HHS Public Access

Author manuscript

Proteins. Author manuscript; available in PMC 2025 January 01.

Published in final edited form as:

Proteins. 2024 July ; 92(7): 854–864. doi:10.1002/prot.26679.

Post-translational modification sites are present in hydrophilic cavities of alpha-synuclein, tau, FUS and TDP-43 fibrils: A molecular dynamics study

Noah Nathan Kochen[1,*], **Darren Seaney**[1], **Vivek Vasandani**[1], **Marguerite Murray**[1], **Anthony R. Braun**[1,‡], **Jonathan N. Sachs**[1,‡]

[1]Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455

Abstract

Hydration plays a crucial role in the refolding of intrinsically disordered proteins into amyloid fibrils; however, the specific interactions between water and protein that may contribute to this process are still unknown. In our previous studies of alpha-synuclein (aSyn), we have shown that waters confined in fibril cavities are stabilizing features of this pathological fold; and that amino acids that hydrogen bond with these confined waters modulate primary and seeded aggregation. Here, we extend our aSyn molecular dynamics (MD) simulations with three new polymorphs and correlate MD trajectory information with known post-translational modifications (PTMs) and experimental data. We show that cavity residues are more evolutionarily conserved than non-cavity residues and are enriched with post-translational modification sites (PTM). As expected, the confinement within hydrophilic cavities results in more stably hydrated amino acids. Interestingly, cavity PTM sites display the longest protein-water hydrogen bond lifetimes (HBL), three-fold greater than non-PTM cavity sites. Utilizing the deep mutational screen dataset by Newberry et. al. and the Thioflavin T aggregation review by Pancoe et. al. parsed using a fibril cavity/non-cavity definition, we show that hydrophobic changes to amino acids in cavities have a larger effect on fitness and aggregation rate than residues outside cavities, supporting our hypothesis that these sites are involved in the inhibition of aSyn toxic fibrillization. Finally, we expand our study to include analysis of fibril structures of tau, FUS, TDP-43, prion, and hnRNPA1; all of which contained hydrated cavities, with tau, FUS and TDP-43 recapitulating our PTM results in aSyn fibril cavities.

Corresponding Authors: brau0123@umn.edu and jnsachs@umn.edu.

Author Contributions

N.N.K., A.R.B., and J.N.S. conceived of and directed the study; N.N.K., V.V and D.S. built and simulated MD systems. N.N.K. and M.M. conducted MD trajectory analysis; N.N.K. conducted computational correlations to published PTMs, fitness, and aggregation datasets ; N.N.K., A.R.B., and J.N.S wrote and edited the manuscript.

*First author;

‡co-corresponding author

Conflict of Interest

The authors declare no competing interests.

Supplemental Material

Supplemental data includes summary table of structures simulated, summary table and plots of fibril cavity properties, summary table of PTMs included in analysis, fibril cavity renders, distribution of KD hydropathicity of mutations included in linear regressions and global RMSD throughout MD production runs.

Code Availability

Hydrogen bond lifetime analysis was conducted using a biexponential fitting script written in MATLAB R2018b. All scripts are available from the corresponding author on reasonable request. Statistical analysis was done in GraphPad Prism 9.

Keywords

Neurodegeneration; Amyloid Fibrils; Molecular Dynamics

Introduction

Amyloid fibril formation and deposition is correlated with the onset and progression of several neurodegenerative disorders including but not limited to Parkinson's, Alzheimer's, and prion disease[1–6]. Fibrils from various proteins associated with neurodegeneration have been found in post-mortem brain histological samples and are neurotoxic when added exogenously in both cell and animal models[7–10]. Recent studies have established the connection between fibril polymorphism—the variety of conformations that arise from a unique or similar protein sequence—and pathology[11,12]. One example of this are the unique tau fibril conformations identified from brain samples from patients with chronic traumatic encephalopathy, corticobasal degeneration, progressive supranuclear palsy and argyrophilic disease[11].

Despite the vast evidence linking amyloid fibrils to pathology, few studies have characterized the differences and similarities across fibril structures from different protein sources and diseases using molecular dynamics (MD). Sawaya et al's seminal article reviewed structures and summarized common amyloid motifs, folding patterns and predicted stabilities based on calculated solvation energies from the structures alone[13]. This study shed light on the commonalities between amyloid folds and the possibility of investigating them using computational methods. Recently, using MD simulations, we showed that waters that infiltrate cavities in both the NAC nanocrystal and Greek-key fold of aSyn are highly stabilizing features that play a crucial role in maintaining the amyloid conformation[14]. Furthermore, we demonstrated this effect in a set of experiments with aSyn fibrils, where we found that the amino acids that interact with cavity waters (most notably threonine 72) are capable of eliminating monomer seeding onto wildtype (WT) fibrils and enhancing primary fibril growth extent when mutated to more hydrophobic residues[15]. Other groups show that a subset of threonine residues that are located within the cavities of these fibril structures are actively O-GlcNACylated in the human brain, suggesting a role of modifying enzymes on the modulation of aSyn conformation and aggregation propensity. Indeed, the recent Parkinson's disease familial mutant T72M demonstrates both loss of post-translational modification (PTM) capability and enhanced aggregation relative to WT aSyn, leading to pathology[16–18]. Interestingly, cavities have been observed in many other amyloid structures (not just aSyn) and its been suggested that these are packing defects that can be alleviated by ligands[19]. The presence of PTM sites within fibril cavities suggests an alternative hypothesis, whereby modification of local hydration at cavity sites can lead to the modulation of conformation, aggregation propensity and toxicity.

In a seminal paper by Newberry et. al.[20], they performed a deep mutational scan of full-length human aSyn, in which every single amino acid in the WT sequence was mutated to the other 19 natural amino acids, revealing the full sequence-fitness landscape of aSyn. Using an unbiased, model-independent approach, the periodicity in the sequence-fitness

landscape suggested an alpha-helical structure-fitness relationship[20]. The publishing of this thorough dataset provides the framework for the investigation of alternative structure-fitness relationships; allowing us to test our hypothesis that cavity sites are more sensitive to chemical modifications (probed via mutagenesis) than non-cavity sites.

Using MD simulations of 8 aSyn fibril structures and published datasets, we tested the hypothesis that amino acids in aSyn fibril cavities are enriched with PTMs and regulate its aggregation and toxicity profile; where we show that cavities contain functionally relevant PTM sites such as O-GlcNAc and are more sensitive to hydrophobic changes that lead to higher toxicity and aggregation rates. We then simulated 5 additional neurodegenerative disease-associated proteins and show that the cavity enrichment of PTM extends to other amyloid fibrils, suggesting a common role of PTMs in highly hydrated fibril cavity sites. In total, we simulated 20 distinct fibril protofilaments from 6 unique proteins (aSyn, tau, FUS, hnRNPA1, TDP-43 and prion protein)[21–36] for a total of 6 μs of all-atom explicitly solvated simulations.

Results

aSyn fibril cavities are present across 8 distinct polymorphs

In our recent paper, we determined the presence of hydrophilic fibril cavities in the recombinant rod (polymorph 1a, PDB: 6CU7) , twister (polymorph 1b, PDB: 6CU8) and A53T (PDB: 6LRQ) fibril structures, as well as the brain-derived MSA Type IA and IB polymorphs (PDB: 6XYO)[15]. Here, we expanded our simulations to three new recombinant aSyn fibril structures: H50Q (PDB: 6PES) and E46K (PDB: 6UFR) mutant fibrils, and the WT aSyn polymorph 2b (PDB: 6SST). As shown in Figure 1, we detected hydrophilic cavities in the new aSyn polymorphs (Figure 1F–H) as indicated by the volumes in gray, green and red, which were identified using the last 25 ns of simulation with the Analysis of Null Areas (ANA2) package developed by Barletta et. al[37].

Amino acids in fibril cavities of aSyn are evolutionarily conserved and contain PTM sites

The persistence of these hydrophilic cavities across aSyn fibril polymorphs from multiple genotypes (WT, A53T, H50Q and E46K) suggests that this motif is conserved. Using the ConSurf-DB server[38,39] (which determines evolutionary conservation using sequence homologues from several species), we tested this hypothesis by comparing the degree of evolutionary conservation of amino acids inside and outside cavities of the 8 distinct aSyn fibrils (Figure 2A). The average conservation score of amino acids in aSyn cavities is approximately ~14% higher than amino acids outside of cavities (~5.6 to ~6.4).

Evolutionarily conserved motifs of mammalian proteins have been shown to contain PTM sites that regulate function and stability; and are capable of modulating disease phenotypes when mutated[40]. Thus, we proceeded to map known aSyn PTM sites to the resolved residues in the fibril structures to determine whether residues within these cavities are enriched in PTMs. All tabulated PTMs are summarized in Supplemental Table T1 and were identified by literature mining of published reviews (a list of pathogenic mutations obtained from the NCBI ClinVar Database and reviews is also included)[18,41–50]. For each structure, we

calculated the baseline percentage of residues that are PTM sites, which accounted for ~27% of fibril residues. When parsed by cavity or non-cavity residues, ~33% of cavity residues are PTM sites, whereas only ~20% of non-cavity residues are PTM sites (Figure 2B). Notably, we observe PTMs such as ubiquitination (K43), O-GlcNAcylation (T54, T59, T64, T72, T81, S87) phosphorylation (S87) and truncation (K58, V74, K80, G84, A89) in cavity sites.

PTM sites within aSyn fibril cavities have largest hydrogen bond lifetimes to solvent water

One of the most notable PTM sites within aSyn fibril cavities is T72 (present in cavities across all structures except polymorph 2b, PDB: 6sst), which is known to be O-GlcNAcylated by the O-GlcNAc transferase (OGT) enzyme *in vivo*. Interestingly, via all-atom MD simulations, we have shown that this site has incredibly long hydrogen bond lifetimes (HBLs) with confined solvent waters. Using REMD simulations and free energy calculations, we show that the removal of these waters, which are cocrystallized and interact with T72 and T75, lead to the de-stabilization of aSyn's amyloid core[14,15].

This led us to hypothesize that PTM sites within cavities have longer HBLs than other residues in the fibril structure. First, we calculated residue-water HBLs for the last 25 ns of each of the triplicate simulations. As expected, amino acids in cavities have on average larger HBLs than amino acids outside cavities (Figure 3A), displaying approximately a 4-fold increase in protein-water HBL (from ~10 ps to ~38 ps). Interestingly, when we parse amino acids by both cavity and PTM, we found that PTM sites within cavities have significantly longer HBLs (>50 ps) than all other residues (Figure 3B). We validated our previous results on T72 and T75, two sites of O-GlcNAcylation found in human brains known to inhibit the aggregation propensity of aSyn *in vitro* and in primary neuronal cultures[16,17]. As mentioned previously, T72 is also a site of mutation (T72M) linked to Parkinson's disease[18]. In addition, we detected T59 and S87, another two O-GlcNAcylation sites found in human brain samples and known to reduce the aggregation capacity of aSyn *in vitro*[17,51,52]. Interestingly, S87 can alternatively be phosphorylated, and a recent study showed that modifying this site with either O-GlcNAc or phosphorylation leads to two unique amyloid conformations, highlighting the importance of cavity PTM sites in determining the amyloid fold and stability[53].

Since most PTMs occur at hydrophilic amino acids such as threonines, we calculated the average Kyte-Doolittle hydropathicity of each cavity and non-cavity group (where hydrophobic > 0 and hydrophilic < 0)[54]. Figure 3C shows that PTM sites in cavities have very similar average hydropathicities to residues in the other three groups, suggesting that overall hydropathicity is not the defining characteristic for this effect.

Hydrophobic and not hydrophilic mutations in aSyn fibril cavities cause enhanced toxicity

Newberry et. al. performed an impressive study employing a deep mutational scan of aSyn in yeast and determined a fitness metric to quantify aSyn induced toxicity relative to WT for all 19 other amino acid at each residue position[20]. In Newberry's analysis, they utilized the α-helical structure of aSyn to explore the structure-fitness relationship, acknowledging that although the α-helical conformation is the dominant species in the yeast synucleinopathy

model, the amyloid and disordered oligomeric conformations may be present and contribute to toxicity[20].

Using this dataset, we imposed an amyloid fibril structure-fitness model and compared the effects of mutating amino acids inside and outside of aSyn's fibril cavities. As shown in Figure 4A–B, we observe a lower overall relative fitness for mutations at cavity residue sites than non-cavity sites when mutated to hydrophobic amino acids (W, F, L, I, V, M, C, A) (Figure 4A–B), supported by an odds ratio of observing lower fitness of ~12 (Supplemental Figure S1A). In contrast, there is an increase in fitness relative to WT when sites are mutated to polar residues (Y, G, S, T, N, Q, H, R, K, D, E) for both cavity and non-cavity residues, with no difference observed between the two groups.

As an important comparison to understand the relevance of cavities to the overall structure-fitness relationship, we calculated and compared the fitness of non-amyloid beta component (NAC) and non-NAC amino acids. The NAC domain (residues 61–95)[55] has a high degree of overlap with cavities and is known to be sufficient for robust aSyn aggregation and toxicity[56], making it a useful reference scale to explore our hypothesis. When we group residues by non-NAC and NAC, no statistical difference for either hydrophobic or polar mutations (Figure 4C–D) was observed. However, we found an odds ratio of ~3 of observing lower fitness in NAC residues relative to non-NAC, further confirming the known importance of this domain (Supplemental Figure S1A).

Despite the overlap between cavity and the NAC residues, we observed a 4-fold higher odds ratio of observing low fitness upon hydrophobic mutations in cavity/non-cavity amino acids relative to NAC/non-NAC (~12 vs. ~3), suggesting a unique contribution of specific amino acids that localize to cavities. In fact, when we compare the fitness of cavity unique (n=9), NAC unique (n=8) and cavity/NAC overlapping (n=27) residues, we observe that cavity unique and cavity/NAC overlapping residues have lower average fitness relative to NAC unique residues (Supplemental Figure S2). Interestingly, the residues with lowest fitness in the cavity unique group are K43 and T54, known to undergo ubiquitination and O-GlcNAcylation, respectively; and G51, known to be mutated to G51D in Parkinson's disease[41,49]. These results suggest that, similar to the NAC domain, aSyn fibril cavities are relevant structural motifs that contribute to toxicity.

Mutations within aSyn fibril cavities are more sensitive to increases in hydrophobicity

We then grouped all residues that were defined as cavity in at least 1 of the 8 aSyn fibril structures and performed a global linear regression of KD hydropathicity vs. fitness. Figure 5A shows that amino acids in cavities have a stronger negative correlation between KD hydropathicity and fitness. Across the hydropathicity scale, mutating an amino acid in a cavity results in a greater reduction in fitness relative to similar mutations in non-cavity residues. In addition, the hydrophobicity threshold at which a mutation becomes more toxic than WT aSyn is lower for cavity residues compared to non-cavity residues (x-intercept for cavity regression is roughly ~2 compared to ~4.2 for non-cavity). This result agrees with the observed increase in HBL for PTM sites within cavities; and supports our hypothesis that amino acids in fibril cavities are regulatory of the toxic fibrillization of aSyn potentially via hydration and steric effects.

Recently, Pancoe et. al. summarized the Thioflavin T aggregation rate of all published aSyn mutations (both familial and artificial) as of 2022[51]. Mutants were scored on a discrete scale ranging from −2 to +1 where increasing scores correspond to faster aggregation rates relative to WT aSyn. Using this dataset, we determined the correlation between KD hydropathicity and aggregation rate for both cavity and non-cavity site aSyn mutants. Figure 5B shows that amino acids in cavities have a positive correlation between hydropathicity and aggregation rate. In contrast, amino acids outside cavities had a slope close to zero. It is important to note that the number of mutants reported by Pancoe et. al. is far less than those in the Newberry et. al. dataset—only covering ~5% of the mutational landscape for aSyn—resulting in the increased 95% confidence interval for the aggregation global linear regression. Supplemental Figure S3 provides the distributions of mutations by amino acid type included in the toxicity and aggregation linear regressions presented in Figure 5A–B.

Cavities with PTM sites are present in other neurodegenerative disease-related amyloid fibrils

Our findings with aSyn led us to investigate whether these features are present in other amyloidogenic proteins associated with neurodegenerative disease. To test this, we simulated an additional 12 structures from the neurodegeneration-related proteins tau, prion, FUS, TDP-43 and hnRNPA1. All of these proteins contain hydrophilic cavities of varying sizes that possess either PTMs or disease associated familial mutation (FM) sites (Figure 6A–E and Supplemental Figure S4). Three of the five additional proteins (tau, FUS and TDP-43) contained PTM sites within the identified cavities. For both FUS and TDP-43, the average HBL of cavity residue PTM sites were comparable to the average HBL of PTM sites in aSyn fibril cavities (Figure 6F). In contrast, tau cavity PTM sites had a lower average HBL than what was observed in aSyn, but still much larger than water-water HBL (~2 ps). A full cataloging of the PTM/FM sites and fibril cavity properties (e.g., volumes, average KD hydropathicity, average number of hbonds, average HBLs and amino acid content are shown in detail in Supplemental Tables T1 and T2 and Supplemental Figures S5 and S6).

In the case of tau, phosphorylation sites S356, T263, T361; methylation site K267 and caspase-2 proteolytic cleavage site D314 have stable HBLs with water. Cleavage at D314 by caspase-2 has been shown to yield a tau fragment that resists fibrillation and misdirects tau towards dendritic spines causing memory deficits and neurodegeneration in mice[57]. Methylation at K267 has been reported to be present in paired helical filament (PHF) isolates from Alzheimer's disease (AD) brains using both mass spectrometry and immunohistochemistry on hippocampal AD brain sections[58]. Phosphorylation at T263 has been not studied extensively, but it has been detected in seeding competent tau aggregates isolated from AD brains, and it is commonly found to be phosphorylated concomitantly with S262, a well-defined site that decreases binding affinity to microtubules upon modification[59]. Phosphorylation at S356 leads to a decrease in the Thioflavin S positive aggregation of tau and is also implicated in the early onset frontotemporal dementia familial mutation S356T[60–62]. Phosphorylation at T361 has been detected in post-mortem AD brains and is modified by casein kinase 1 lambda (CK1δ), which is elevated in AD patients[63,64].

In FUS, we observe elevated HBLs at S57, S127 and S135, three known sites of phosphorylation, which have been shown to increase the liquidity of FUS condensates and prevent the formation of irreversible aggregates[42,65–67]. In addition, S57 and S135 are sites of ALS familial mutations (S135G and S57del), two modifications that eliminate the phosphorylation sites[68,69]. Finally, we observe a single PTM site with long HBL inside the TDP-43 fibril cavities, which corresponds to S409, a site of phosphorylation in the C-terminus of the protein. Modification at this site has been shown to be increased in pathological TDP-43 inclusions in brain tissue samples of patients with FTD, ALS and other neurodegenerative diseases[70].

Discussion

This study builds upon our previous work with aSyn and identifies a relationship between PTMs and aSyn fibril cavity amino acids. We show that aSyn fibril cavities are hydrated with confined waters and that changing these sites to hydrophobic amino acids which eliminate potential PTMs and reduce hbond propensity results in higher toxicity and aggregation rate relative to similar changes to residues outside cavities. It is important to note that waters within hydrophilic cavities experience elevated HBLs due to confinement. However, at the sites of cavity PTMs we observed even higher HBLs, suggesting that changes in the solvation of these residues will have a larger contribution to fibril stability.

Our analysis from simulations of FUS, TDP-43 and tau fibrils suggest that this phenomenon may extend to other neurodegenerative disease associated proteins, as we revealed at least one PTM site with exceptional solvent hbonding properties within the cavities of each one of these amyloids. Although we did not observe PTM sites in the fibril cavities of prion and hnRNPA1, the presence of familial mutation sites supports the hypothesis that cavity amino acids influence the stability and toxicity of amyloidogenic proteins. Our results warrant future *in vitro* and *in cellulo* experimental studies to validate the relevance of the identified PTM sites within hydrophilic cavities.

The results here support our hypothesis that cavities in disease-associated fibrils of aSyn are locations where PTMs are capable of inhibiting the fibrillization cascade and are conserved to reduce the possibility of toxic protein aggregation. In the cell, aSyn adopts multiple conformations depending on function (e.g., alpha-helix classically associated with membrane binding) or pathology (e.g., oligomers and fibrils associated with misfolding). As shown by Newberry et. al.[20], mutants that modulate aSyn's toxicity under the alpha-helix structure-fitness model may correspond to loss of native function; whereas those that modulate toxicity under a fibril structure-fitness model could represent a gain of toxic function. The crossover between these two models may inform on residues that are relevant to both loss and gain of function events and should be conserved at all costs. For example, a fibril cavity residue that is also solvated in the alpha-helix (e.g., T72) likely plays an essential role in regulating aSyn's conformational state; whereby deviation from WT (e.g., T72M or T72-O-GlcNAc) can bias towards toxic or non-toxic ensembles.

From a disease etiology standpoint, it will be crucial to study whether the disruption of PTM enzymes or their regulatory pathways are dysfunctional upstream of protein aggregation.

From a therapeutic standpoint, these results could be leveraged to rationally activate or inhibit PTM enzymes with small molecules as to globally modulate the modification of amyloid-forming proteins. Tools such as proximity-directed engineered PTM enzymes could be employed to modify sites predicted to have high water h-bonding activity with bulky hydrophilic modifications as to efficiently block aggregation and evade functional monomer conversion to toxic oligomers and fibrils. Proximity directed modification has been successfully utilized to introduce highly polar O-GlcNAc modifications to aSyn *in cellulo* in a specific manner by tethering the O-GlcNAc transferase (OGT) to a C-terminal aSyn-specific nanobody[71]. Alternatively, small molecules or peptides modified with proteasomal, chaperone-mediated autophagy or PROTAC signals could be designed to initially engage fibrils by binding to solvent accessible cavities, and subsequently directing aggregates to protein clearance pathways.

Although we only focused on neurodegeneration-related proteins, preliminary inspection of fibril structures of proteins associated with non-neurological disorders, such as transthyretin (ATTR amyloidosis), lambda light chain 1 (AL amyloidosis) and serum amyloid A protein (AA amyloidosis)[72–75], suggest that these cavities may be a common feature of fibril structures— independent of source protein, sequence, tissue type or associated disease.

Methods

MD simulations of amyloid fibrils

MD simulations were conducted following the exact workflow detailed previously[15]. Briefly, we simulated fibrils in NAMD version 2.12 under the CHARMM 36 force field with 10mer fibrils that we replicated along the z-axis[76–78]. Each structure was fully solvated in TIP3P water with a final concentration of 150 mM NaCl, which includes counter-ions for charge neutralization.

Each system was energy minimized and equilibrated via sequential 0.1 ns minimizations with fixed backbone constraints, followed by a 0.1 ns of NVT ensemble run, 0.5 ns harmonic constraint applied to backbone atoms, 0.2 ns harmonic constraint release, and a 1.0 ns NPT equilibration run. All simulations were conducted in triplicate with random initial velocities seeded from a Maxwell distribution at 310 K and simulated for a production run of 100 ns with a 1 fs time step. A full detail of protein, PDB entry, origin (recombinant vs. brain derived), genotype (wildtype vs. mutant), fibril formation conditions, structure determination method and resolved residues for each protein is presented in Supplemental Table T3. Upon completion, trajectories were re-centered and checked for the absence of self-interactions across the periodic boundary. In addition, root mean squared deviation (RMSD) convergence was verified, (shown in Supplemental Figure S7). Global and per residue RMSDs were calculated in VMD using the original 10mer structure as reference.

Cavity amino acid detection, volume, KD hydropathicity, hbonds and RMSD

Fibril cavities were detected using the Analysis of Null Areas 2's (ANA2)[37] pocket discovery feature on the last 25 ns of simulation which we then used to build the cavity amino acid content heatmap shown in Supplemental Figure S5A and Supplemental Table

T2. Using the detected cavity residues, ANA2 was used to calculate volumes by sampling 5 equally spaced timepoints of the last 25 ns production run (75–100 ns of simulation) in triplicate.

Average cavity RMSD was calculated by averaging the per residue RMSD of lining amino acids in a given cavity. Similarly, average KD hydropathicity cavity scores were calculated by averaging the KD hydropathicity of lining residues in a given cavity.

Hydrogen bond analysis

Cavity hydrogen bond counts were calculated using the VMD Hydrogen Bonds tool by selecting waters within 5 Å of cavity residues and counting the number of unique hbonds between water and residues with distance and angle cutoffs of 3 Å and 30°. Average cavity hbonds were calculated by averaging the hbond contribution from each cavity residue. For hbond counting we sampled the last 25 ns production runs every 0.2 ns for a total of 125 frames analyzed per simulation replicate.

Hydrogen bond lifetimes between residues and solvent water molecules were calculated in MDAnalysis as previously described[15,79]. Briefly, we used a water sphere of 5 Å around each residue as the cutoff for water-residue interactions. We calculated time autocorrelations for residue-water hbonds using a 30 ps sliding window across the entire last 25 ns production run and recovered HBLs using Equations 1–3 in MATLAB. Cavity HBLs were calculated by averaging HBLs from cavity residues.

$$C(t) = \left\langle \frac{h_{ij}(t_0)h_{ij}(t_0+t)}{h_{ij}(t_0)^2} \right\rangle$$

Equation 1.

$$C(t) = A_1 e^{-t/\tau_1} + (1 - A_1)e^{-t/\tau_2}$$

Equation 2.

$$HBL = \int_0^\infty C(t)dt$$

Equation 3.

Evolutionary conservation scoring using ConSurf-DB

Conservation scores for resolved amino acids in fibril structures were obtained using ConSurf-DB[38,39]. Briefly, ConSurf-DB utilizes published homologue sequences to build a phylogenetic tree that is used to estimate the evolutionary rates at protein sites (considering both natural substitution rates for amino acids and the phylogenetic tree branch length between species). These rates are mapped into discrete conservation scores (1–9) that reflect

the rate of substitutions per site per evolutionary time, with low conservation scores being the most variable sites and vice versa.

Newberry et. al. aSyn deep mutational scan in yeast data mining

We downloaded the deep mutational scan fitness data published in Newberry et. al.[20], and calculated average fitness scores for hydrophobic (W, F, L, I, V, M, C, A) and polar (Y, G, S, T, N, Q, H, R, K, D, E) mutations at cavity and non-cavity sites. In our analysis for Figure 4A–B, we included tyrosine as part of polar residues due to the hydroxyl group that readily hbonds with solvent waters (tyrosine was counted as a hydrophobic residue in Newberry et.al.[20]). Furthermore, we excluded proline from hydrophobic and polar definitions as done by the authors of the original paper. For the linear regression analysis, we correlated the Kyte-Doolittle hydropathicity scores of the 20 amino acids relative to their average fitness across non-cavity and cavity residues using GraphPad Prism 9.

Pancoe et. al. aSyn mutant aggregation rate data mining

Aggregation rate scores of mutations at cavity and non-cavity amino acids (as assigned by Pancoe et. al.[51]) were correlated to KD hydropathicity by fitting a linear regression model in GraphPad Prism 9.

Statistical analysis

Conservation scores and PTM enrichment in fibril cavities were tested using paired Student T tests. Global HBL comparison between cavity and non-cavity residues was tested using an unpaired two-tailed Student T test. HBL and KD hydropathicity comparison between cavity PTM sites to other residues in the structures was tested with a one-way ANOVA with post-hoc Bonferroni correction for multiple testing. Newberry et. al.[20] fitness data comparisons (both cavity/non-cavity and NAC/non-NAC groupings) were conducted with two-tailed unpaired Student T tests. Statistical significance for variance differences in fibril cavity properties grouped by protein type in Supplemental Figure S3 were analyzed by conducting one-way ANOVAs.

Additionally, we calculated odds and likelihood ratios with 95% confidence intervals (CIs) for the key findings shown in Figures 2A, 2B, 3A, 3B, 4A and 4C, and referenced the results in corresponding figure captions. We parsed the data by non-cavity/cavity (2A, 2B, 3A and 4A) or NAC/non-NAC (4C) and dichotomized properties of interest (low/high evolutionary conservation, PTM site count, low/high hydrogen bond lifetimes and low/high fitness). Dichotomization was conducted by calculating global averages for conservation scores, HBLs and fitness, which were then used to parse residues into low (< global average) and high (> global average) groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Molecular dynamics simulations were performed at the University of Minnesota's Minnesota Supercomputing Institute (MSI). The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by U.S. National Institutes of Health (NIH) to J.N.S. (NINDS R01NS117968).

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Spillantini MG et al. Alpha-Synuclein in Lewy bodies. Nature 388, 839–840 (1997). 10.1038/42166 [PubMed: 9278044]
2. Frontzkowski L et al. Earlier Alzheimer's disease onset is associated with tau pathology in brain hub regions and facilitated tau spreading. Nature Communications 13 (2022). 10.1038/s41467-022-32592-7
3. Mori F et al. Maturation process of TDP-43-positive neuronal cytoplasmic inclusions in amyotrophic lateral sclerosis with and without dementia. Acta Neuropathologica 116, 193–203 (2008). 10.1007/s00401-008-0396-9 [PubMed: 18560845]
4. Kovacs G & Budka H Molecular Pathology of Human Prion Diseases. International Journal of Molecular Sciences 10, 976–999 (2009). 10.3390/ijms10030976 [PubMed: 19399233]
5. Marrone L et al. FUS pathology in ALS is linked to alterations in multiple ALS-associated proteins and rescued by drugs stimulating autophagy. Acta neuropathologica 138, 67–84 (2019). 10.1007/s00401-019-01998-x [PubMed: 30937520]
6. Bampton A, Gittings LM, Fratta P, Lashley T & Gatt A The role of hnRNPs in frontotemporal dementia and amyotrophic lateral sclerosis. Acta Neuropathologica 140, 599–623 (2020). 10.1007/s00401-020-02203-0 [PubMed: 32748079]
7. Andorfer C et al. Hyperphosphorylation and aggregation of tau in mice expressing normal human tau isoforms. Journal of Neurochemistry 86, 582–590 (2003). 10.1046/j.1471-4159.2003.01879.x [PubMed: 12859672]
8. Lee MK et al. Human α-synuclein-harboring familial Parkinson's disease-linked Ala-53 → Thr mutation causes neurodegenerative disease with α-synuclein aggregation in transgenic mice. Proceedings of the National Academy of Sciences 99, 8968–8973 (2002). 10.1073/pnas.132197599
9. Igaz LM et al. Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. Journal of Clinical Investigation 121, 726–738 (2011). 10.1172/jci44867 [PubMed: 21206091]
10. Sharma A et al. ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function. Nature Communications 7, 10465 (2016). 10.1038/ncomms10465
11. Vaquer-Alicea J, Diamond MI & Joachimiak LA Tau strains shape disease. Acta Neuropathologica 142, 57–71 (2021). 10.1007/s00401-021-02301-7 [PubMed: 33830330]
12. Shrivastava AN et al. Differential Membrane Binding and Seeding of Distinct α-Synuclein Fibrillar Polymorphs. Biophys J 118, 1301–1320 (2020). 10.1016/j.bpj.2020.01.022 [PubMed: 32059758]
13. Sawaya MR, Hughes MP, Rodriguez JA, Riek R & Eisenberg DS The expanding amyloid family: Structure, stability, function, and pathogenesis. Cell 184, 4857–4873 (2021). https://doi.org/10.1016/j.cell.2021.08.013 [PubMed: 34534463]
14. Romo TD, Lewis AK, Braun AR, Grossfield A & Sachs JN Minimal Nucleation State of α-Synuclein Is Stabilized by Dynamic Threonine–Water Networks. ACS Chemical Neuroscience 8, 1859–1864 (2017). 10.1021/acschemneuro.7b00171 [PubMed: 28677385]

15. Nathan Kochen N et al. Threonine Cavities Are Targetable Motifs That Control Alpha-Synuclein Fibril Growth. ACS Chemical Neuroscience 13, 2646–2657 (2022). 10.1021/acschemneuro.2c00327 [PubMed: 36001084]
16. Marotta NP et al. O-GlcNAc modification blocks the aggregation and toxicity of the protein α-synuclein associated with Parkinson's disease. Nature chemistry 7, 913–920 (2015). 10.1038/nchem.2361
17. Levine PM et al. α-Synuclein O-GlcNAcylation alters aggregation and toxicity, revealing certain residues as potential inhibitors of Parkinson's disease. Proceedings of the National Academy of Sciences of the United States of America 116, 1511–1519 (2019). 10.1073/pnas.1808845116 [PubMed: 30651314]
18. Fevga C et al. A new alpha-synuclein missense variant (Thr72Met) in two Turkish families with Parkinson's disease. Parkinsonism & Related Disorders 89, 63–72 (2021). 10.1016/j.parkreldis.2021.06.023 [PubMed: 34229155]
19. Sawaya MR, Hughes MP, Rodriguez JA, Riek R & Eisenberg DS The expanding amyloid family: Structure, stability, function, and pathogenesis. Cell 184, 4857–4873 (2021). https://doi.org/10.1016/j.cell.2021.08.013 [PubMed: 34534463]
20. Newberry RW, Leong JT, Chow ED, Kampmann M & DeGrado WF Deep mutational scanning reveals the structural basis for α-synuclein activity. Nat Chem Biol (2020). 10.1038/s41589-020-0480-6
21. Li B et al. Cryo-EM of full-length alpha-synuclein reveals fibril polymorphs with a common structural kernel. Nature Communications 9, 3609 (2018). 10.1038/s41467-018-05971-2
22. Sun Y et al. Cryo-EM structure of full-length α-synuclein amyloid fibril with Parkinson's disease familial A53T mutation. Cell Research 30, 360–362 (2020). 10.1038/s41422-020-0299-4 [PubMed: 32203130]
23. Schweighauser M et al. Structures of α-synuclein filaments from multiple system atrophy. Nature 585, 464–469 (2020). 10.1038/s41586-020-2317-6 [PubMed: 32461689]
24. Boyer DR et al. Structures of fibrils formed by α-synuclein hereditary disease mutant H50Q reveal new polymorphs. Nature Structural & Molecular Biology 26, 1044–1052 (2019). 10.1038/s41594-019-0322-y
25. Boyer DR et al. The α-synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. Proceedings of the National Academy of Sciences 117, 3592 LP-3602 (2020). 10.1073/pnas.1917914117
26. Guerrero-Ferreira R et al. Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. eLife 8, e48907–e48907 (2019). 10.7554/eLife.48907 [PubMed: 31815671]
27. Murray DT et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell 171, 615–627.e616 (2017). https://doi.org/10.1016/j.cell.2017.08.048 [PubMed: 28942918]
28. Lee M, Ujjayini G, R. TK, Masato K & Robert T Molecular structure and interactions within amyloid-like fibrils formed by a low-complexity protein sequence from FUS. Nature communications 11, 1–14 (2020).
29. Sun Y et al. The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure. Nature Communications 11 (2020). 10.1038/s41467-020-20227-8
30. Wang L-Q et al. Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. Nature Structural & Molecular Biology 27, 598–602 (2020). 10.1038/s41594-020-0441-5
31. Glynn C et al. Cryo-EM structure of a human prion fibril with a hydrophobic, protease-resistant core. Nature Structural & Molecular Biology 27, 417–423 (2020). 10.1038/s41594-020-0403-y
32. Falcon B et al. Structures of filaments from Pick's disease reveal a novel tau protein fold. Nature 561, 137–140 (2018). 10.1038/s41586-018-0454-y [PubMed: 30158706]
33. Zhang W et al. Novel tau filament fold in corticobasal degeneration. Nature 580, 283–287 (2020). 10.1038/s41586-020-2043-0 [PubMed: 32050258]
34. Falcon B et al. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. Nature 568, 420–423 (2019). 10.1038/s41586-019-1026-5 [PubMed: 30894745]

35. Shi Y et al. Structure-based classification of tauopathies. Nature 598, 359–363 (2021). 10.1038/s41586-021-03911-7 [PubMed: 34588692]
36. Fitzpatrick AWP et al. Cryo-EM structures of tau filaments from Alzheimer's disease. Nature 547, 185–190 (2017). 10.1038/nature23002 [PubMed: 28678775]
37. Barletta GP, Barletta M, Saldaño TE & Fernandez-Alberti S Analysis of changes of cavity volumes in predefined directions of protein motions and cavity flexibility. Journal of Computational Chemistry 43, 391–401 (2022). 10.1002/jcc.26799 [PubMed: 34962296]
38. Ben Chorin A et al. ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Sci 29, 258–267 (2020). 10.1002/pro.3779 [PubMed: 31702846]
39. Goldenberg O, Erez E, Nimrod G & Ben-Tal N The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. Nucleic Acids Res 37, D323–327 (2009). 10.1093/nar/gkn822 [PubMed: 18971256]
40. Reimand J, Wagih O & Bader GD Evolutionary constraint and disease associations of post-translational modification sites in human genomes. PLoS Genet 11, e1004919 (2015). 10.1371/journal.pgen.1004919 [PubMed: 25611800]
41. Landrum MJ et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46, D1062–D1067 (2018). 10.1093/nar/gkx1153 [PubMed: 29165669]
42. Rhoads SN, Monahan ZT, Yee DS & Shewmaker FP The Role of Post-Translational Modifications on Prion-Like Aggregation and Liquid-Phase Separation of FUS. Int J Mol Sci 19 (2018). 10.3390/ijms19030886
43. Eck RJ, Kraemer BC & Liachko NF Regulation of TDP-43 phosphorylation in aging and disease. Geroscience 43, 1605–1614 (2021). 10.1007/s11357-021-00383-5 [PubMed: 34032984]
44. Fan T-S, Chi-Hao Liu S & Ruey-Meei W Alpha-Synuclein and Cognitive Decline in Parkinson Disease. Life 11 (2021). 10.3390/life11111239
45. Hoffman-Zacharska D et al. Novel A18T and pA29S substitutions in α-synuclein may be associated with sporadic Parkinson's disease. Parkinsonism Relat Disord 19, 1057–1060 (2013). 10.1016/j.parkreldis.2013.07.011 [PubMed: 23916651]
46. Kumar ST et al. A NAC domain mutation (E83Q) unlocks the pathogenicity of human alpha-synuclein and recapitulates its pathological diversity. Sci Adv 8, eabn0044 (2022). 10.1126/sciadv.abn0044 [PubMed: 35486726]
47. Mishra R et al. Impact of N-glycosylation site variants during human PrP aggregation and fibril nucleation. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1867, 909–921 (2019). https://doi.org/10.1016/j.bbapap.2019.03.010 [PubMed: 30935958]
48. Rossi G & Tagliavini F Frontotemporal lobar degeneration: old knowledge and new insight into the pathogenetic mechanisms of tau mutations. Front Aging Neurosci 7, 192 (2015). 10.3389/fnagi.2015.00192 [PubMed: 26528178]
49. Zhang J, Li X & Li J-D The Roles of Post-translational Modifications on α-Synuclein in the Pathogenesis of Parkinson's Diseases. Frontiers in Neuroscience 13 (2019). 10.3389/fnins.2019.00381
50. Alquezar C, Arya S & Kao AW Tau Post-translational Modifications: Dynamic Transformers of Tau Function, Degradation, and Aggregation. Front Neurol 11, 595532 (2020). 10.3389/fneur.2020.595532 [PubMed: 33488497]
51. Pancoe SX et al. Effects of Mutations and Post-Translational Modifications on α-Synuclein In Vitro Aggregation.
52. Lewis YE et al. O-GlcNAcylation of α-Synuclein at Serine 87 Reduces Aggregation without Affecting Membrane Binding. ACS Chem Biol 12, 1020–1027 (2017). 10.1021/acschembio.7b00113 [PubMed: 28195695]
53. Hu J et al. Phosphorylation and O-GlcNAcylation at the same α-synuclein site generate distinct fibril structures. bioRxiv, 2023.2006.2027.546682 (2023). 10.1101/2023.06.27.546682
54. Kyte J & Doolittle RF A simple method for displaying the hydropathic character of a protein. J Mol Biol 157, 105–132 (1982). 10.1016/0022-2836(82)90515-0 [PubMed: 7108955]

55. Uéda K et al. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proceedings of the National Academy of Sciences 90, 11282–11286 (1993). 10.1073/pnas.90.23.11282
56. Rodriguez JA et al. Structure of the toxic core of α-synuclein from invisible crystals. Nature 525, 486–490 (2015). 10.1038/nature15368 [PubMed: 26352473]
57. Zhao X et al. Caspase-2 cleavage of tau reversibly impairs memory. Nat Med 22, 1268–1276 (2016). 10.1038/nm.4199 [PubMed: 27723722]
58. Thomas SN et al. Dual modification of Alzheimer's disease PHF-tau protein by lysine methylation and ubiquitylation: a mass spectrometry approach.
59. Wesseling H et al. Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's Disease. Cell 183, 1699–1713.e1613 (2020). 10.1016/j.cell.2020.10.029 [PubMed: 33188775]
60. Haj-Yahya M et al. Site-Specific Hyperphosphorylation Inhibits, Rather than Promotes, Tau Fibrillization, Seeding Capacity, and Its Microtubule Binding.
61. Xia Y, Bell BM, Kim JD & Giasson BI Tau mutation S356T in the three repeat isoform leads to microtubule dysfunction and promotes prion-like seeded aggregation. Front Neurosci 17, 1181804 (2023). 10.3389/fnins.2023.1181804 [PubMed: 37304025]
62. Momeni P et al. Familial early onset frontotemporal dementia caused by a novel S356T MAPT mutation, initially diagnosed as schizophrenia. Clin Neurol Neurosurg 112, 917–920 (2010). 10.1016/j.clineuro.2010.07.015 [PubMed: 20708332]
63. Hanger DP et al. Novel phosphorylation sites in tau from Alzheimer brain support a role for casein kinase 1 in disease pathogenesis. J Biol Chem 282, 23645–23654 (2007). 10.1074/jbc.M703269200 [PubMed: 17562708]
64. Yasojima K, Kuret J, DeMaggio AJ, McGeer E & McGeer PL Casein kinase 1 delta mRNA is upregulated in Alzheimer disease brain. Brain Res 865, 116–120 (2000). 10.1016/s0006-8993(00)02200-9 [PubMed: 10814741]
65. Ranganathan S, Dasmeh P, Furniss S & Shakhnovich E Phosphorylation sites are evolutionary checkpoints against liquid-solid transition in protein condensates. Proc Natl Acad Sci U S A 120, e2215828120 (2023). 10.1073/pnas.2215828120 [PubMed: 37155880]
66. Monahan Z et al. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity.
67. Owen I et al. The prion-like domain of Fused in Sarcoma is phosphorylated by multiple kinases affecting liquid- and solid-phase transitions.
68. Belzil VV et al. Mutations in FUS cause FALS and SALS in French and French Canadian populations.
69. Morgan S et al. A comprehensive analysis of rare genetic variation in amyotrophic lateral sclerosis in the UK. Brain 140, 1611–1618 (2017). 10.1093/brain/awx082 [PubMed: 28430856]
70. Neumann M et al. Phosphorylation of S409/410 of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. Acta Neuropathol 117, 137–149 (2009). 10.1007/s00401-008-0477-9 [PubMed: 19125255]
71. Ramirez DH et al. Engineering a Proximity-Directed O-GlcNAc Transferase for Selective Protein O-GlcNAcylation in Cells. ACS Chem Biol 15, 1059–1066 (2020). 10.1021/acschembio.0c00074 [PubMed: 32119511]
72. Iakovleva I et al. Structural basis for transthyretin amyloid formation in vitreous body of the eye. Nature Communications 12 (2021). 10.1038/s41467-021-27481-4
73. Schmidt M et al. Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis. Nature Communications 10 (2019). 10.1038/s41467-019-13038-z
74. Radamaker L et al. Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. Nature Communications 10 (2019). 10.1038/s41467-019-09032-0
75. Liberta F et al. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. Nature Communications 10 (2019). 10.1038/s41467-019-09033-z
76. Phillips JC et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. The Journal of chemical physics 153, 44130–44130 (2020). 10.1063/5.0014475

77. Huang J et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nature Methods 14, 71–73 (2017). 10.1038/nmeth.4067 [PubMed: 27819658]

78. Berendsen HJC, van der Spoel D & van Drunen R GROMACS: A message-passing parallel molecular dynamics implementation. Computer Physics Communications 91, 43–56 (1995). https://doi.org/10.1016/0010-4655(95)00042-E

79. Michaud-Agrawal N, Denning EJ, Woolf TB & Beckstein O MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. J Comput Chem 32, 2319–2327 (2011). 10.1002/jcc.21787 [PubMed: 21500218]

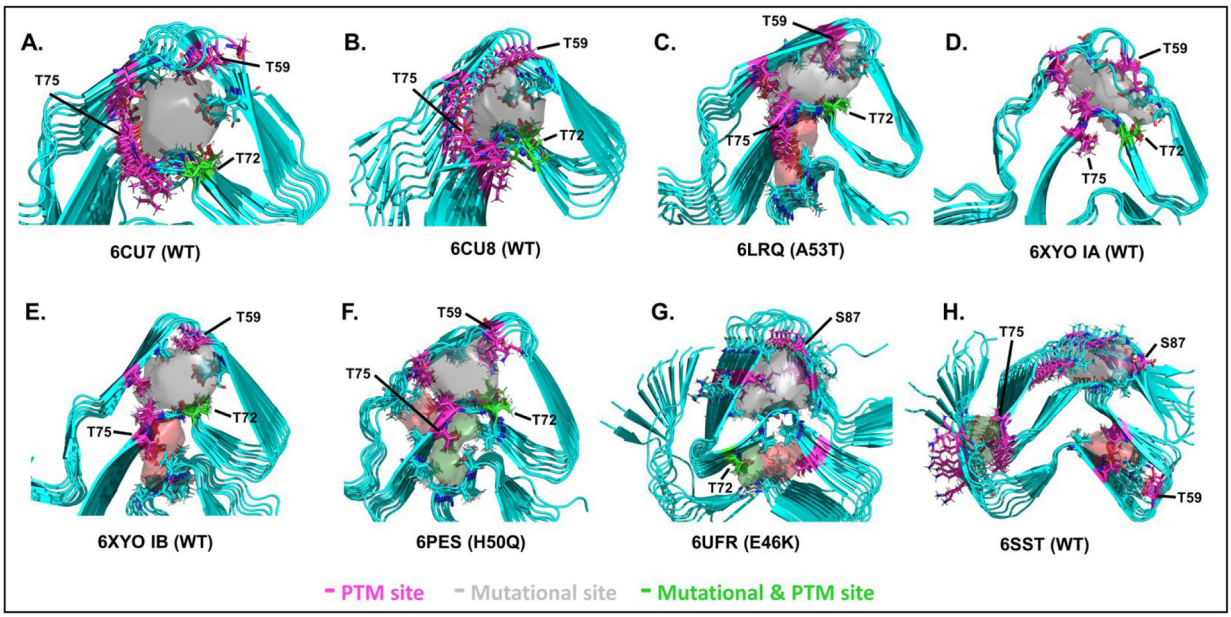

Figure 1. Hydrophilic cavities across 8 unique aSyn amyloid fibril structures/polymorphs.

(A-H) Fibril renders taken from the last 25 ns of MD simulation for each aSyn polymorph (PDB ID and genotype indicated below each fibril). Cavity residues are shown in stick representation with the indicated PTMs (magenta), familial mutation (FM; gray), and FM/PTM (green) sites. Cyan residues represent cavity sites that are neither a PTM nor FM. Different color densities indicate distinct cavities identified using ANA2.

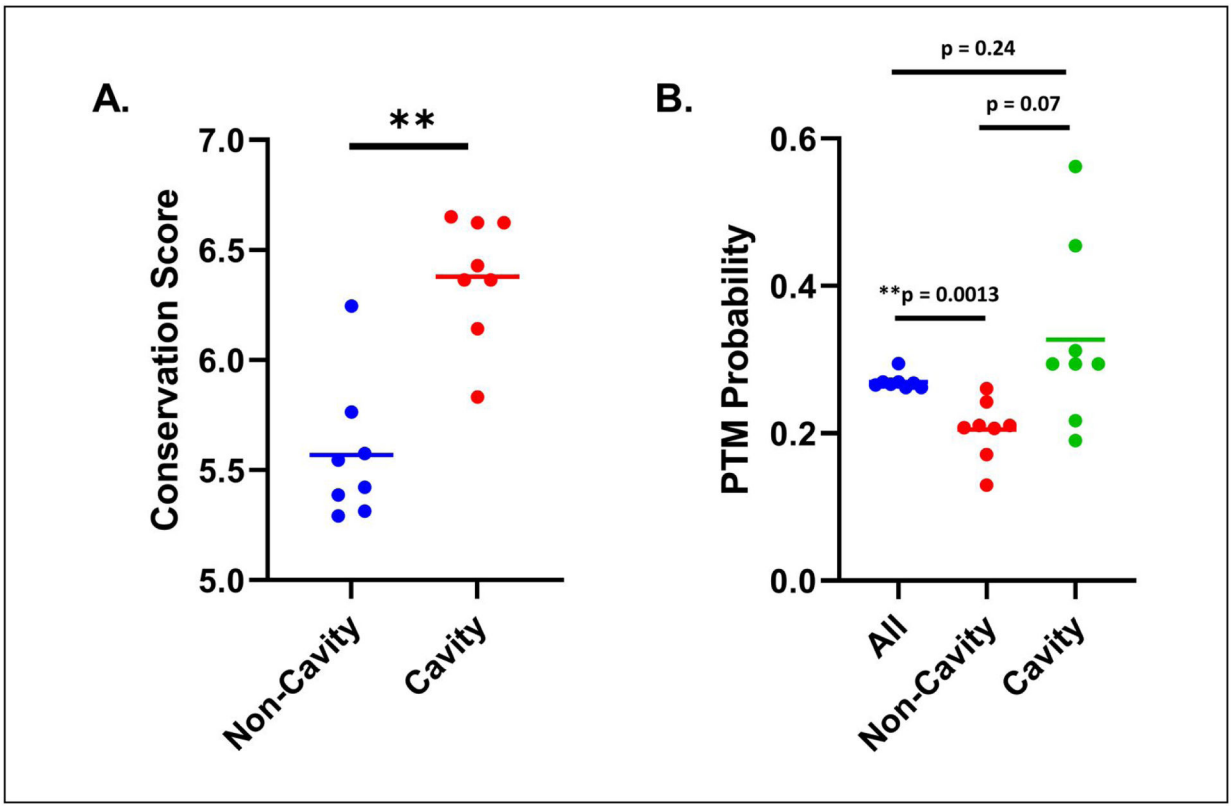

Figure 2. aSyn fibril cavity residues are evolutionarily conserved and contain an elevated percentage of PTM sites.

(A) ConSurf-DB average conservation scores of cavity and non-cavity residues. Each data point represents a unique fibril structure. Statistic is a two-tailed paired Student T test (**p = 0.0011). For this analysis, 117 different homologous sequences of aSyn were utilized by ConSurf-DB to construct multiple sequence alignments and to calculate the evolutionary conservation scores. **(B)** PTM probability in whole structure sequences, inside and outside fibril cavities. Each data point represents a unique aSyn fibril structure. Statistics shown are paired two-tailed Student T-tests. Odds and likelihood ratios for comparisons in (A) and (B) are shown in Supplemental Figure S1.

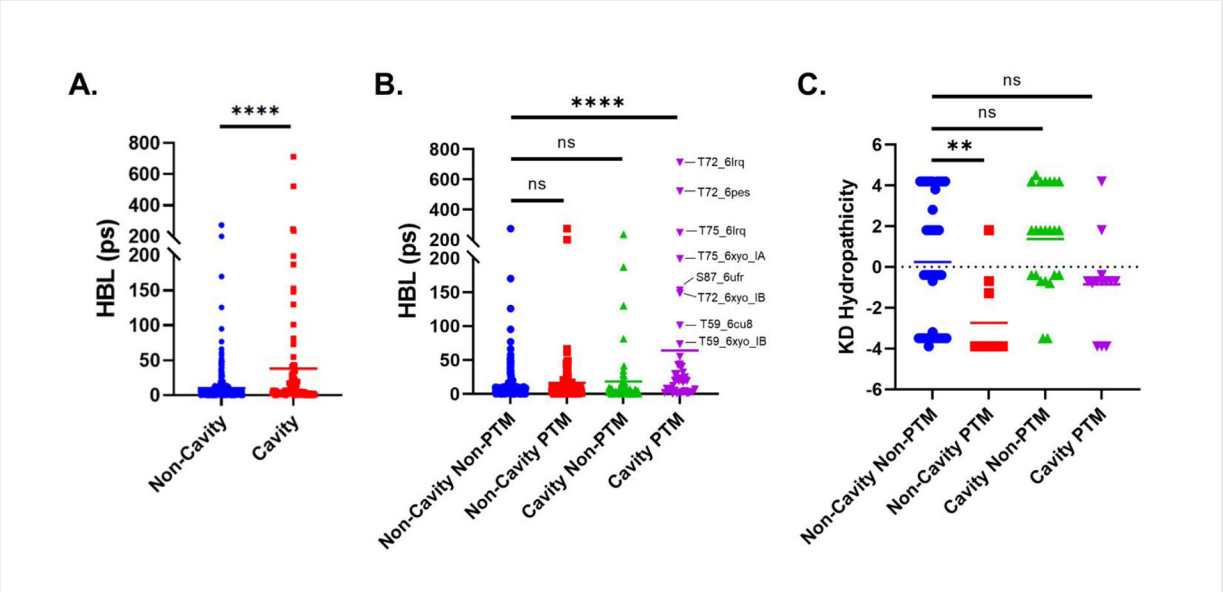

Figure 3. Protein-water HBLs are enhanced at fibril cavity PTM sites.

(A) Residue-water HBLs of cavity and non-cavity residues across all aSyn structures simulated. Each data point represents the HBL between a unique amino acid and solvent water. Statistic shown is an unpaired two-tailed Student T-test (****$p < 0.0001$). **(B)** Residue-water HBL grouped by cavity/non-cavity and PTM/non-PTM categories. Statistic shown is a one-way ANOVA with comparisons to the non-cavity non-PTM group with Bonferroni correction (****$p < 0.0001$). **(C)** KD hydropathicity of amino acids in subgroups shown in (B). Statistic shown is a one-way ANOVA with comparisons to the non-cavity non-PTM group with Bonferroni correction (**$p = 0.0076$). Odds and likelihood ratios for comparison between cavity vs. non-cavity PTM site HBLs are shown in Supplemental Figure S1.

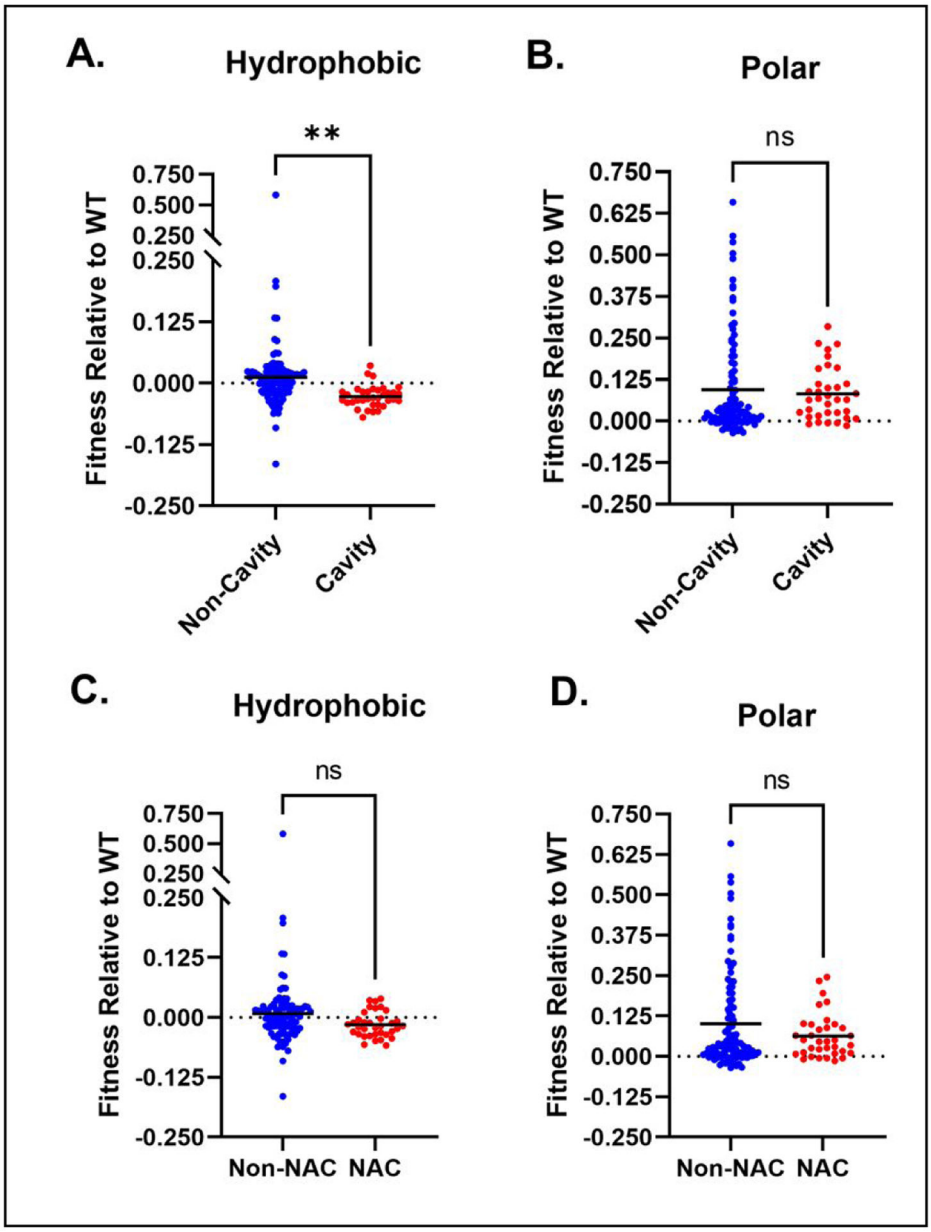

Figure 4. Hydrophobic mutations in aSyn fibril cavity amino acids reduce fitness in yeast cells.

(A) Relative average fitness of non-cavity and cavity residues mutated to hydrophobic residues (W, F, L, I, V, M, C, A). **(B)** Relative average fitness of non-cavity and cavity residues mutated to polar residues (Y, G, S, T, N, Q, H, R, K, D, E). Each data point represents a single amino acid. Statistics shown are two-tailed Student T-tests (**p = 0.0022). **(C-D)** Comparison of fitness using non-NAC and NAC amino acid grouping. Odds and likelihood ratios for comparison in (A) and (C) are shown in Supplemental Figure S1.

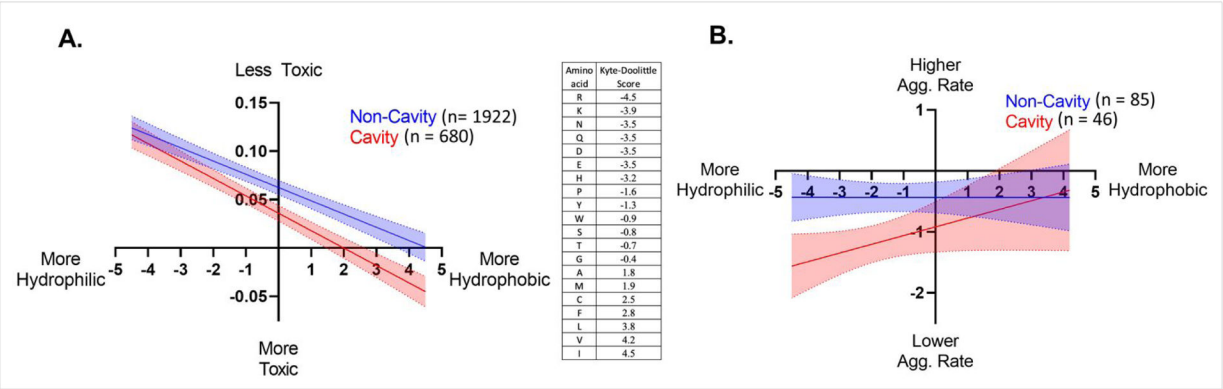

Figure 5. Hydropathicity changes strongly dictate toxicity and aggregation rate at cavity sites.

(A) Linear regression of Kyte-Doolittle hydropathicity and fitness for non-cavity and cavity residues (slope difference between lines p = 0.0606; intercept difference between lines ***p = 0.003). **(B)** Linear regression of Kyte-Doolittle hydropathicity and aggregation rate for non-cavity and cavity residues (only slope of cavity residues is statistically significant at *p = 0.018).

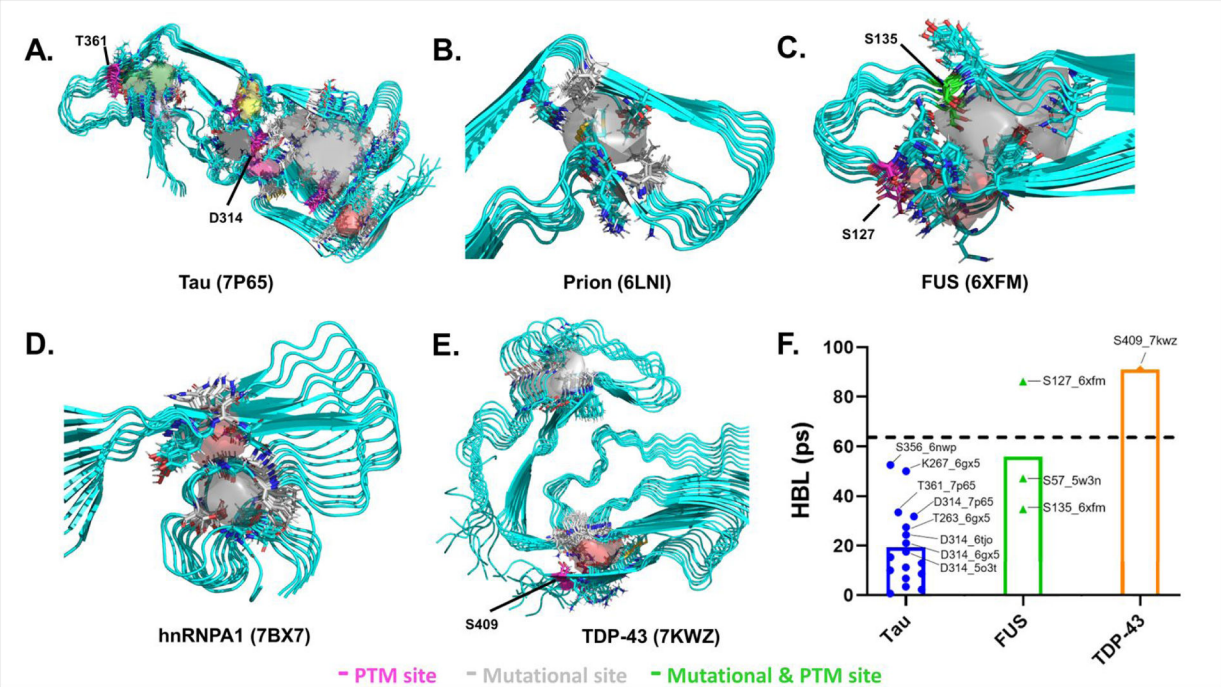

Figure 6. Hydrophilic fibril cavities are present in other neurodegeneration-related amyloidogenic proteins.

(A-E) Renders of a representative tau, prion, FUS, hnRNPA1 and TDP-43 fibril simulated. **(F)** Residue-water HBL for PTM site amino acids in fibril cavities of tau, FUS and TDP-43. Black dashed line indicates average HBL of PTM sites in aSyn fibril cavities. Prion and hnRNPA1 cavities lacked sites with known PTMs. Data point labels indicate residue number and fibril structure.