamic, (ii) with both subunits being compact and globular, and (iii) with both subunits being expanded and flexible.

Here, we demonstrate αSyn exhibits predominantly a monomer-dimer equilibrium in the nanomolars (nM) and up to a few micromolars (μM), which is below the αSyn concentrations in which αSyn exhibits aggregation that leads to fibril formation[13]. Using single-molecule photo-isomerization related fluorescence enhancement (smPIFE; commonly known as protein-induced fluorescence enhancement)[28–31], we demonstrate dimer formation is associated with structural changes that lead to specific changes in amino acid steric hindrances. Nevertheless, at a high dimer fraction, the smPIFE results still exhibit several different conformations should exist, with transitions between them as slow as milliseconds.

Encouraged by the results, we performed hybrid structural modeling of the αSyn dimer, using hetero-isotopic cross-linking mass-spectrometry (CL-MS)[32–34] as experimentally-derived restraints for discrete molecular dynamics (DMD) simulations, to elucidate structure models of the αSyn dimer. We report the features of the structure models of different conformations of the αSyn dimer.

Results

MALS: αSyn monomer-dimer equilibrium

We perform analytical anion exchange (AIEX) chromatography coupled to multi-angle light scattering (MALS), yielding a direct readout of the molecular mass of the different species in solution with high detection sensitivity and minimal chemical modifications or perturbations[35]. AIEX-MALS of *wt*-αSyn yields a single elution peak with an average molecular mass of 21.6±0.8 kDa, which is larger than that of a monomer (14.4 kDa) and smaller to that of a dimer (< 28.8 kDa; Fig. 1). We propose that the elution peak contains a mixture of primarily αSyn monomers and dimers. Possible reasons for a single elution peak with an average molecular mass between a monomer and a dimer are: i) both monomer and dimer may have a small difference in their ionic strength values, or ii) monomers and dimers undergo rapid dynamic association-dissociation transitions. We perform further validation of the monomer-dimer suggestion using size exclusion chromatography (SEC)-MALS. The chromatogram reports on similar results: after a first elution peak of early-formed aggregates, we record one elution peak with an average molecular mass (identified directly from MALS) larger than a monomer and smaller than a dimer (Fig. S1). Therefore, the ensemble of αSyn could be described as a mixture of primarily monomers and dimers.

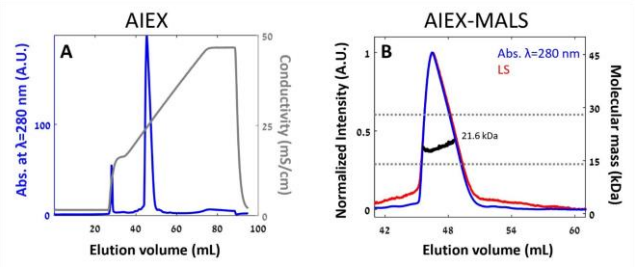

Fig. 1: **AIEX-MALS of WT-αSyn elutes as a single peak with a molecular mass larger than αSyn monomer.**
A - AIEX using 1 mL mono Q column with the gradient shown in grey, has been performed on a sample of the WT αSyn. The sample has eluted as a single elution peak (absorption at a wavelength of 280 nm, blue). **B** – the absorption at 280 nm (blue) and the light scattering intensity (red) were used for the calculation of MALS-derived molecular weight of 21.6±0.8 kDa. The elutant had a molecular mass value between that expected for an αSyn monomer and that expected for a dimer (dotted grey lines).

Steady-state inter-molecular FRET: small-scale self-association equilibrium

For assessing αSyn dimerization as a self-association process, we perform bulk inter-molecular FRET measurements between αSyn labeled with either ATTO 488 (donor) or ATTO 647N (acceptor) at the thiol introduced by a cysteine at position 39.

Freshly-prepared αSyn should exhibit mainly the monomeric form and if any oligomeric form exists in the sample, they should yield a minimal inter-molecular FRET signature. We measure the acceptor excitation spectrum of a mixture of 100 nM donor- and acceptor-labeled αSyn molecules focusing on the acceptor fluorescence intensity (see *Methods*). As a control, we measure the acceptor excitation spectrum of a sample containing only 100 nM acceptor-labeled αSyn, hence exhibiting no FRET signatures, and solely direct acceptor excitation.

Our results exhibit a peak in the acceptor excitation spectrum at λ~500 nm (Fig. S2), which reports on the occurrence of FRET. The FRET signal confirms the presence of self-associated αSyn species, which are primarily dimers, according to the MALS-based measurements.

Time-resolved inter-molecular FRET: small-scale self-association equilibrium

Although the steady-state inter-molecular FRET measurements clearly indicate interactions occurring between donor- and acceptor-labeled αSyn molecules at 100 nM, we strive to test these interactions at lower concentrations, for which a standard steady-state fluorimeter is not sensitive enough. To do so, we use a confocal microscope setup equipped with pulsed lasers and point detectors with time-correlated single photon counting (TCSPC) detection. This allows us to record photon arrival times relative to the moments of excitation, hence fluorescence decays. Performing measurements similar to the steady-state donor- and acceptor-labeled measurements, but at lower concentrations of α-syn, we can test for dimers. If a fraction of the donor- and acceptor-labeled αSyn molecules under measurement represents self-associated species, the acceptor fluorescence following donor excitation should exhibit a delayed decay, owing to the donor

excited-state lifetime in addition to the acceptor excited-state lifetime. At a high αSyn concentration, for which the fraction of dimers exhibiting FRET should be high, the acceptor fluorescence decay following donor excitation are expected to decay slower (longer lifetime) relative to the same decays measured at lower αSyn concentrations, for which the fraction of self-associated species is low. Figure 2 demonstrates exactly this trend. One can observe that the normalized acceptor fluorescence decay following donor excitation pertains, even at low αSyn concentrations. This is due to the leakage of the donor fluorescence decay to the acceptor detection channel, which is low (Fig. S3). Nevertheless, the number of molecules exhibiting FRET is low at this concentration. Overall, the delayed acceptor fluorescence decays following donor excitation exhibit that the lower αSyn concentration is, the faster the decay curves are (Fig. 2). The slow decay component is a signature of the fraction of inter-molecular FRET occurring at higher αSyn concentrations.

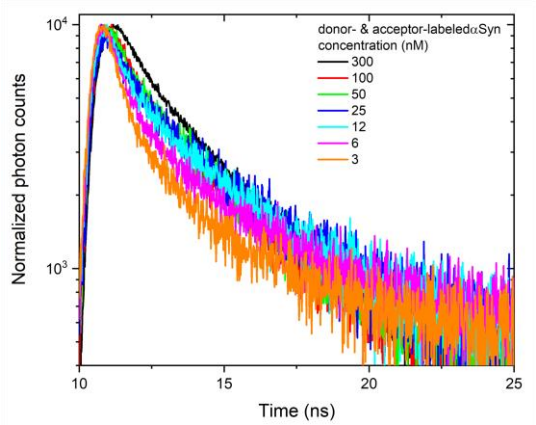

Fig. 2: **A signature of intermolecular FRET in acceptor fluorescence decay, decreases with αSyn concentration.**
The acceptor fluorescence decays, following donor excitation, were measured in donor- and acceptor-labeled αSyn mixtures at different nominal αSyn concentration (see inset). The fluorescence decay curves exhibit two components: a fast component, that is most probably due to leakage of the donor fluorescence red edge into the acceptor detection channel (see Fig. S3) and a slow component, indicative of acceptor excitations that were delayed in the donor excited state prior to the FRET event. The amplitude of the slow component is decreasing as the nominal αSyn concentration decreases.

Fluorescence correlation spectroscopy inter-molecular FRET: small-scale self-association equilibrium

To further test the abundance of the αSyn self-associated species at lower concentrations, we perform fluorescence correlation spectroscopy[36] (FCS)- FRET (FCS-FRET) measurements. These experiments measure the donor and acceptor fluorescence as a function of time of mixtures of donor- and acceptor- labeled αSyn in decreasing concentrations (50, 12 and 3 nM), and then calculate the donor fluorescence autocorrelation and donor-acceptor fluorescence cross-correlation curves. If two or more monomers are in a sufficient proximity with each other while diffusing through the confocal

spot, then exciting the donor will lead to energy transfer to the acceptor and photons will be detected on both the donor and acceptor fluorescence detection channels. The donor-acceptor fluorescence cross-correlation curve reports on such species that traverse the effective excitation volume. If αSyn is in its monomeric state, donor fluorescence will mostly be detected. The donor fluorescence autocorrelation curve reports on both donor-labeled species: monomers and self-associated species. The ratio of the amplitudes of the donor-acceptor cross-correlation and the donor autocorrelation curves is proportional to the fraction of self-associated αSyn species labeled with both donor and acceptor[37]. The results reported in Fig. 3 exhibit a fraction of αSyn self-associated species exists even at 3 nM concentration of αSyn. Overall, bulk steady-state, time-resolved and FCS-FRET results point to an equilibrium between monomers and self-associated species of αSyn in the nM range. Taken together with the results from chromatography coupled to MALS, this equilibrium is primarily between monomers and dimers of αSyn.

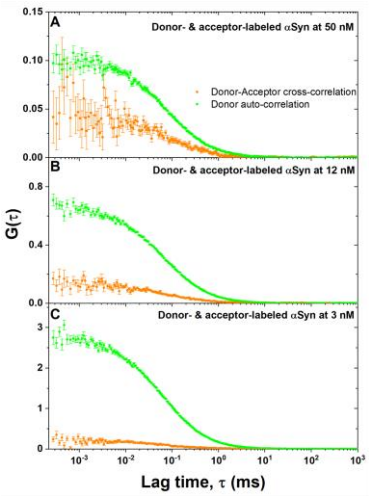

Fig. 3: **The abundance of a self-associated species of αSyn is decreasing with decreasing αSyn concentration.**
Donor fluorescence auto-correlation curves (green) and donor-acceptor fluorescence cross-correlation curves (orange) of donor- and acceptor-labeled αSyn are shown at decreasing nominal concentrations (**A** - 50 nM; **b** – 12 nM; **c** – 3 nM). The amplitude of the donor-acceptor fluorescence cross-correlation relative to the amplitude of the donor fluorescence auto-correlation decreases with decreasing αSyn concentration, indicative of a decreasing concentration of species having coincidence of both donor and acceptor fluorescence (namely self-associated species).

smPIFE: aSyn dimer formation involves structural changes

To investigate the structural changes associated with the formation of an αSyn dimer, we perform sulfo-Cy3 (sCy3)-based intra-molecular time-resolved smPIFE measurements[30,31]. The uniqueness of this dye is that it exhibits excited-state isomerization between a bright *trans* isomer and a dark *cis* isomer, leading to an overall low fluorescence quantum yield. However, if the excited-state isomerization of sCy3 is obstructed due to steric hindrance, for instance by a protein surface in the vicinity of the dye, then the isomerization rate will decrease leading to more fluorescence from the *trans* isomer, and hence to a higher fluorescence quantum yield.

We hypothesize that like in the case of the free-form monomer[30,31] the αSyn dimer may exhibit distinct subpopulations with different characteristic fluorescence lifetimes. To gain

insights on the dimer formation we performed smPIFE measurements. We measured 25 pM of freely diffusing sCy3-labeled αSyn in the presence of increasing concentrations of *wt*-αSyn, up to few μM to induce the formation of dimers (Fig. 4). Additionally, fitting of the mean nanotime histograms to a sum of Gaussian model may assist in the quantification of the mean lifetimes and fractions of each sub-population (Table S1). From observing how the burst histogram of mean nanotimes change with increasing the *wt*-αSyn concentration, it is clear that: 1) sCy3 senses the local structural reorganization in the vicinity of residues 56 and 140, upon dimer formation, with an enhanced low mean nanotime sub-population, hence with minimal steric obstruction pointing towards more solvent exposure of the residues, 2) sCy3 senses the local structural reorganization in the vicinity of residues 39, upon dimer formation, with both a low and intermediate mean nanotime sub-population, hence with steric obstruction levels pointing towards high and intermediate levels of solvent exposure, 3) sCy3 fails to sense the local structural reorganization in the vicinity of residue 26, upon dimer formation, and 4) some local structural changes when shifting from the monomer to the dimer include gradual changes at multiple *wt*-αSyn concentration concentrations and not necessarily a single transition. Importantly, in all sCy3-labeled residues, the results with the highest dimer fraction, at the highest *wt*-αSyn concentration exhibit dynamic heterogeneity between multiple mean nanotime subpopulations, with transitions between them slower than diffusion times through the confocal spot, hence slower than a few ms.

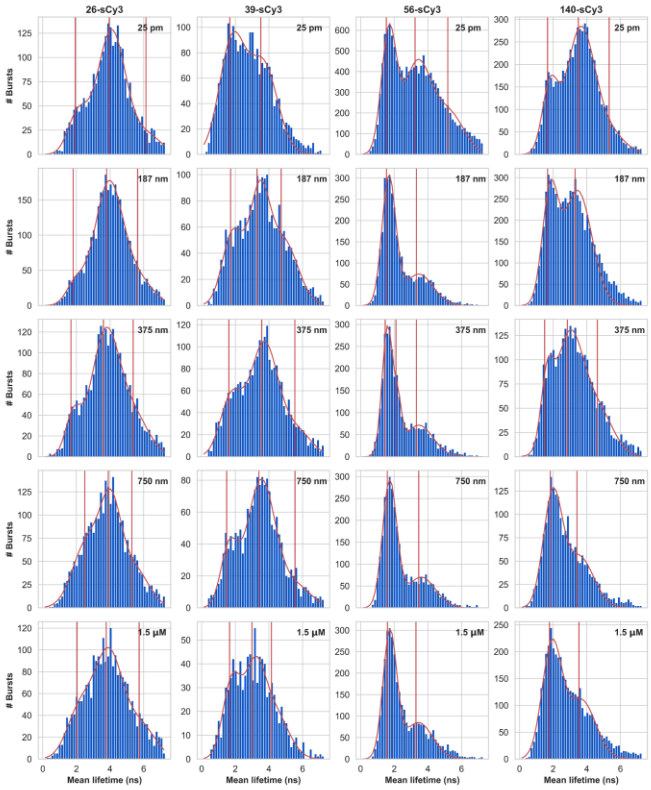

Fig. 4: **smPIFE data exhibiting how different sCy3-labeled residues sense a local structural change when changing from a monomer to a dimer.**
Mean fluorescence lifetime histograms of sCy3 labeling different residue positions (26, 39, 56, 140; columns) with 25 pM sCy3-labeled αSyn, in the presence of increasing concentrations (0, 187, 375, 750, 1,500 nM; rows) of unlabeled *wt*-αSyn. The concentrations shown at the top of panels signify the approximate overall αSyn concentration.

Our steady-state, time-resolved and FCS-FRET measurements demonstrated a monomer-dimer equilibrium exists in the hundreds of nM, and hence the smPIFE results are expected to be related to the monomer-dimer changes. Nevertheless, to ensure that the *wt*-αSyn titration results in dimer formation and not higher order self-associated species, we correlated the sCy3 mean fluorescence lifetimes with their burst durations (Fig. S4). Burst durations are indicative of the molecular diffusion through the confocal spot, and hence are proportional to the hydrodynamic radii of the measured molecules. Indeed, our measurements report on changes in sCy3 fluorescence lifetime sub-populations that are correlated with a 30-50% increase in the burst durations in each labelled residue, which agrees with the previously-reported increase in the hydrodynamic radius of monomeric αSyn versus dimeric αSyn[26]. Therefore, we expect the αSyn dimer to be represented by multiple distinct conformations.

Hetero-isotopic CL-MS: retrieving multiple spatial restraints to support hybrid structural modelling

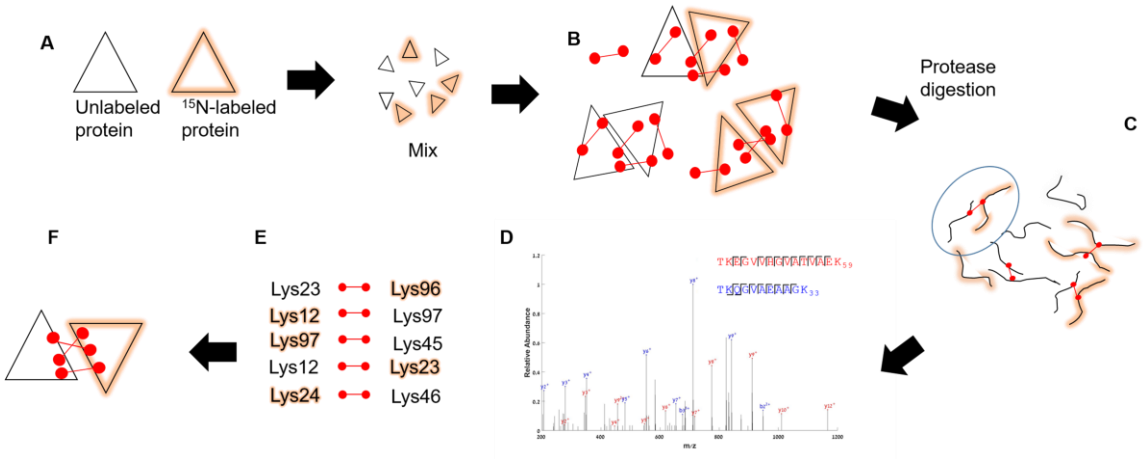

Fig. 5: **Hetero-isotopic CL-MS workflow.**
The ^{14}N-^{15}N CL-MS workflow. **(A)** ^{14}N- and ^{15}N-labeled α-syn are mixed. **(B)** Incubation with cross-linkers (red dumbbells). Covalent inter and intra molecular cross-links in the homodimer (within the cross-linker length) **(C)** Complexes are digested via a protease, ending with peptides linked with the cross-linker. Pairs of hetero-isotopic peptides can originate only from inter-molecular cross-links. Similarly, cross-links between a pair of homo-isotopic peptides that are not identified as inter-molecular, are intra-molecular **(D)** The digested peptides analyzed by mass-spectrometry and the spectra help identify pairs of cross-linked residues. Differently cross-linked pairs of peptides with different nitrogen isotope contents yield different distinguishable peaks. **(E)** Hundreds of cross-links can be identified from a large complex. **(F)** The cross-links are converted into distance restraints/constraints, which drive computational modeling.

To the best of our knowledge, a high-resolution structure model of the αSyn dimer is not available. The main reason that hinders solving the structure by NMR or X-ray crystallography is the oligomeric and structural heterogeneity of αSyn. Therefore, there is a need to use a different approach to overcome these difficulties. Here, we introduce cross-linking mass-spectrometry (CL-MS) followed by discrete molecular dynamics (DMD) as an integrative structural biology approach to model the structure of the αSyn dimer.

CL-MS has emerged as a robust method for studying protein structures and assemblies[38–41]. Recently, CL-MS has been used to study small disordered proteins, such as the free-form αSyn monomer and assisted in elucidating its ensemble structure[31]. One of the limitations of CL-MS is the inability to distinguish between inter- and intra-molecular cross-links in homodimers. One way to circumvent this limitation is by metabolic labeling of one subunit with the heavy isotope, ^{15}N, while the other monomer remains as the natural light ^{14}N isotope[32–

[34]. The mass shift of the heavy isotope allows the distinction between inter- or intra-molecular cross-links (Fig. 5). In this approach, we mix the 'heavy' and the 'light' monomers and introduce the cross-linking reagent. Beside the heavy-heavy and light-light peptides that may originate from inter- or intra-molecular interactions, we will obtain heavy-light cross-linked peptides that can only originate from inter-molecular interactions (Fig. 5). Additionally, light-light or heavy-heavy cross-linked residues, X and Y, which have not appeared in the list of light-heavy cross-linked residues, are suggestive of intra-molecular cross-links. Nevertheless, since the homodimer is in equilibrium with the dissociated monomer, it is hard to tell which intra-molecular cross-link comes from the dimer subunit and which comes from the dissociated monomer. Therefore, it is important to work in conditions in which the homodimer is enriched, and the amount of dissociated monomer is minimal (see *Methods*).

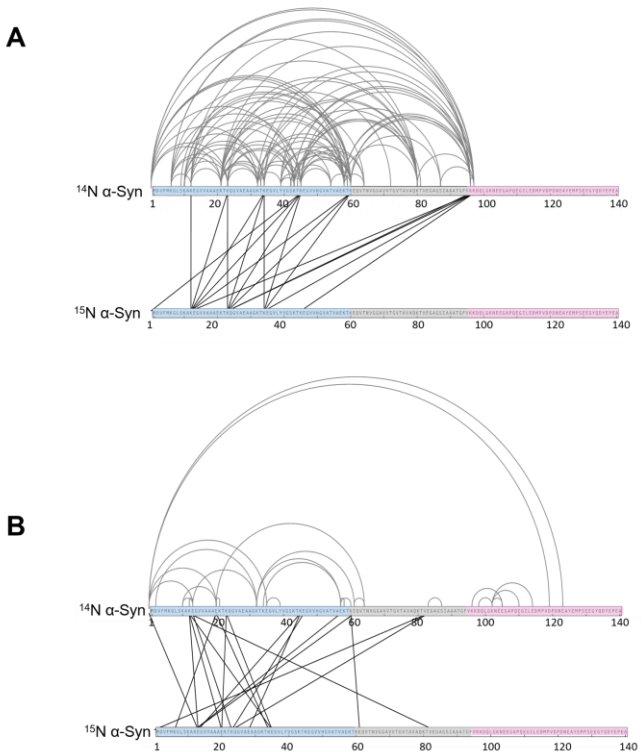

Fig. 6: **Hetero-isotopic CLMS of αSyn dimer.**
Grey – intra-molecular cross-links; Black – inter-molecular cross-links. (**A**) BS[3] with Trypsin digestion. (**B**) DMTMM with either Trypsin or GluC digestion

We conduct three different experiments on an equimolar mixture of either ^{14}N-^{14}N, ^{15}N-^{15}N, or ^{14}N-^{15}N α-Syn with a total concentration of 2.5 µM. For the experiment we used two bifunctional cross-linkers: i) the homo-bifunctional amine-reactive linker BS[3] with <30Å C_α-C_α distance, and ii) the hetero-bifunctional amine-carboxyl reactive coupler DMTMM, a zero-length linker <16Å C_α-C_α distance[42]. We prepared the sample by standard preparation protocols[43], digested the proteins with trypsin or GluC, which cleaves after lysine or after glutamate and

aspartate, respectively. We then perform the mass spectrometry analyses (see *Methods*).

Using BS^3, we identify a total of 121 cross-links, of which 103 are homo-isotopic and 18 are hetero-isotopic (see *supplementary spreadsheet CL-MS*). The estimated false discovery rate (FDR) is 5% (see methods). The majority of BS^3 inter-molecular cross-links are between the NTD and the NAC segments, due to the lack of lysines in the CTD (Fig. 6A). To gain better cross-link coverage of the C-terminal we apply DMTMM and obtain 23 inter-molecular cross-links (Fig. 6B).

To gain short range spatial information, potentially regarding proximities to the CTD segment, we use DMTMM and digest using GluC. However, we could not identify any inter-molecular cross-link with the CTD segment (Fig. 6B). This indicates that there are no short-range proximities between one subunit and the CTD segment of the other subunit in the dimer. In contrast, we identify intra-molecular cross-links between the CTD segment and the other segments within the subunit.

Remarkably, the pattern of the inter-molecular interactions between the NAC and NTD segments within the dimers captured by cross-linking include interactions between the NTD segment of one subunit and the NAC segment of the other subunit. These results indicate a specific fold is expected between the NTD and NAC segments in the dimer.

In addition to understanding the residue-residue proximities in the homodimer, gained from CL-MS information, we are able to acquire multiple non-redundant inter-molecular spatial restraints. The CL-MS attained spatial restraints are sufficient to direct the experimentally-restrained DMD simulations to form dimers in the simulation and to result in structure models, which exhibit partial agreement with the experimental data, and ones not biased by the cross-linking distance ranges (see restraints designed for DMD simulations in Fig. S5).

Experimentally-restrained DMD simulations: structure models of αSyn dimer conformations

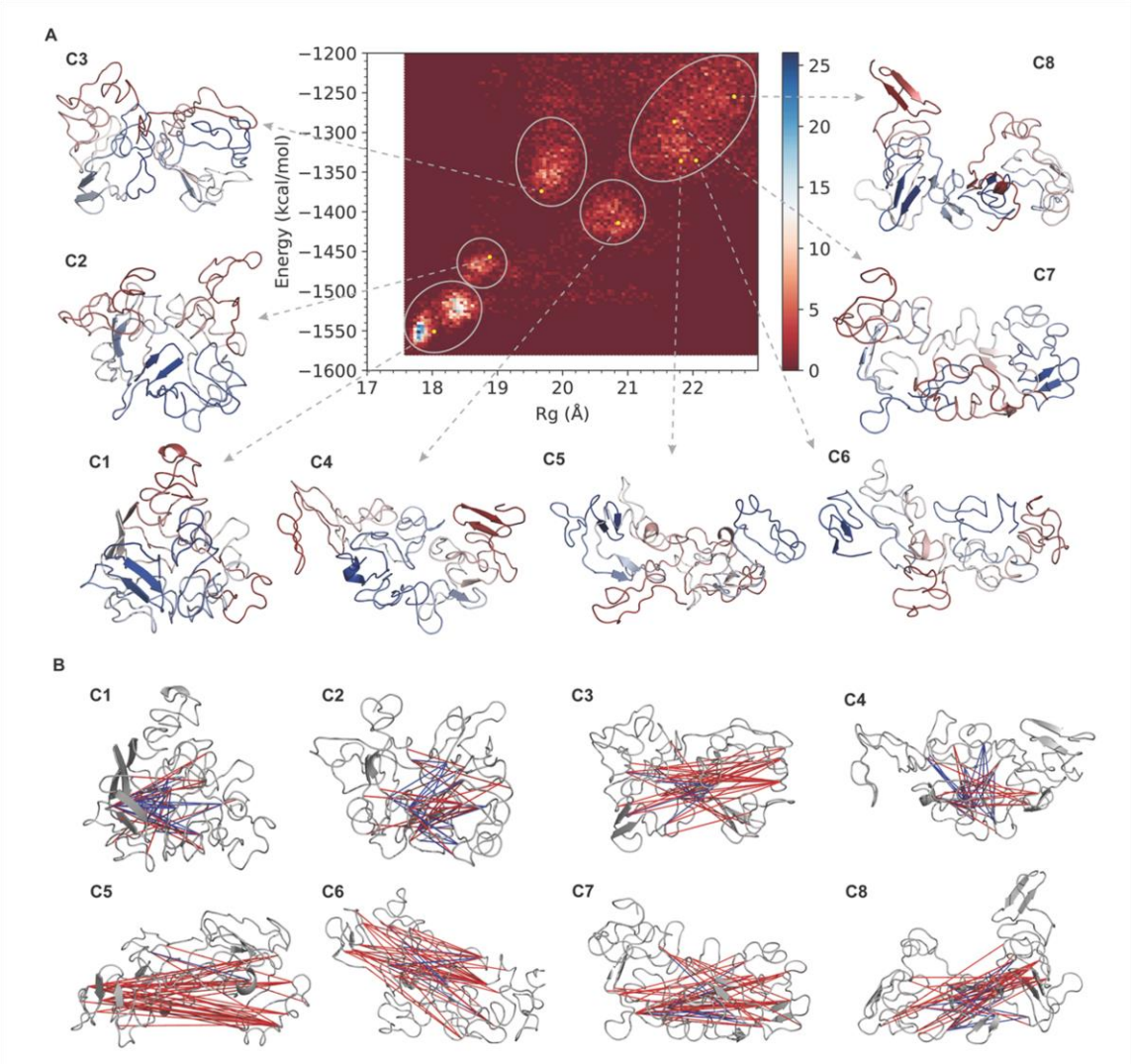

Fig. 7: **The eight structural sub-populations or clusters of the αSyn dimer predicted by cross-linking-restrained DMD simulations.**
(**A**) The hexbin plot of energy versus radius of gyration (R_g) for simulated dimer structures and the centroid structure of each cluster (C1-C8). The structures are marked from blue (NTD segment) to red (CTD segment). (**B**) The centroid structures of the eight clusters. Each line in the structure indicates the inter-molecular cross-links used for DMD simulations. Blue line means that over 50% of the structures within the corresponding cluster are validated by this constraint, while red line implicates that this constraint is satisfied by less than 50% of the structures for each cluster. The exact values of the levels of agreement of structure models with cross-linking data can be found in table S4.

We elucidate αSyn dimer structure models by cross-linking restraint-guided DMD simulations (for details, see *Methods*). To analyze the dimer structures, we choose the most-populated conformations and perform agglomerative hierarchical clustering (see *Methods*), which divides the structural ensemble into eight clusters

of different compactness, as judged by radii of gyration (R_g), and different energies (Fig. 7A). Structures in each cluster are highly similar as the pairwise root mean square deviation (RMSD) in each cluster is <9 Å (Table S2). The RMSD between clusters (Table S3) is >14 Å, indicating that structures among different clusters are distinct from each other. Then, we extract the dimer structure models and validate them against the cross-linking data (Fig. 7B). Specifically, we calculate the fraction of structures within each cluster, which satisfy the inter-molecular cross-linking restraints (Table S4; cluster-based distance distribution histograms deposited over Zenodo[44]). If a majority of structures in a cluster exhibit the inter-residue distances values within the C_α-C_α distance range the cross-linkers cover (BS^3 with C_α-C_α distance <30 Å; DMTMM with C_α-C_α distance <16 Å)[42], these cross-links are confirmed as agreeing with the structure model of the cluster. All 121 intra-molecular and 36 inter-molecular cross-links are tested against our dimer structure models (almost all intra-molecular cross-links agree well with the structural data; we focus on the inter-molecular cross-linking data). Following this, clusters 1, 2, and 4 contain more structures, which are validated by inter-molecular cross-links, than the other clusters (Fig. 7B). That is because we design the restraints to energetically favor distances covered by the cross-linkers, but not to abolish potential longer distances, and hence not to bias the simulations.

The results of our smPIFE experiments identify several subpopulations of the αSyn dimer (Fig. 4). To investigate if the simulation-predicted dimer structures include different subpopulations or not, we calculate the solvent accessible surface area (SASA) of the accessible volumes of Cy3 when it is attached to residues 26, 39, 56, and 140 within each monomer for all the eight clusters. The higher the SASA value is, the more accessible is the residue. Importantly, clusters with no histograms point to residues that are fully buried and inaccessible. Residue 26 has three subpopulations in both αSyn subunits (Fig. S6). Residue 39 contains three subpopulations with different degrees of accessibility (SASA values centered at ~1,000, 1,500, and 2,500) in chain A and B (Fig. S7). Residue 56 exhibits two SASA subpopulations in one subunit and three in the other subunit (Fig. S8). Residue 140 exhibits three SASA subpopulations in either one of the subunits (Fig.

S9). Overall, the different subpopulations represented by our simulated dimer structures are in accordance with the smPIFE data.

Two tyrosine residues could form a dityrosine covalent linkage, if they are spatially close enough and if their hydroxyl groups are radicalized. This bond formation stabilizes homo-oligomeric αSyn complexes. Moreover, a Y39-Y39 dityrosine linkage has previously been reported as occurring in αSyn and stabilizing its later

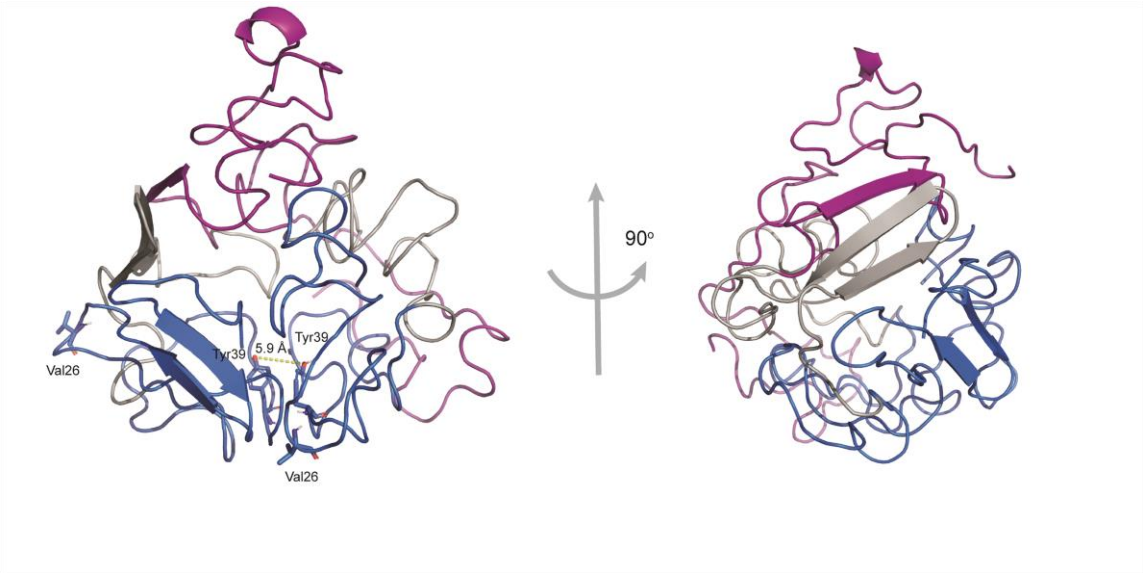

Fig. 8: **The cartoon representation of the centroid structure of cluster 1**.
The NTD, NAC, and CTD segments are highlighted in blue, grey, and magenta, respectively. Note that while the NTD and NAC segments of both subunits in close proximity, the CTD segments of both subunits are far from each other, in line with the hetero-isotopic CL-MS results in Fig. 6. Tyrosine 39 of both subunits are shown, with the short proximity between their hydroxyl groups. Additionally, it is clear that Tyrosine 39 is not accessible and buried, and is well characterized by the high mean nanotime sub-population for this residue in Fig. 4. Residue 26, one of the other hand is exposed and accessible in both subunit, and hence is well characterized by the low mean nanotime sub-population for this residue in Fig. 4.

oligomeric species, some of which in the aggregation *on-pathway*, potentially acting as stable seeds of aggregation[17,19]. Another dityrosine covalent linkage has been reported for αSyn in Y125-Y125[18]. To find out if αSyn dimers have pairs of tyrosines from the two subunits that are within interaction distances to form this covalent bond, we calculate the distance distributions between hydroxyl groups of all permutations of pairs of such tyrosine groups (e.g., involving tyrosine residues 39, 125, 133 and 136) in each of the clusters. Out of these permutations, the only pair of tyrosines, which exhibits the short distances required for interaction, is the Y39-Y39 pair, and only in cluster 1 (Fig. S11). The distances between hydroxyl

groups of tyrosine residues in cluster 1 are significantly smaller than those in the other clusters, with >50% of its structures exhibiting distances <6 Å. These dimer structures have the potential to form a covalent bond between tyrosine residues, if at least one of their hydroxyl groups have undergone radicalization, which can covalently stabilize the dimer structure. Therefore, cluster 1 is of particular interest. In addition, cluster 1 has the lowest energy compared to the other clusters (Fig. 7A) and is the most compact (Fig. 7A), the most populated cluster and it is the cluster that exhibits the highest level of agreement with the CL-MS data. The NTD segment as well as the NAC and CTD segments in one subunit of the centroid of cluster 1 form β-sheet structures (Fig. 8), with a partial degree of surface exposure in both NTD and NAC segments (Fig. S10). Therefore, out of all structure models of the αSyn, the one described by cluster 1 is of most interest.

Discussion

We performed AIEX-MALS, steady-state-, time-resolved- and FCS-FRET, to demonstrate a monomer-dimer equilibrium exists in αSyn at the concentration range of nM and until a few μM. Therefore, when probing a solution of freshly-prepared recombinant αSyn with a nominal concentration in a few μM, a significant number of dimers are formed. To gain insight into dimer structures, we perform hetero-isotopic CL-MS, which retrieves inter- and intra-molecular spatial information on the dimer, and use the cross-linking data as restraints in DMD simulations[45].

Our experimentally-restrained DMD simulations retrieves an ensemble of structure models clustered in eight structural sub-populations or clusters (C1-8). It is noteworthy that since the RMSD values within clusters C1-8, perhaps except for C4, are lower than 6 Å (Table S2), it is relatively safe to infer structural features from the inspection of the centroid structures of the retrieved clusters. The energies of the centroid structures of these clusters increase (Fig. 7A) as the number of amino acids in the dimerization interface decreases (Table S5). Additionally, This trend appears after analyzing the non-covalent interactions at the dimer interface, using the PPCheck[46] server (see Table S5 and *supplementary spreadsheet*

PPCheck). Among all the clusters, the centroid structure model of C1 has several important features being the most populated cluster in the simulation (17.3%) with the lowest energy (Fig. 7A), the lowest stabilizing energy (Table S5), and it appears to have the highest number of residues in the interface between the subunits (155 residues; Table S5). Furthermore, the centroid structure of C1 is the most compact (Fig. 7A). Overall, it seems that among all eight clusters, C1 is the most stable αSyn dimer species, with multiple inter-molecular interactions (Table S5). Interestingly, the structure of one subunit of the dimer in C1 exhibits partially buried NTD and NAC segments and a more solvent-exposed CTD segment. Additionally, the structure of the other subunit in the C1 dimer exhibits NTD and NAC segments that are even more buried than their counterparts in the first subunit, and again a relatively solvent-exposed CTD segment. This observation indicates that either C1 is not on the NAC-induced aggregation pathway, which requires solvent exposure of the NAC segment, or that C1 is an intermediate, where later changes in conformation, upon further oligomerization or covalent stabilization may expose the NAC. Additionally, β-sheet structures are found in this subunit of C1, in its NTD segment and in its NAC and CTD segments. These β-sheet structures of the stable and abundant dimer conformation C1 exhibit partial accessibilities, and therefore might serve in aggregation pathways alternative to the one promoted by the solvent-exposed NAC segment β-sheet structure.

Covalent stabilization of αSyn dimers has been previously reported to enhance the formation of stable aggregates and seeds of aggregation[16,19]. Specifically, the formation of Y39-Y39 dityrosine linkages between two subunits was implicated in that respect[17]. However, other dityrosine covalent linkages, such as Y125-Y125 were also implicated[18]. We examine the inter-molecular distances between the hydroxyl groups of all tyrosine pairs against the structures of each cluster out of the eight identified clusters. Only one cluster, in only one tyrosine pair, Y39-Y39, exhibits distances within interaction distances (Fig. S11), which upon hydroxyl radicalization can generate such dityrosine linkages. Such tyrosine hydroxyl radicalization can occur as a result of oxidative or nitrative stress. All the characteristics of C1 raise the question of whether the dimer will stay in this

conformation, or will it undergo conformational changes after dityrosine linkages covalently stabilize the dimer in such a way that the NAC will later be fully exposed to the solvent to promote aggregation.

C2 is low in the energy, however not as low as C1 (Fig. 7A), and its centroid structure includes fewer residues in the dimerization interface than C1 (Table 5). Additionally, the centroid structure of C2 is less compact than that of C1 (Fig. 7A). Like in C1, the centroid structure of C2 exhibits β-sheet structures, some of which include multiple long β strands (Table S6). C3 and C4 are higher in energy relative to C1 and C2 (Fig. 7A), their centroid structures include fewer residues within the dimerization interface (Table S5), and they are more expanded relative to the centroid structures of C1 and C2 (Fig. 7A). However, the centroid structure of C4 is expected to be more stable than that of C3, with lower total stabilizing energy values (Table S5). None of the structure models of C2-4 exhibit Y39-Y39 proximity to accommodate the potential formation of a dityrosine covalent linkage.

The existence of multiple αSyn dimer conformations is warranted by the existence of multiple mean nanotime sub-populations in time-resolved smPIFE experiments (Fig. 4), signifying the existence of dimers with different local structures in the vicinity of the labeling sites[47]. Moreover, these different local-structure related sub-populations undergo structural changes at times slower than single-molecule burst durations, hence at milliseconds or slower. Nevertheless, it is noteworthy that at the moment we cannot exclude the possibility that each mean nanotime sub-population is the result of time-averaging of transitions faster than typical burst duration timescales, better known as *within-burst dynamics*. We have recently developed an analysis approach such dynamics in single-molecule FRET experiments[48], and will soon extend this approach to analyze time-resolved smPIFE within-burst dynamics.

Although the restraints derived from hereto-isotopic CL-MS play a pivotal role in the experimentally-restrained DMD simulations of the αSyn dimer, one major limitation is the inability to report whether the intra-molecular cross-links originate from an αSyn monomer or the subunit of the dimer, since both are found in

equilibrium. In that respect, it is noteworthy that the concentrations of αSyn used in hetero-isotopic CL-MS are ones, in which the fraction of the dimer is high, as can be seen from our chromatography coupled to MALS and inter-molecular FRET titration measurements. Additionally, the DMD restraints were designed to give less energetic weight for the intra-molecular cross-links and more for the inter-molecular cross-links (Fig. S7, compare panels A and B).

The structure models of the αSyn dimer contribute to the understanding of the oligomerization pathway of αSyn, which has been extensively studied[26,49–52]. Nevertheless, to date, the mechanism is not fully understood. To enhance our understanding of the first step in αSyn oligomerization, we studied the smallest known αSyn oligomer, which marks the first inter-molecular self-association encounter - the dimer. αSyn dimers have been previously studied using SDS-PAGE, atomic force microscopy (AFM)[25], and reporting on the formation of the dityrosine covalent linkage in dimers[18,19].

Altogether, we identify structural models of the dimer, some of which are interesting from the functional point of view, whether they are or are not in the NAC-inducing aggregation pathway. These structural models will be the input for a follow-up work, in which we intend to design point mutations that will selectively stabilize each of the dimer structures. Additionally, we intend to form the dityrosine covalent linkages to see if we indeed stabilize a dimer with the structure covered in C1. Then, we will test the functional routes of these αSyn variants and by that gain better understanding as to their potential function.

Acknowledgements

We would like to thank Drs. Asaf Grupi, Dan Amir, and Elisha Haas from the Mina & Everard Goodman Faculty of Life Sciences in Bar Ilan University for sharing the plasmids of αSyn bearing single cysteine mutations. We would like to thank Dr. Yuval Garini for inviting us to perform the bulk inter-molecular time-resolved- and FCS-FRET measurements on his experimental setup. Additionally, we would like to thank Dr. William Breuer from the Proteomics and Mass Spectrometry Unit in

the Alexander Silberman Institute of Life Sciences, the Hebrew University of Jerusalem, for assisting us with intact protein mass spectrometry. Finally, we would like to thank Drs. David Eliezer, Philipp Selenko, Hagen Hofmann, Yifat Miller, Deborah Toiber and Rina Rosenzweig for fruitful and insightful discussions. We acknowledge support from the National Institutes of Health (NIH) 1R35 GM134864 and the Passan Foundation (to N.V.D.). The project described was supported by the National Center for Advancing Translational Sciences, NIH, through grant UL1 TR002014. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH (to N.V.D.). In addition, this project was supported by the Israel Science Foundation (grant 1768/15 to N.K.; grant 3565/20 to E.L., within the KillCorona – Curbing Coronavirus Research Program), the NIH (grant R01 GM130942 to E.L. as a subaward), by the Milner Fund (to E.L.), and by the Hebrew University of Jerusalem (start-up funds to E.L.).

Author contributions

J.Z. performed hetero-isotopic CL-MS experiments, S.Z. and E.L. performed inter-molecular FRET and S.Z. performed time-resolved smPIFE experiments, with the assistance of P.D.H. in data analysis, S.Z., J.Z., P.D., M.L. and E.L. performed chromatography coupled to MALS experiments, J.C. and N.V.D. performed CL-MS-restrained DMD simulations. J.Z., S.Z., J.C., N.V.D. and E.L. wrote the paper and all co-authors assisted in refining it.

Competing Interests

The authors declare no competing interests.

Materials & Correspondence

The authors declare no competing interests.

Methods

Expression and purification of recombinant αSyn variants

The expression and production of recombinant αSyn variants (*wt* and Cys mutants given to us as a gift; see Acknowledgements) is performed following previous existing protocols[30,53]. Briefly, αSyn is expressed from a pT7-7 plasmid encoding either the WT or the single Cys mutant proteins. Plasmids are transformed into BL21(DE3) component cells. The purification of the *wt* and mutant αSyn proteins are done by cells lysis, followed by osmotic shock and amino-acid precipitation, ending with proteins precipitation by ammonium sulfate. To further separate and purify a pure monomeric αSyn, the protein samples are loaded into 1 mL MonoQ column using FPLC system. The eluted fractions are loaded into 12% SDS-PAGE gel, relevant fractions, which exhibit a single band with molecular mass of αSyn (Fig. S12) are unified and dialyzed by 8 kDa MWCO dialysis bags at 4 ºC against buffer A. Protein samples are stored a -20 ºC.

Chromatography coupled to MALS

AIEX- and SEC-MALS experiments are performed using the same setup as was previously reported[35,54]. Briefly, we use a miniDAWN TREOS multi-angle light scattering detector, with three angle detectors (43.6°, 90°, and 136.4°) and a λ=658.9 nm laser beam (Wyatt Technology, Santa Barbara, CA) with a Wyatt QELS dynamic light scattering module for determination of hydrodynamic radius and an Optilab T-rEX refractometer (Wyatt Technology) set in-line with either a 1 mL Mono-Q analytical column (GE, Life Science, Marlborough, MA) for AIEX-MALS or a Superdex 75 column (GE, Life Science, Marlborough, MA) for SEC-MALS. Experiments are performed using an Äkta Pure M25 system with a UV-900 detector (GE) adapted for analytical runs. All AIEX-MALS experiments are performed at room temperature (25 °C) at 1.5 mL/min, using 30 mM Tris-HCl buffer pH = 8, as equilibration buffer, and 30 mM Tris-HCl buffer pH = 8 with 500 mM NaCl as elution buffer. 1 mg of αSyn at 0.2 mg/mL are injected to the column. Wash with 20 column volumes (CVs) of 0% elution buffer and 7 CVs of 30% elution buffer and eluted by 30 CVs gradient of 30-100% elution buffer collecting 1 mL fractions. Data collection and AIEX-MALS analysis are performed using the

ASTRA 6.1 software (Wyatt Technology). The refractive index (RI) of the solvent is defined as 1.331, and the viscosity is defined as 0.8945 cP (common parameters for PBS buffer at λ=658.9 nm). dn/dc (the RI increment with protein concentration) value for all samples is defined as 0.185 mL/g (a standard average value for proteins). The molecular mass is calculated by both the absorption at λ=280 nm and RI. We use blank baseline reduction for calculations according to RI. We identify similar molecular mass values using the two types of calculations, and present the molecular mass values calculated using the absorption at λ=280 nm.

αSyn dye labeling

αSyn dye labeling with either ATTO 488, ATTO 647N (ATTO-TEC, GmbH) or sulfo-Cy3, linked to either iodoacetamide or maleimide thiol-reactive groups, are coupled specifically to single Cys residues in αSyn mutants via a thiol coupling reaction, as previously described[30].

Steady-state FRET measurements

Acceptor excitation spectra are recorded using a spectrofluorometer (Jasco FP8200ST, Japan), scanning excitation wavelength in the range 400-650 nm and focusing on acceptor emission at λ=665 nm. Acceptor excitation spectra are normalized to the excitation spectrum part, in the wavelength range for direct acceptor excitation (550-650 nm; see Fig. S4). Then, the excitation spectra part in the wavelength range that covers donor excitation (440-550 nm; see Fig. S4) serves as a reporter for FRET.

Time-resolved FRET and FCS-FRET measurements

Measurements are performed on a confocal fluorescence microscope (MicroTime200, PicoQuant GmbH Berlin, Germany) assembled on top of a modified Olympus IX71 inverted microscope stand, the same as previously reported[55]. Briefly, pulsed picosecond diode lasers are used for the excitation of the donor and acceptor (LDH-P-C-470B and LDH-P-C-640B, respectively, PicoQuant GmbH Berlin Germany) with output wavelengths of 470 and 640 nm, respectively (60 and 40 μW average-power at the sample, respectively). The

conservation of polarization modes before reaching the sample is verified for both laser sources. Bandpass filters ensured that only light within the desired excitation band reached the sample. The repetition rate is set to 20 MHz. A dichroic beam splitter with high reflectivity at 470 and 635 nm reflects the light through the optical path to a high numerical aperture (NA) super Apo-chromatic objective (60X, NA=1.2, water immersion, Olympus, Japan), which focuses the light onto a small confocal volume. Fluorescence from the excited molecules is collected through the same objective and focused with an achromatic lens (f = 175 mm) onto a 50 µm diameter pinhole. The fluorescence emissions of the donor and the acceptor are separated using a dichroic long pass filter with the dividing edge at 605 nm (FF605-Di01, Semrock Rochester NY, USA). Emission that passes through the donor channel is selected with a 520/35 nm band-pass filter (FF01-520/35-25, Semrock Rochester NY, USA); the emission passed to the acceptor channel is selected by a 690/70 nm band-pass filter (HQ690/70, Chroma Technology Corp. Rockingham Vermont, USA). The detectors used are single-photon avalanche photodiodes (SPAD; Donor Channel: 100 µm, MPD PD1CTC, Acceptor Channel: 170 µm, Perkin Elmer SPCM-AQRH 13). The data is correlated by the HydraHarp400 time correlated single photon counting (TCSPC) system and collected with SymphoTime version 5 (both by PicoQuant GmbH Berlin, Germany). The fluorescence decay curves are recorded as TCSPC histograms with bin widths of 16 ps.

Recorded data is stored in the time-tagged time-resolved (TTTR) format, allowing each detected photon to be recorded together with its individual timing relative to the beginning of the measurement, and relative to the time of the excitation pulse, as well as the detection channel. The presented measurements are performed at an approximate depth of 70 µm inside a drop of solution, enclosed between two glass coverslips, sealed by a silicon double sticky sheet (Sigma). The total acquisition times are several minutes.

smPIFE measurements

All measurements are performed on the same setup and are analyzed exactly as recently explained[30], just in the presence of varying *wt*-αSyn concentrations.

Intact protein mass determination

We characterize the molar mass of the purified αSyn by intact protein mass spectrometry and the results report on the exact molecular mass of the αSyn (Fig. S13). The procedures is hereby described: Proteins are dissolved in 40 % acetonitrile, 0.3 % formic acid (all solvents were MS-grade) at a concentration of 2-5 mg/mL. Then, the dissolved proteins are injected directly via a HESI-II ion source into a Q Exactive Plus (Thermo Fisher Scientific) mass spectrometer and a minimum of three scans lasting 30 seconds are obtained. The scan parameters are: i) scan range 1,800 to 3,000 m/z without fragmentation; ii) resolution 140,000; iii) positive polarity; iv) AGC target 3×10^6; v) maximum inject time 50 ms; vi) spray voltage 4.0 kV; vii) capillary temperature 275 °C; and viii) S-lens RF level 55. Scan deconvolution is done using Mag Tran version 1.0.3.0 (Amgen).

Hetero-isotopic CL-MS

Testing the ^{15}N coverage in ^{15}N-αSyn

We digest ^{14}N- and ^{15}N-αSyn separately and observe the peptides. We examine the spectrum of a specific peptide's mass in the ^{15}N-αSyn and find that the mass precisely equals the mass of ^{14}N-αSyn of the exact same peptide plus ~1 Da multiplied the number Nitrogen atoms present in the peptide.

If the labeling is not full, we will obtain overlapping between the series of the peaks with a difference in the mass that is corresponding to the number of nitrogen atoms that are not labeled in the peptide. The analysis takes into consideration the ^{13}C isotope series and in each peptide, the expected mono-isotopic mass of the examined peptides is the first peak in the isotopic series without preceding peaks that could be unlabeled nitrogen atoms.

Cross-linking of α-synuclein with BS3

Equimolar ratio mixtures of ^{14}N, ^{15}N (AlexoTech) or ^{14}N-^{15}N α-Syn are prepared to a final concentration of 2.5 μM. We dissolve BS^3 powder (Sigma Aldrich) in HEPES buffer (pH 7.18) to a concentration of 10 mM. We add the prepared BS^3 to each of the α-syn mixtures to a final concentration of 1 mM and incubated at 4 °C for 1.5 h with shaking at 600 rpm. We quenched the cross-linking reaction by the addition of 20 mM ammonium bicarbonate for 20 min[56,57].

Cross-linking of α-synuclein with DMTMM

Equimolar ratios mixtures of ^{14}N, ^{15}N (AlexoTech) or ^{14}N-^{15}N α-Syn are prepared to a final concentration of 2.5 μM. We dissolve DMTMM powder (Sigma Aldrich) in HEPES buffer (pH 7.18) to a concentration of 70 mM. We add the prepared DMTMM to each of the α-syn mixtures to a final concentration of 7 mM and incubated at 4 °C for 1.5 h with shaking at 600 rpm. We quenched the cross-linking reaction by the addition of 20 mM ammonium bicarbonate for 20 min[56,57].

Preparation of samples for mass spectrometry

The proteins are precipitated in 1 mL of acetone (-80 °C) for one hour, followed by centrifugation at 14,000 g. The pellet is resuspended in 20 μL of 8 M urea. The urea is diluted by adding 200 μL of digestion buffer (25 mM TRIS, pH=8.0; 10% Acetonitrile). We add 0.5 μg of trypsin (Promega) or 0.5 μg GluC (Promega) for the DMTMM to the diluted urea and digest the protein overnight at 37°C under agitation. Following digestion, the peptides are desalted on C18 stage-tips and eluted by 55% acetonitrile. The eluted peptides are dried in a SpeedVac, reconstituted in 0.1% formic acid, and measured in the mass spectrometer[58].

Mass spectrometry

The samples are analyzed by a 120 minute 0-40% acetonitrile gradient on a liquid chromatography system coupled to a Q-Exactive HF mass spectrometer. The analytical column is an EasySpray 25 cm heated to 40°C. The method parameters of the run are: i) Data-Dependent Acquisition; ii) Full MS resolution 70,000 ; iii) MS1 AGC target 1e6 ; iv) MS1 Maximum IT 200 ms ; v) Scan range 450 to 1800 ; vi) dd-MS/MS resolution 35,000 ; vii) MS/MS AGC target 2e5 ; viii) MS2 Maximum

IT 600 ms ; ix) Loop count Top 12; x) Isolation window 1.1 ; xi) Fixed first mass 130 ; xii) MS2 Minimum AGC target 800 ; xiii) Peptide match - off ; xiv) Exclude isotope - on ; and xv) Dynamic exclusion 45 seconds. Each cross-linked sample is measured twice in two different HCD energies (NCE): 26, and stepped 25, 30, and 35. All cross-linked samples are measured with the following charge exclusion: unassigned,1,2,3,8,>8. Proteomics samples are measured with the following charge exclusion: unassigned,1,8,>8.

Identification of cross-links

The RAW data files from the mass spectrometer are converted to MGF format by Proteome Discoverer (Thermo). The identification of BS^3 or DMTMM cross-links uses a search application (FindXL: http://biolchem.huji.ac.il/nirka/software.html) that exhaustively enumerates all the possible peptide pairs. The search parameters are as follows: i) Sequence database- ^{14}N α-synuclein and ^{15}N α-synuclein that differs in the molecular weight of the amino acids by 1 Da multiplied by the number of Nitrogen in the amino acid; ii) Protease – trypsin or GluC, allowing up to three mis-cleavage sites; iii) Variable modifications: methionine oxidation, lysine with hydrolyzed mono-link; iv) Cross-linking must occur between two lysine residues; v) Cross-linker is not cleaved; vi) MS/MS fragments to consider: b-ions, y-ions; vii) MS1 tolerance – 6 ppm; viii) MS2 tolerance – 8 ppm; and ix) Cross-linker mass – one of three possible masses: 138.0681, 138.0681 + 1.00335, and 138.0681 + 2.0067 for the BS^3 and -18 Da for the DMTMM that represent the release of water molecule. The three masses address the occasional incorrect assignment of the mono-isotopic mass by the mass spectrometer.

Restrained DMD simulations

Preparation of restraints for DMD simulations from CL-MS results

BS^3 and DMTMM lead to the cross-linking of specific pairs of amino acids and cover these pairs if they were in a structure with C_α-C_α distances as long as 30 and 16 Å, respectively. To maximize the efficiency of using these results in structural modeling, we used the C_α-C_α distance of DMTMM, to design DMTMM-related

discrete potential functions as restraints, however for BS^3, we chose instead the inter-residue end-to-end distance to be within the distances covered by the length of the cross-linker (max. 11.1 Å). Unlike in rigid protein complexes, where the cross-linking data can be used as if only these distances occur within the complex, in a protein as flexible and dynamic as αSyn, inter-residue distances longer than the range covered by the cross-link are possible. We took this into account when designing discrete restraining potentials for DMD simulations, by allowing all inter-residue distances to occur, but with an energetic preference towards the range covered by the cross-linker (Fig. S5).

CL-MS-restrained DMD simulations and validations

Cross-linking-restrained discrete molecular dynamics (DMD) simulations were developed to solve protein structures successfully[59,60]. Based on the same strategy, we apply 121 intra-molecular and 36 inter-molecular cross-links to restrain DMD simulations of the αSyn dimer. Intra-molecular cross-links restrain the distance between residues within a single subunit, while inter-molecular cross-links restrain the distance between residues from two different subunits. We convert the cross-link distance ranges to square-well potential functions (Fig. S5), which are incorporated with the Medusa force field during simulations[61–66]. We start simulations by using two fully extended conformations of αSyn that are separate from each other. We conduct an all-atom replica exchange DMD simulation where 26 replicas with temperatures ranging from 0.400 to 0.800 kcal/(mol•k_B) are run for 4×10^6 time steps (50 ps per time step)[67].

After simulations, we validate the resulting dimer structure models by comparing them with CL-MS and smPIFE experimental data. We calculate the distances between cross-linked residues within all the simulated dimer structures using GROMACS and plot the distance distribution histogram for each cross-link. If the maximum length of the cross-linker falls within or above the distance distribution that indicates this cross-link is satisfied by our simulation. On the other hand, if the maximum length is smaller than the overall distance distribution that means this cross-link is not captured by our simulation results. Moreover, we calculate the

percentage of structures within each cluster that are validated by cross-linking constraints. To compare our dimer structure models with smPIFE data, we calculate the accessible volumes of residues 26, 39, 56, and 140 within each monomer for each one of the eight clusters using Python libraries MDTraj (https://www.mdtraj.org/1.9.5/index.html) and AvTraj (https://github.com/Fluorescence-Tools/avtraj). Then, we measure the solvent accessible surface area (SASA) of each volume using PyMOL (https://pymol.org/2/#download). The subpopulations indicated by the surface area of all the structures within each cluster are compared with the number of subpopulations identified by smPIFE experiments.

Analysis of αSyn dimer structure models

We extract all the dimer structure models from the 26 trajectories. We calculate the radius of gyration (R_g) of all the dimers and plot R_g versus energy. We perform agglomerative hierarchical clustering of the most populated structures using the ward linkage algorithm from TTClust[68]. We optimize the clustering cutoff until the average RMSD values between clusters are much larger than the average RMSD values within each cluster (Tables S2 and S3). We calculate the average SASA of each residue for all the clusters using the GROMACS SASA algorithm[69]. The SASA values are further divided by the maximum surface area of the corresponding residue to generate normalized SASA. A threshold of 40% is applied to determine the buried and exposed states[70,71].

Data Availability

The raw photon-HDF5[72] files of all smPIFE measurements, and the data analysis pipeline using the FRETbursts[73] software are available publicly as Jupyter Notebooks over Zenodo[74].

The mass spectrometry proteomics data have been deposited to the ProteomeXchange[75] Consortium via the PRIDE[76] partner repository with the dataset identifier PRIDE: PXD030299.

The models of the a-syn ensemble structure have been deposited in the PDB-Dev Database with the accession number PDB-Dev: PDBDEV_00000098[77].

References

1. Recchia, A. *et al.* α-Synuclein and Parkinson's disease. *FASEB J.* **18**, 617–626 (2004).

2. Burré, J. *et al.* α-Synuclein Promotes SNARE-Complex Assembly in Vivo and in Vitro. *Science (80-.).* **329**, 1663 LP – 1667 (2010).

3. Eliezer, D., Kutluay, E., Bussell, R. & Browne, G. Conformational properties of α-synuclein in its free and lipid-associated states11Edited by P. E. Wright. *J. Mol. Biol.* **307**, 1061–1073 (2001).

4. Uéda, K. *et al.* Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl. Acad. Sci.* **90**, 11282 LP – 11286 (1993).

5. Hashimoto, M. *et al.* The Role of NAC in Amyloidogenesis in Alzheimer's Disease. *Am. J. Pathol.* **156**, 734–735 (2000).

6. Ulmer, T. S., Bax, A., Cole, N. B. & Nussbaum, R. L. Structure and dynamics of micelle-bound human alpha-synuclein. *J. Biol. Chem.* **280**, 9595–603 (2005).

7. Ferreon, A. C. M., Gambin, Y., Lemke, E. A. & Deniz, A. A. Interplay of alpha-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5645–5650 (2009).

8. Bodner, C. R., Maltsev, A. S., Dobson, C. M. & Bax, A. Differential Phospholipid Binding of α-Synuclein Variants Implicated in Parkinson's Disease Revealed by Solution NMR Spectroscopy. *Biochemistry* **49**, 862–871 (2010).

9. Bartels, T., Choi, J. G. & Selkoe, D. J. α-Synuclein occurs physiologically as

a helically folded tetramer that resists aggregation. *Nature* **477**, 107–110 (2011).

10. Wang, W. *et al.* A soluble α-synuclein construct forms a dynamic tetramer. *Proc. Natl. Acad. Sci.* **108**, 17797 LP – 17802 (2011).

11. Binolfi, A., Theillet, F.-X. & Selenko, P. Bacterial in-cell NMR of human α-synuclein: a disordered monomer by nature? *Biochem. Soc. Trans.* **40**, 950 LP – 954 (2012).

12. Longhena, F., Faustini, G., Spillantini, M. G. & Bellucci, A. Living in Promiscuity: The Multiple Partners of Alpha-Synuclein at the Synapse in Physiology and Pathology. *International Journal of Molecular Sciences* vol. 20 (2019).

13. Wördehoff, M. M. & Hoyer, W. α-Synuclein Aggregation Monitored by Thioflavin T Fluorescence Assay. *Bio-protocol* **8**, e2941 (2018).

14. Levitan, K. *et al.* Conserved C-terminal charge exerts a profound influence on the aggregation rate of α-synuclein. *J. Mol. Biol.* **411**, 329–333 (2011).

15. Coelho-Cerqueira, E., Carmo-Goncalves, P., Pinheiro, A. S., Cortines, J. & Follmer, C. alpha-Synuclein as an intrinsically disordered monomer--fact or artefact? *FEBS J.* **280**, 4915–4927 (2013).

16. Pivato, M. *et al.* Covalent α-Synuclein Dimers: Chemico-Physical and Aggregation Properties. *PLoS One* **7**, e50027 (2012).

17. Souza, J. M., Giasson, B. I., Chen, Q., Lee, V. M.-Y. & Ischiropoulos, H. Dityrosine Cross-linking Promotes Formation of Stable α-Synuclein Polymers: Implications of Nitrative and Oxidative stress in the Pathogenesis of Neurodegenerative Synucleinopathies. *J. Biol. Chem.* **275**, 18344–18349 (2000).

18. Takahashi, T., Yamashita, H., Nakamura, T., Nagano, Y. & Nakamura, S. Tyrosine 125 of α-synuclein plays a critical role for dimerization following nitrative stress. *Brain Res.* **938**, 73–80 (2002).

19. van Maarschalkerweerd, A. *et al.* Formation of covalent di-tyrosine dimers in recombinant α-synuclein. *Intrinsically Disord. proteins* **3**, e1071302–e1071302 (2015).

20. Gaeta, A. & Hider, R. C. The crucial role of metal ions in neurodegeneration: the basis for a promising therapeutic strategy. *Br. J. Pharmacol.* **146**, 1041–1059 (2005).

21. Tsigelny, I. F. *et al.* Dynamics of α-synuclein aggregation and inhibition of pore-like oligomer development by β-synuclein. *FEBS J.* **274**, 1862–1877 (2007).

22. Illes-Toth, E., Ramos, M. R., Cappai, R., Dalton, C. & Smith, D. P. Distinct higher-order α-synuclein oligomers induce intracellular aggregation. *Biochem. J.* **468**, 485–493 (2015).

23. Mane, J. Y. & Stepanova, M. Understanding the dynamics of monomeric, dimeric, and tetrameric α-synuclein structures in water. *FEBS Open Bio* **6**, 666–686 (2016).

24. Lv, Z. *et al.* Direct Detection of α-Synuclein Dimerization Dynamics: Single-Molecule Fluorescence Analysis. *Biophys. J.* **108**, 2038–2047 (2015).

25. Zhang, Y. *et al.* High-speed atomic force microscopy reveals structural dynamics of α-synuclein monomers and dimers. *J. Chem. Phys.* **148**, 123322 (2018).

26. Li, X. *et al.* Early stages of aggregation of engineered α-synuclein monomers and oligomers in solution. *Sci. Rep.* **9**, 1734 (2019).

27. Kayed, R., Dettmer, U. & Lesné, S. E. Soluble endogenous oligomeric α-synuclein species in neurodegenerative diseases: Expression, spreading, and cross-talk. *J. Parkinsons. Dis.* **10**, 791–818 (2020).

28. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc. Natl. Acad. Sci.* **108**, 7414 LP – 7418 (2011).

29. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. *Chem. Soc. Rev.* **43**, 1221–1229 (2014).

30. Zaer, S. & Lerner, E. Utilizing Time-Resolved Protein-Induced Fluorescence Enhancement to Identify Stable Local Conformations One α-Synuclein Monomer at a Time. *JoVE* e62655 (2021) doi:doi:10.3791/62655.

31. Chen, J. *et al.* The structural heterogeneity of α-synuclein is governed by several distinct subpopulations with interconversion times slower than milliseconds. *Structure* (2021) doi:https://doi.org/10.1016/j.str.2021.05.002.

32. Walker, R. G. *et al.* The Structure of Human Apolipoprotein A-IV as Revealed by Stable Isotope-assisted Cross-linking, Molecular Dynamics, and Small Angle X-ray Scattering*. *J. Biol. Chem.* **289**, 5596–5608 (2014).

33. Lima, D. B. *et al.* Characterization of homodimer interfaces with cross-linking mass spectrometry and isotopically labeled proteins. *Nat. Protoc.* **13**, 431–458 (2018).

34. Gaber, A., Gunčar, G. & Pavšič, M. Proper evaluation of chemical cross-linking-based spatial restraints improves the precision of modeling homo-oligomeric protein complexes. *BMC Bioinformatics* **20**, 464 (2019).

35. Amartely, H., Avraham, O., Friedler, A., Livnah, O. & Lebendiker, M. Coupling Multi Angle Light Scattering to Ion Exchange chromatography (IEX-MALS) for protein characterization. *Sci. Rep.* **8**, 6907 (2018).

36. Elson, E. L. & Magde, D. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers* **13**, 1–27 (1974).

37. V, Buschmann; Kapusta, P; Koberling, F; Konig, M; Kramer, B; Ruttinger, S; Tannert, S; Veiga, M; Erdmann, R; Ortmann, U. Fluorescence lifetime correlation spectroscopy (FLCS) - a powerful tool for measuring diffusion, concentrations and interactions. https://www.picoquant.com/images/uploads/downloads/webtalk_flcs.pdf

(2014).

38. Kolesnikova, O. *et al.* Molecular structure of promoter-bound yeast TFIID. *Nat. Commun.* (2018) doi:10.1038/s41467-018-07096-y.

39. Yang, Q. fen & Tang, C. On the necessity of an integrative approach to understand protein structural dynamics. *J. Zhejiang Univ. Sci. B* **20**, 496–502 (2019).

40. Na, S. & Paek, E. Computational methods in mass spectrometry-based structural proteomics for studying protein structure, dynamics, and interactions. *Comput Struct Biotechnol J* **18**, 1391–1402 (2020).

41. Koukos, P. I. & Bonvin, A. M. J. J. Integrative Modelling of Biomolecular Complexes. *J. Mol. Biol.* **432**, 2861–2881 (2020).

42. Leitner, A. *et al.* Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc. Natl. Acad. Sci.* **111**, 9455 LP – 9460 (2014).

43. Slavin, M. & Kalisman, N. Structural analysis of protein complexes by cross-linking and mass spectrometry. in *Methods in Molecular Biology* (ed. Marsh, J. A.) vol. 1764 173–183 (Springer New York, 2018).

44. Lerner, E. *et al.* Additional analysis data of α-synuclein dimer. (2021) doi:10.5281/zenodo.5784461.

45. Dokholyan, N. V. Experimentally-driven protein structure modeling. *J. Proteomics* **220**, 103777 (2020).

46. Sukhwal, A. & Sowdhamini, R. Oligomerisation status and evolutionary conservation of interfaces of protein structural domain superfamilies. *Mol. Biosyst.* **9**, 1652–1661 (2013).

47. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chem. Soc. Rev.* **43**, 1221–1229 (2014).

48. Harris, P. D. *et al.* Multi-parameter photon-by-photon hidden Markov modeling. *bioRxiv* (2021) doi:10.1101/2021.04.08.439035.

49. Andreasen, M., Lorenzen, N. & Otzen, D. Interactions between misfolded protein oligomers and membranes: A central topic in neurodegenerative diseases? *Biochim. Biophys. Acta - Biomembr.* **1848**, 1897–1907 (2015).

50. Ingelsson, M. Alpha-Synuclein Oligomers-Neurotoxic Molecules in Parkinson's Disease and Other Lewy Body Disorders. *Front. Neurosci.* **10**, 408 (2016).

51. Salveson, P. J., Spencer, R. K. & Nowick, J. S. X-ray Crystallographic Structure of Oligomers Formed by a Toxic β-Hairpin Derived from α-Synuclein: Trimers and Higher-Order Oligomers. *J. Am. Chem. Soc.* **138**, 4458–4467 (2016).

52. Du, X., Xie, X. & Liu, R. The Role of α-Synuclein Oligomers in Parkinson's Disease. *Int. J. Mol. Sci.* **21**, (2020).

53. Grupi, A. & Haas, E. Segmental Conformational Disorder and Dynamics in the Intrinsically Disordered Protein α-Synuclein and Its Chain Length Dependence. *J. Mol. Biol.* **405**, 1267–1283 (2011).

54. Some, D., Amartely, H., Tsadok, A. & Lebendiker, M. Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *JoVE* e59615 (2019) doi:doi:10.3791/59615.

55. Lerner, E. *et al.* Preparation of homogeneous samples of double-labelled protein suitable for single-molecule FRET measurements. *Anal Bioanal Chem* **405**, 5983–5991 (2013).

56. Kalisman, N., Adams, C. M. & Levitt, M. Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl. Acad. Sci.* **109**, 2884–2889 (2012).

57. Tayri-Wilk, T. *et al.* Mass spectrometry reveals the chemistry of

formaldehyde cross-linking in structured proteins. *bioRxiv* 820779 (2019) doi:10.1101/820779.

58. Kaledhonkar, S., Fu, Z., White, H. & Frank, J. *Protein Complex Assembly: Methods and Protocols*. *Methods Mol Biol* vol. 1764 (2018).

59. Brodie, N. I., Popov, K. I., Petrotchenko, E. V, Dokholyan, N. V & Borchers, C. H. Conformational ensemble of native α-synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations. *PLoS Comput. Biol.* **15**, e1006859–e1006859 (2019).

60. Brodie, N. I., Popov, K. I., Petrotchenko, E. V., Dokholyan, N. V. & Borchers, C. H. Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. *Sci. Adv.* **3**, (2017).

61. Ding, F. & Dokholyan, N. V. Emergence of Protein Fold Families through Rational Design. *PLOS Comput. Biol.* **2**, e85 (2006).

62. Dokholyan, N. V. Studies of folding and misfolding using simplified models. *Curr. Opin. Struct. Biol.* **16**, 79–85 (2006).

63. Yin, S., Biedermannova, L., Vondrasek, J. & Dokholyan, N. V. MedusaScore: An Accurate Force Field-Based Scoring Function for Virtual Drug Screening. *J. Chem. Inf. Model.* **48**, 1656–1662 (2008).

64. Proctor, E. A., Ding, F. & Dokholyan, N. V. Discrete molecular dynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 80–92 (2011).

65. Shirvanyants, D., Ding, F., Tsao, D., Ramachandran, S. & Dokholyan, N. V. Discrete molecular dynamics: An efficient and versatile simulation method for fine protein characterization. *J. Phys. Chem. B* **116**, 8375–8382 (2012).

66. Wang, J. & Dokholyan, N. V. MedusaDock 2.0: Efficient and Accurate Protein–Ligand Docking With Constraints. *J. Chem. Inf. Model.* **59**, 2509–2515 (2019).

67. Ding, F., Tsao, D., Nie, H. & Dokholyan, N. V. Ab Initio Folding of Proteins with All-Atom Discrete Molecular Dynamics. *Structure* **16**, 1010–1018 (2008).

68. Tubiana, T., Carvaillo, J. C., Boulard, Y. & Bressanelli, S. TTClust: A Versatile Molecular Simulation Trajectory Clustering Program with Graphical Summaries. *J. Chem. Inf. Model.* **58**, 2178–2182 (2018).

69. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **16**, 273–284 (1995).

70. Stultz, C. M., White, J. V & Smith, T. F. Structural analysis based on state-space modeling. *Protein Sci.* **2**, 305–314 (1993).

71. Ullman, O., Fisher, C. K. & Stultz, C. M. Explaining the Structural Plasticity of α-Synuclein. *J. Am. Chem. Soc.* **133**, 19536–19546 (2011).

72. Ingargiola, A., Laurence, T., Boutelle, R., Weiss, S. & Michalet, X. Photon-HDF5: An Open File Format for Timestamp-Based Single-Molecule Fluorescence Experiments. *Biophys. J.* **110**, 26–33 (2016).

73. Ingargiola, A., Lerner, E., Chung, S. Y., Weiss, S. & Michalet, X. FRETBursts: An open source toolkit for analysis of freely-diffusing Single-molecule FRET. *PLoS One* **11**, e0160716–e0160716 (2016).

74. Lerner, E. *et al.* Time-resolved smPIFE data - alpha-Synuclein dimer. (2021) doi:10.5281/zenodo.5760034.

75. Deutsch, E. W. *et al.* The ProteomeXchange consortium in 2020: enabling ‘big data’ approaches in proteomics. *Nucleic Acids Res.* **48**, D1145–D1152 (2020).

76. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

77. Vallat, B., Webb, B., Westbrook, J. D., Sali, A. & Berman, H. M. Development of a Prototype System for Archiving Integrative/Hybrid Structure Models of Biological Macromolecules. *Structure* **26**, 894-904.e2 (2018).