

Unraveling the Intricate Structural Transitions in α -Synuclein: Conformations Governing Fibril Formation and the Influence of Osmolytes

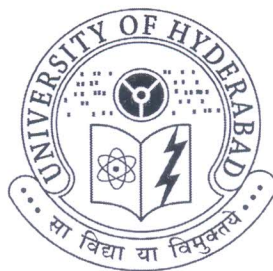
*A Thesis submitted
for the degree of
Doctor of Philosophy*

By

ARCHI SAURABH



Department of Biotechnology & Bioinformatics
School of Life Sciences
University of Hyderabad
Hyderabad-500046
Telangana, India
June, 2023



UNIVERSITY OF HYDERABAD

(A central university established in 1974 by an act of Parliament)

Department of Biotechnology and Bioinformatics

School of Life Sciences

University of Hyderabad, Hyderabad 500 046, India

STATEMENT

I, Archi Saurabh, hereby declare that the work presented in this thesis, entitled as “**Unraveling the Intricate Structural Transitions in α -Synuclein: Conformations Governing Fibril Formation and the Influence of Osmolytes**” has been carried out by me under the supervision of *Dr. N. Prakash Prabhu* at Department of Biotechnology and Bioinformatics. To the best of my knowledge, this work has not been submitted for the award of any degree or diploma to this or any other university or institute. A report on plagiarism statistics from the University Librarian is enclosed.

Place: Hyderabad

Date: 30/06/2023

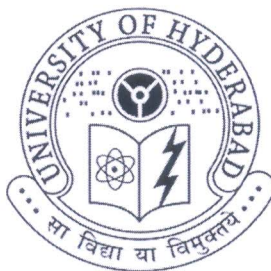
Archi Saurabh

15LTPM06

Supervisor

Dr. N. Prakash Prabhu

N. PRAKASH PRABHU PhD
Associate Professor,
Dept. Biotechnology & Bioinformatics
School of Life Sciences
University of Hyderabad
Hyderabad-500 046



UNIVERSITY OF HYDERABAD

(A central university established in 1974 by an act of Parliament)

Department of Biotechnology and Bioinformatics

School of Life Sciences

University of Hyderabad, Hyderabad 500 046, India

CERTIFICATE

This is to certify that the thesis entitled “**Unraveling the Intricate Structural Transitions in α -Synuclein: Conformations Governing Fibril Formation and the Influence of Osmolytes**” submitted by **Mrs. Archi Saurabh** bearing registration number **15LTPM06** in partial fulfilment of the requirements for the award of Doctor of Philosophy in the Department of Biotechnology and Bioinformatics, School of Life Sciences is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other university or institution for award of any degree or diploma

Further, the student has the following publication(s) before submission of the thesis/monograph for the adjudication and has produced evidence for the same in the form of acceptance letter or the reprint in the relevant area of his research:

1. **Saurabh, A., & Prabhu, N. P.** (2022). Concerted enhanced-sampling simulations to elucidate the helix-fibril transition pathway of intrinsically disordered α -Synuclein. *International Journal of Biological Macromolecules*, 223, 1024-1041. <https://doi.org/10.1016/j.ijbiomac.2022.11.079>. (ISSN Number- 0141-8130)

This publication appears in chapter number 2 of this dissertation.

And has made presentations in the following conferences:

1. **Saurabh, A., & Prabhu, N. P.** (2021, October). Conformational states guiding the monomer-fibril transition of α -Synuclein. In *Protein Science* (Vol. 30, pp. 65-65). 111 river St, Hoboken 07030-5774, NJ USA: Wiley. <https://doi.org/10.1002/pro.4191> (ABS# 156) (International)
2. **Saurabh, A., and N. Prakash Prabhu.** "Effect of charged amino acids on the α -synuclein fibril formation: spectroscopic and computational analysis" In *90th annual meeting of the Society of Biological Chemists of India (SBCI)*, 2021 (National)

Further, the student has passed the following courses towards the fulfilment of the course work requirement of the Ph.D. degree.

S.No.	Course code	Name	Credits	Pass/Fail
1	BT 801	Research Methodology/ Analytical Techniques	4	PASS
2	BT 802	Research Ethics, Biosafety, Data Analysis and Biostatistics	4	PASS
3	BT 803	Research Proposal and Scientific Writing	4	PASS

N. Prakash Prabh
30/06/23
Supervisor

Dr. Anurag
31/7/23
Head of Department
Biotechnology and Bioinformatics

Naveen Kumar
03/07/23
Dean
School of Life Science
संकाय अध्यक्ष / Dean
जीव विज्ञान संकाय / School of Life Sciences
हैदराबाद विश्वविद्यालय / University of Hyderabad
हैदराबाद / Hyderabad-500 046.

N. PRAKASH PRABHU PhD
Associate Professor,
Dept. Biotechnology & Bioinformatics
School of Life Sciences
University of Hyderabad
Hyderabad-500 046

HEAD
Dept. of Biotechnology & Bioinformatics
University of Hyderabad
Hyderabad.

Dedicated to my Mother and Father



ACKNOWLEDGEMENT

I feel extremely privileged to express my respect for my supervisor **Dr. N. Prakash Prabhu**, Associate Professor, Department of Biotechnology & Bioinformatics, School of Life Sciences, University of Hyderabad, Hyderabad. Working under him gave me a true sense of freedom, broadened my scientific outlook and inculcated a spirit of positive attitude in me. Without his constant support and active guidance, the research work presented in the thesis would not have been possible. As I wrap up this work, I really do not know how far I have met his expectations but I can't forget his acute scientific criticism and enthusiasm for social responsibility.

I am equally grateful to my Doctoral Committee member **Prof. J.S.S. Prakash** for encouragement and insightful comments throughout my research work. I am indebted to him for allowing me to continue part of my research in his laboratory with all the freedom. I am also very thankful to my Doctoral Committee member **Dr. M. Venkatramana** for his precious time and suggestions. I would like to thank the present and former heads of Department of Biotechnology and Bioinformatics and also present and former Deans of School of Life sciences for allowing me to use the Departmental/School facilities.

I am thankful to all the departmental and school non-teaching staffs for their help in doing all the official works in the department/school throughout this period of work. I would like to thank University for providing me fellowship and also DST, SERB, IOE and CSIR for funding support to the lab. I would also like to thank UGC-SAP and DST-FIST for their support to the department. I would like to acknowledge School of Chemistry, UoH for TEM facility and CMSD, UoH for the computational facilities.

I would like to thanks all my present (Subhasree Ghosh, Sanjay Kumar, Dr. Pratibha Kushwaha, Muzafar Ahmad and G. Priyanka) and former (Dr. Neshatul Haque, Dr. Kiran Kumar, Dr. Bramhini Anumalla and Dr. K. Tejaswi Naidu) lab Ph.D. students for their valuable discussions and for providing a happening environment in our lab. I would like to thank master student Vani, Harshita, Mouly, Salik and members of our synuclein family (Umair, Priyanka, Namrata and Saquib) for their love and discussion.

Also, I thank my friends Euphinia, Subhasree, Yasaswi, Pritikana, Sarita, Hema, Ritika, Tanay Shasanka, Priyansha, Shubhangi, Ramana, Anamya and Sharada. I will be unvirtuous, if I do not acknowledge my UoH constant, reason for my smile everyday here, Neha, Suraj, Satyajit and Aradhana. I would like to thanks my best friend nonchalant Neha Kumari for listening all my non-relevant “vakawas” thoughts.

Everyone goes through a good as well as rough time and there is always a pillar to provide strength, for me that’s my parent. First of all, I would like to thanks Ma for showing me how independent, caring, strong and beauty with brain looks like and Papa for describing me fearless passion and fighter’s attitude. Thank you, Ma and Papa, for believing and supporting me always. I would like to thanks my very own “shield of Hercules”, my brothers, Apurv and (Balak) Ankit. Thank you for making me free bird, bhaiya and let me chippers, balak. I am very grateful to my husband “Tejaswi Sir” for motivating and supporting me. I don’t have to worry much because of your wonderful disaster management skill. Thank you for tolerating my crankiness. I am grateful to my Atya, Mama Garu and Vadina for their continuous prays and support.

All that I cherish today is the grace of God. I thank the Almighty for granting me the strength, wisdom, knowledge and showering his blessings.

.....*Archi Saurabh*

ABBREVIATIONS

1D	-	One dimension
3D	-	Three dimension
α -syn	-	Alpha-synuclein
ΔG_{sol}	-	Solvation energy
ρ_{AA}	-	Number density of AA
Γ_{P-AA}	-	Preferential interaction coefficient
χ_{hyd}	-	Hydration coefficient
AA	-	Amino acid used as osmolytes
Amp	-	Ampicillin
Arg	-	Arginine
Asn	-	Asparagine
Asp	-	Aspartate
CD	-	Circular dichroism
Chl	-	Chloramphenicol
CMD	-	Conventional molecular dynamics
CNS	-	Central nervous system
COM	-	Center of mass
Coul	-	Coulombic
CTR	-	C-terminal region
C27FF	-	CHARMM27 force field
C36IDPSFF	-	CHARMM36 force field modified for IDPs
DLB	-	Dementia with Lewy bodies
FC	-	Fibril core
FEL	-	Free energy landscape
FL	-	Full-length protein
$g(r)$	-	Integral RDF values
$G(r)$	-	Kirkwood-Buff integral
Gln	-	Glutamine

Glu	-	Glutamate
H1`	-	Helix-1
H2	-	helix-2
HFiP	-	Hexafluoroisopropanol
HPR	-	Hydrophobic region
$I_{(X)}$	-	Interaction energy in pure water
$I_{(X,\emptyset)}$	-	Interaction energy in protein-water
IDP	-	Intrinsically disordered protein
IDPR	-	Intrinsically disordered protein region
IPTG	-	Isopropyl- β - d-1-thiogalactopyranoside
IVR	-	Isovolumic regulation
k_B	-	Boltzmann constant
k_{el}	-	Rate of fibril elongation
K_p	-	Partition coefficient
LB	-	Luria-Bertani
L-J	-	Lennard-Jones
Lys	-	Lysine
M _F	-	A single protein chain of α -syn taken from fibril structure
M _S	-	Monomeric α -helical form of α -syn
MD	-	Molecular dynamics
n_w	-	Average number of waters within cutoff
n_{AA}	-	Average number of AAs within cutoff
N_w^{tot}	-	Total number of water in the system
N_w^{tot}	-	Total number of AAs in the system
NAA	-	N-Acetyl aspartate
NAAG	-	N-acetylasparylglutamate
NAC	-	Non-amyloid core
NAG	-	N-acetyl-L-glutamic acid
NAL	-	N-acetyl-L-lysine

NMR	-	Nuclear magnetic resonance
NTR	-	N-terminal region
PCR	-	Polymerase chain reaction
PD	-	Parkinson's disease
PDB	-	Protein data bank
PMSF	-	Phenylmethanesulphonyl fluoride
RDF	-	Radial distribution function
REMD	-	Replica exchange molecular dynamics
RMSD	-	Root mean square deviation
RMSF	-	Root mean square fluctuation
R _g	-	Radius of gyration
RTD	-	Pseudo-repeat domains
RVD	-	Regulatory volume decrease
T	-	Temperature
SASA	-	Solvent accessible surface area
SMD	-	Steered molecular dynamics
<i>SNCA</i>	-	Human synuclein- α gene
t_o	-	Time required to attain half of the total fluorescence change
t	-	Time
T_{app}	-	Apparent fibrillation time
T_{lag}	-	Lag time
TEM	-	Transmission electron microscope
TFE	-	Trifluoroethanol
TMAO	-	Trimethylamine N-oxide
US	-	Umbrella sampling

CONTENTS

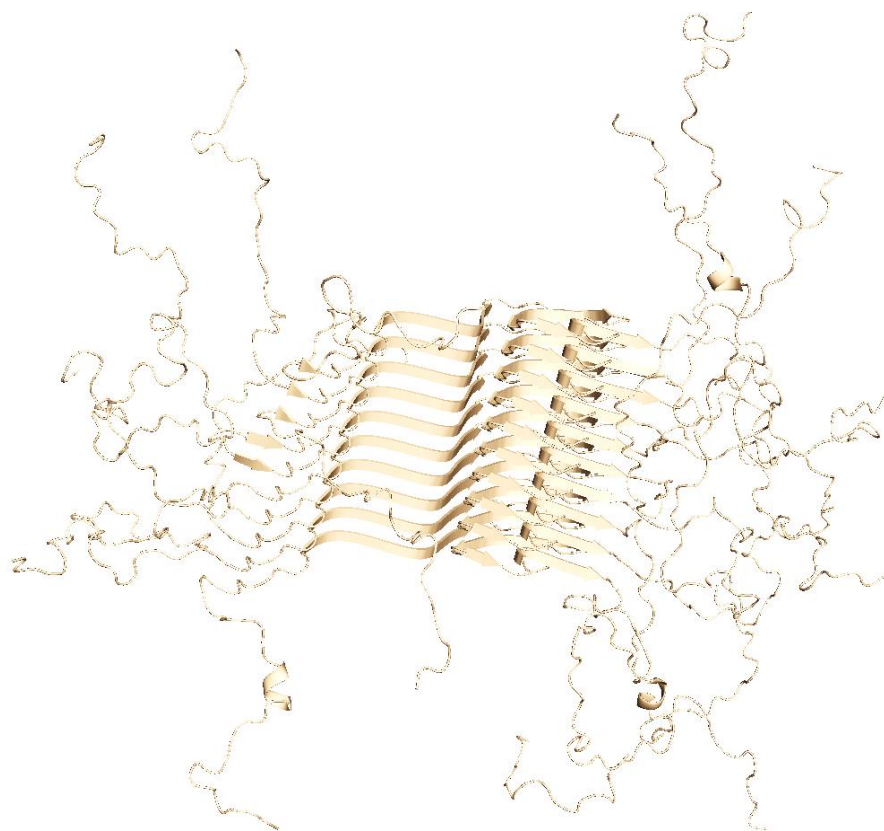
Chapter 1	1
Introduction	1
1.1 Evolution of Proteins	3
1.2 Protein folding	4
1.3. Protein fibrillation	5
1.4. Energy landscape: globular proteins vs intrinsically disordered proteins	6
1.5. α -Synuclein and fibril formation	8
1.6. Neural osmolytes and α -synuclein	10
Chapter 2	13
Elucidating the helix-fibril transition pathway of human α -synuclein by concerted enhanced-sampling simulations	13
2.1. Abstract	15
2.2. Introduction	16
2.3 Methods	19
2.3.1. Steered molecular dynamics (SMD) and umbrella sampling (US)	19
2.3.2. Replica exchange molecular dynamics (REMD)	20
2.3.3. Conventional molecular dynamics simulation (CMD)	22
2.3.4. Analysis	22
2.4. Results	24
2.4.1. Conformational sampling of fibril dissociation by SMD-US	24
2.4.2. Intermediate conformations from FEL	28
2.4.3. REMD simulation of a single chain from the fibrillar structure (MF):	31
2.4.4. REMD simulation of monomeric α -syn (MS):	33
2.4.5. Structural characteristics of IR1-IR8 conformations	35
2.4.6. Conventional-MD of MF and MS conformations of α -syn	37
2.4.7. Structural characteristics of IC1-IC3 intermediates	40
2.4.8. Fibril dissociation pathway(s)	42

2.4.9. Solvation in different steps of fibril-dissociation	49
2.4.10. Dihedral angle distribution during transitions	56
2.5. Discussion	59
2.5.1. Sampling of fibril dissociation pathway:	59
2.5.2. Monomer to fibril - the probable reverse transition pathway	61
2.5.3. Limitations	62
2.6. Conclusion	63
Chapter 3	65
Effect of charged amino acids and their derivatives on the fibril formation of α -synuclein	65
3.1 Abstract	67
3.2 Introduction	68
3.2.1 Osmolytes in neural environment	68
3.2.2 Osmolytes on protein fibrillation	69
3.3 Materials and Methods	70
3.3.1 Materials	70
3.3.2 Cloning of α -synuclein	70
3.3.3 Expression of α -synuclein	71
3.3.4 Purification of α -synuclein	71
3.3.4 Western blot analysis	71
3.3.5 Structure and fibrillation studies	72
3.3.6 Data analysis	73
3.3.7 Microscopic studies	73
3.4 Results	73
3.4.1 Confirmation of clone	73
3.4.2 Fibrillation of α -Syn	74
3.4.3 α -Syn fibrillation in the presence of amino acids	75
3.5 Discussion	78
3.5.1 Amino acid osmolytes	79

3.5.2 Fibrillation of α -syn in cosolvents	79
3.6 Conclusion	81
Chapter 4	83
Structural insights of the interaction between α -synuclein and amino acid osmolyte	83
4.1 Abstract	85
4.2 Introduction	86
4.2.1 Osmolytes	86
4.2.1 Osmolytes in protein aggregation	87
4.3 Methods	89
4.3.1 Conventional molecular dynamics (MD) simulation	89
4.3.2 Preliminary analysis	90
4.3.3 Preferential interaction	91
4.3.4 Cosolvent clustering	92
4.4 Results	93
4.4.1. Global analysis	93
4.4.2. Surface solvation of α -syn	96
4.4.3 Dynamics of Cosolvent	101
4.4.4 Conformational characteristics of α -syn in presence of AA	102
4.5 Discussion	105
4.5.1 AAs accelerating fibrillation with lag-dependent mechanism	106
4.5.2 AAs accelerating fibrillation with lag-independent mechanism	107
4.5.3 Fibrillation inhibiting AA with α -syn	107
4.5.4 Molecular interactions and fibrillation propensity	108
4.6 Conclusion	110
Overall summary	111
Bibliography	113
List of Publication	130
Plagiarism Certificate	131

Chapter 1

Introduction



Chapter 1

“Biology is the study of complicated things that give the appearance of having been designed for a purpose.” - Richard Dawkins.

In nature, protein molecules exhibit an extraordinary interconnection between their structure and function. Proteins play diverse roles with immense specificity and accuracy dictated by their three-dimensional (3D) structure. This 3D structure is determined by the sequence of amino acids that make up the protein's 1D strings. In 1961, Anfinsen et al. showed that the folding of RNase A is driven by spontaneous folding free energy.^[1] This observation laid a foundation for structural biologists to explore the relationship between the sequence, structure and function of proteins.

1.1.Evolution of Proteins

The interest in protein structure and folding began way before Anfinsen. Langmuir, in 1938^[2] introduced the hydrophobic concept in protein stability based on Traube's surface tension studies^[3] in the late 19th century.^[4] Hydrophobic concept or a factor play a major role in the structural organization of protein polymers by reducing water molecules interaction from the non-polar residues on the protein surface. Although the hydrophobic principle was recognized in the early 20th century,^[5,6] it was popularized among protein chemists only after Kauzmann took it forward.^[7]

During the same period, Fischer and Fischer proposed that enzymes are proteins and their catalytic activities depend on their 3D-structure.^[8] Nevertheless, the biggest challenge was how to resolve the 3D-structure of a protein. The revolutionary article by Pauling and Corey in 1951 deduced the two major structural features of proteins: α -helix and β -sheet.^[9] In their opening remarks, they stated “The problem we have set ourselves is that of finding all hydrogen-bonded structures for a single polypeptide chain, in which the residues are equivalent (except for the differences in the side chain R).” which is indeed the basic framework for a protein secondary structure till date. Although the discovery of α -helix and β -sheet was significant, the discovery of sophisticated X-ray crystallographic methods lead to the depiction of complete 3D-conformation of the globular protein, myoglobin by Kendrew and Perutz in 1958.^[10,11]

Chapter 1

1.2. Protein folding

Over the following decades, researchers used X-ray crystallography, NMR spectroscopy, and other techniques to investigate the structures of numerous proteins. After understanding the role of H-bonding and hydrophobic concept the major concern was how proteins attain their generic 3D-conformation. In 1960s, Anfinsen^[1] introduced the concept that the native state (N) can be attained by its amino acid sequence which will be unique, stable, active and highly accessible which is later known as Anfinsen's dogma.

“It struck me recently, that one should really consider the sequence of a protein molecule about to fold into a precise geometric form as a line of melody written in a canon form and so designed by the Nature to fold back into itself, creating harmonic chords of interaction consistent with biological function.” - Christian Anfinsen

In 1964, based on the folding equilibrium data of chymotrypsin, Brandts proposed two-state model for protein folding mechanism. He suggested that the native (N) and unfolded (U) conformations stay in an equilibrium.^[12] This concept is still valid for many of the single-domain proteins.^[13,14] The two-state protein folding mechanism guided the musing of Levinthal's article “Are There Pathways for Protein Folding?”^[15] where he stated “a native protein exists in some kind of thermodynamic configurational equilibrium, with the biologically active state being the one with the lowest configurational energy” which is popularly known as Levinthal's paradox. This indicates that protein folding is not a random search, it should have a folding pathway which is energetically favorable.^[16,17] Thereafter, many protein chemists proposed different methods to isolate or identify folding intermediates^[18–20] including molten globules. Molten globules from different proteins shares a common property, like tertiary structure, non-rigid side chains and compact conformations.^[21] Though the presence of folding intermediates is widely accepted, its role is still controversial among researchers. Folding intermediates shed light on energy landscape and folding kinetics, but it also arises the question of whether these intermediates lead toward folding or merely represent misfolded conformations.

1.3. Protein fibrillation

In the course of understanding the function and kinetics of folding intermediates or molten globules, researchers encountered a natural phenomenon where proteins fail to fold properly and form large highly stable oligomeric states resulting in a dramatic decrease in their native conformations.^[22–24] In 1851 Virchow observed an aggregation in cerebral corpora and named it “amyloid”.^[25] He predicted that it could be a cellulose accumulation. Later in 1859, Friedrich and Kekule identified it as protein aggregates.^[25] Despite of varying amino acid sequences and native conformations, the amyloids formed by proteins have a common structural arrangement. Amyloid fibrils are an outcome of homotypic polymerization of protein monomers resulting in protofibrils intertwined with each other and this process is termed as fibrillation or amyloidosis. Amyloid fibrils are several micrometers long and 5-15 nm width filamentous structural architect to form cross- β as tertiary contacts in which multiple β -strands run perpendicular to the fibrillar axis as illustrated in Fig. 1.1A.^[26–29]

The fibrillation process of a protein might include three stages, lag phase, exponential phase and stationary phase, and occurs by two different pathways, lag-independent and lag-dependent as shown in Fig. 1.1B

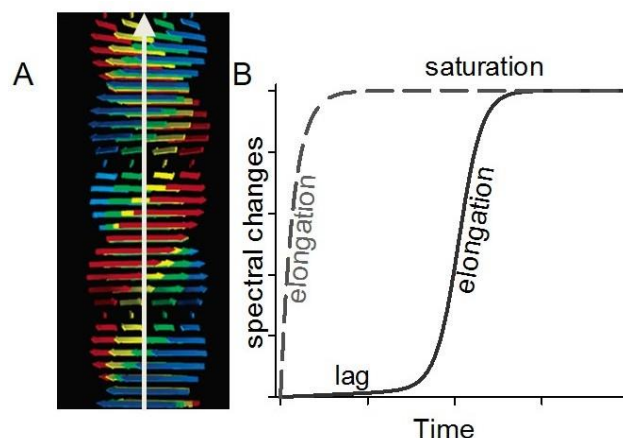


Fig. 1.1. Protein fibrillation. (A) Pictorial representation of protein fibrils adapted from *Anthony, 2006*^[30] showing multiple β -strands stacked perpendicular to the fibrillar axis represented as white colour arrow. The protofibrils are represented in different colours as red, yellow, blue and green intertwined with each other to form mature fibrils. (B) Representative fibrillation kinetics of a protein following a lag-dependent (solid line) and lag-independent (dashed line) fibril formation pathway monitored by a spectral probe. Different phases of fibril formation, lag, elongation and saturation are marked in the graph.

Chapter 1

During lag phase, soluble protein monomers undergo oligomerization which further acts as a nucleus to form protofibrils in lag-dependent pathway. In case of lag-independent pathway, soluble protein monomers do not undergo nucleation, but form fast-intermediate oligomers. These protofibrils or intermediate oligomers rearrange themselves to form fibrils during the elongation phase. If the fibrils formed are stable, then this process is known as primary process. Otherwise, in the secondary process, these fibrils undergo fragmentation and reassociation or twisting and branching to form mature fibrils. The probable steps involved in fibril formation are schematically explained in Fig. 1.2.^[31,32]

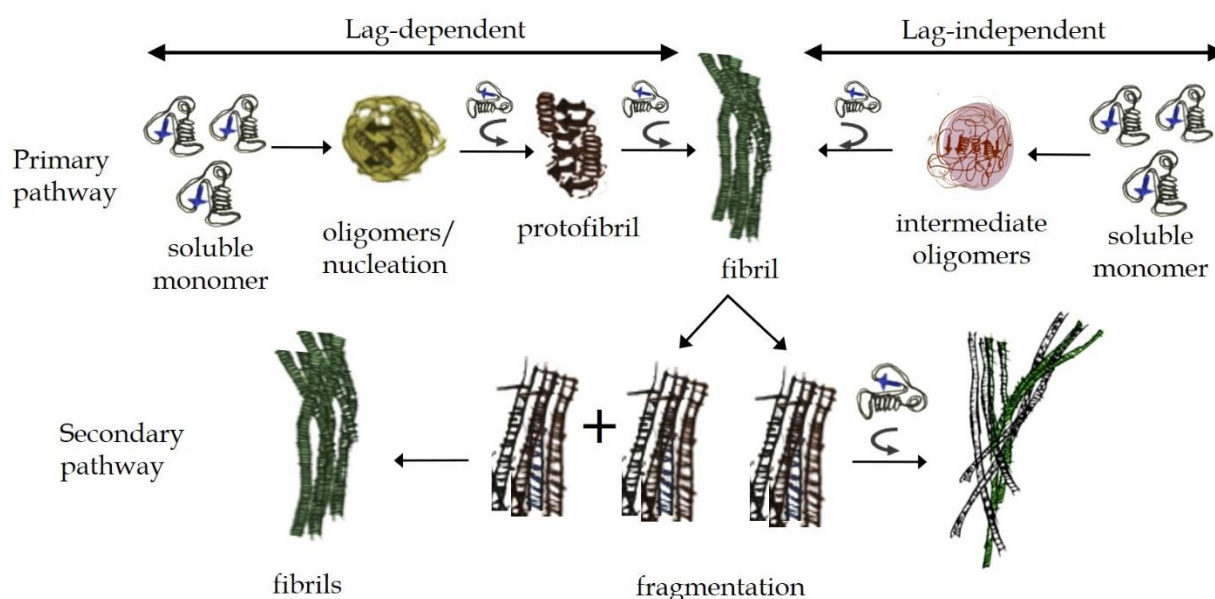


Fig. 1.2. Protein fibrillation occurs through two pathways:^[32] the primary pathway (top row) involving lag-dependent mechanism (left side) which requires protofibril formation or lag-independent (right side) intermediate oligomers which directly form fibrils. The secondary pathway (bottom row) involving fragmentation and reassociation (left side) or branching (right side) of existing fibrils.

1.4. Energy landscape: globular proteins vs intrinsically disordered proteins

In order to understand the protein folding and aggregation kinetics, it is important to know the intermediate states and how they guide proteins to attain their native structure in living cells within a fraction of a second. It is known that proteins can attain multiple conformations which makes finding the intermediates “a needle in the haystack”. In a protein, the compact and low energy conformational ensembles have less entropy distribution.^[33–37] Here, ensembles refer to various microscopic states occurring simultaneously and even contribute to many microscopic

processes together subsequently. This idea converted the sequential folding pathway to funnel-based parallel events. This concept was a solution to Levinthal's paradox as "folding funnel, not tunnel".^[37] Let's consider a monomeric linear sequence of homopolymer that can attain all the spatial conformations available at the top of a folding funnel and referred as unfolded states due to its high entropy. The minimum energy search from these conformations may lead the polypeptide chain to different minima states in the funnel such as folding intermediates, partially unfolded states, the native conformation or aggregates by random condensation followed by chain reconfiguration with entropy reduction^[34] as shown in Fig. 1.3A.

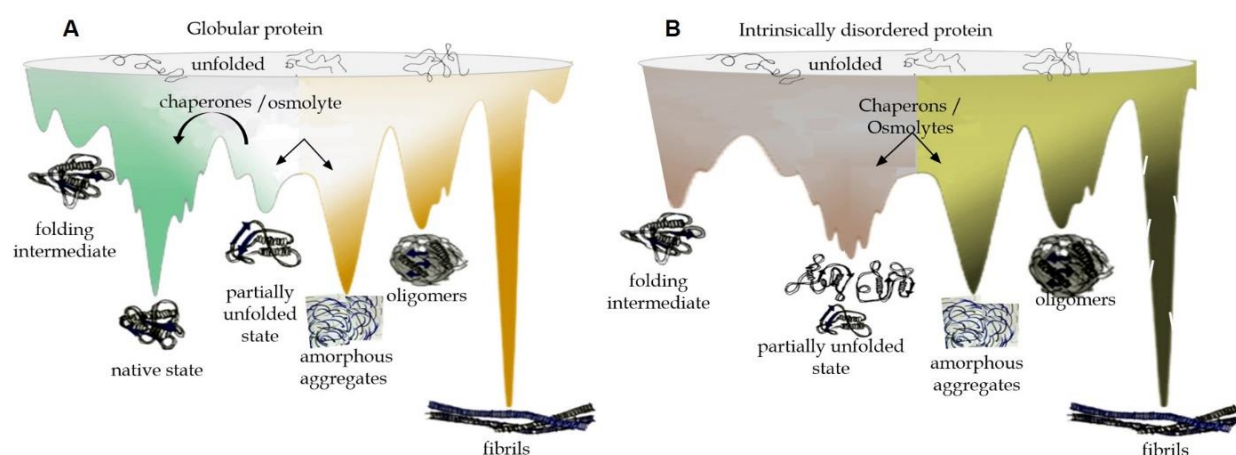


Fig. 1.3. Energy landscape: (A) Globular protein folds to the native conformation with stable intramolecular interactions from different unfolded states. This could be disturbed by a stress condition leading to a higher entropy conformation such as folding intermediate or partially unfolded state. (B) Intrinsically disorder protein has ensembles of partially unfolded conformations due to its highly rough lower energy surface for intramolecular interactions. In presence of osmolytes, these partially unfolded states might undergo oligomerization to form amorphous aggregates or highly stable amyloid fibrils through the processes described in Fig. 1.2 or with the help of chaperons it could be refolded to prevent aggregation.

“Orgel's First Rule: Whenever a spontaneous process is too slow or too inefficient, a protein will evolve to speed it up or make it more efficient.” - Leslie Eleazer Orgel^[38]

Earlier studies on globular proteins were focused on protein structure-function paradigms that a protein requires a native folded structure to perform its biological function. This considered the proteins as compact and in a packed conformation.^[39,40] After Koshland's “induced-fit model”, the presence of flexibility in a protein to perform a biological function was widely accepted.^[41] Protein structure evolves based on its function inside the cell and also due to its stability, environmental properties, aggregation propensity and the interacting chaperones.^[42,43,52–54,44–51]

Chapter 1

For a long time, the disordered characteristics of some of the proteins was neglected due to missing electron density, difficulty in purification and lack of structure and considered as artifacts or experimental error. However, the structure-independent functions were so noticeable due to its abundance in all the organisms^[55–61] which resulted in the discovery of many disordered proteins. It was also observed that the complexity of organisms increases the fraction of disordered protein or region, but the reason is yet to be clearly established. In some studies, it is linked with increase in multidomain functions in proteins in higher organisms^[49,56,67,68,57,58,60,62–66].

These proteins or unstructured regions in proteins are named as intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs), respectively. Hydrophobic concept and electrostatic repulsion were well understood, but IDPs introduce the classic chaos by these forces which set a new arena for biochemists.^[69–73] IDPs and IDPRs can attain multiple conformations by interacting with different substances and may act on completely unrelated targets for biological functions (termed as ‘moonlighting’)^[70,74–77] Thus, it plays wider roles in cell signaling, maintaining cell integrity and transporters etc. The ensemble of conformations accessed by a single polypeptide chain of IDP even at lower energy states called as partially unfolded states results in a more roughed landscape compared to globular proteins^[78,79] (Fig. 1.3B). This roughed landscape drives many IDPs to undergo fibrillation which is reported to be involved in many disease conditions.^[80,81]

1.5. α -Synuclein and fibril formation

Synucleinopathies are neurodegenerative disorders with a characteristic accumulation of fibrillar aggregates of the protein, α -synuclein (α -syn).^[82–85] α -syn is a 140 amino acids protein expressed abundantly in the brain primarily in the synaptic termini.^[86,87] The presence of an aggregated form of α -syn brought major attention as a mediator of Parkinson’s disease (PD) and dementia with Lewy bodies (DLB) pathogenesis.^[86] The monomeric form of α -syn structure (protein data bank^[88] id: 1XQ8) consists of two α -helices ranging from V3 to V37 (helix-1) and from K45 to T92 (helix-2)^[89] and an unstructured C-terminal domain (Fig. 1.4A). The fibrillar form of the protein (PDB id: 2N0A) has fibril structure with parallel in-register β -sheets between polypeptide chains having Greek-key topology formed by the residues V37 to D98 form (Fig. 1.4B).^[90] This region consists of five β -strands from each monomer (Fig. 1.4.L5), namely strand-A (L38-V55), strand-B (E61-V66), strand-C (V70-A78), strand-D (T81-E83) and strand-E (I88-

K97). Based on fibril association, the α -syn sequence can be divided into three regions, (Fig. 1.4.C)^[91]:

1. *N*-terminal (NTR: residues 1-36) - an amphipathic membrane binding domain.^[92]
2. Fibrillar core (FC: residues 37-98) – consists of non-amyloid β -component (61-95 amino acids) a hydrophobic region essential for fibrillation.^[93] The amino acid residues 37-98 has the potential to form fibrils by self-association.
3. *C*-terminal (CTR: 99-140) - a disordered region mostly consists of acidic residues.^[92]

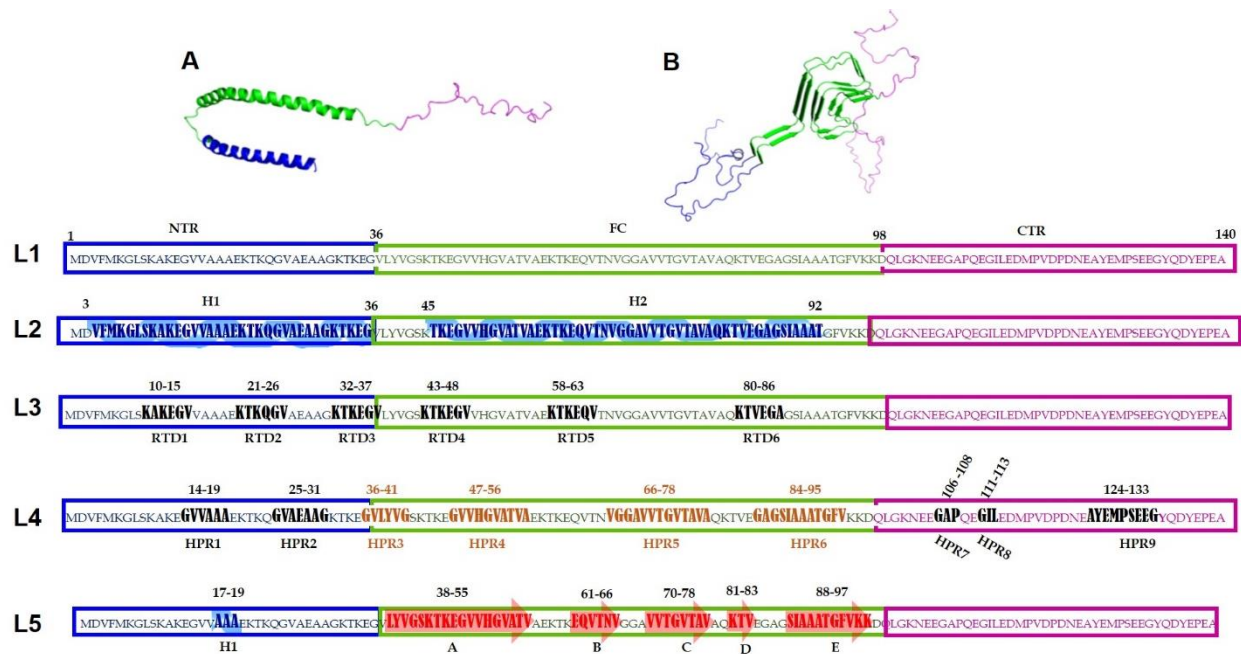


Fig. 1.4. Structure of α -synuclein in its (A) monomeric (PDB id: 1XQ8) and (B) fibrillar form (PDB id: 2N0A). The blue, green and pink colors represent the N-terminal, fibril core and C-terminal regions, respectively. (C) Structural characteristics of α -syn: (L1) The polypeptide chain of α -syn is divided into three regions: N-terminal region (NTR: M1-G36), fibril core (FC: V37-D98) and C-terminal region (CTR: Q99-A140). (L2) Monomeric form α -syn (PDB id: 1XQ8) with helix-1 and helix-2 represented as blue colour helices. (L3) Six pseudo repeat domains in the protein are marked as RTDs. (L4) Nine hydrophobic regions are marked as HPRs. (L5) Monomer from fibril (PDB id: 2N0A) with five red colour arrows representing β -strands.

The protein has six repeat domains named as RTD, three in NTR as RTD1-RTD3 and three in FC as RTD4-RTD6) demonstrate in 1.4.L3 as RTD1 (K10-V15), RTD2 (K21-V26), RTD3 (K32-V37), RTD4 (K43-V48), RTD5 (K56-V63) and RTD6 (K80-A86). In general IDPs have high electrostatic repulsion and low hydrophobicity in the sequence^[94] but exceptionally in α -syn there

Chapter 1

are also nine HPR (Hydrophobic regions) demonstrate in 1.4.L4 as two in NTR named as HPR1 (G14-A19) and HPR2 (G25-G31), four in FC as HPR3 (G36-G41), HPR4 (G47-A56), HPR5 (V66-A78) and HPR6 (G84-V95), and three in CTR as HPR7 (G106-P108), HPR8 (G111-L113) and HPR9 (A124-G133).^[91]

The intrinsically disordered nature of α -syn makes it unique and challenging to depict its aggregation mechanism. Due to its chameleon's nature, it may either hold an unfolded or partially folded amyloidogenesis conformation or may even fold into an α -helical or a β -rich structure in both monomeric or oligomeric forms.^[95] Conformational ensembles of α -syn generated using various computational and experimental methods have provided certain insights into its fibrillation mechanism, but a clear correlation between the structures and the fibrillation initiation is still a topic of debate. In chapter 2, we demonstrated the major intermediates structures of α -syn during its monomeric^[89]to fibril^[90] transition by concerted enhanced sampling methods and classical dynamics simulation (CMD) of monomeric and fibrillar forms of the protein.

1.6. Neural osmolytes and α -synuclein

During hyperosmotic stress in the hippocampus, there have been observations of cell-dependent fibril formation of α -syn.^[96] α -Syn is partially unfolded in nature, which allows its residues to have increased access to the solvent, and their interactions become crucial in conformational transitions. In hyperosmotic conditions, neural cells respond by releasing both organic and inorganic osmolytes in high concentrations.^[97] Osmolytes are a significant class of smaller molecules that cells accumulate to counteract external stresses and protect macromolecules from the stress environment.^[98,99] These osmolytes also play a role in regulating molecular chaperones within cells.^[100] When cells experience combined salt and heat stresses, osmolytes are found to stabilize or destabilize the protein surface by altering the nature of the hydration shell and affect aggregation and folding of proteins demonstrate in Fig. 1.5.^[101–104] Organic osmolytes can be categorized into different groups such as carbohydrates (polyols and sugars), amino acids (arginine, glutamate, and glutamine), methylamine (TMAO), and methyl sulphonium (urea).^[105] The release of charged amino acids and their derivatives, such as glutamate and N-acetylated aspartic acid, in an isovolumic state regulates protein folding and aggregation within the cell.^[97,99,106] Previous research has suggested that charged amino acids can impede the

formation of protein fibrils that are involved in neuronal diseases.^[107–109] Therefore, in chapter 3 we are trying to analyze effect of charge distribution on polar amino acids and its derivatives (AAs) on the fibrillation of α -syn (recombinant human α -syn) and in chapter 4 the molecular mechanism behind the AA- α -syn interactions are depicted with the help of CMD.

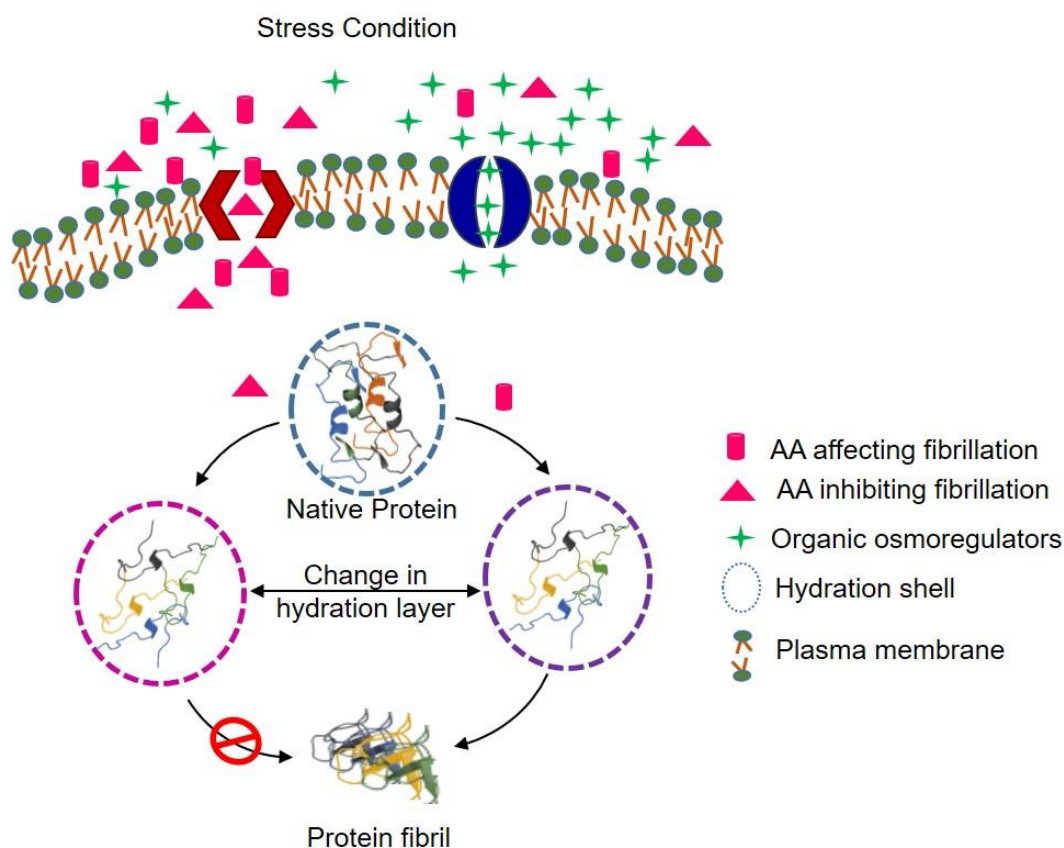
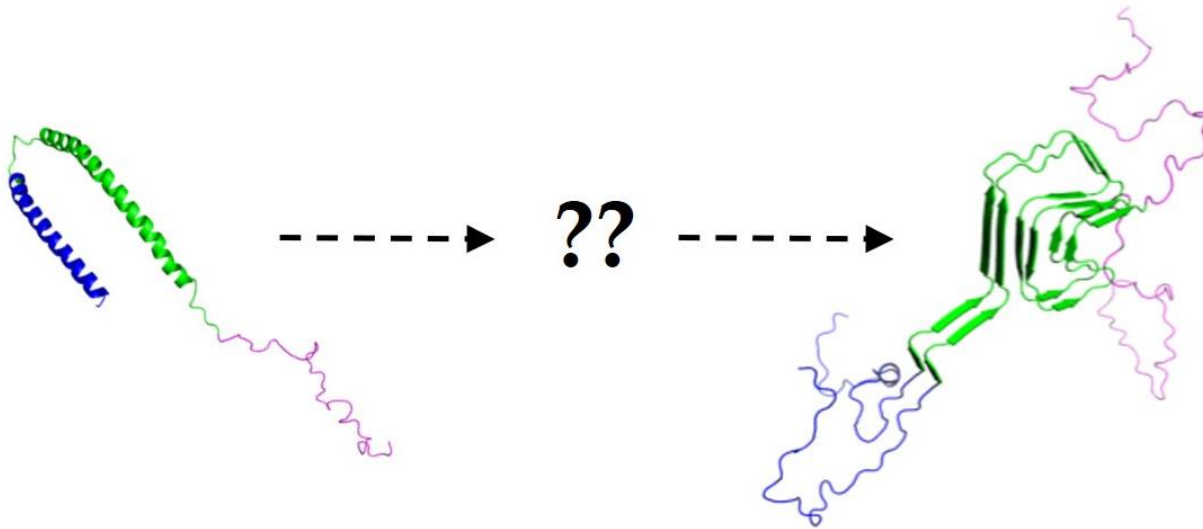


Fig. 1.5. AA as osmolytes: In stress environment cells release organic (green coloured) and inorganic (pink coloured) osmolytes. Here, AA from class of inorganic osmolytes affects partially unfolded state shown in Fig.1.3 by altering its hydration shells and could promotes or inhibits the fibrillation process.

Chapter 1

Chapter 2

Elucidating the helix-fibril transition pathway of human α -synuclein by concerted enhanced-sampling simulations



Chapter 2

2.1. Abstract

Fibril formation of α -synuclein (α -syn) is directly linked with Parkinson's disease. Due to its intrinsically disordered nature, α -syn shows extensive conformational plasticity. Hence, it is quite impossible to identify all the intermediate structures of α -syn during its transition from the native monomeric to disease-associated fibril form by any single sampling method. Therefore, we implemented different simulation methods such as steered molecular dynamics (SMD) followed by umbrella sampling, and replica-exchange and conventional MD simulations of both the native (M_S) and a single chain from the fibril state (M_F). Nineteen distinct intermediate structures were obtained from these simulations either by plotting free energy landscape or cluster analysis. These structures were then sorted based on their secondary structural content, radius of gyration and solvent exposure to depict the fibril dissociation pathway from cross β -sheets to α -helical conformation. The analysis showed that following the initial dissociation of the polypeptide chain from the fibril, the protein might either form compact conformations by long-range interactions or form extended conformations stabilized by local interactions. This could lead the protein to adapt two different pathways which would further converge into a helical-conformational pathway. The secondary structure content, solvent accessibility, contact distance, backbone dihedral angles of sixty selected residues, and interaction energies were analyzed for all the 19 intermediate structures. The results suggested that the formation of β -turn in the fibril core region, reorganization of salt bridges in the pseudo-repeat domains, dihedral changes in the hydrophobic regions are the major driving forces for the conversion of α -syn from monomer to fibril state. Further, the long-range interactions between *N*- and *C*-terminal regions are found to alter the stability of the α -helical structure. The interactions in the identified structures also correlated with the earlier experimental and computational studies. Thus, the present study provides critical information about the essential intermediate structures and pathways leading to the fibril formation of α -syn.

Chapter 2

2.2. Introduction

α -Synuclein (α -syn) is an intrinsically unfolded 140 amino acid residue protein. It is abundantly expressed in the brain, primarily in the synaptic termini.^[110–112] The aggregated form of α -syn is associated with Parkinson's disease (PD) and dementia with Lewy bodies (DLB) pathogenesis.^[86,110,113,114] The protein has three major domains (Fig. 2.1): (i) membrane-binding *N*-terminal domain (residues 1-60) consists of four pseudo tandem repeats of KTKXGV (RTD1 to RTD4) responsible for making the *N*-terminal positively charged, (ii) highly hydrophobic central domain (residues 61-95) known as non- amyloid component (NAC) which also contains a repeat domain (RTD6),and (iii) highly acidic *C*-terminal domain (residues 96-140) with 5 aspartate and 10 glutamate residues.^[115–117] Further, a repeat domain, RTD5 overlaps with both *N*-terminal and NAC regions (residues K58-V63).

Compared to structurally well-defined globular proteins, intrinsically disordered proteins (IDPs) exhibit larger conformational fluctuations and are sensitive to the solvent environment.^[118,119] α -Syn adapts different conformational states in water and in surfactants or lipids. The *N*-terminal of α -syn forms an extended single helix or broken helices with a non-helical linker region depending on the protein to surfactant ratio. The broken helical form of α -syn consists of two α -helices ranging from V3 to V37 (helix-1) and from K45 to T92 (helix-2), respectively.^[89] Helix-1 interacts with membranes and transforms helix-2 into a β -sheet that facilitates the oligomerization process.^[120,121] The toxic oligomeric intermediates formed by α -syn are invariably β -rich in NAC region.^[122,123] The *C*-terminal has disordered conformation in either of the structural states, monomeric or fibril; however, it maintains long-range interactions with the rest of the protein.^[124,125] The residues D115-D119 and Y125-S129 in *C*-terminal might form interactions with the residues in NAC region which could influence the fibril formation of α -syn.^[125] Also, the interaction of hydrophobic region in CTR with helix-1 in *N*-terminal delays the fibrillation process.^[126,127]

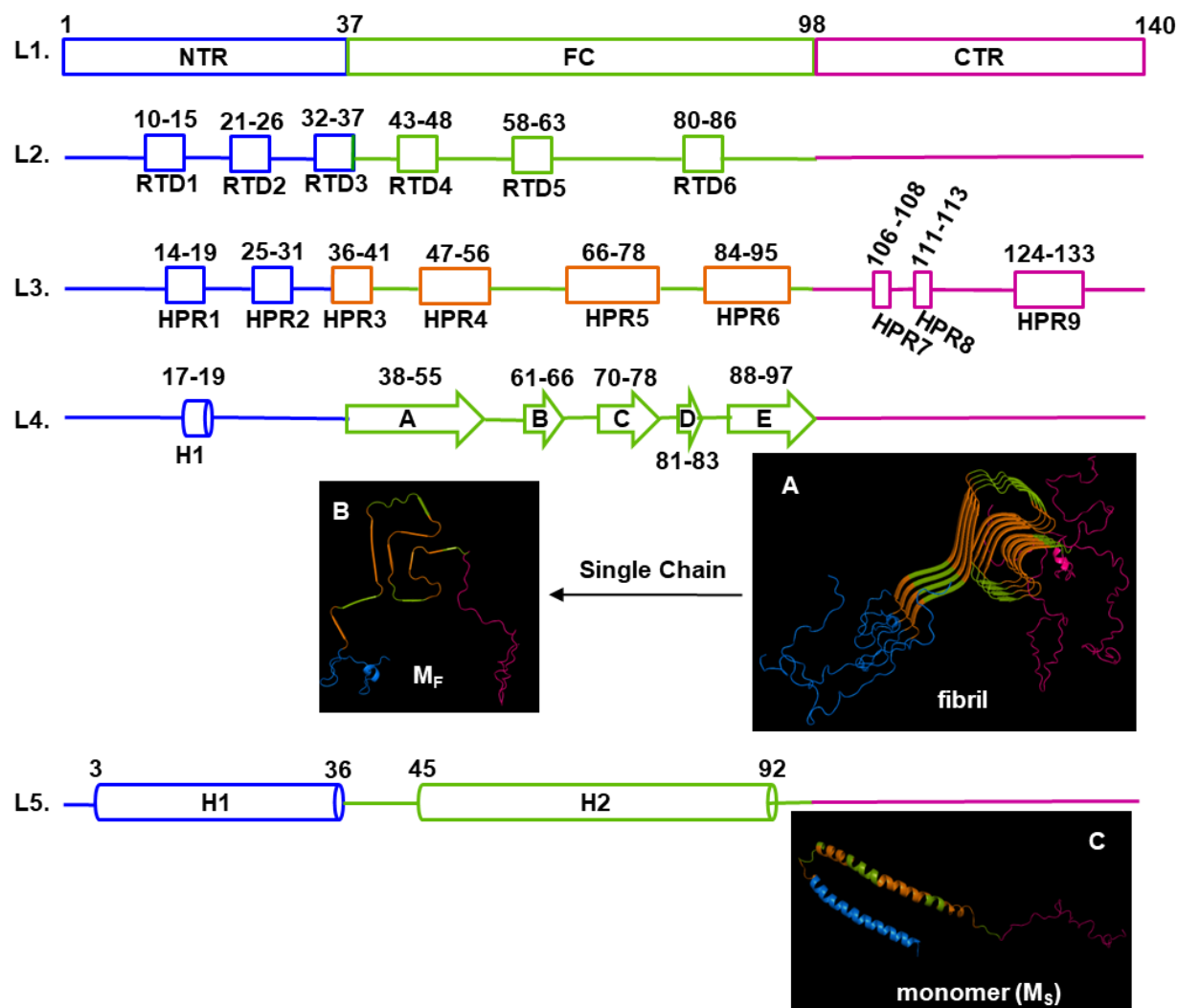


Fig. 2.1. Structural characteristics of α -synuclein: (L1) The polypeptide chain of α -syn is divided into three regions: *N*-terminal region (NTR: M1-G36), fibril core (FC: V37-D98) and *C*-terminal region (CTR: Q99-A140). (L2) Six pseudo repeat domains in the protein are marked as RTDs. (L3) Nine hydrophobic regions are marked as HPRs. (L4) α -helix (cylinder) and β -sheet (arrows) regions in the fibril form of α -syn. The panels show (A) the fibril assembly of α -syn obtained from PDB id: 2N0A, and (B) a single chain from the fibril assembly, M_F . (L5) Monomeric form α -syn (PDB id: 1XQ8) with helix-1 and helix-2 represented as cylinders. Panel (C) shows the monomeric helical form of α -syn, M_S . In all the lines, L1-L5, blue, green and pink colors represent the NTR, FC and CTR, respectively, and the orange color represents HPRs in the FC.

The matured fibrils of α -syn are formed by cross- β interactions of the residues in the central NAC region. A stretch of amino acid residues prior to NAC region known as pre-NAC segment (37-60) is also reported to involve in various familial mutants linked to PD and has implications in the fibrillation of α -syn.^[87,128] The β -hairpin formed by residue 36-55 of α -syn promotes

Chapter 2

nucleation-dependent fibril formation.^[129,130] It is also reported that β -wrapin designed with the residues in pre-NAC region (residues 37-54) interferes with the nucleation process and inhibits the fibril formation.^[131,132] The familial mutants in the *N*-terminal of α -syn might affect the fibril elongation process and lead to polymorphism in the fibrils.^[93,133] The fibril structure resolved from solid-state NMR experiments shows that the residues V37 to D98 form the fibril core (FC) with parallel in-register β -sheets having Greek-key topology.^[90] This FC consists of five β -strands from each monomer, namely strand-A (L38-V55), strand-B (E61-V66), strand-C (V70-A78), strand-D (T81-E83) and strand-E (I88-K97).^[90] From H/D exchange mass spectrometric experiments, two types of oligomeric states are identified. The fibril core sequence Y39–T75 is protected from hydrogen exchange in both type-I and type-II oligomers, but the flanking regions (A18-L38 and A76-A89) are protected only in type-I. Oligomer of type-I forms straight fibrils, whereas oligomer type II forms amorphous aggregates.^[134] Further, there could be differences between the β -sheet geometry of the intermediate oligomeric species (anti-parallel β -sheet) and the fibrillar form of the protein (mainly parallel β -sheet).^[130,135]

Different mechanisms have been proposed for the initiation of structural changes in the fibrillation of α -syn. It is suggested that fibrillation begins with the destabilization of helical tetrameric aggregates leading to misfolded α -syn monomer followed by aggregation,^[136,137] while the presence of monomeric α -syn has been reported to be harmless in mouse brain and mammalian cell lines.^[138] A few studies suggest that monomeric α -syn bound to a membrane via helix-2 region loses its helicity and transforms into a β -rich conformational state. This might further lead to fibril formation. It is also suggested that these β -sheets may be different in conformation than the fibril state ones.^[139] The disruption of long-range interactions between the CTR with NTR and NAC regions could be one of the major steps in fibril initiation. NTR retains its helical character in oligomeric states.^[140] In the absence of any lipids, the initial monomeric state might show short β -sheet-like structures. The β -sheet to oligomeric conversion might be the major energy barrier during fibril formation.^[120]

The conformational ensembles of α -syn generated using various computational and experimental methods illustrate different structures adopted by the protein. However, the intrinsic conformational plasticity of α -syn makes it more challenging to depict its aggregation mechanism.^[141–145] Most of the computational studies are carried out with the truncated forms of α -syn covering only the NAC or pre-NAC region to analyze the monomeric or fibrillar structures

of the protein.^[131,146–148] However, the experimental studies suggest the presence of long-range interactions between CTR and the other regions.^[149] Disruption of these long-range interactions might facilitate the aggregation of the protein by solvent exposure of the hydrophobic NAC region.^[125,150] Further, *in vitro* studies have demonstrated that CTR truncation alters the rate of fibrillation and morphology.^[151,152] The computational studies carried out with full-length protein, however, encounter conformational sampling limits and metastable intermediates during the fibrillation process could not be identified.^[93,153–156] A comprehensive study using different sampling methods might overcome this limitation and enable us to identify the possible intermediate conformations involved in the fibrillation pathway of α -syn.

In the present study, steered molecular dynamics (SMD) was implemented to pull out a single chain from the fibril assembly of α -syn followed by umbrella sampling (US) of the selected conformations from the fibril dissociation pathway. Further, replica exchange molecular dynamics (REMD) simulations were performed for a single chain taken from the fibril conformation (M_F) and for the monomeric form of α -syn (M_S). Conventional molecular dynamics (CMD) simulations were also performed for both conformations. The combined analysis of all these simulations demonstrated that there could be at least two different pathways for fibrillation distinguishable by the long-range interactions between the domains. Moreover, the major intermediate states during fibril dissociation and their structural features such as β -turns in pre-NAC and NAC regions, essential salt bridges and broken α -helix are identified.

2.3. Methods

2.3.1. Steered molecular dynamics (SMD) and umbrella sampling (US)

The fibrillar structure of α -syn was retrieved from the protein databank (PDB)^[88] having the ID: 2N0A.^[90] The structure consists of 10 chains. Out of this, the first three chains (A, B and C) were selected to represent the fibrillar assembly for steered molecular dynamics (SMD) study. CHARMM27 force field^[157] implemented in GROMACS 5.1.4 package^[158] was used for all the simulations. The protein chains were placed in a rectangular box and solvated with TIP3P water model. The system was energy minimized with the steepest descent algorithm. The system was then equilibrated at 310 K and 1 atm using V-rescale thermostat^[159] and Parrinello–Rahman barostat,^[160] respectively with non-bonded interaction cut-off of 1.2 nm. From the fibril assembly,

Chapter 2

the polypeptide chain at the one end (chain-A) was subjected to a center of mass (COM) based pulling^[161] along the Z-axis direction till the chain was completely dissociated from the other two chains. The weighted histogram analysis method (WHAM)^[162] was used to calculate the force on harmonic spring from a single SMD of 1450 ps. During this, a position-restrained was applied on the chains B and C.

The conformations for umbrella sampling (US) were selected from the trajectory of SMD based on the distance between the center of mass of chain-A and chain-B. For the distance ranging between 1.6 and 4.0 nm, the conformations were selected at every 0.1 nm which resulted in 24 conformations. This was considered as phase-I of SMD-US. From the distances ranging from 4.0 to 12.0 nm, the conformations were selected at every 0.4 nm and for the range 12.0 to 14.6 nm, every 0.2 nm interval was used. This provided 20 and 13 conformational states, respectively, which were labeled as phase-II and phase-III of SMD-US. The 57 selected conformations were subjected to independent conventional MD simulation (CMD) with NPT equilibration for 100 ps followed by 10 ns of production simulations. The potential of mean force from the trajectories was evaluated using WHAM^[162] analysis.

2.3.2. Replica exchange molecular dynamics (REMD)

The full-length monomeric structure of α -syn was obtained from PDB (ID:1XQ8) which has two anti-parallel helices in the *N*-terminal and a disordered *C*-terminal region.^[89] Monomeric α -syn (M_S) placed in a dodecahedron box was solvated with TIP3P water. The temperature intervals to run REMD between the range of 303 K to 328 K were chosen using REMD temperature generator server^[163] with an acceptance ratio of 0.2. There were 54 replicas generated and the temperatures are listed in Table 1. NPT and NVT equilibrations of M_S were performed for 2 ns at each temperature. It was followed by 10 ns of production runs with every 2 ps of exchange attempt between the replicas. Also, a single chain from the fibril structure of α -syn (M_F) was retrieved from the solid-state NMR structure of α -syn fibril (Chain-A from PDB id: 2N0A) and REMD simulation was performed for the same temperature range using 32 replicas following a similar procedure. The temperatures of replicas are presented in Table 1.

Table 2.1. Temperatures used for the replicas of REMD simulation.

Monomer from fibrillar α-Syn, M_F (2N0A)	Replica	Temperature (K)	Monomer Native α-Syn, M_S (1XQ8)	Replica	Temperature (K)
	1	303.00		1	303.00
	2	303.80		2	303.46
	3	304.61		3	303.93
	4	305.41		4	304.39
	5	306.22		5	304.86
	6	307.02		6	305.32
	7	307.83		7	305.79
	8	308.65		8	306.26
	9	309.46		9	306.72
	10	310.27		10	307.19
	11	311.09		11	307.64
	12	311.91		12	308.11
	13	312.73		13	308.58
	14	313.55		14	309.05
	15	314.38		15	309.53
	16	315.20		16	310.00
	17	316.03		17	310.47
	18	316.86		18	310.94
	19	317.69		19	311.42
	20	318.52		20	311.89
	21	319.36		21	312.36
	22	320.19		22	312.84
	23	321.03		23	313.31
	24	321.87		24	313.79
	25	322.71		25	314.27
	26	323.55		26	314.75
	27	324.39		27	315.22
	28	325.24		28	315.70
	29	326.09		29	316.18
	30	326.94		30	316.66
	31	327.79		31	317.14
	32	328.00		32	317.62
				33	318.10
				34	318.58
				35	319.07
				36	319.55
				37	320.03
				38	320.52
				39	321.00
				40	321.49
				41	321.97
				42	322.46
				43	322.95
				44	323.44 continued...

Chapter 2

45	323.93
46	324.42
47	324.90
48	325.39
49	325.89
50	326.38
51	326.87
52	327.36
53	327.85
54	328.00

2.3.3. Conventional molecular dynamics simulation (CMD)

CMD of monomeric form of α -syn (M_S) was performed in a dodecahedron box with two force fields CHARMM27 (C27FF)^[158] and CHARMM36-modified IDP specific force field (C36IDPSFF)^[164,165] using TIP3P and TIP4P-D^[164] water models. Initially, NVT and NPT equilibrations of the structures were performed for 1 ns each at 310 K and 1 atm pressure. This was followed by 200 ns of production simulation. CMD of the single chain obtained from fibril structure (M_F) was also carried out with the same experimental conditions.

2.3.4. Analysis

2.3.4.1. Preliminary analysis

C α -RMSD of full-length (FL) protein and only the residues in the fibril core (FC), Radius of gyration (Rg) of FL, and solvent accessible surface area of FL and FC residues were calculated using `gmx_rms`, `gmx_gyrate` and `gmx_sasa` modules of GROMACS, respectively. The contact map between the residues was obtained using `gmx_mdmat` module with a distance cut-off of 1.5 nm. The secondary structure components were analyzed using `gmx_do_dssp` module. The distance calculations were carried out with `gmx_distance` module. Radial distribution function (RDF) for oxygen atom of water around the heavy atoms of the protein was evaluated by `gmx_rdf`. The number of hydrogen bonds was calculated using `gmx_hbond` with a distance cut-off of 3.5 Å and an angle cut-off 30°. Changes in the torsion angles, phi (Φ) and psi (Ψ), were calculated using `gmx_angle` for the selected residues. Secondary structural contents were analyzed using the DSSP

module in VMD^[166]. The salt bridge interactions were analyzed in VMD using a built-in package with a distance cutoff of 0.6 nm.

2.3.4.2. Selection of intermediate structures

To identify the plausible intermediate structures during the conformational transitions from a monomeric α -helical form to a cross- β fibrillar structure, the metastable structures from different simulations were evaluated. Free energy landscape (FEL) was constructed by projecting the RMSD of fibrillar core (FC) residues against the native contacts using the 57,000 conformations obtained from SMD-US. Eight distinct energy minimum structures representing eight different umbrellas in the SMD-US were extracted and used for further analysis. The conformational distribution in REMD simulations were analyzed by clustering for both the simulations of M_F and M_S with a cut-off of 1.5 nm and 1 nm, respectively, using `gmx_cluster` module. The trajectories of CMD simulation were also analyzed by clustering with a cut-off of 1 nm. To exclude the selection of structurally close conformations obtained from different simulation methods, RMSD matrix between the selected structures were constructed using the `gmx_rms`. Only the conformations having RMSD values >1 nm was considered as distinct structures.

2.3.4.3. Energy calculations

The solvent rearrangements around the three different regions of α -syn were evaluated using solvation energy (ΔG_{sol}) defined as, ^{[167][168]}:

$$\Delta G_{sol} = -k_B T \log \left(\frac{\sum e^{(-\beta\{I_{(X)} + I_{(X,\emptyset)}\})}}{\sum e^{-\beta I_{(X)}}} \right) \quad (2.1)$$

Here, $I_{(X)}$ is water-water interaction energy in a pure water simulation and $I_{(X,\emptyset)}$ is protein-water interaction energy. $\beta = 1/k_B T$, where k_B is Boltzmann constant and $T = 310$ K. The solvent interaction energies ($I_{(X)}$ and $I_{(X,\emptyset)}$) were calculated by the `gmx_energy` module.

The non-covalent interaction energy between the residues in the protein was defined as the sum of Lennard-Jones (L-J) potential and coulombic interaction energy.^[167] These values were also calculated using `gmx_energy` module.

Chapter 2

2.4. Results

2.4.1. Conformational sampling of fibril dissociation by SMD-US

Identification of all the accessible conformations of IDPs is an uphill task due to their high plasticity. The structural transformation of α -syn from a monomeric α -helix to interchain cross- β -sheets involves many conformational changes across the protein chain. Characterization of the structural intermediates during this transition using conventional MD simulation alone is unachievable. Therefore, steered molecular dynamics (SMD) with umbrella sampling (US) was employed for the initial analysis of the fibril dissociation pathway. SMD was performed by applying a pull force to detach a polypeptide chain from one end of the fibril assembly (chain-A)

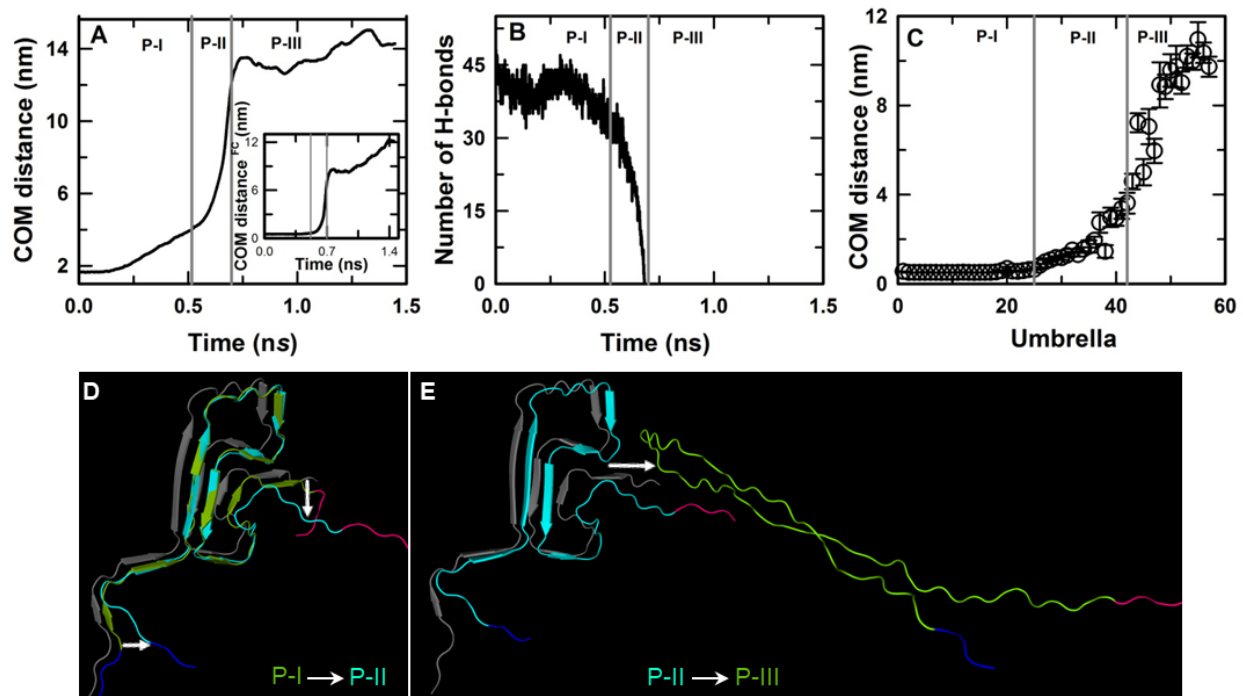


Fig. 2.2. Steered molecular dynamics-umbrella sampling (SMD-US): (A) The distance between the center of masses (COM) of chain-A and chain-B of α -syn during SMD. The inset shows the change in COM distance calculated for the residues in the fibril core (FC) region alone. (B) Number of interchain H-bonds between chain-A and chain-B. The time points of three different phases, P-I, P-II, and P-III, are marked with gray lines. From these phases, 24, 20 and 13 conformations were selected for US, respectively. (C) Distance between the COM of FC residues of chain-A and chain-B in each umbrella. (D and E) Representative conformations of α -syn during the transitions from P-I to P-II and P-II to P-III extracted from SMD simulation. The white arrows indicate the direction of dissociation of chain-A (green to cyan in panel D and cyan to green in panel E) from chain-B (gray ribbon).

from its neighboring chain (chain-B). The resultant SMD trajectory was divided into three different phases based on the center of mass (COM) distance between the chains and on the number of interchain H-bonding interactions formed by the FC residues (Fig. 2.2A and B).

In phase-I, until 0.53 ns of SMD, the COM between the chains was increased from 1.6 nm to 4 nm; however, the distance between the residues in the FC was not significantly altered (Fig. 2.2A-inset). After 0.53 ns, the interchain distance was steeply increased and the cross- β interactions were completely lost at 0.70 ns. This time window was represented as phase-II. In the last phase, from 0.70 to 1.45 ns, there was no interaction between chains A and B which was confirmed by the loss of interchain H-bonds (Fig. 2.2B). From these three phases of SMD, 57 representative conformations were collected (as mentioned in Section 2.3.1) to analyze the different stages of the chain dissociation and were subjected to umbrella sampling (US).

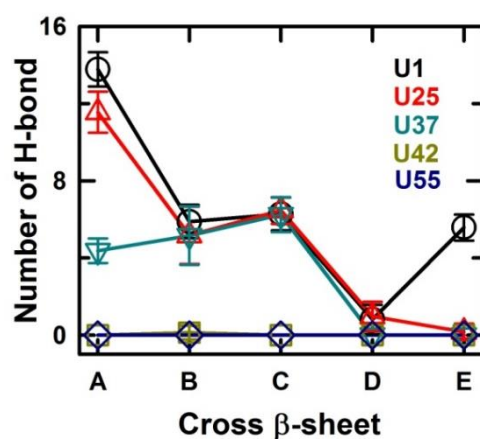


Fig. 2.3. Number of H-bonds between each cross β -strand of chain A and chain B of representative umbrellas from different phases of SMD P-I (U1), P-II (U25, U37 and U42), and P-III(U55).

The COM distance between the residues in FC of chain-A and chain-B calculated for each umbrella (Fig. 2.2C) was similar to the distances calculated from SMD. In phase-I, the cross β -sheets in the FC were mostly retained. At the beginning of phase-II, the β -sheets A and E in the FC (for labeling, refer Fig. 2.1) lost a few of their interchain H-bonding whereas the other β -sheets did not show any significant change (Fig. 2.3). Further in phase-II, the interchain H-bonding of β -sheet E was completely lost followed by the loss of β -sheets D. At the end of phase-II, the conformations obtained from US did not show any cross β -sheet interactions between the chains. Moreover, the formation of β -turns was observed during this phase, particularly in the hydrophobic region V66-V70. In phase-III, the chain-A was completely dissociated from the fibril assembly

Chapter 2

and showed formation of a short α -helical structure in H2 region around the residues K45-H50, a part of RTD4. Further, the structural transition from cross β -sheet to α -helix (fibril to monomer) was analyzed by secondary structural changes in each umbrella. The representative umbrellas from each phase of SMD are presented in Fig. 2.4.

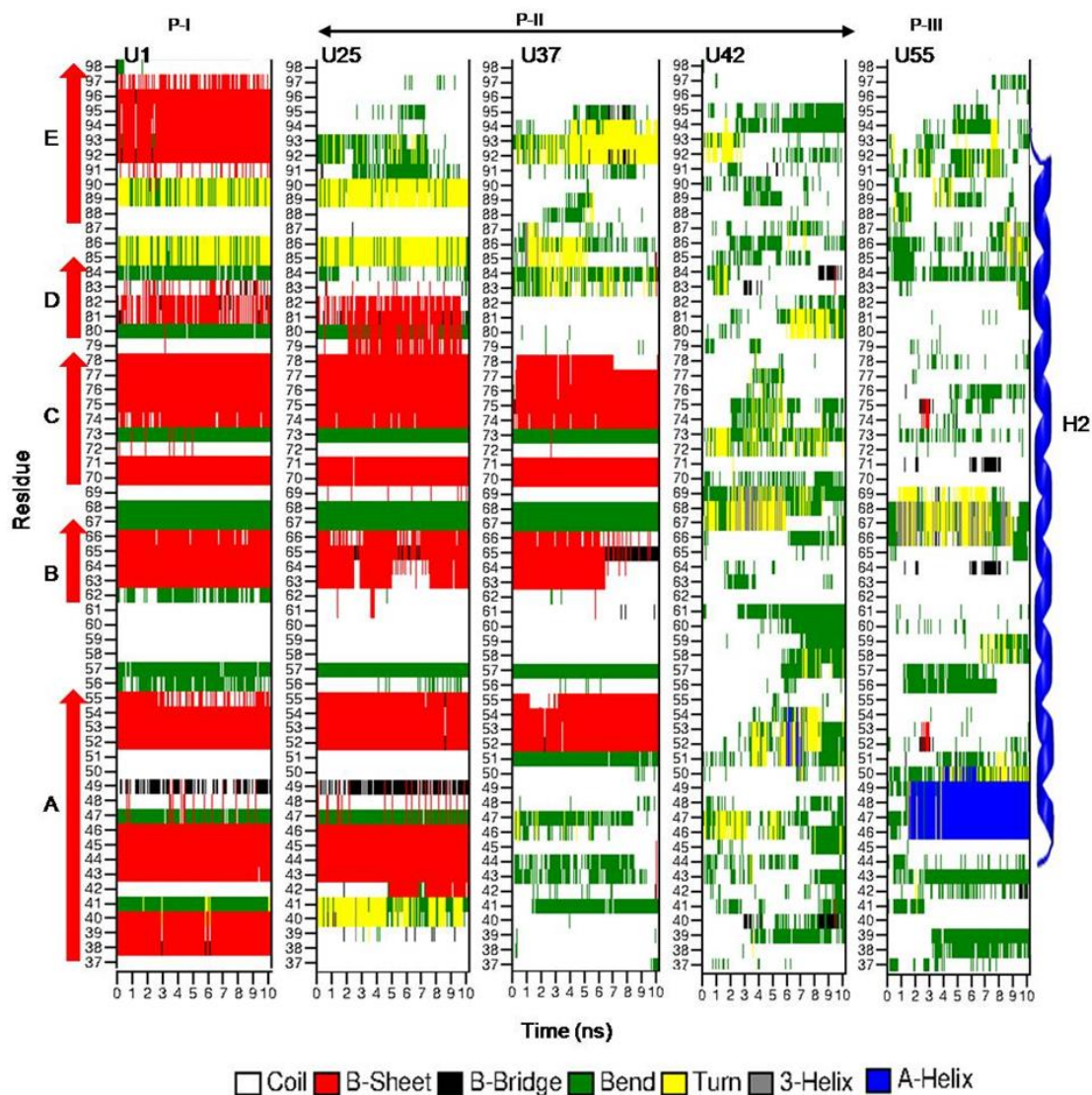


Fig. 2.4. Secondary structural changes in FC residues during SMD-US calculated by DSSP analysis. Representative umbrellas from phase-I (U1), phase-II (U25, U37 and U42) and phase-III (U55) are shown. The colors corresponding to each secondary structural elements are given in the legend. Red arrows on the left side represent the residues forming β -sheets in the FC and the blue spiral on the right side represents the residues forming α -helix in the monomeric form of α -syn.

To understand the overall structural changes, the trajectories obtained from the US simulations were concatenated and analyzed. Root mean square deviation (RMSD) of chain-A

The helix to fibril transition pathway of α -syn

showed an increase up to ~ 2.38 nm in which nearly 1.67 nm of RMSD was from FC residues alone (Fig. 2.5A). The radius of gyration (R_g) of chain-A increased during phase-I (Fig. 2.5B) suggesting the formation of an extended conformation of the chain. In the following phases, R_g value decreased which indicated that chain-A attained a compact conformation after its dissociation from chain-B. Solvent accessible surface area (SASA) of chain-A also exhibited a similar trend as R_g (Fig. 2.5C) that the SASA values slightly increased in phase-I, but decreased during phase-II. The SASA of FC residues showed a notable increase during phase-II indicating that the FC was more solvent accessible in these umbrellas which could be attributed to the dissociation of chain-A from the fibril assembly. This was complemented by the gradual loss of native contacts in FC residues during phase-II (Fig. 2.5D). In phase-III, chain-A showed a compact conformation with a decrease in R_g and SASA, and a loss in most of the native fibril contacts. Overall, from SMD-US simulation, it was observed that chain-A had extended conformation with less solvent exposure in the fibril state which was exposed to the solvent and become compact during the dissociation phase. The separated chain-A, in phase-III, had a compact structure with the loss of all cross β -interactions and showed a short α -helical formation as well.

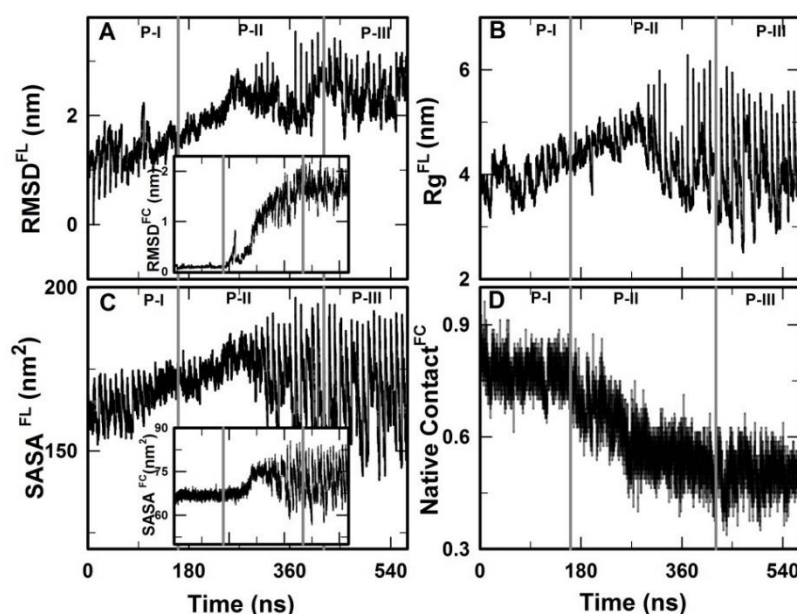


Fig. 2.5. Analysis of cumulative trajectory of 57 US simulations each run for 10 ns. (A) Root mean square deviation (RMSD) of $C\alpha$ atoms of full-length α -syn. The inset shows the RMSD of fibril core residues alone. (B) Radius of gyration (R_g) changes in the protein. (C) Solvent accessible surface area (SASA) of full length α -syn and the inset shows the SASA of fibril core residues alone. (D) Fraction of native contacts of fibril core residues of chain-A. The vertical grey lines in each panel represent the umbrellas from phase-I, phase-II and phase-III of SMD.

2.4.2. Intermediate conformations from FEL

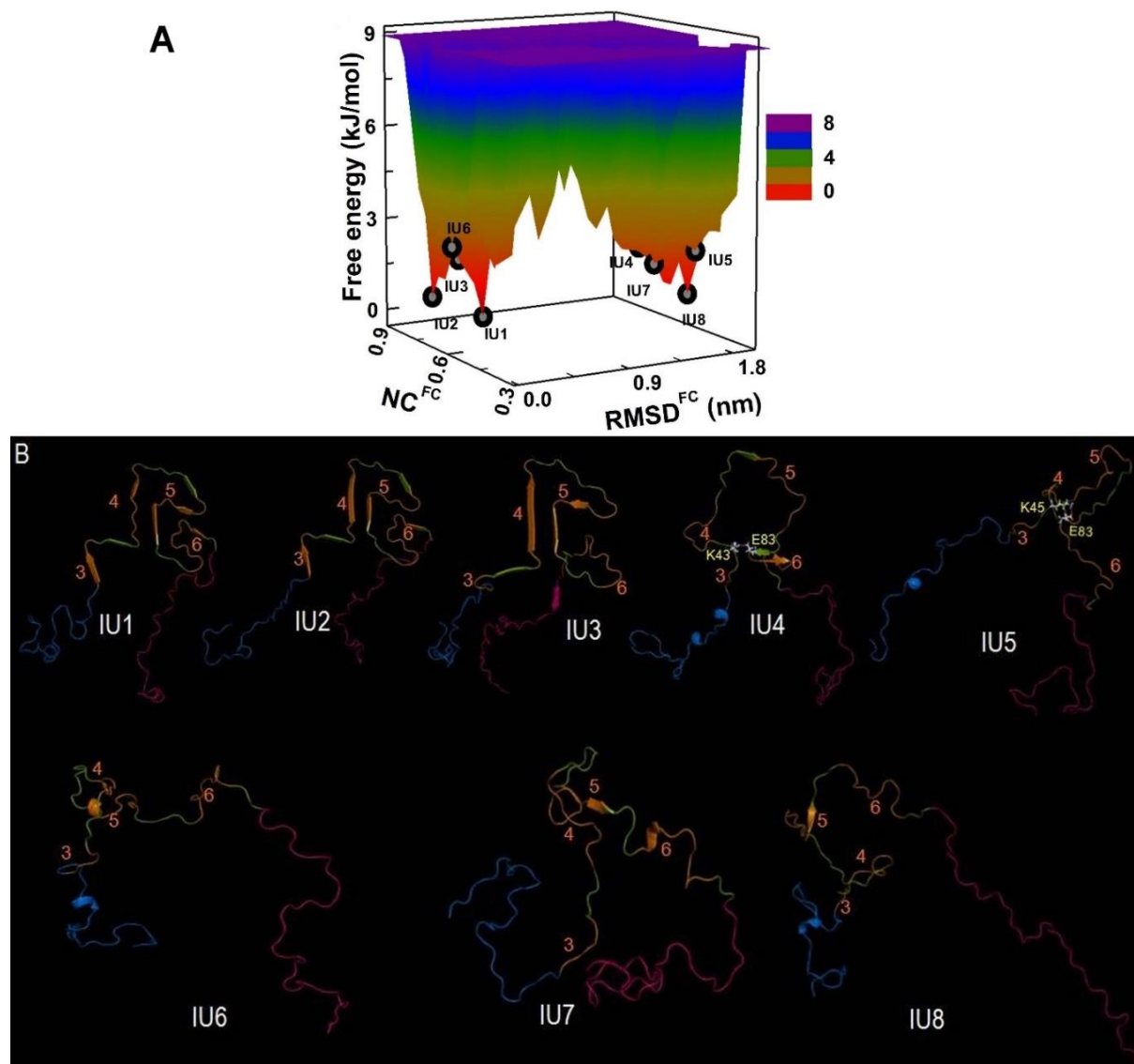


Fig. 2.6. Free energy landscape constructed by projecting the fraction of native contacts in FC against $\text{C}\alpha$ -RMSD of FC residues for the conformations obtained from the umbrella sampling. The eight minimum energy structures identified are marked as **IU1** to **IU8**. (B) Structures of low-energy conformations extracted from FEL. Blue, green and pink indicate NTR, FC and CTR, respectively. The four orange-colored regions represent the hydrophobic regions, HPR3, HPR4, HPR5 and HPR6 in the FC. The intrachain salt bridges in **IU4** and **IU5** are shown as sticks.

To identify major structural intermediates during the fibril dissociation process, FEL was constructed by projecting the native contacts against $\text{C}\alpha$ -RMSD of the FC residues (Fig. 2.6) using 57,000 conformations obtained from US. Eight distinct minimum energy conformations from different energy basins were identified from the FEL. Two of these conformations (**IU1** and **IU2**)

belonged to phase-I of SMD which were traced to umbrellas 8 and 17. In these structures, FC region of chain-A was intact with chain-B and the number of H-bonds in the cross β -sheets was comparable with the fibril structure (Fig. 2.7A). However, in **IU2**, the helix-1 (H1) formed by the hydrophobic residues V16-A20 was disrupted. Such conformational change in the *N*-terminal region was observed in the fibril structure characterized by solid-state NMR experiment as well.^[90]

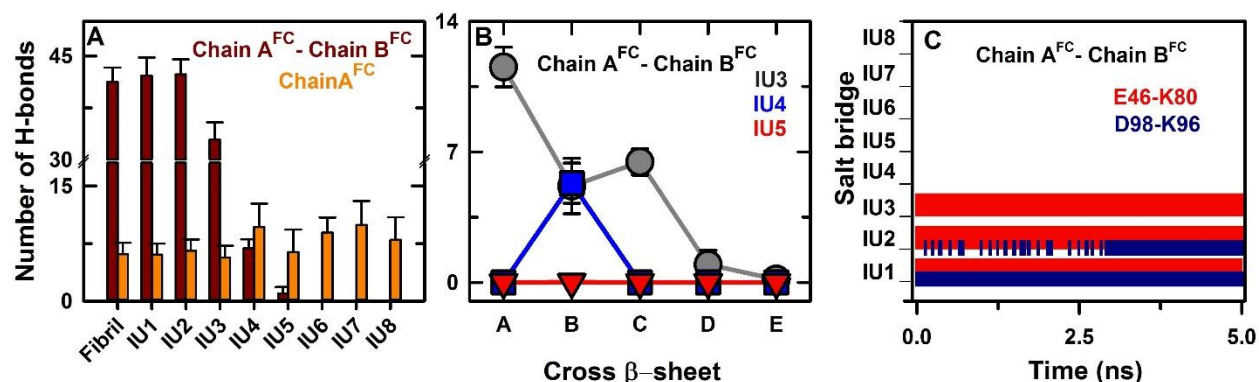


Fig. 2.7. (A) Number of interchain (between chain-A and chain-B, brown) and intrachain (within chain-A, yellow) H-bonds formed by FC residues in the intermediate structures identified from FEL analysis, **IU1** to **IU8**. (B) Number of H-bonds in each cross β -sheet of selected conformations from phase-II, **IU3** (gray), **IU4** (blue) and **IU5** (red). (C) The existence map of interchain salt bridges formed by E46 and D98 of chain-A with K80 (red) and K96 (blue) of chain-B, respectively in the intermediate structures, **IU1** to **IU8**. The vertical lines represent the presence of salt bridge interactions.

Three of the conformations from FEL (**IU3**, **IU4** and **IU5**) fell within phase-II of SMD and the corresponding umbrellas were 25, 37 and 39. These structures gradually lost the interchain H-bonding interactions between chain-A and chain-B (Fig. 2.7A). However, the number of H-bonds formed within chain-A was slightly increased in **IU4** and **IU5** compared to **IU3**. The analysis of each cross β -sheet interaction in the FC (Fig. 2.7B) showed that β -sheet E was first disrupted in **IU3** followed by the loss of interactions in β -sheets A, C and D in **IU4**. This could be mainly attributed to the breaking of the interchain salt bridge between D98 of chain-A (part of β -sheet E) and K96 of chain-B which might initiate the fibril dissociation (Fig. 2.7C). At the later stages, the salt bridge between E46 in chain-A and K80 in chain-B present in β -sheets A and D, respectively, was also disrupted. This could facilitate the further dissociation of FC. The secondary structural analysis of the intermediates (Fig. 2.8) showed that **IU3** formed a short β -turn in strand-A and strand-E regions whereas in **IU4**, β -turns were formed between the residues of strand-B and strand-C. In the case of **IU5**, the interactions between the chains were completely lost and β -turns were

Chapter 2

observed in the regions of strands A, C and E. For further insight, the distance between each cross β -sheet in the FC was evaluated (Fig. 2.9). This also suggested that initially, the residues in strand-E moved away followed by the residues in the strands A, C, and D. However, in **IU5**, the distance between the residues in strand-C was reduced. This did not lead to any increase in the number of interchain H-bonds. Since this region (V70-A78) is dominated by hydrophobic residues (Fig. 2.9C), these conformational changes could be attributed to hydrophobic interactions between the chains.

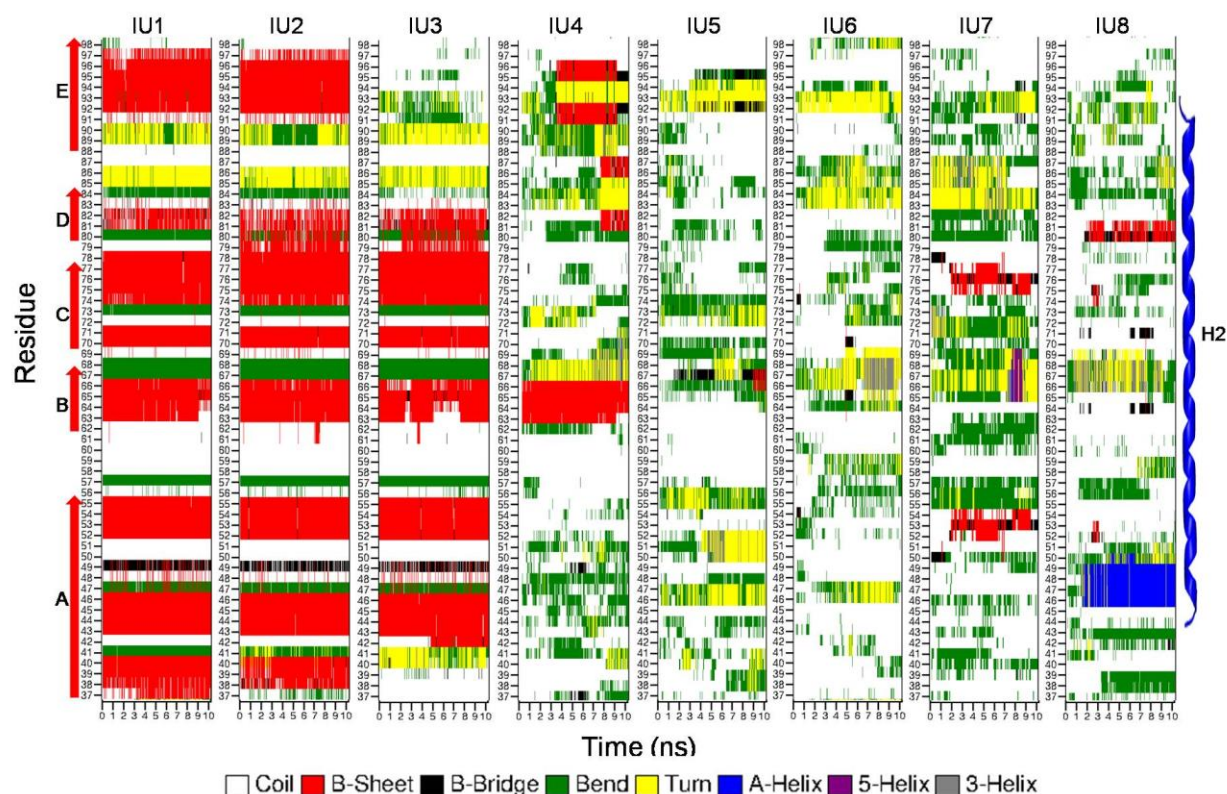


Fig. 2.8. Change in the secondary structure of FC residues of selected conformations from the free energy landscape analysis of SMD-US, **IU1** to **IU8**. The color codes of the secondary structural elements are given in the legend. The red arrows on the left side show the residues belonging to strands-A to E in the fibril form of α -syn and the blue spiral on the right side shows the residues belonging to helix-2 in the monomeric form of α -syn.

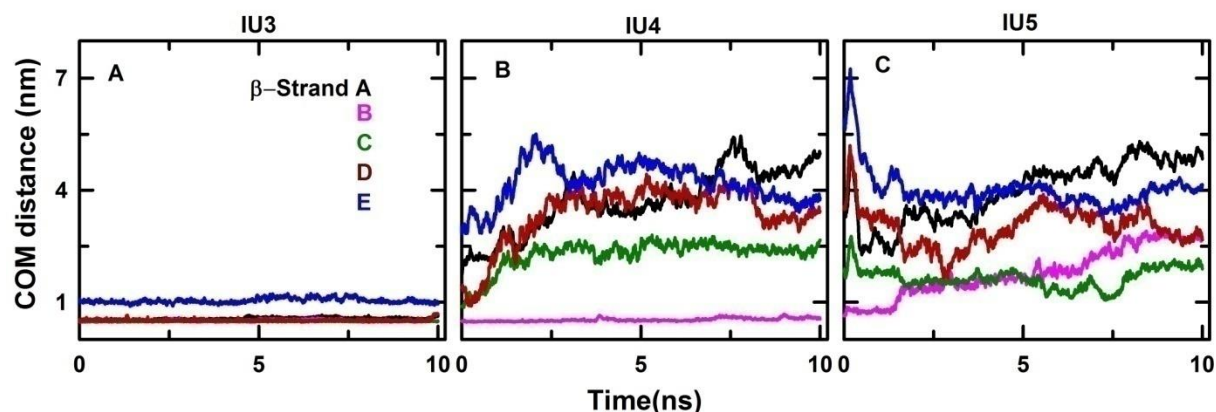


Fig. 2.9. Centre of mass (COM) distance between each strand, A to E, in chain-A and chain-B for the selected conformations from phase-II of umbrella sampling, (A) **IU3**, (B) **IU4**, and (C) **IU5**. It may be noted that in phase-I, all the strands have cross- β interactions and in phase-III the cross β -sheets are completely broken.

The remaining three conformations (**IU6**, **IU7** and **IU8**) were found in phase-III of SMD corresponding to the umbrellas 46, 48 and 55, respectively. In all these conformations, chain-A completely dissociated from the fibril assembly which was evident from the absence of any interchain H-bonding interactions (Fig. 2.7A). Also, the H-bonding interactions within chain-A gradually increased in **IU6** and **IU7** and there was a marginal decrease in **IU8**. Both **IU6** and **IU7** showed β -turn conformation with the residues corresponding to strands B, D and E (Fig. 2.9). In addition, **IU7** also showed a β -sheet formation between the hydrophobic regions G51-V55 and V74-A78. In **IU8**, there was an α -helical structure in the RTD-4 which was a part of helix-2 (H2) in the native monomeric conformation of α -syn. In all three structures in phase-III, β -turn was observed between the residues V66-V73, a part of hydrophobic region-5 (HPR5). There were no notable interactions between the three domains in all the selected conformations.

2.4.3. REMD simulation of a single chain from the fibrillar structure (M_F):

The conformational states accessible by the protein from its β -sheet state were analyzed by REMD simulation of a single chain obtained from the fibrillar state, M_F (Fig. 2.1, L4-B), using 32 replicas generated in the temperature range of 303 to 328 K. Nearly 32,000 conformations were obtained from the total simulation of 320 ns. The overall analysis of all the conformations obtained from REMD (Fig. 2.10) showed that $C\alpha$ -RMSD of the protein varied up to 3.1 nm from the initial

Chapter 2

structure at 303 K, while the fluctuation in FC alone was nearly 2.3 nm. The R_g value of the protein was ranging from 1.74 to 4.74 nm. The predicted R_g value for a protein with 140 amino acids is 1.5 nm in a globular state and 5.2 nm in a fully-extended state.^[169] The obtained R_g values, thus, suggest that the REMD simulation could access varying conformations from compact to extended. This was complemented by the differences in the SASA values which was from 122.7 nm² to 175.7 nm² for the full-length of α -syn and from 46.8 nm² to 76.70 nm² for the residues in FC alone.

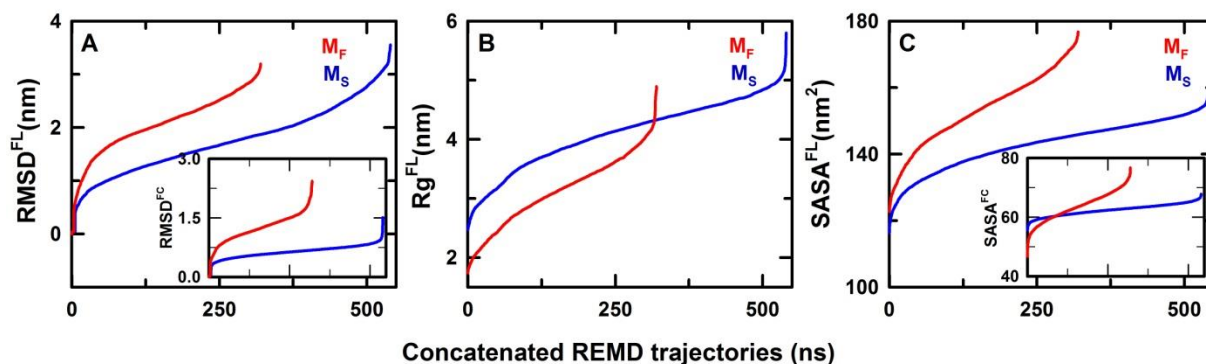


Fig. 2.10. Change in (A) C α -RMSD of full-length α -syn and only FC residues (inset), (B) Radius of gyration (R_g) of full-length α -syn, and (C) SASA of full-length protein and only FC residues (inset) during the REMD simulations of a single-chain from fibril assembly, M_F (red) and the native α -helical monomer, M_S (blue). The trajectories of all the replicas are concatenated and sorted to represent the overall change in the conformations obtained from REMD.

For further evaluation of conformational distribution, clustering analysis was performed with a cut-off value of 1 nm of C α -RMSD. Though 143 clusters were obtained in total, the frequency of occurrence of many of the conformations was found to be less (Fig. 2.11A). This could be attributed to the plasticity of C-terminal residues of the protein. Therefore, the conformations appearing for at least 1% of the conformations in the total trajectory were extracted. This resulted in 41 clusters covering nearly 73% of the total conformations (Fig. 2.11A). Among the selected conformations, it was noticed that the fibril core region was similar in a few of the conformations. The mean structure from each cluster was collected and compared for their structural differences in the FC region based on their C α -RMSD values (Fig. 2.11B). The conformations showing RMSD values >1 nm in the FC region were identified. These 14 structures were further compared with the intermediate structures identified from umbrella sampling (IU1 to IU9) for their similarity using the above-mentioned criteria that C α -RMSD of fibril core residues was >1 nm. Out of 14, eight structures were similar to the conformations obtained from US (Fig.

2.11C). The other six distinct structures were chosen for further analysis and labeled as **IR1** to **IR6** (Fig. 2.13A).

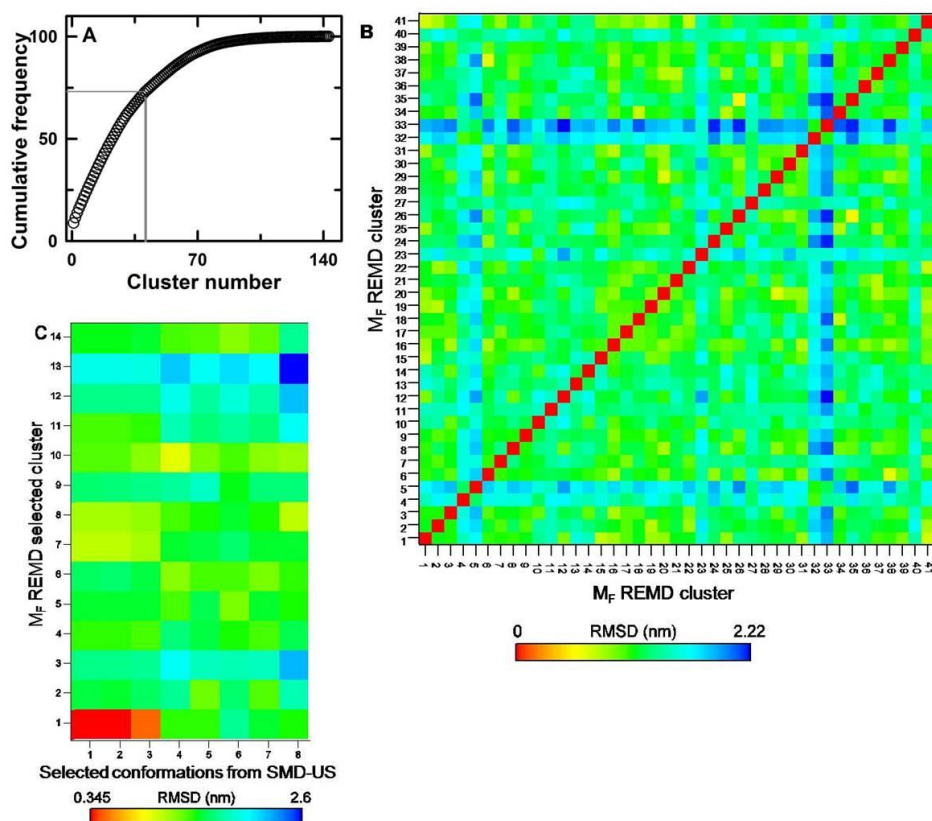


Fig. 2.11. (A) Cumulative percentage of the conformations covered by the clusters obtained from REMD simulation of a single chain from fibril structure of α -syn, M_F. (B) C α -RMSD matrix of FC residues for the mean structures obtained from the first 41 clusters each covering at least 1% of the total conformations. Out of 41, only 14 structures showed RMSD values of >1 nm. (C) C α -RMSD matrix between the 14 structures obtained from REMD simulation and the structures identified from SMD-US (**IU1** to **IU8**).

2.4.4. REMD simulation of monomeric α -syn (M_S):

To evaluate the conformational states accessible by the native-monomeric form of α -syn, REMD simulation of M_S (Fig. 2.1.L5-A) was performed for the temperature range of 303 to 328 K using 54 temperature replicas. From the total simulation run, 54,000 conformations were obtained. The combined analysis of the trajectories (Fig. 2.10) showed that the RMSD values of C α -atoms increased up to 3.55 nm compared to the initial conformation, whereas the C α -RMSD of FC residues alone was up to 1.5 nm. The R_g values varied from 2.5 to 5.8 nm suggesting that the simulation assessed most of the extended conformations. The SASA values of the protein varied from 116 to 157 nm² and for the FC alone the changes were from 55.6 to 67.8 nm². To

Chapter 2

extract the frequently accessed conformations, clustering analysis was performed using a cut-off value of 1 nm of C α -RMSD. Among the 317 clusters obtained from the trajectory, the conformations appearing for at least 1 % of the total conformations were extracted. This resulted in 20 structures which covered nearly 35% of the total simulation (Fig. 2.12A). The mean structure from each cluster was collected and compared for the structural differences within their FC residues (Fig. 2.12B). Only two conformations showed RMSD values >1 nm in the FC region suggesting that the other structures might have more fluctuations in the disordered C-terminal region. The selected conformations were compared with the structures obtained from SMD-US and found to have RMSD values > 1.5 nm. These conformations were named as **IR7** and **IR8** (Fig. 2.13) and considered for further analysis.

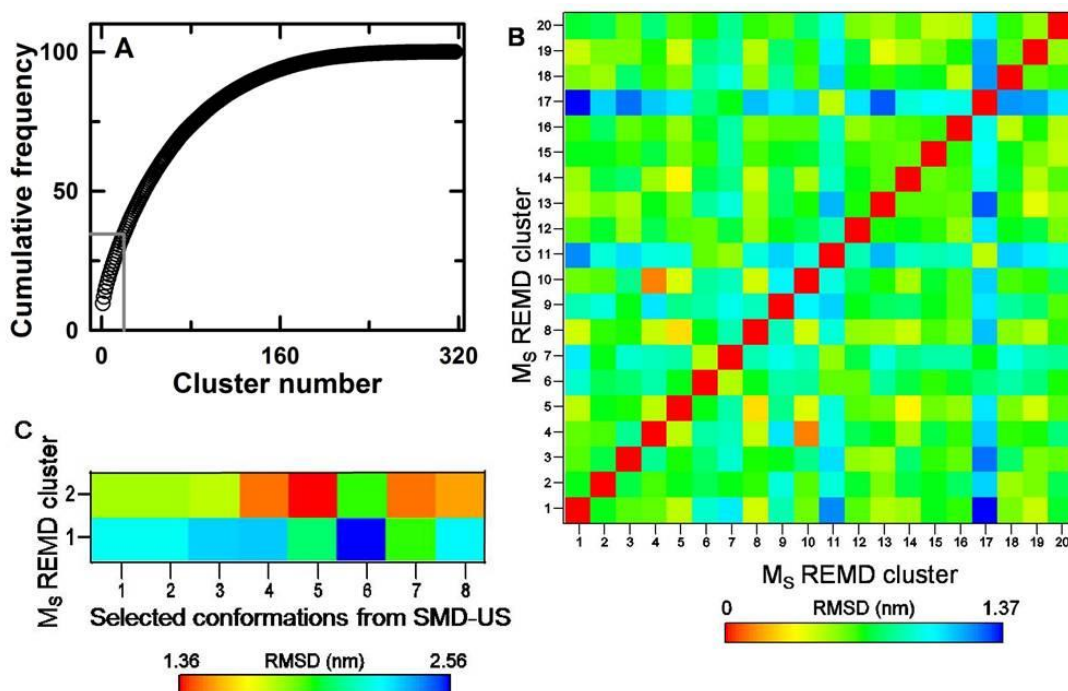


Fig. 2.12. (A) Cumulative frequency of conformations covered by the clusters obtained from REMD simulation of monomeric α -helical form of α -syn, M_S. (B) C α -RMSD matrix of FC residues for the mean structures obtained from the first 20 clusters each covering at least 1% of the total conformations. Out of 20, only 2 structures showed RMSD values of > 1 nm. (C) C α -RMSD matrix between the 2 structures obtained from REMD simulation and the structures identified from SMD-US (IU1 and IU8).

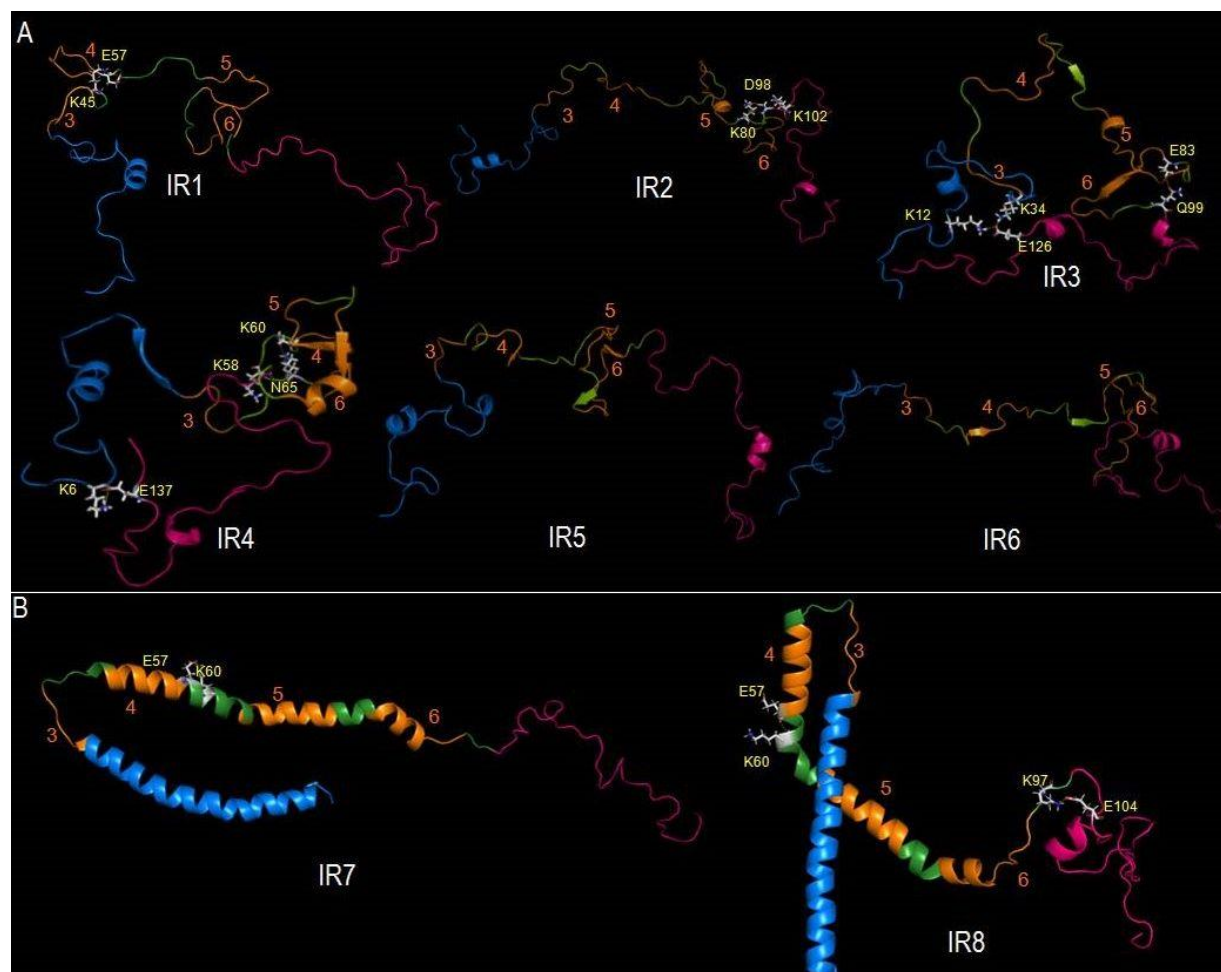


Fig. 2.13. Selected structures from REMD simulations of (A) a single chain from the fibril structure of α -syn, M_F (**IR1** to **IR6**) and (B) native α -helical structure of the protein, M_S (**IR7** and **IR8**). In the cartoons, blue, green and pink indicate NTR, FC and CTR, respectively. The four orange-colored regions in the FC represent the hydrophobic regions, HPR3 to HPR6. The intrachain salt bridges are shown as sticks.

2.4.5. Structural characteristics of IR1-IR8 conformations

Secondary structural analysis of the conformations selected from REMD simulations were analyzed (Fig. 2.14). **IR1-IR6** showed a helical structure between the residues 14-21 in NTR that consistently appeared in the fibril structure of α -syn. A transient helix formation in CTR, between the residues 119-124 was also observed in these conformations. In CTR, only **IR8** showed a helical structure formed by the residues 108-115. In the FC region, a transient appearance of a short-helix (mostly with the length of four residues) was noted in **IR2-IR6**.

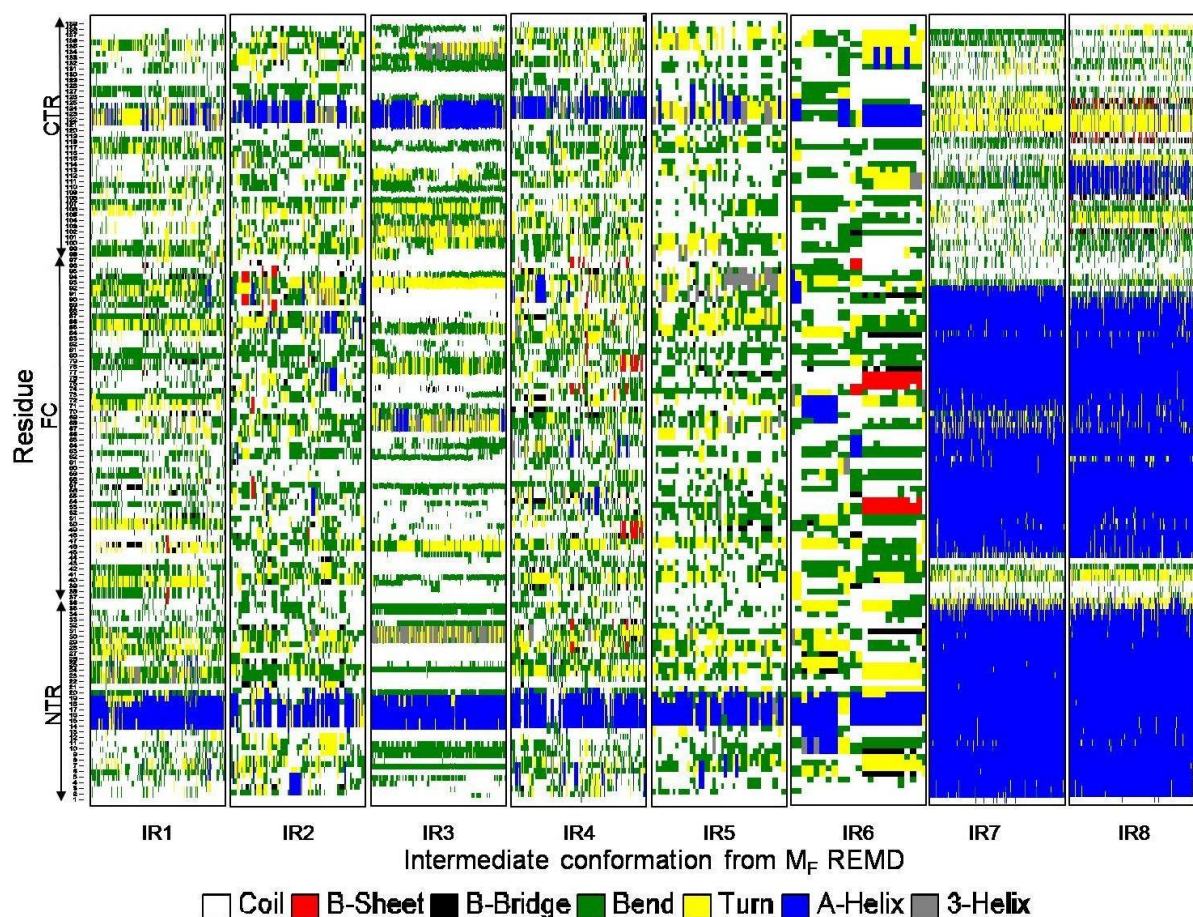


Fig. 2.14. Change in the secondary structure of the selected intermediate conformations obtained from the REMD simulations of M_F (IR1-IR6) and M_S (IR7 and IR8). The color codes of the secondary structural elements are given in the legend.

The hydrophobic regions in the FC (HPR3 to HPR6) showed major differences in their conformations and changes in the long-range interdomain contacts (Fig. 2.13). In the fibrillar state, the regions HPR4, HPR5 and HPR6 are packed together to form a Greek-key motif which is facilitated by a salt bridge between E46 and K80, the flanking residues of HPR4 and HPR5 regions, respectively. In any of the conformations, **IR1-IR8**, this salt bridge and the Greek-key topology were not observed. In **IR1**, the HPR4 and HPR5 moved away from each other (Fig. 2.13.A). HPR4 formed a long loop with a formation of a salt bridge between the flanking residues K45 and E57. HPR5, however, retained the hydrophobic packing with HPR6. In **IR2**, the packing of HPR5 and HPR6 regions was lost, whereas there was a formation of a short helix in HPR5 (V74-V77). This region is a part of helix-2 in the monomeric state of α -syn. Salt bridge formed by D98 with K80/K102 could stabilize the loop formed by HPR6.

The helix to fibril transition pathway of α -syn

In **IR3**, there were short loops formed by the residues partly from HPR5 and HPR6. Residue E83 formed a salt bridge or H-bonding interactions with K97/Q99 which facilitated the loop structure. Further, **IR3** showed significant long-range interactions between NTR and CTR (K32-E131 and K12/K34-E126, Fig 2.13A). In **IR4**, each of the hydrophobic regions formed loops with no significant packing between the HPRs. H-bonding interaction between the residues K58/K60 and N65 was noted. Also, a short-helix formation by the residues in HPR-6 (A89- F94) was observed. **IR5** showed a significant rearrangement of residue conformations with a formation of a long loop covering the residues from HPR-5 and HPR-6 together. Also, a part of HPR-3 and HPR-4 residues formed short loops. In **IR6**, along with the loop formation by HPRs, a short helix formation (T59-E61) was also noted. **IR7** and **IR8** had two helical structures covering the residues 2-34 and 45-91; however, the orientation of helices was different in these two states (Fig. 2.13B). **IR7** showed an anti-parallel H1-H2 conformation similar to the monomeric α -syn whereas **IR8** showed an orthogonal bend in helix-2. The bend in helix-2 region has been earlier reported in the membrane-bound forms of the protein.^[93,153] Further, **IR7** and **IR8** showed a salt bridge interaction between E57 and K60 which was similar to the monomeric α -helical structure.

2.4.6. Conventional MD simulations of M_F and M_S conformations of α -syn

Conventional MD (CMD) simulations of a single-chain obtained from the fibril structure (M_F) and the monomeric α -syn (M_S) were carried out at 310 K for 200 ns each. The simulations were carried out with C27FF and with C36IDPSFF as well which was specifically developed and tested for intrinsically disordered proteins.^[164] Initially, the convergence of simulations was analyzed for both the force field trajectories by “time- dependent changes of conformational clusters”.^[164] The trajectories of C27FF converged at 20 and 6 clusters for M_S and M_F , respectively, whereas the trajectories of C36IDPSFF showed 2 or 3 clusters (Fig. 2.15A and B).

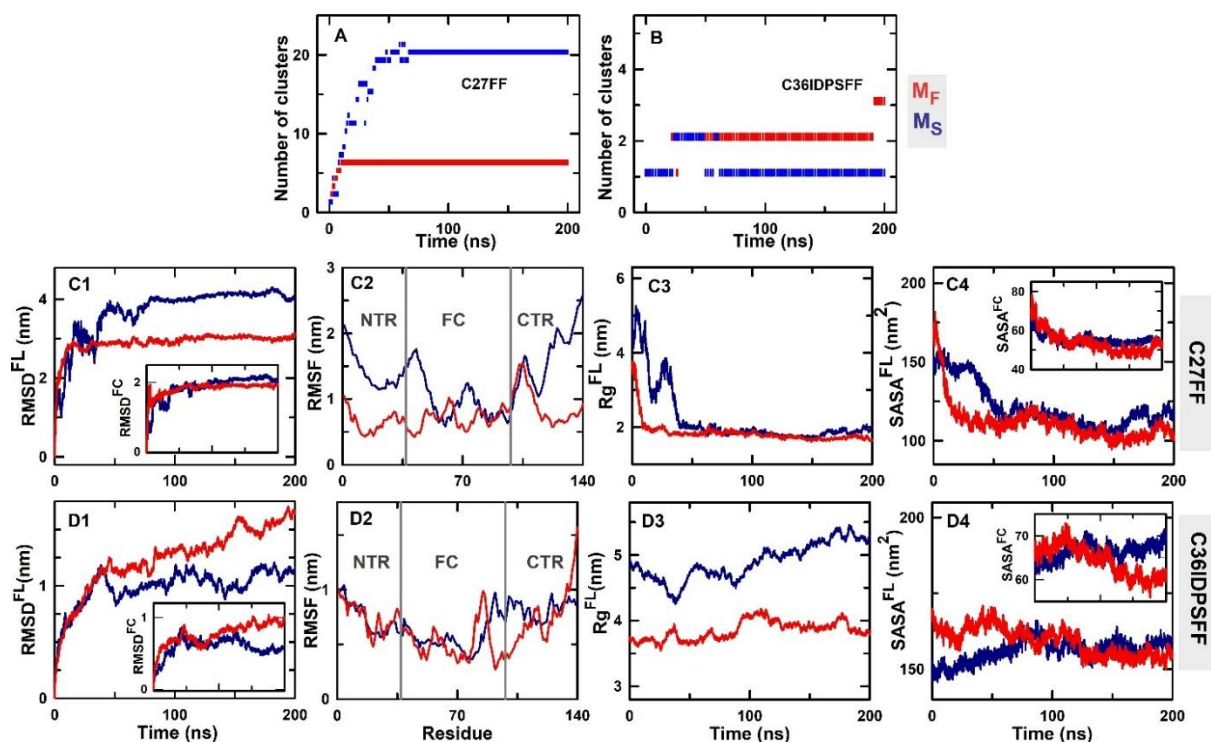


Fig. 2.15. The convergence of CMD simulations of M_F (red lines) and M_S (blue lines) analyzed by time-dependent conformational clusters for the trajectories obtained using (A) C27FF and (B) C36IDPSFF. Panels (C1 to C4) and (D1-D4) present the analysis of trajectories of C27FF and C36IDPSFF, respectively. (C1&D1) RMSD of $C\alpha$ atoms of all the residues in the protein and the inset shows RMSD of the FC residues alone. (C2&D2) Root mean square fluctuations (RMSF) of each residue. (C3&D3) Radius of gyration (Rg) of the protein chain. (C4&D4) Solvent accessible surface area (SASA) for the full-length protein and the inset shows SASA of the FC residues alone.

The initial structural analysis of the trajectories is presented in Fig. 2.15. The RMSD changes suggested that the residue fluctuations were more in M_S than in M_F for C27FF trajectories whereas the fluctuations in the FC residues were similar in both the conformations. This suggested that the residues in NTR and CTR of M_S might have larger fluctuations compared to M_F . This was confirmed by the RMSF values calculated for each residue in both the conformations (Fig. 2.15-C2). However, the RMSD of M_F was higher than M_S for the C36IDPSFF trajectories and the fluctuations in the FC residues were also slightly higher for M_F . The corresponding RMSF of the individual residues suggested that the difference in the fluctuations between M_F and M_S were around the residues 80 to 110 covering a part of FC residues and CTR as well. The Rg and SASA of both the conformations decreased and converged to similar values for C27FF trajectories (Fig. 2.15-C3 and C4), whereas the Rg of M_S was consistently higher than M_F in the C36IDPSFF simulation (Fig. 2.15-D3). Also, the SASA values showed significant variations. Though they

converged to a similar value for the full-length, the SASA value of FC residues was slightly increased for M_S , whereas it was decreased for M_F (Fig. 2.15-C4 and D4). Overall, the amplitude of fluctuations, thus the conformational changes, were more in C27FF compared to C36IDPSFF for α -syn.

Further, to analyze the major conformational states sampled in the CMD simulations, cluster analysis was performed for each trajectory using the same conditions mentioned for REMD simulations. The trajectories of M_F and M_S from C27FF produced 7 and 22 clusters, respectively (Fig. 2.16A). Out of these, only first 3 and first 11 were found to cover at least 1 % of the conformations in their respective trajectories. M_F and M_S trajectories obtained from C36IDPSFF simulations produced only 3 clusters in each case. The mean structures of these clusters were extracted and compared. The six structures obtained from C36IDPSFF were found to be similar (RMSD <1 nm) to one of the 14 structures obtained from the cluster analysis of M_F and M_S of C27FF (Fig. 2.16B). Therefore, the conformations obtained from the trajectories of C27FF were alone used for further analysis. The structures obtained from M_F and M_S simulations with C27FF were compared among themselves to identify the changes around the FC residues. All the three structures of M_F were distinct in their fibril core (Fig. 2.16C) whereas only three out of 11 were distinct in the case of M_S (Fig. 2.16D). These six structures were then compared with the already selected structures from SMD-US and REMD simulations (**IU1** to **IU8** and **IR1** to **IR8**). Only one structure from M_F (**IC1**) and two structures from M_S (**IC2** and **IC3**) were distinct from the other selected conformations (Fig. 2.16E). These three conformations (Fig. 2.17) were used for further analysis.

Chapter 2

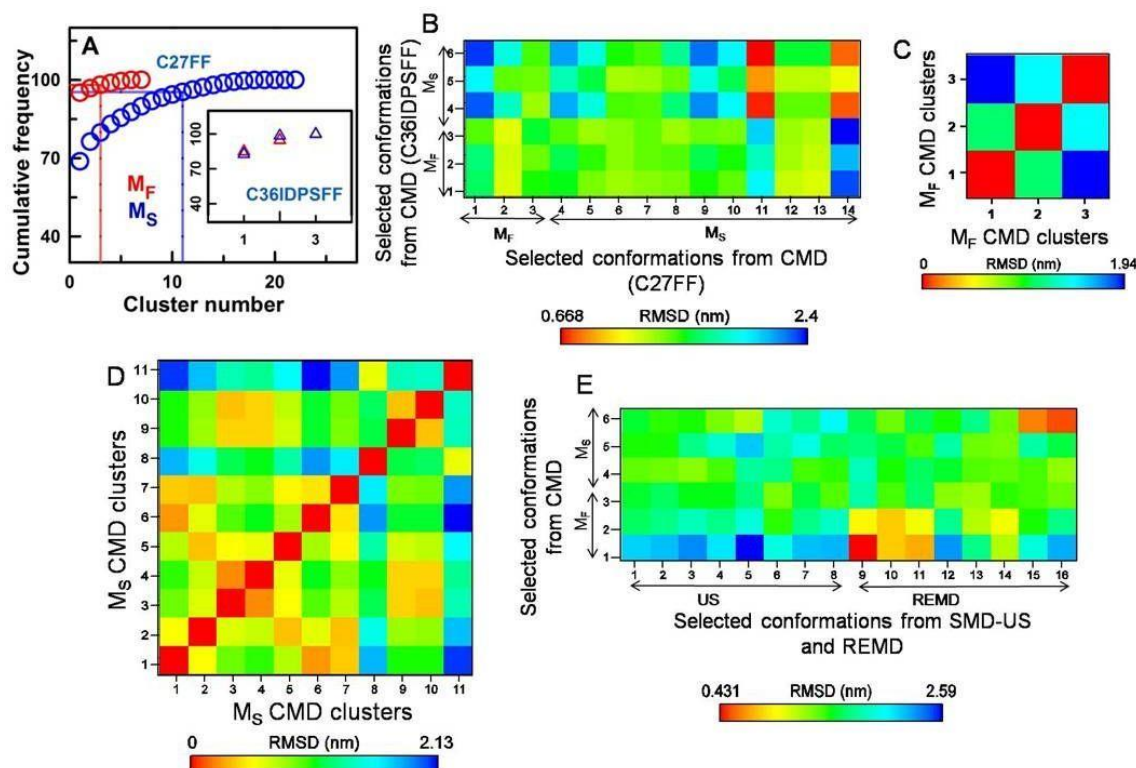


Fig. 2.16 (A) Cumulative frequency of conformations covered by the clusters obtained from conventional MD simulations of M_F and M_S structures of α -syn using C27FF. The first three and first 11 clusters of M_F and M_S simulations cover at least 1% of the total conformations in their respective trajectories. The inset presents the cumulative frequency for the simulations run with C36IDPSFF. Only three clusters were obtained for each M_F and M_S and their mean-structures were extracted. (B) $C\alpha$ -RMSD matrix of fibril core residues of the six conformations selected from the clustering of C36IDPSFF trajectories against the conformations selected from C27FF trajectories. All the six structures from C36IDPSFF trajectories are similar to one of the structures from C27FF trajectories with RMSD <1 nm. (C and D) $C\alpha$ -RMSD matrix of fibril core residues for the mean structures obtained from the cluster analysis of M_F and M_S simulations with C27FF. All three from M_F and three out of 11 from M_S have RMSD >1 nm within themselves. (E) These six structures are further compared with the structures obtained from SMD-US (IU1 to IU8) and REMD (IR1 to IR8) simulations. One structure from M_F (IC1) and two structures from M_S (IC2 and IC3) are distinct from the others with $C\alpha$ -RMSD >1 nm.

2.4.7. Structural characteristics of IC1-IC3 intermediates

The secondary structural analysis of the selected conformations was performed using DSSP (Fig. 2.17B). The FC region of IC1 showed a short helix in HPR5. The HPRs in FC region were well packed in this conformation with significant interactions between the HPRs (Fig. 2.17A). Notably, there was a salt bridge formation between the residues E46- K58 and E28-K45. IC1 also had two

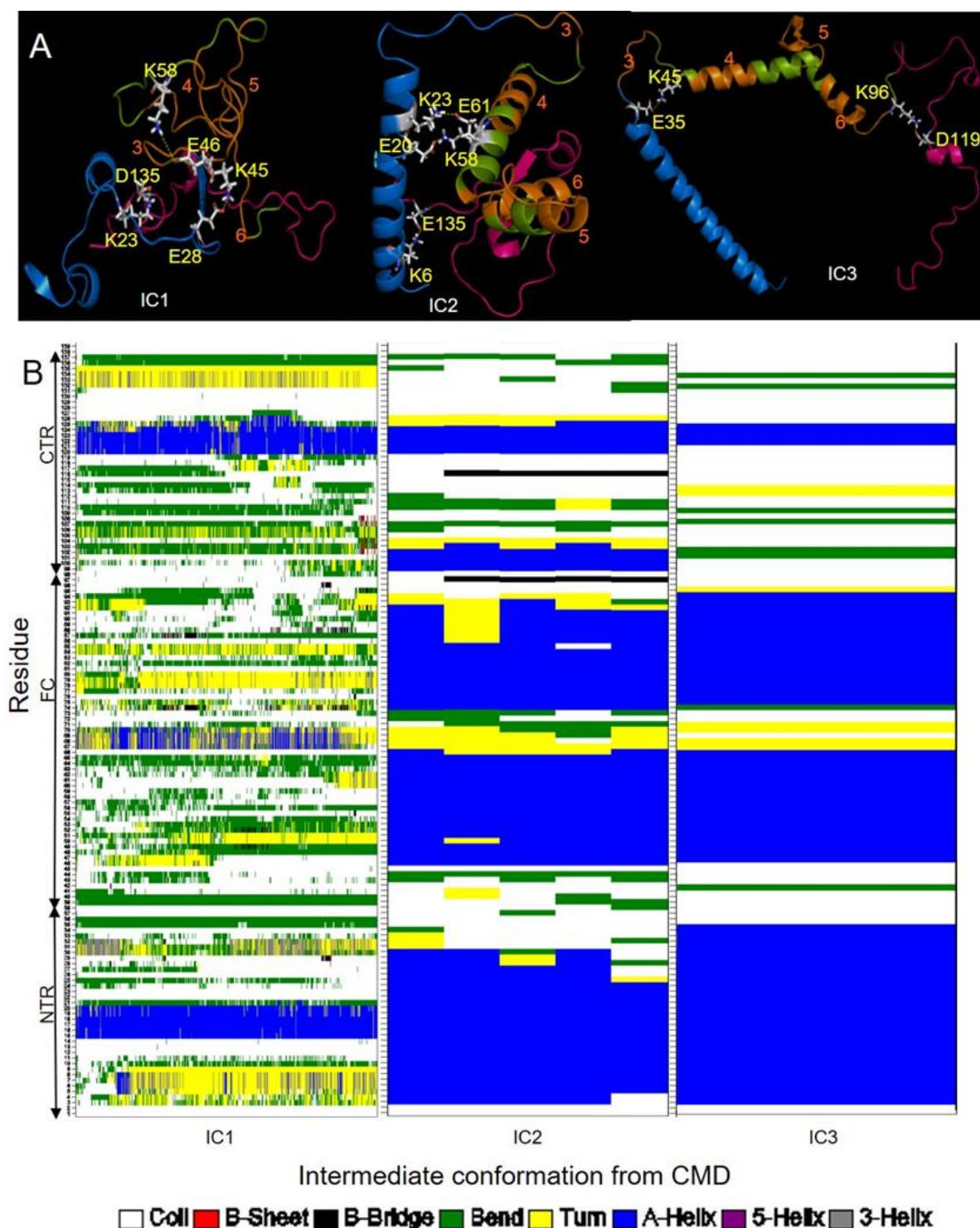


Fig. 2.17. (A) Selected conformations from CMD simulations, **IC1**, **IC2** and **IC3** are shown as cartoons. The blue, green and pink colors indicate NTR, FC and CTR, respectively. Four orange-colored regions in FC represent the hydrophobic regions, HPR3 to HPR6. The intrachain salt bridges are shown as sticks. (B) Change in the secondary structure of the selected intermediate conformations **IC1**, **IC2** and **IC3** obtained from the conventional MD simulations of M_F and M_S . The color codes of the secondary structural elements are given in the legend.

Chapter 2

short helices in NTR which is a part of helix-1 in the native monomeric α -syn. In **IC2**, helix-2 region (K45-T92) was broken at HPR-5 and formed two separate helical structures which were anti-parallel to each other and orthogonal to helix-1 (Fig. 2.17A). A break in the helix-2 region was observed in **IC3** as well; however, overall, it showed a different conformational state (Fig. 2.17A). HPR4 and HPR6 regions were closer to each other in **IC2** whereas they were relatively far away in **IC3** state. In all three conformations, a short-helical structure by the residues D119-Y125 was noted in CTR similar to the helix observed in the conformations **IR1-IR8**. There was another short helical structure in **IC2** covering the residues Q99- N103. Further, **IC1** and **IC2** showed long-range interdomain interactions mainly stabilized by salt bridges, whereas the interdomain interactions in **IC3** were almost null.

2.4.8. Fibril dissociation pathway(s)

From SMD-US, REMD and CMD simulations, eight, eight, and three distinct structures were obtained, respectively. The C α -RMSD of FC residues between these structures were calculated and found to be >1 nm for each other (Fig. 2.18). The overall procedure followed for the selection of these conformations is presented in Fig. 2.19. These 19 structures were arranged based on their secondary-structural content, accessible surface area, and intra and intermolecular H-bonding interactions to predict the probable fibril dissociation pathway (Fig. 2.20). The analysis predicted that after the initial dissociation of chain-A from the fibril (fibril-dissociation stage), there could be at least two different pathways. In one of the pathways, the protein chain maintained compactness through interdomain (NTR, FC and CTR) interactions that was named as “compact-conformational pathway”. In another pathway, the α -syn had an elongated chain with little or no long-range interactions that was named as “extended-conformational pathway”. In further steps, these conformations followed similar structural changes to attain a monomeric α -helical structure (helical-conformational stage).

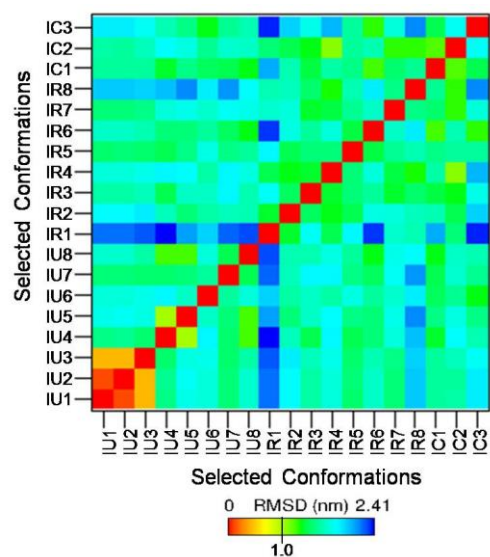


Fig. 2.18. $C\alpha$ -RMSD of fibril core residues calculated between all 19 selected structures: **IU1** to **IU8** from SMD-US, **IR1** to **IR8** from REMD and **IC1** to **IC3** from CMD.

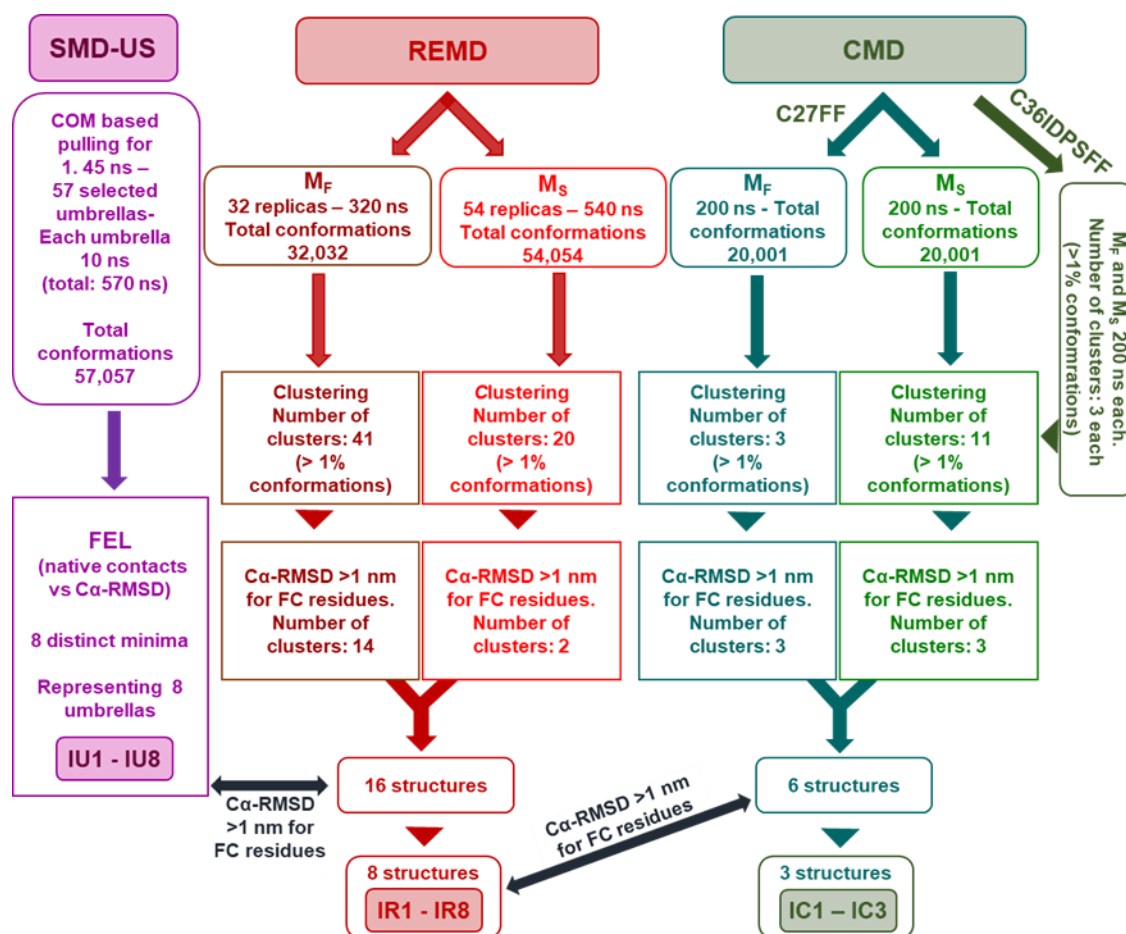


Fig. 2.19. A flowchart of the methodology followed to select 19 distinct conformations (**IU1** to **IU8**, **IR1** to **IR8**, and **IC1** to **IC3**) from different sampling methods, steered MD-umbrella sampling (SMD-US), replica exchange MD (REMD) and conventional MD (CMD) simulation. M_F and M_S indicate a single chain from α -syn fibril chain and monomeric helical form of α -syn, respectively. COM – center of mass, FC – fibril core, and FEL – free energy landscape.

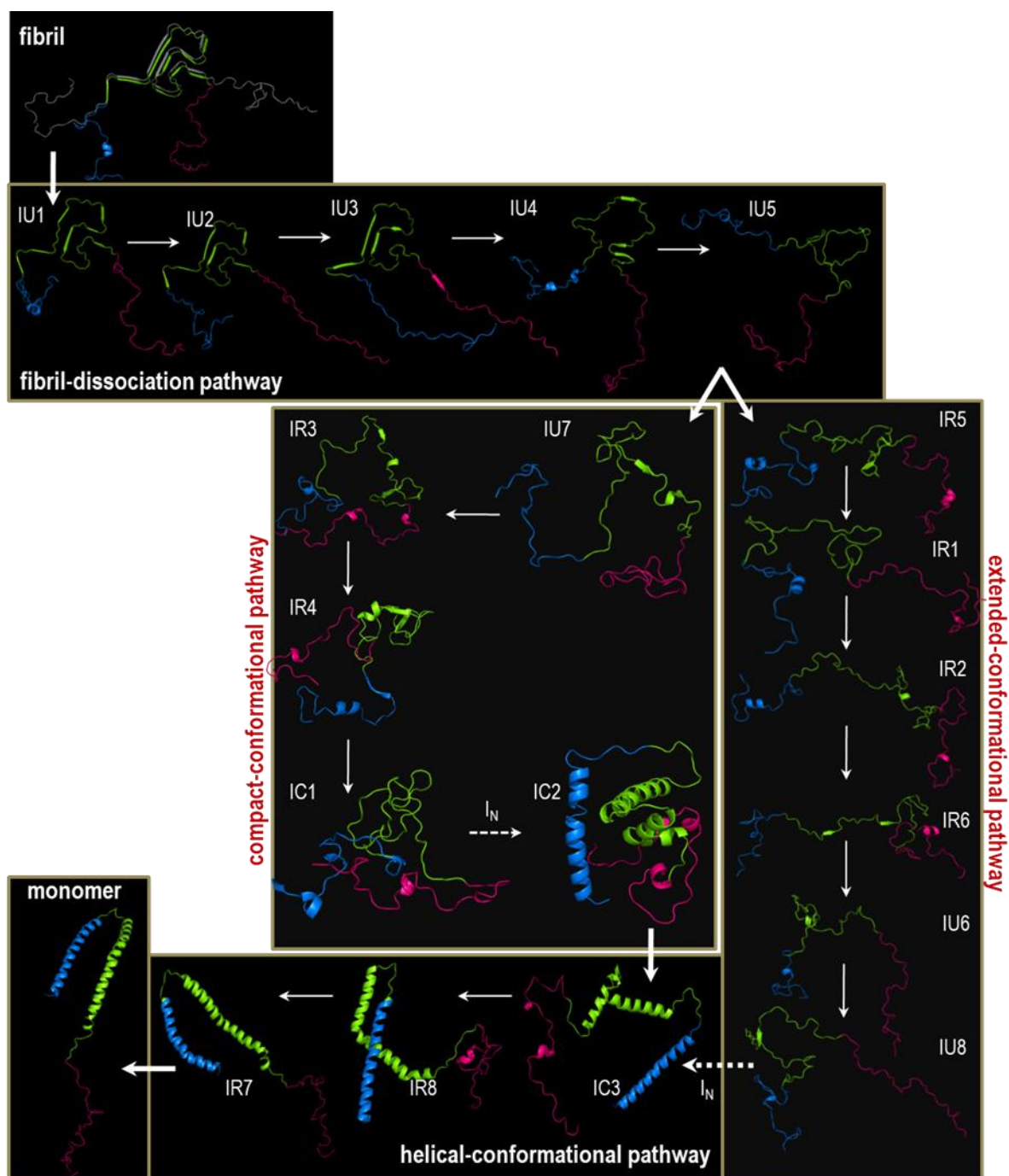


Fig. 2.20. The schematic presentation of transition of cross β -sheet rich fibril form of α -syn to α -helical monomer. The transition is divided into three different stages. Stage-1: fibril-dissociation stage where chain A separates from chain B of the fibril assembly. Stage-2: After dissociation, the protein might adopt two different pathways, compact-conformational pathway or extended-conformational pathway. During this stage, α -syn undergoes conformational changes to attain helical structure. Stage-3: helical-conformational pathway where rearrangement of helix-1 and helix-2 regions occur. The dotted arrows suggest that there are unidentifiable intermediate states between those structures. In the cartoon representations of α -syn, blue, green and pink colors indicate NTR, FC and CTR, respectively.

2.4.8.A. Intermediates in fibril-dissociation stage

The initial conformational changes during the fibril dissociation could be identified from SMD-US simulations. The first five structures selected from these simulations showed a gradual disruption of cross β -structure between the polypeptide chains A and B. Initially, the interchain salt bridge between chain A:K96 - chain B:D98 was broken (**IU2**). This further led to the complete loss of β -sheet E formed by the residues I88-D98 along with a partial loss of interactions in β -sheet A (**IU3**) consisting of the residues V37-G41. These residues were part of HPR6 and HPR3, respectively. In the next stage (**IU4**), the salt bridge between chain A:K80 - chain B:E46 was broken (Fig. 2.7C) which destabilized the β -sheets A, C, and D (Fig. 2.8). This was facilitated by the formation of intrachain salt bridges between the residues in the β -sheet regions A and D (E46-K80, K43-E83 and E57-K60). A long loop was formed by the hydrophobic residues in HPR4 and HPR5 which were part of the β -sheets B and C. Also, a short α -helical structure was noted in HPR6 region. The salt bridges in **IU4** were labile and they might form interactions with the charged residues in the adjacent regions as well. This altered the conformation of the loop formed by HPR4 and HPR5 (**IU5**). Further from this stage, two possible pathways for the transition into α -helix could be identified from the selected structures based on their inter-domain contacts (Fig. 2.21).

2.4.8.B. Intermediates in compact-conformational pathway:

At the initial stage, the salt bridge interactions between RTD4-RTD6 (both E46-K80 and K43-E83) were lost. This increased the distance between HPR4 and HPR6, but there was a formation of β -hairpin by the residues in HPR4 and HPR5 (**IU7**). Further, the formation of short α -helix in HPR6 was also noted in this conformation. In the next stage, all the four HPRs in the fibril core forming dependent loops having no interaction with other HPRs except a few found between HPR5 and HPR6 (**IR3**). This could be due to the H-bonding between the sidechains of E83 and Q99. Apart from this, a short helix in HPR5 and long-range interactions between *N*- and *C*-terminal regions and NTR with FC were also found (Fig. 2.21). Particularly, salt bridge interactions between the terminal regions (K12/K34-E126) and hydrophobic packing of HPR1 from NTR with HPR3 from FC were noted.

Chapter 2

A compaction of fibril core region occurs in the further steps which were mostly stabilized by H-bonding and salt bridge interactions. In **IR4**, the HPR4 structure was stabilized by the H-bonding interactions of K58/ K60-N65. The structure also showed a helix formation in HPR6 region. The next compact conformation (**IC1**) was also stabilized by the interactions between RTD4 and RTD5. Further, long-range interactions between the domains were noticeable (Fig. 2.20 and 2.21). For instance, NTD and CTD had interactions via, K23-D135/E139 while the FC region (K97) interacted with both NTD (E35) and CTD (E114/ D119). Also, K45 from FC region interacted with E28 of NTD. In **IC1**, a helical structure in HPR5 was also observed. Long-range interactions between CTR residues 124-138 and NTR are noted from paramagnetic relaxation experiments.^[127] The experiments also showed that the CTR interacts with pre-NAC region of the fibril core and with HPR5 and HPR6 as well.

The next possible conformation obtained from the analysis was **IC2** which contained two helices in the FC region, one by the residues in HPR4 and another by HPR5 and HPR6, which were part of helix-2 in the native monomeric state of α -syn. A long helix in NTD (F4 to E28) similar to that of helix-1 in the monomeric form of α -syn was also found in the structure. In this conformation, there were long-range interactions between NTD and CTD by salt bridge interactions of K6/K10 with D135, and NTD and FC region by E20-K58 & K23-E61. The paramagnetic relaxation enhancement (PRE) experiments have demonstrated that long-range interactions between NTR and CTR might be stabilized by columbic interactions whereas NAC and CTR interactions could be mostly hydrophobic.^[125] We propose that there could be a few more intermediate structures between **IC1** and **IC2** that could not be sampled in the present analysis. Specifically, the forces or interactions inducing the formation of α -helical structure by HPR4 which might be part of helix-2 and the extension of helix-1 in the native monomeric-like conformations could not be resolved.

The helix to fibril transition pathway of α -syn

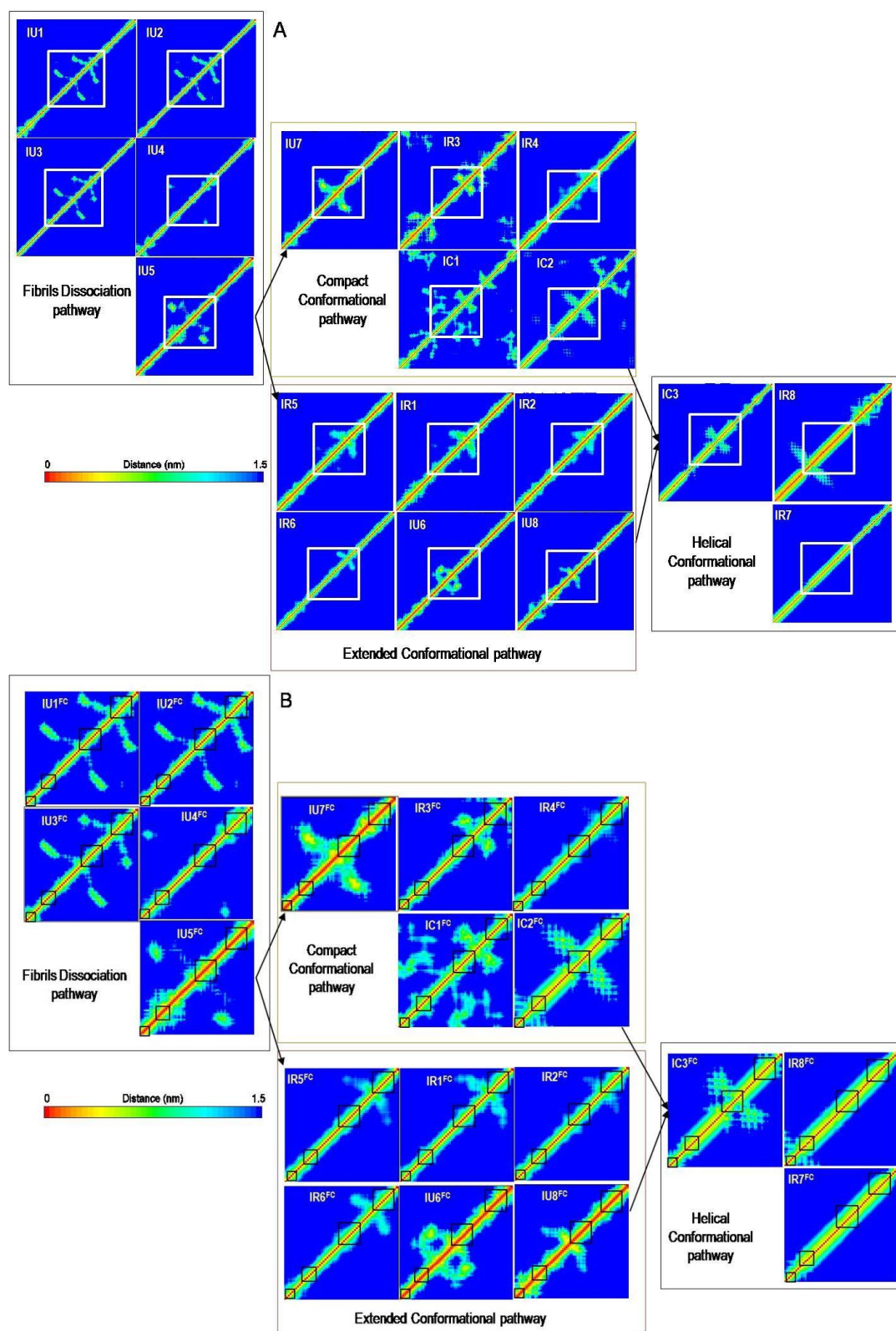


Fig. 2.21. Contact maps of (A) full length and (B) FC region of α -syn shown in the order of identified-intermediate structures (refer Fig. 10). White boxes in (A) and black boxes in (B) represent the fibril core and hydrophobic (HPR3 to HPR6) regions, respectively

Chapter 2

2.4.8.C. Intermediate structures in extended-conformational pathway:

In an alternate pathway, the protein chain might attain α -helical conformation from the fibril dissociation stage (**IU5**) without adapting compact structures. In the extended conformational pathway, no long-range interactions between the HPRs or between the domains were observed (Fig. 2.21). In the initial conformation (**IR5**), the protein had three major loops in the FC region formed individually by HPR4, HPR5 and HPR6. The loops formed by HPR5 and HPR6 were packed together by hydrophobic interactions. The interactions between the HPRs in the monomeric α -syn are observed in NMR experiments. They are denoted as fibril-like contacts, since they are part of β -sheets in the fibril state of the protein.^[127] The loop region was more flexible and might adapt different conformations (**IR1**). In the further stages, the loop region was disrupted and formed a short helical structure(s). Such a helix was observed in HPR5 in **IR2**, which also showed less hydrophobic packing with HPR6. Also, a salt bridge was found in HPR6 formed between K80 and D98. The fibril core further adopted an extended conformation with the loss of loop structure formed by HPR4 (**IR6**). In this conformation, the HPR5 showed more flexibility and a transient conversion between a short helix and β -hairpin α -helix and β -hairpin (Fig. 2.13). In the following step (**IU6**), a loop formation by HPR4 and HPR5 with a β -bridge within the residues V66-A69 was noted. In **IU8**, a helical structure was formed by the residues in HPR4 and HPR5 formed a long loop whereas HPR6 showed an extended conformation.

The formation of helices by HPR4 and HPR5 could be sampled in the extended conformational pathway in which long-range interactions between NTD and CTD were completely absent (Fig. 2.21). We propose that the next step would be the extension of these helices as seen in **IC3** in the helical-conformational stage (discussed below) which might further lead to native-like helical conformation. However, within the sampling limits, we could not detect the possible intermediate structures between the conformations **IU8** and **IC3**.

2.4.8. D. Intermediates in helical-conformational stage:

In this stage, the early conformation (**IC3**) had a fully extended helix-1 as in the native form of α -syn (residues F4 to E35) and two interdomain salt bridge interactions E35-K45 and K96-D119. However, the FC region still retained the two helices formed by the residues of HPR4 and the residues of HPR5 and HPR6 as similar to **IC2** conformation. The formation of broken α -helices in

α -syn is reported in the presence of micelles. The hydrophobic residues G84-G86 and G67-A69 are noted as linker region between the helices.^[125,170] Computational studies on monomeric α -syn have also shown the broken helical forms in FC region.^{[93][153]} In the next step, the conformation of helices in the FC region was altered and showed an extended single α -helix from the residue K45 to A91 which was anti-parallel orientation with the *N*-terminal helix formed by the residues V3-G36 (**IR8**). The helix in FC region showed a bend between HPR4 and HPR5 with the inner surface mostly hydrophobic in nature which was similar to the conformation earlier reported for α -syn while binding to the membranes.^[89] The helical conformation was mostly stabilized by the salt bridges formed within the repeat domain (RTD) residues. Also, in CTD, a transient helical formation in HPR7 and salt bridge interactions (K97-E104 and K102-E123) were also noticed. This could reduce the solvent exposure of the residues in CTD compared to preceding structures. This conformation further led to the native-like monomeric conformation of α -syn (**IR7**) which consisted of two long helices formed by the residues F4 to V27 from NTR and K45 and T92 from FC region.

2.4.9. Solvation in different steps of fibril-dissociation

The hydration shell around the protein chain in different simulations was analyzed using radial distribution function, RDF (Fig. 2.22). The positions of the peaks represent the thickness of the coordination sphere of solvent molecules around the protein. For all three types of simulations, SMD, REMD and CMD, the first and second peaks appeared at the distance of 0.27 and 0.38 nm. Also, the RDFs calculated from the CMD simulations run with C36IDPSFF also showed the peaks at the same distance. The values were comparable with the reported values derived from the experimental methods.^[171] These observations suggested that the thickness of solvation sphere around the protein is similar in all the cases. The peak height, however, varied between the simulations (the insets in Fig. 2.22) indicating that the local density of water around the protein was different. This could be due to the differences in the conformation of α -syn in SMD, REMD and CMD which had differences in the hydrophobic exposure and inter-residue interactions during different stages of their simulations. In all the cases, the convergence of RDF values was almost the same implying that the bulk water behavior was also similar in all the simulations.

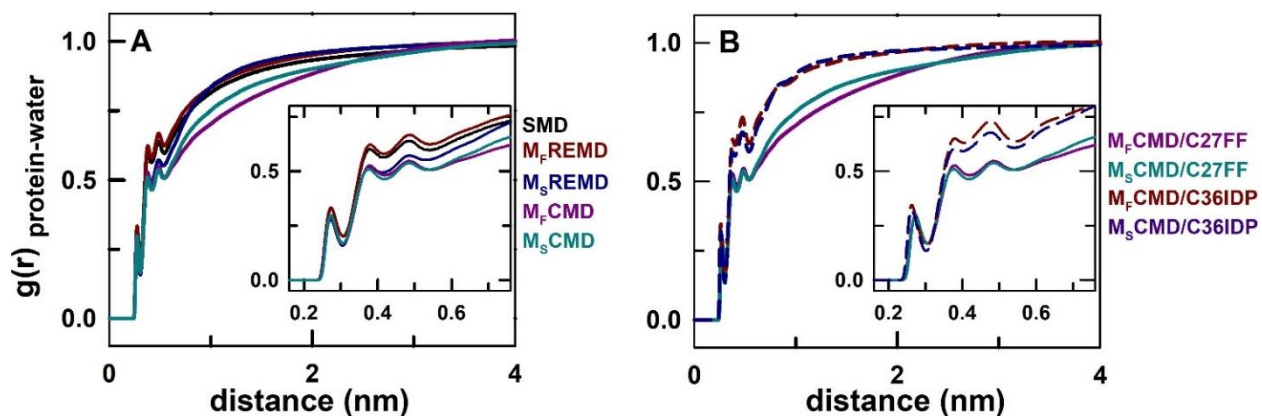


Fig. 2.22. The water around the protein surface was analyzed by the radial distribution function. (A) The RDF values of water evaluated from the steered MD (SMD) and REMD of both M_F and M_S simulations and CMD of M_F and M_S simulations performed using C27FF. (B) The RDF values of water evaluated from the CMD of M_F and M_S simulation performed using C27FF (solid lines) and C36IDPSFF (dashed lines). The colors of the lines are presented in the legends of each panel.

Further, to understand the role of solvation at each intermediate state, radius of gyration, SASA (Fig. 2.23), RDF of water around each intermediate, the number of H-bonds formed by the residues within the protein (intramolecular), and the number of protein-water H-bonds (Fig. 2.24) were calculated. Further, the free energy of solvation and the non-bonded interaction potential for each conformation was also evaluated (Fig. 2.25). The analyses were done for full-length and individual domains (NTR, FC and CTR) to understand the local and global solvation effects.

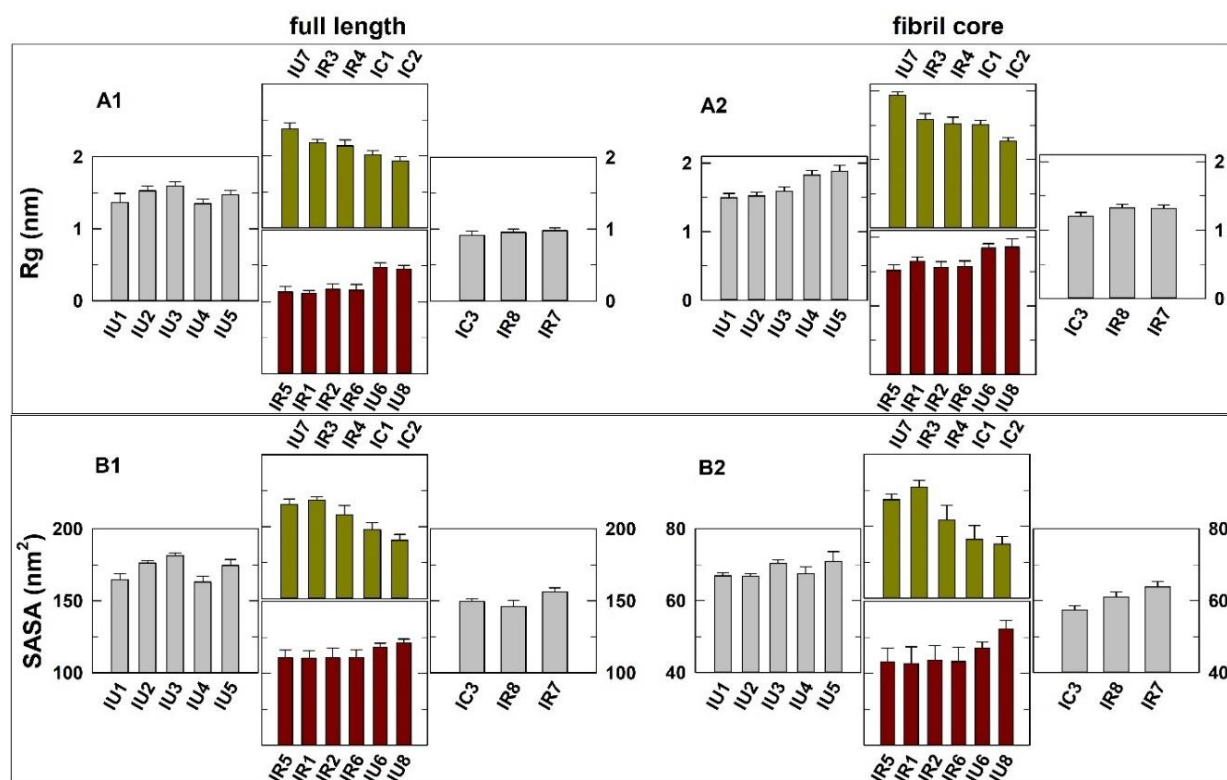


Fig. 2.23. (A1-A2) Rg and (B1-B2) SASA for full length protein and the fibril core of α -syn. In each panel moving from left to right represents the transition of fibril state to monomeric helical state of α -syn with each panel representing the identified intermediates (refer Fig 2.19 and Fig. 2.20).

In the fibril-dissociation stage, the initial conformations (**IU1** to **IU3**) showed an increase in the Rg and SASA of the whole protein and in FC region as well (Fig. 2.23). This could be due to the loss of interchain cross β -sheets which resulted in the exposure of FC residues. The change in RDF values suggested that during the dissociation of chain-A from chain-B, the fraction of solvent molecules around the protein did not change in the first hydration shell, whereas the RDF values slightly increased in the second hydration shell (Fig. 2.24A and 2.24B). Further, the solvation free energies calculated for individual regions (Fig. 2.25) indicated that NTR and FC regions had increased solvation energy from **IU1** to **IU3** which corroborated with the observation that the number of protein-water H-bonding interactions was higher in **IU3** (Fig. 2.24-D).

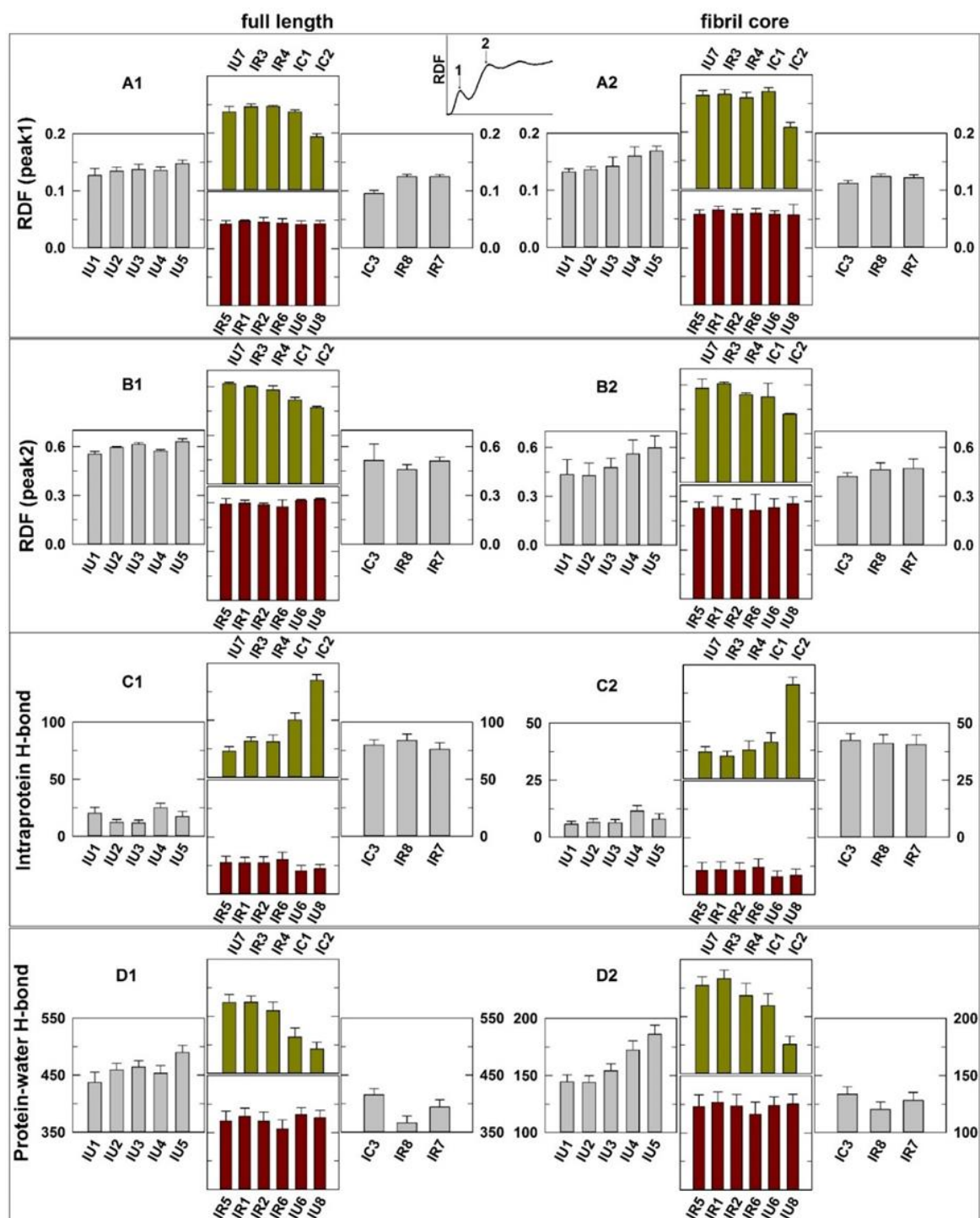


Fig. 2.24. The values of (A1-A2) first maxima and (B1-B2) second maxima in the radial distribution function (RDF) of water around the protein. A representative first and second maxima in an RDF plot is shown in the inset. (C1-C2) the number of intramolecular H-bonds in the protein, and (D1-D2) the number of H-bonds between the protein and water for full length (left panels) and fibril core residues (right panels) of α -syn. In each panel moving from left to right represents the transition of α -syn from fibril state to monomeric helical state with each panel representing the identified intermediates (refer Fig 2.20).

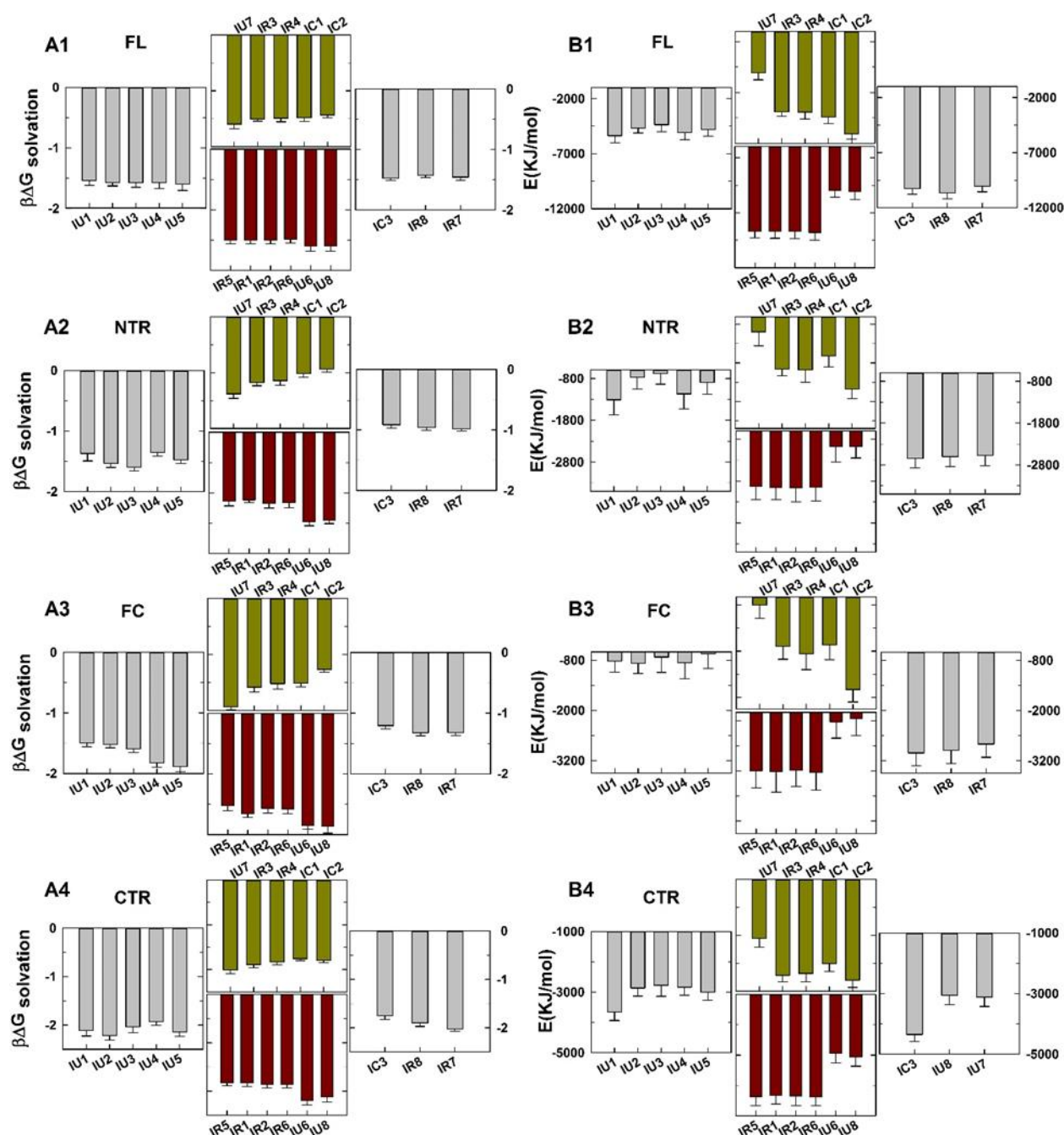


Fig. 2.25. (A1-A4) Solvation free energy and (B1-B4) intrachain non-bonded interaction energy for full length (FL), N-terminal (NTR), fibril core (FC) and C-terminal (CTR) of α -syn calculated for the selected 19 intermediate structures. In each panel, moving from left to right represents the transition of the protein from fibril state to monomeric α -helical state following different pathways (for the transition pathways refer Fig. 2.20).

Chapter 2

In the next steps (**IU4** and **IU5**), where chain-A almost completely lost its interaction with chain-B, the protein showed a decrease in R_g and SASA values. The RDF values suggested that the fraction of solvent molecules around the complete protein chain remained similar to that of earlier stages, but the fraction around the FC region increased. A significant increase in the number of protein-water H-bonds compared to the change in the number of intramolecular H-bonds in the FC region (Fig. 2.24C) suggested that the compaction could be due to the water network around the FC. An increase in solvation energy of FC region compared to NTR and CTR (Fig. 2.25) also emphasized that the water structure around the FC region would drive the chain to compaction. Also, the residues in HPR1 formed a short helix which could contribute to the reduction in the solvation free energy of N-terminal, whereas the solvation of C-terminal almost remained the same.

When α -syn follows the compact-conformational pathway, the SASA of the protein was gradually decreased (Fig. 2.23B) with a decrease in the number of water molecules around the protein, particularly the FC region as observed by the decrease in RDF peak values and the number of protein-water H-bonding interactions (Fig. 2.24). Decreased solvation free energy and increased non-bonded interaction potential suggested that intrachain interactions might stabilize these conformations. The increased intrachain H-bonding interactions in the FC region also evidenced for the same. Further, in these conformational states, the interdomain interaction of FC with both NTR and CTR was significantly increased along with the interactions between NTR and CTR (Fig. 2.26). These long-range interactions are known to influence the fibril formation propensity of α -syn:^[125,127,150,169,172] thus, contributing directly to the alteration of fibril pathway.

In the extended-conformational pathway, the R_g and SASA values of **IU6** and **IU8** were notably higher while the other four conformations did not show significant changes (Fig. 2.23). The increase in solvent accessibility of **IU6** and **IU8** could also be related with the decreased intrachain H-bonding interactions (Fig. 2.24C) and an increase in the protein-water hydrogen bonds (Fig. 2.24D). An increase in the solvation energy of all three regions, NTR, FC and CTR, in the case of **IU6** and **IU8** (Fig. 2.25) clearly indicated the expansion of the polypeptide chain. This could be due to weak long-range interactions within the domains and between three domains as well. Though other extended conformations showed intradomain interactions, the long-range

The helix to fibril transition pathway of α -syn

interactions between the domains (Fig. 2.26) were significantly reduced compared to the compact-conformational pathway.

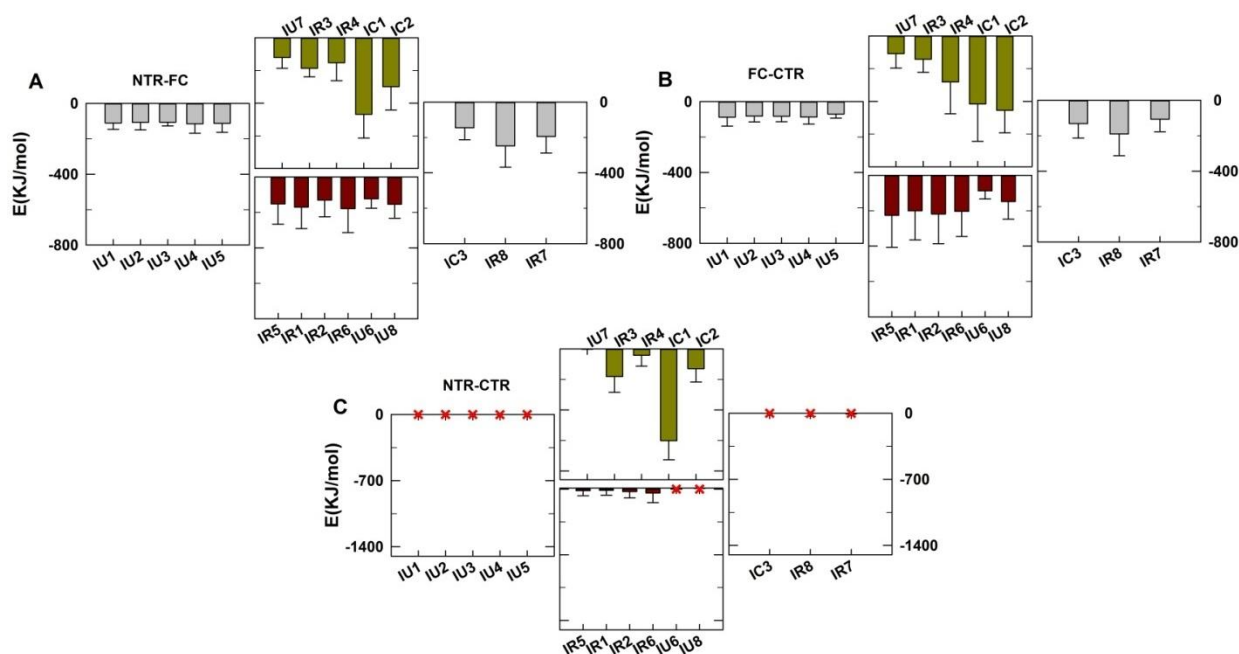


Fig. 2.26. Non-bonded interaction potential between the domains, (A) NTR-FC, (B) FC-CTR, and (C) NTR-CTR calculated to represent the long-range interaction between the regions.

In the later steps, which was the helical-conformational stage, NTR and FC region formed three helical structures (**IC3** and **IR8**) which reduced the overall SASA and Rg values of α -syn (Fig. 2.23). The helix formation increased intrachain H-bonds and the non-bonded interaction potential, but the protein-water interactions were reduced. Also, a decrease in SASA of CTR was observed for **IR8** in which formation of two salt bridges K97-E104 and K102-E123 were observed (Fig. 2.13B). This intradomain interaction could specifically reduce the solvent accessibility of the hydrophobic regions HPR-7 and HPR-8. At the end, the formation of a native-like monomer structure (**IR7**), two long-helices formed by NTR and FC, increased the Rg and SASA of the protein chain. In these conformational states, FC showed a few long-range interactions with NTR due to helix-helix interaction whereas CTR did not show any significant interaction with NTR (Fig. 2.26).

2.4.10. Dihedral angle distribution during transitions

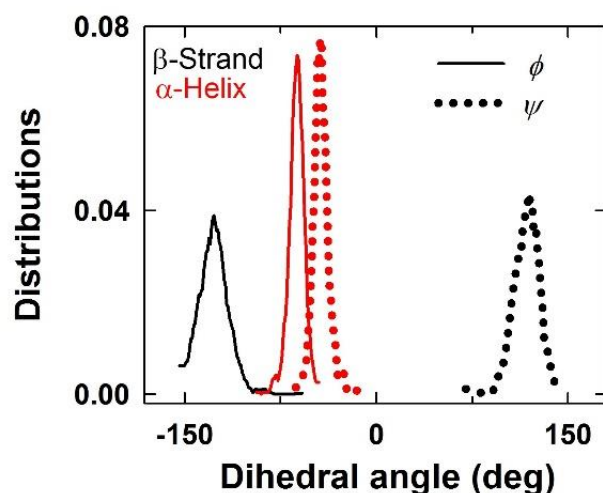


Fig. 2.27. Representative Φ and Ψ dihedral distributions of an α -helix (red) and a β -strand (black) in a protein.

Helices and sheets have specific distributions of Φ and Ψ angles (Fig. 2.27). Any secondary structural transition can be traced by following these distributions. To evaluate the changes in the backbone dihedral angles of α -syn during its helix-sheet transition, the Φ and Ψ angles of 32 residues, such that two residues from each of six RTDs and 20 residues from HPR3 to HPR6 in the fibril core, were analyzed. RTD1 to RTD3 which are part of helix-1 showed large conformational plasticity (Fig. 2.28). These regions mostly formed β -turns of type I and II^[173] with an intermittent occurrence of a short α -helical conformation when the protein was in the fibril form. During the dissociation via the compact conformational pathway, the helical formation was observed at late intermediates whereas in the extended-conformational pathway only a short α -helical formation was observed. In the helical-conformational stage, the backbone dihedrals of RTD1 and RTD2 were mostly α -helical, RTD3 showed a considerable fraction of β -turn as well. The repeat domains RTD4 to RTD6 which were part of FC showed the dihedral angles corresponding to β -sheet contents or type-II β -turns during the initial dissociation stage. In the compact-conformational pathway, these regions showed structural transformations between type-I and type-II β -turns whereas the dihedrals with helical structural arrangements were observed in RTD4 only at the later stages. In the extended-conformational pathway, the RTD4 and RTD5 mostly showed type-II turns and in RTD6 notable fraction of helical structures were also observed.

The helix to fibril transition pathway of α -syn

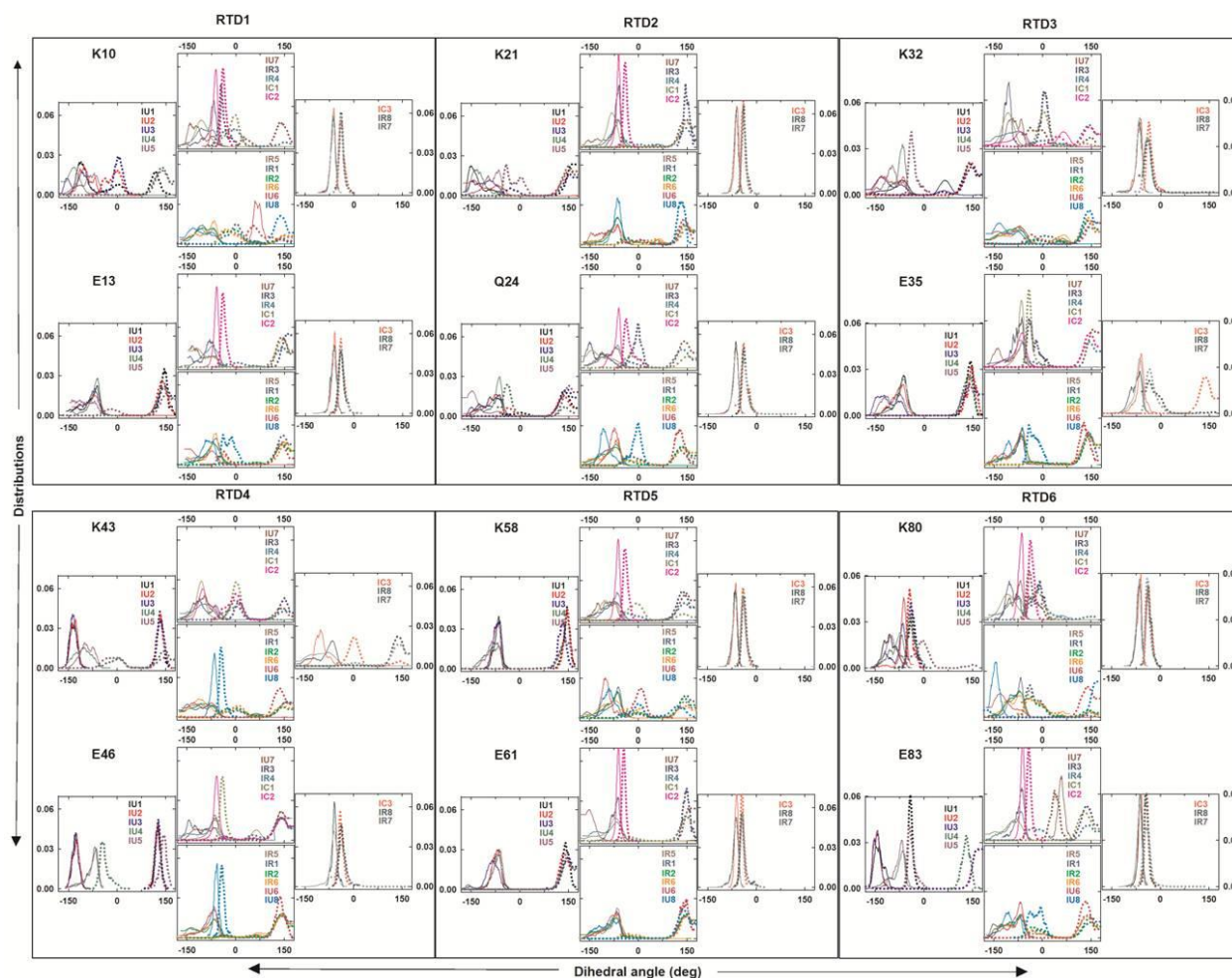


Fig. 2.28. Distribution of Φ and Ψ dihedral angles, of 12 residues of α -syn selected such that two residues (K and E or Q) from each of six pseudo-repeat domains, RTD1 to RTD6. For the transition pathway and the intermediate structures, refer Fig 2.20.

The HPR3 and HPR5 regions which were part of FC, mostly remained as β -sheets or type-II β -turn whereas HPR4 and HPR6 showed larger fluctuations in the backbone dihedral angles during the fibril dissociation stage (Fig. 2.29 and Fig. 2.30). In the compact-conformational pathway, the frequency of β -sheet occurrence was reduced; however, the α -helical dihedrals were sparsely observed only in HPR5 and HPR6 regions. Though the number of residues forming α -helices was short during this stage, the dihedral distributions indicate that these regions might have more conformational plasticity to form α -helices. In the extended-conformational pathway, along with β -sheet, the dihedrals corresponding to both type-I and type-II were also noticed for the residues in the HPRs. Along with HPR5 and HPR6, the residues in the HPR4 region also showed α -helical dihedral distribution, though the fraction was less. In the helical-conformational

Chapter 2

state, the sampling of the residues in HPR3 and end of HPR6 for α -helical dihedral angles was less.

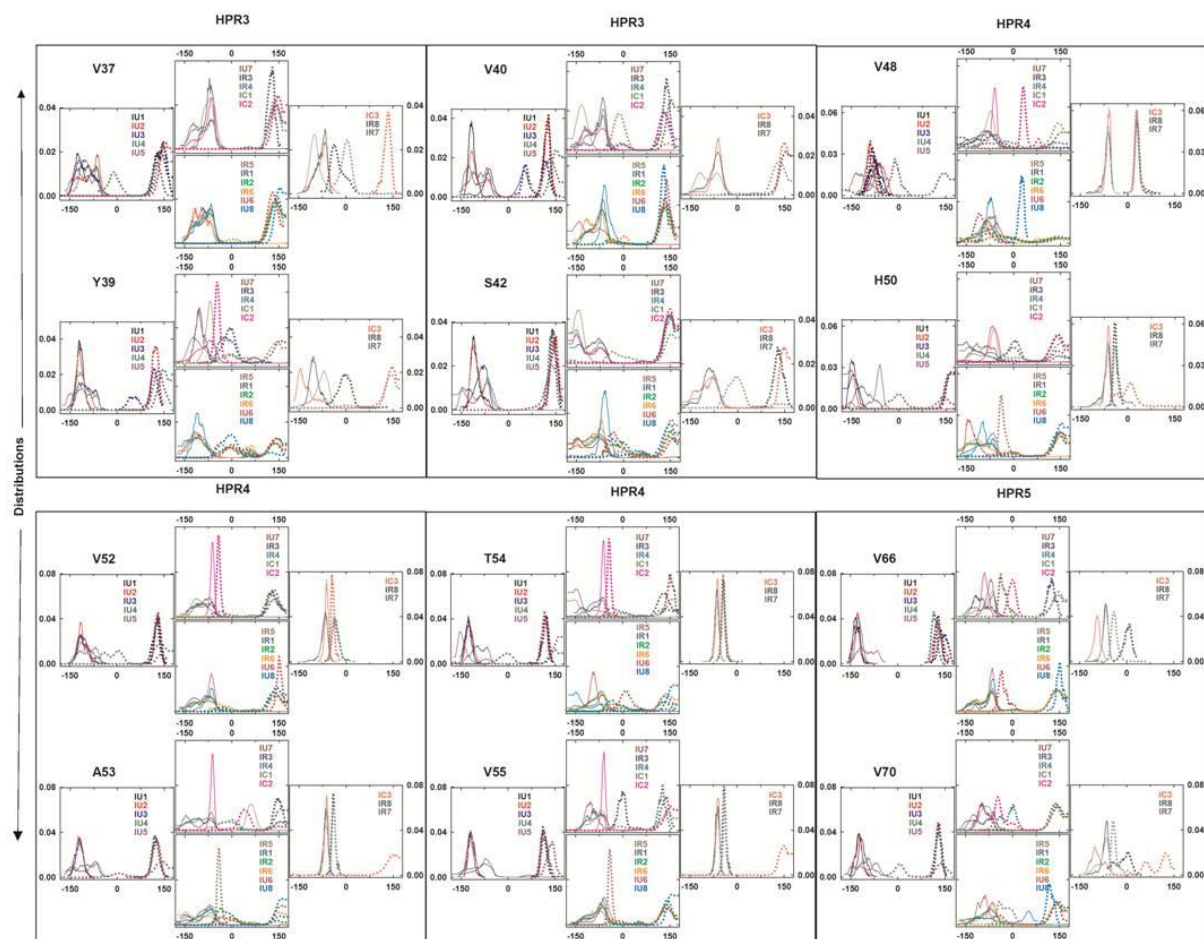


Fig. 2.29. Distribution of Φ and Ψ dihedral angles, of 20 residues from hydrophobic regions in the fibril core (four each from HPR3, six from HPR4 and three from HPR5) calculated for each intermediate structure identified for fibril-helix transition. For the transition pathway and the intermediate structures, refer Fig 2.20.

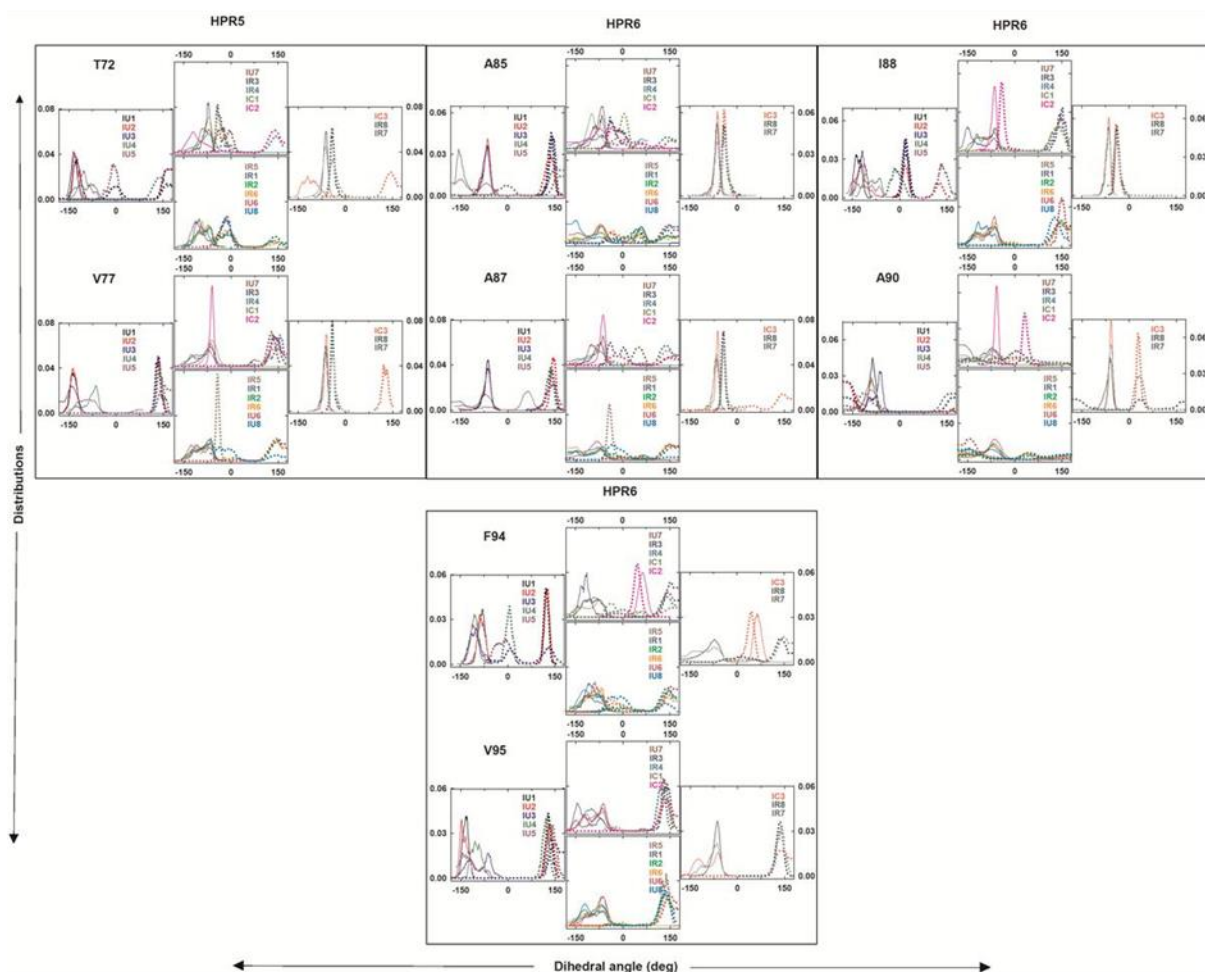


Fig. 2.30. Distribution of Φ and Ψ dihedral angles, of 20 residues from hydrophobic regions in the fibril core (one from HPR5, six each from HPR 6) calculated for each intermediate structure identified for fibril-helix transition. For the transition pathway and the intermediate structures, refer Fig 2.20.

2.5. Discussion

The structural transformation in α -syn from α -helical conformation to cross- β sheets involves many intermediate states. The combination of different conformational sampling methods enabled the identification of some of the major intermediate states that might show a definite influence on driving the fibril formation in the protein.

2.5.1. Sampling of fibril dissociation pathway:

From the enhanced sampling methods, most sampled conformations were obtained and ordered based on the structural properties to characterize the probable fibril dissociation

Chapter 2

pathways. The SASA, Rg values and secondary structural changes clearly indicated that there could be two different pathways. The initial dissociation of fibril monomers might be driven by the loss of salt bridges, then the hydrophobic residues in HPRs are finally released. The residues V71-V83, the center of the fibril core region, dissociates only at the end of steered-MD which are found to form rigid β -sheets by the experiments.^[134] The flanking “helix-fraying” residues show more conformational flexibility in the fibril state as well compared to the central region. The dissociated chain might either transfer into a compact state by long-range interactions or helical transition might occur through the extended conformation itself driven by local interactions. Structural intermediates with varying solvent exposures in the FC region are observed in HX-mass experiments.^[134]

In the compact-conformational pathway, the appearance of helical structure is more frequent compared to the structures in the extended-conformational pathway. This emphasized the experimental observations that long-range interactions between the NTR, NAC and CTR regions could stabilize the helices in α -syn.^[127,174] The changes in the intermediate structures are mostly driven by the reorganization of β -turns in the FC residues which are mostly facilitated by the rearrangement of salt bridges between the repeat domains. Particularly, the Gly/Ala-rich regions G67-A69, G84-G86 and G89-A91 are more populated with the β -turns as suggested in the earlier solid-state NMR experiments^[149] and by computational methods.^[93] Formation of β -hairpin structures in HPR3 and HPR4 (pre-NAC region) is also observed in computational studies.^[175] It may also be noted that HPR5 (residues: 66-78) and HPR6 regions (residues: 84-95) show the early formation of α -helical conformation in both pathways. Further, β -turn around the Tyr-39 residue, similar to earlier studies,^[127,146] could be noted at the early fibril dissociation stage and at the intermediates in the helical-conformational pathway. The conformational change in Tyr-39 plays a crucial role in fibril assembly and has been the target site for the binding of anti-amyloidogenic agents.^[131]

In RTD2-RTD5 regions, most of the dihedral changes are associated with change in Φ -angle during initial fibril dissociation states (**IU1-IU4**), whereas Ψ -angles show only little variations (except V48, K80 & E83, I88 & A90). HPR6 shows variation in both which shows a considerable population of β -turn formation as well. This suggests that the formation of β -turn could be an important conformational-switch during the helix-fibril transition in α -syn. In experiments, differences in the buffer conditions are found to influence the final morphology of

the fibrils by altering the fibrillation pathway.^[109,176] The present study could identify at least two different pathways, though more might be possible.

2.5.2. Monomer to fibril - the probable reverse transition pathway

From the major intermediates identified, the probable transition pathway from helix to cross β -sheet could be predicted (Fig. 2.20). The initial steps could be bending of the long-helices (**IR7**), particularly through the dihedral changes around the RTD2 in helix-1 and HPR5 in helix-2. Further, the loop region connecting the helices (HPR3) undergoes large conformational changes along with the flanking repeat domains, RTD3 and RTD4. In the further step, the bend region of HPR5 may lose its α -helical content and form loops which would convert the linear helix-2 into nearly two orthogonal helices (**IR8** and **IC3**). However, helix-1 almost retains its complete α -helical content. During these steps, the protein might become compact by the changes in intramolecular H-bonds in the FC region and the interdomain interactions between NTR and FC. From these conformations, α -syn may adapt different pathways to attain the cross β -sheet structure.

In a compact-conformational pathway, helix-2 might be disrupted further at HPR6 region and helix-2 becomes three short-helices (**IC2**). Then, the helical structures might be destabilized by protein-water H-bonding interactions and the loss of intramolecular salt bridges formed in the RTDs. However, they might possess long-range interdomain interactions between NTR, FC and CTR. During these conformational changes, helix-1 might also be disrupted due to the loss of interactions in RTD1 and RTD2 (**IC1**). In the further steps, the HPRs 3 to 6 (covering the FC region residues 37-98) form loops which might gradually lead to a decrease in the intramolecular H-bonding interactions and an increase in protein-water interactions (**IR4**). Also, the long-range interactions between the domains are also gradually lost. Initially, HPR3-HPR4 residues and HPR5-HPR6 residues form two loops with a transient appearance of short α -helical structure across the FC region. With the rearrangement of salt bridge interactions in the RTDs, HPR3 forms a short-independent loop whereas HPR4-HPR6 residues form a long loop. These conformational changes happen through the formation of transient β -turns. The loop interactions are slowly lost by the loss of intramolecular H-bonding interactions and the FC region become more solvent-exposed and adopts a conformation similar to the Greek-key motif. These transformations also disrupted the long-range interactions between the domains (**IU7**).

Chapter 2

In an alternate, extended-conformational pathway, there are no notable interactions between the residues in HPRs. Each region independently adopts its own loop conformation. In these conformations (**IU8**, **IR6** and **IR2**), the H-bonding interactions among the protein residues are less in FC region and the water around the FC region is consistently higher compared to the intermediates in the initial stage of compact-conformational pathway. Further, there are no significant long-range interactions between NTR and CTR, but a few interactions by FC region with the terminal regions, NTR and CTR, might occur. At later stages, the HPR regions might pack together, particularly the residues in HPR3-HPR4, and HPR5-HPR6 (**IR1** and **IR5**). This might be driven by hydrophobic packing rather than the H-bonding interactions within the protein residues. In the further steps, the FC region may adapt the conformation similar to the Greek-key motif by loop rearrangements.

Once the Greek-key motif-like conformation is attained by the loops in FC region, the interaction between the chains may be induced by the packing of hydrophobic residues in the FC region (**IU3**, **IUR**, and **IU5**) and interchain salt bridges. This might further stabilize the formation cross β -sheets between the chains (**IU2** and **IU1**). In these steps, no long-range interactions between NTR and CTR regions could be found.

2.5.3. Limitations

This study mainly illustrates the conformational changes during helix-fibril transition with the ensembles obtained from different simulation methods. There could be more intermediate structures between any two states proposed here. For instance, the probable intermediate states between $IC1 \rightarrow IC2$ and $IU8 \rightarrow IC3$ are not sampled by any of the methods. Since different sampling methods are used, the rate of transition between the states could not be derived. Also, the transition energy between the states could not be evaluated.

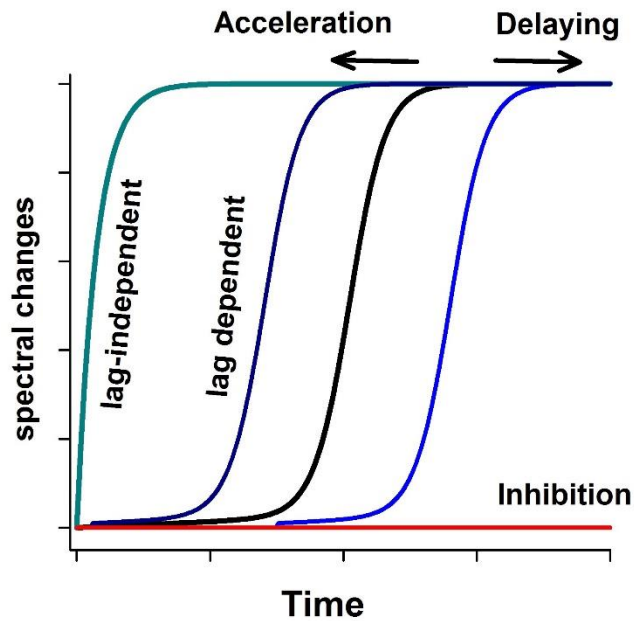
2.6. Conclusion

Fibril dissociation pathway of full-length α -syn was followed by steered-MD simulation of pulling a single chain from the fibril conformation of α -syn. From three different stages of the SMD, 57 representative conformations were chosen and subjected to umbrella sampling. Further, REMD and CMD simulations of a single chain from the fibril (M_F) and α -helical monomer (M_S) of the protein was performed. From all these simulations, 19 distinct structures were chosen and arranged based on their R_g , SASA and secondary structure content to predict the fibril dissociation pathway. It was observed that the fibril dissociation might occur either through a compact-conformational pathway or an extended-conformational pathway. The secondary structural content, salt bridge interactions, long-range contacts and solvent accessibility of each intermediate state was analyzed. From these analyses, probable fibril formation pathways were also predicted. This concerted approach helps to characterize the intermediate structures in the fibril formation pathway(s) of the intrinsically unfolded protein which otherwise would be an impossible task with any single method. Also, understanding the interactions in these intermediates including solvation energies and long-range contact might also aid in designing anti-fibrillation agents.

Chapter 2

Chapter 3

Effect of charged amino acids and their derivatives on the fibril formation of α -synuclein



Chapter 3

3.1 Abstract

Aggregation of α -synuclein, a 140 residue intrinsically disordered protein, is a hallmark of Parkinson's and dementia with Lewy bodies (DLB) pathogenesis. The high plasticity and lack of stable tertiary structure make α -Syn highly susceptible to its surrounding environment. Under stress conditions, small organic molecules, called osmolytes might accumulate in the cells which could affect the conformational state and fibrillation pathways of proteins. In this chapter, the effect of charged amino acids (L-glutamate, L-aspartate, L-lysine, and L-arginine), amino acids with amide side chains (L- glutamine and L-asparagine), and *N*-acetylated amino acids (N-acetyl-L-glutamic acid (NAG) and N-acetyl-L-lysine (NAL)) on the fibrillation of recombinant human α -syn was examined. Arg and NAL inhibited the fibrillation of α -syn at concentrations above 0.2 M and 0.4 M, respectively. In the presence of Lys, Asn and Glu, the protein attained the fibrillar state faster than in buffer which could be attributed to the reduction of lag time by these amino acids. NAG induced a lag-independent fibrillation, whereas Gln and Asn exhibited concentration-dependent effects on α -syn fibrillation such that with an increase in their concentration, the lag time was reduced.

Chapter 3

3.2 Introduction

‘If you make monomers wet, they don’t turn into polymers’ - Bryson, B. ^[177]

The popular literature explains the role of hydration in monomer to polymer transition. As an intrinsically disordered protein, α -syn has more unstructured regions in which folding or fibrillation is driven by its solvation properties.^[118,119,178] These IDPs are involved in many neurodegenerative diseases by forming protein aggregates. The aggregated form of α -syn is associated with Parkinson's disease (PD) and dementia with Lewy bodies (DLB) pathogenesis.^[110,113,114,179] It is reported that α -syn aggregates under hyperosmotic stress caused by different osmolytes. These aggregates appear only in dead cells suggesting that they are toxic.^[96]

3.2.1 Osmolytes in neural environment

The earlier studies reported the activation of efflux pathways for amino acids, K^+ , and Cl^- during a regulatory volume decrease (RVD) in cultured cerebellar granule neurons that are exposed to hyperosmotic conditions. During hyponatremia, an influx of water inside the neural cell across the plasma membrane in response to reduced plasma osmolarity is less than expected.^[180,181] Franco et.al demonstrated the efflux of amino acids in the isovolumic regulation (IVR) inside the hippocampus slices to decrease the osmolarity of the cell.^[97] The mechanism of cell volume restoration is different for different cells; however, it should be precise for the central nervous system (CNS) due to the physical restriction of the skull. Minor alterations in the volume of brain cells can significantly impact the spatial arrangements among neurons, astrocytes and the extracellular space. When brain cells swell, causing a decrease in the extracellular space, it leads to greater lateral diffusion and elevated levels of neurotransmitters outside the cells.^[182,183] While inorganic osmolytes make up the majority (65%) of the contribution to cell volume adaptation, organic osmolytes are highly concentrated in the central nervous system (CNS).^[184] The use of organic osmolytes helps to minimize the changes in membrane potential that can result from the outflow of inorganic osmolytes such as K^+ or Cl^- . Nevertheless, it is important to note that organic osmolytes have a "neuroactive" property and can activate their respective receptors on nearby neurons and glia leading to potential neurobiological effects.^[185–189] Therefore, the release of organic osmolytes during volume correction has significant neurobiological consequences.

3.2.2 Osmolytes on protein fibrillation

Inside a cell, osmolytes can behave in two different ways on globular proteins. Firstly, as protecting osmolytes such as trimethylamine N-oxide (TMAO) which pushes the free energy of protein towards the native conformation. Secondly, denaturing osmolytes such as urea lower the free energy to favour the unfolded states of proteins.^[190–193] α -Syn aggregates inside the cells in the presence of osmoregulatory molecules, but how the presence of these molecules affect the fibril formation is still unclear. An osmolyte's mode of action can change based on its concentration. For example, the addition of 1 M TMAO results in a partially unfolded state of α -syn which leads to fibril formation, whereas fibrillation is inhibited in the presence of higher concentrations of TMAO (>3M) and the protein forms only oligomers.^[194] Charged amino acids and their derivatives act as osmolytes in neural cells.^[195] Majorly, glutamate, glutamine, lysine, arginine and *N*-acetyl aspartate has been reported to affect the fibrillation of proteins inside the cell during neural disease conditions.^[107,196,197]

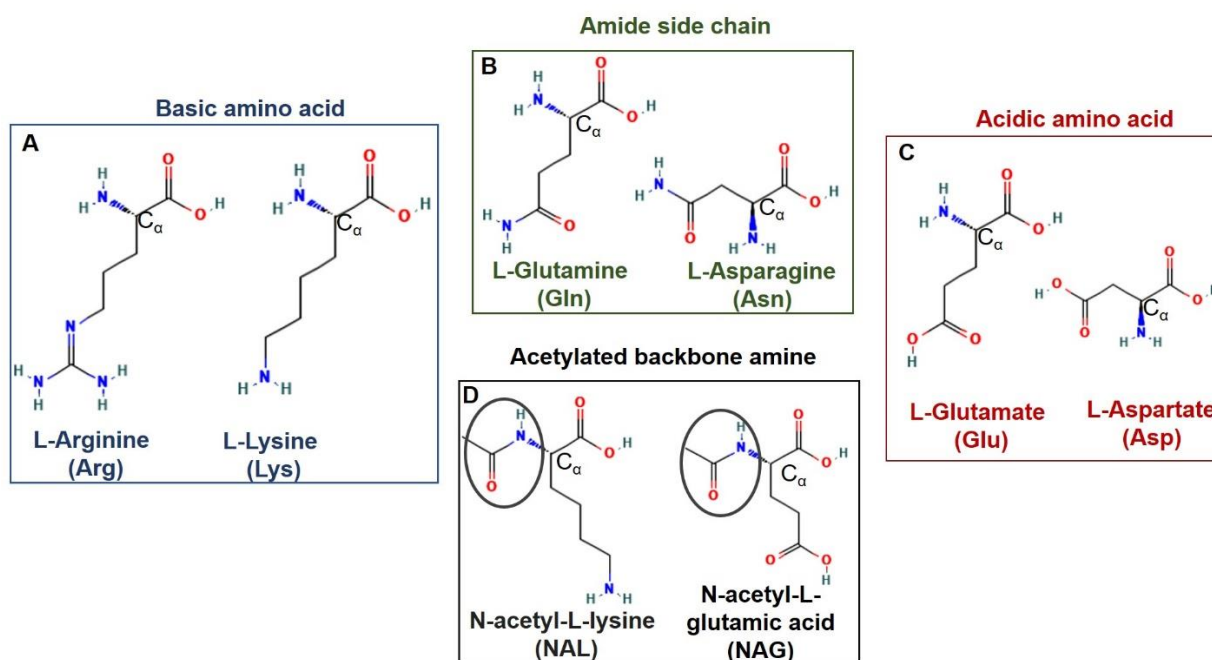


Fig. 3.1 Amino acids chosen for the current study: (A) basic amino acids with net positive charge, Arg and Lys, (B) polar amino acids with amide side chains, Gln and Asn, (C) acidic amino acids with net negative charge, Glu and Asp, and, (D) *N*-acetylated amino acids NAL and NAG. Black circles highlight the *N*-acetylation group.

Chapter 3

These acetyl derivatives are found in the mammalian brain and almost in all cells of eukaryotes. We here investigated the concentration-dependent effect of four charged amino acids, (L-glutamate, L-aspartate, L-lysine, and L-arginine) and amino acids with amide side chains (L-glutamine and L-asparagine) on α -syn fibrillation. It may be noted that *N*-acetyl aspartate (NAA) is insoluble in water at pH 7. Hence, *N*-acetyl glutamate (NAG) and *N*-acetyl lysine (NAL) were chosen to study the effect of acetylated amino acids.

3.3 Materials and Methods

3.3.1 Materials

Nucleo-Spin® Gel and PCR Clean-Up Kit by Macherey-Nagel and GSURE® Plasmid Mini Prep Kit by GCC Biotech were used. DNA ladder, FastDigest restriction digestion enzymes and PCR mix along with their activation buffers were procured from Thermo Scientific. Protein ladder was bought from PUREGENE and Agar was from SRL. All the LB media components, EtBr and antibiotics were from HiMedia. Agarose was from Lonza. His60 Ni superflow resin and pCOLD1 were from Takara, PVDF membrane were from Merck Millipore. Pierce™ ECL western blotting substrate and α -synuclein recombinant rabbit monoclonal antibody were from Thermo Fisher Scientific. All the chemicals required for SDS-PAGE and buffers preparation were obtained from SRL. Glutamate, Aspartate, Arginine, and Lysine were from SRL, and Glutamine, Asparagine, NAG, and NAL were from Sigma Aldrich.

3.3.2 Cloning of α -synuclein

The human synuclein-alpha (*SNCA*) complementary DNA (cDNA) open reading frame was amplified from *pTOB7-SNCA* (gifted by Prof. Chen Ru Ruby, Academia Sinica, Taiwan)^[198] using forward primers SNCA-FP-NdeI (5' CAA CAT ATG ATG GAT GTA TTC ATG AAA GGA CTT TCA AAG G 3') and reverse primer SNCA-RP-HindIII (5' GAC AAG CTT TTA GGC TTC AGG TTC GTA GTC TTG ATA C 3'). The amplified *SNCA* PCR products were digested with NdeI and HindIII. The vector *pCold* was digested with NdeI and HindIII to generate similar sticky ends for cloning the insert *SNCA*. NdeI and HindIII were selected, because they are single cutters for *pCold* I expression vector and zero cutters for *SNCA*. The ligation reaction was performed using T4 DNA ligase with 1:3 molar ratio of vector with insert and transformed into *E. coli* Rosetta (DE3) pLysS expression host strain by heat-shock method. Luria-Bertani (LB) Agar

plates with 0.1 mg/mL ampicillin (Amp) and 0.05 mg/mL chloramphenicol (Chl) were used for the selection and growth of the transformed bacteria.

3.3.3 Expression of α -synuclein

A single colony from the plate was used to inoculate a primary culture in 10 mL LB broth medium (Amp + Chl). The culture was allowed to grow overnight for 16 hours at 37 °C. From the overnight grown primary culture, 800 μ L was used to inoculate secondary culture in 800 mL LB broth medium (without antibiotics). The secondary culture was incubated at 37 °C till the O.D. reached 0.6-0.8 as measured by Cary100 absorbance spectrophotometer at wavelength 600 nm. It was then removed from 37 °C and incubated on ice for 30 minutes. From the leftover primary culture, glycerol stock was prepared and stored at -80 °C refrigerator. Then 1 mM IPTG (isopropyl β - d-1-thiogalactopyranoside) was added to the culture and allowed to incubate at 15 °C for 24 hours. Afterward 800 ml of culture was subjected to centrifugation at 10,000 rpm at 4 °C where supernatant was discarded and pellet was stored at -80 °C.

3.3.4 Purification of α -synuclein

Stored cells were harvested by sonication, where stored pellet was dissolved in lysis buffer (20 mM Tris buffer (pH 6.5), 500 mM NaCl, 20 mM Imidazole and 1mM PMSF) using sonicator with the amplitude 20 kHz. Each pulse was carried for 45 sec with an interval of 1 min cooling on ice. The process was repeated either till 10 cycle or until clear lysate is visible. Then, the lysed of cells were subjected to centrifuge at 10,000 rpm for 30 min. The supernatant was collected and filtered through a 0.45 μ m filter syringe. Filtered supernatant consisted of all the proteins in the cells along with the overexpressed α -syn. 10 ml of the filtered cell lysate was added to the 1 ml of Ni-NTA beads in the column and allowed to incubate by rotating the column for 2 hours on Rotospin by Tarsons for effective binding of the protein on beads. The bound protein was washed and eluted with different concentration of imidazole (20 mM- 400 mM). The purity of the obtained protein was assessed on 15% SDS-PAGE.

3.3.4 Western blot analysis

α -Syn monoclonal antibody was used as a primary antibody, and horseradish peroxidase-linked antibody raised in goat against Rabbit IgG was used as a secondary antibody. After SDS-

Chapter 3

PAGE separation, protein was blotted on to PVDF membrane in a semi-dry transfer apparatus (TE77-PWR semi-dry transfer unit, GE Healthcare). α -Syn protein was detected immunologically with Pierce TM ECL Western blotting substrate using Chemidoc (XRS+) and analyzed with Image Lab software (Bio-Rad).

3.3.5 Structure and fibrillation studies

Circular dichroism (CD) measurements were performed using 10 μ M of purified recombinant α -syn (α -syn) in 20 mM sodium phosphate buffer at pH 7 in far-UV region 190 nm to 250 nm to evaluate the secondary structural content. The spectrum was collected in Jasco J-1500 spectropolarimeter with a 0.02 cm pathlength cuvette. Fibrillation of α -syn was induced by 40 μ M of the protein in 20 mM phosphate buffer at pH 7 and shaken at 180 rpm with four sterile glass beads at 310 K (37°C) for 72 hrs. Fibril formation was followed using Thioflavin T (ThT) fluorescence measured between 460 and 510 nm in Jasco FP8500 after exciting the dye at 440 nm. The kinetics of fibril formation was followed by the change in ThT fluorescence at 485 nm. Fibrils formation was studied in the presence of varying concentration (0.2 M, 0.4 M, 0.6 M, 0.8 M, and 1.0 M) of the selected amino acids under the same condition.

3.3.6 Data analysis

Fibrillation kinetics followed lag-dependent and lag-independent fibrillation pathways. The data were analyzed using the following equations to obtain the rate and lag time.

For lag-dependent kinetics^[32],

$$Y = (y_i - m_i t) + \frac{(y_f + m_f t)}{1 + (\exp^{k_{el}(t - t_o)})} \quad (3.1)$$

In Eqn. 3.1, Y is the change in ThT fluorescence, y_i and y_f are intercepts and m_i and m_f are slopes for the initial and final phases, respectively. k_{el} is the elongation rate of fibril formation, t_o is the time required to attain half of the total fluorescence change and t is time. Lag time (T_{lag}) can be calculated as, $T_{lag} = x_o - (2/k_{el})$.

For lag-independent kinetics^[31]

$$Y = b + a (\exp^{-k_{el}t}) \quad (3.2)$$

in Eqn. 3.2, b represents the exponential maxima and a is amplitude of the fluorescence change.

To understand the cumulative effect of lag time (T_{lag}) and elongation rate (k_{el}) on the fibril formation time, apparent time (T_{app}) was calculated as^[199],

$$T_{app} = \begin{cases} t_{lag} + 2.3/k_{el} - lag \text{ dependent kinetics} \\ 2.3/k_{el} - lag \text{ independent kinetics} \end{cases} \quad (3.3)$$

3.3.7 Microscopic studies

The fibril samples of α -syn placed on a carbon-coated copper grids (from Ted Pella) were dried in ambient condition and stained with 1% uranyl acetate. The images were collected in a JEOL transmission electron microscope (TEM), JEM-F200 at an accelerating voltage of 200 kV.

3.4 Results

3.4.1 Confirmation of clone

Polymerase chain reaction (PCR) was performed using the primers mentioned in the section 3.3.2. PCR product (Fig. 3.2C) and vector were double digested by NdeI and HindIII (Fig. 3.2B) to produce a sticky end for each other and align in the proper direction in a frame. Both vector and insert were ligated using T4 DNA ligase and then transformed into the Top10 strain of *E.coli*. Screening for positive clones was done by colony PCR (Fig. 3.2D). Sequencing was performed to confirm the insertion of *SNCA* gene into the *pCold I* plasmid vector using the same forward and reverse primers. The obtained sequence with insert was aligned with the *Snca* cDNA sequence from NCBI with the help of Snap gene software. There was no mutation found (Fig. 3.2A). The isolated *pCOLD-SNCA* gene was then transformed into a Rosetta cell (expression vector) for overexpression. The protein was then purified using 100 mM-400 mM of imidazole (Fig.3.2E). A Western blot was developed with the purified recombinant α -synuclein to confirm the protein (Fig.3.2F).

Chapter 3

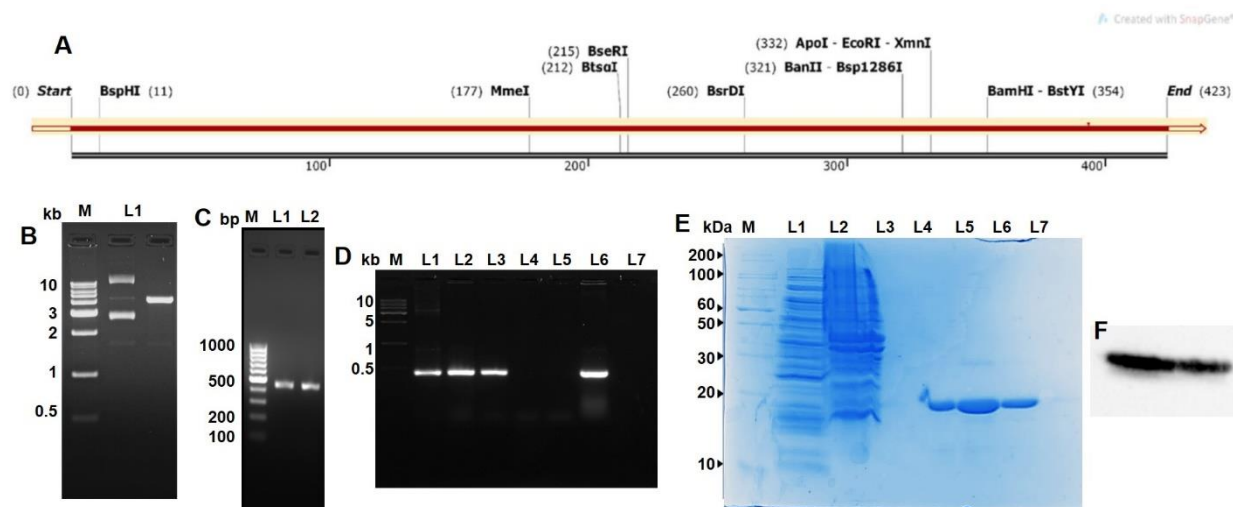


Fig. 3.2 (A) Sequence alignment with cDNA sequence of α -syn with single endonuclease cutter site. Black line represents the cDNA sequence and red line represent the *SNCA* gene in pCOLD1. (B) L1: undigested *pCold* expression vector and L2: digested *pCOLD* I (4407 bp). (C) L1 and L2: PCR Amplified *SNCA* (423 bp). (D) *SNCA* gene from colony PCR of *pCold* I-*SNCA* transformed to the Top10 cell. L1 has PCR amplified *SNCA* gene (positive control), L2, L3, L4, L5 and L6 represent *SNCA* gene from five different colonies from LB-Agar plate along with negative clone in L7 (reaction mix without clone). In all the gels (B-D) first lane represent DNA ladder marked as M. (E) SDS-PAGE of the α -Syn purified from Ni-NTA column where M is protein ladder of 10- 250 kDa, followed by (L1) protein found in pallet of lysed cells and (L2) supernatant flow through. (L3) Unbound proteins washed with 40 mM of imidazole and eluted with different concentrations of imidazole. The protein eluted at (L4) 100 mM, (L5) 200 mM, (L6) 300 mM and (L7) 400 mM of imidazole. (F) The Western blot developed using α -synuclein recombinant rabbit monoclonal antibody with 1:2000 dilution.

3.4.2 Fibrillation of α -Syn

The random coil structure of purified α -syn was verified with far-UV CD spectrum of the protein. The negative ellipticity at 198 nm referring random coil conformation was observed similar to the earlier reports on monomeric state of α -syn (Fig. 3.3A).^[200–202] α -Syn is known to form fibrils through lag-dependent pathway with or without agitation.^[30,198,203] The initial phase of fibrillation is known as nucleation which involves the formation of a homogenous nucleus composed of protein molecules. This step is considered to be an energy barrier during the fibrillation

Following the lag time, the protein begins to assemble into cross- β structures by recruiting additional monomers or protofibrils.^[32] This second phase, known as fibril elongation, typically

occurs at a faster rate than the lag phase and follows an exponential kinetics.^[204–206] In the present study, the fibril formation of α -syn was initiated at pH 7 by shaking at 180 rpm and 37 °C for 72 hours. The rate of fibrillation was followed by increase in ThT fluorescence (Fig. 3.3B). ThT specifically binds with the cross- β structures fibrils,^[207] therefore, during the initial oligomerization it does not show any significant change in its fluorescence. The kinetic curve was fitted using Eqn. 3.1 to evaluate the lag time and rate of fibrillation. Further, the fibrils of α -syn were characterized with TEM micrographs that showed thin and long fibrils (Fig. 3.3C).

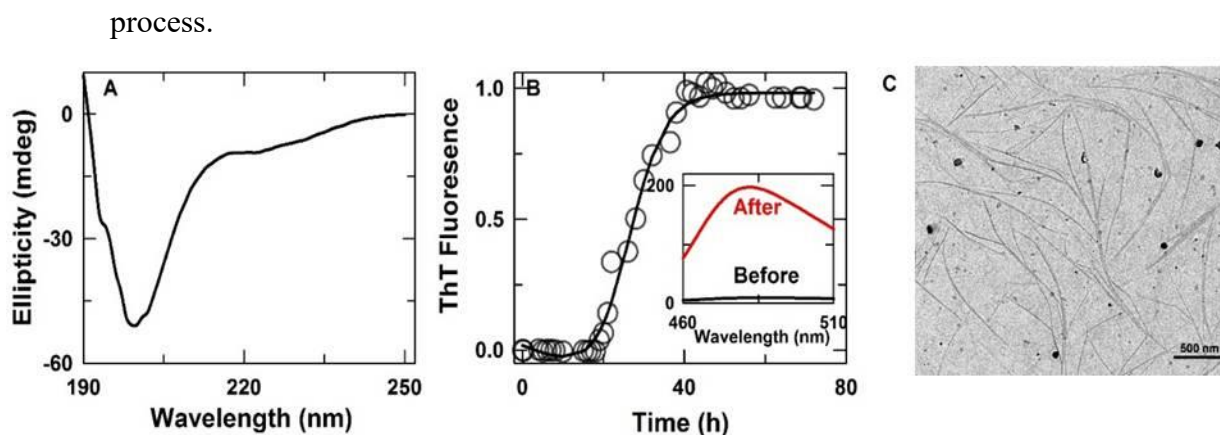


Fig. 3.3. (A) The random coil nature of α -syn confirmed by far-UV CD spectra collected using 10 μ M of the protein. (B) The fibril formation of α -syn followed by change in ThT fluorescence in 20 mM phosphate buffer at pH 7 incubated at 37 °C with a constant shaking at 180 rpm. The solid line represents the data-fit using the Eqn. 3.1. Inset represents the fluorescence spectra at the initiation of fibrillation (black) and after 72 hours (red). (C). Transmission electron microscopic image of the fibrils of α -syn.

3.4.3 α -Syn fibrillation in the presence of amino acids

The fibrillation process of α -syn was monitored in the presence of eight selected amino acids by varying their concentrations from 0.2 M to 1 M (Fig. 3.4). This concentration range could mimic the molar ratio of α -syn:amino acids inside the cellular environment. The rates and lag time were derived from these kinetic traces using Eqn. 3.1 or 2 (Fig. 3.5). Further to understand the total time taken by the protein to reach the saturation phase of the kinetics in the presence of amino acids, irrespective of whether it followed lag-dependent kinetics or not, apparent fibrillation time was evaluated using the Eqn. 3.3.

In the presence of Glu, Asp, and Lys, α -syn formed fibrils earlier than in buffer alone. (Fig. 3.4A-C) This could be due to the reduced lag time combined with faster elongation rate (Fig. 3.5A and C). The lag time in the presence of Glu and Asp was drastically reduced to 3 ± 1 hour from the

Chapter 3

18.7 hour in buffer. Lys showed a concentration-dependent changes in lag time that at the intermediate concentrations (0.4 and 0.6 M) the lag time was longer whereas at 0.2, 0.8 and 1.0 M it was similar to the time observed in the presence of other amino acids. Further, the elongation rate was increased in the presence of Glu, Asp and Lys. The rate was slightly decreased at the intermediate concentrations of Glu (0.4 and 0.6 M) and Lys (0.4 M) whereas in Asp the rate was similar at all the concentrations. Arg inhibited fibril formation at the concentrations higher than 0.2 M (Fig. 3.4D).

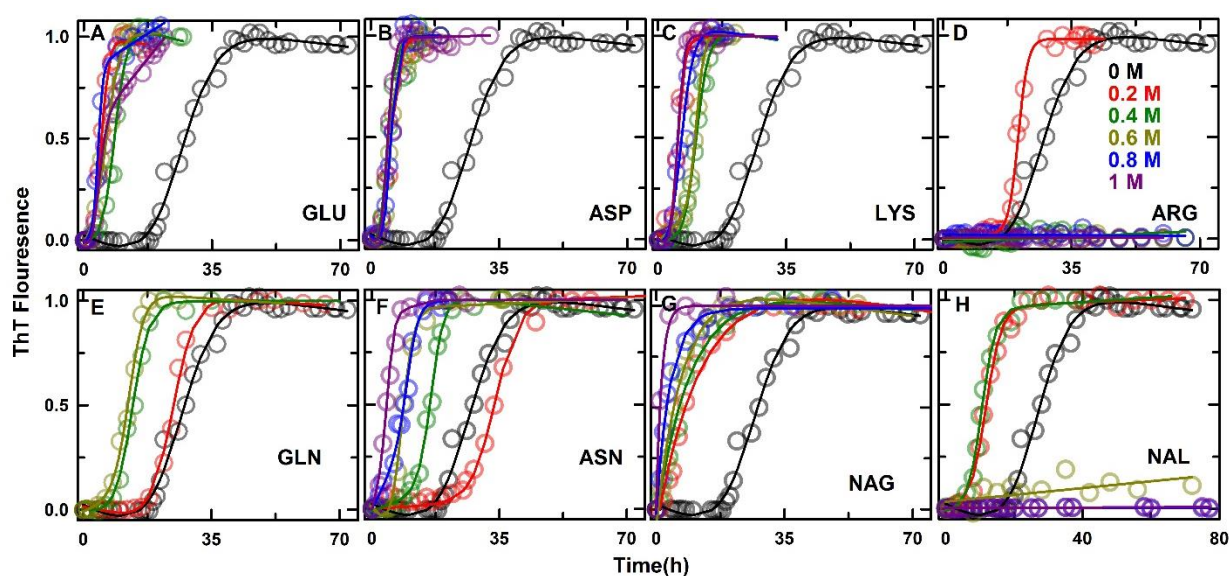


Fig. 3.4. Fibrillation kinetics of α -syn was monitored using ThT fluorescence in the presence of varying concentrations of amino acids and their derivatives, (A) L- glutamate, (B) L-aspartate, (C) L-arginine, (D) L-lysine, (E) L-glutamine, (F) L-asparagine (G) *N*-acetyl glutamate and (H) *N*-acetyl lysine. The circles represent the normalized fluorescence and the solid lines are data-fit using Eqn. 3.1 or Eqn. 3.2 for lag-dependent or lag-independent kinetics, respectively. The presence of different concentrations of the amino acids are represented as black (0 M), red (0.2 M), green (0.4 M), dark yellow (0.6 M), blue (0.8 M), and purple (1.0 M).

In 0.2 M of Arg, the lag time was not significantly altered whereas the elongation rate was increased that reduced the overall fibril formation time (Fig. 3.5A, C and E). The polar amino acids, Gln and Asn exhibited a concentration-dependent effect on α -syn fibrillation (Fig.3.4E and F, and Fig.3.5B and D). As the concentration of these amino acid was increased, the lag time decreased except in 0.2 M of Asn. Experiments could not be performed in the presence of 0.8 and 1.0 M of Gln due the poor solubility of the amino acid. In the presence of NAG, α -syn showed a

lag-independent fibrillation kinetics. When 0.2 M of NAG was added, the fibrillation rate was reduced, however, further increase in the concentration of NAG accelerated the fibrillation. In 1.0 M of NAG, the rate was higher than in buffer. At the concentrations ≤ 0.4 M, NAL induced a faster fibrillation with a short lag phase whereas at higher concentration it inhibited the fibrillation.

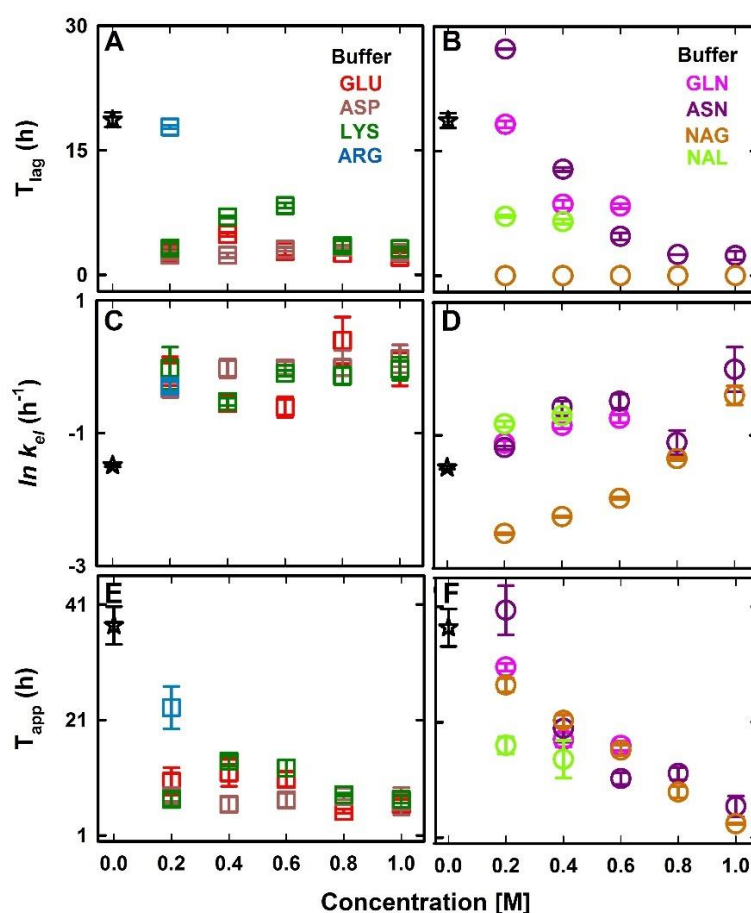


Fig. 3.5. (A and B) Lag time, (C and D) rate of elongation, and apparent fibrillation time to form fibrils (E and F) by α -syn in different concentrations of the amino acids evaluated from Fig. 3.4. Left panels (A, C, and E): α -syn in the presence of Glu (red), Asp (brown), Lys (green), and Arg (blue). Right panels: (B, D and F): α -syn in the presence of Gln (pink), Asn (violet), NAG (orange) and NAL (light green). α -Syn without any cosolvent is shown as black star in every panel.

Multiple cryo-electron microscopic structures are reported for α -syn fibrils in the protein database (www.rcsb.org).^[88,90,208–211] Fibril formation in the presence of amino acids were analyzed with TEM images as described in section 3.3.7. In the absence of any cosolvent, the

Chapter 3

fibrils were thin and long (Fig. 3.3C) whereas fibrils formed in the presence of cosolvents were structurally different (Fig 3.6). In Asp, α -Syn fibrils were thicker and showed end-to-end elongation whereas in Lys they were smaller. In Asn, the protein formed thin fibrils that were entangled. In 1.0 M of Arg no fibril-like structure was detected. The fibrils in the presence of NAG were thin as the fibrils formed in buffer, but shorter in length. α -syn formed thicker fibrils in the presence of NAL with the length similar to that formed in NAG.

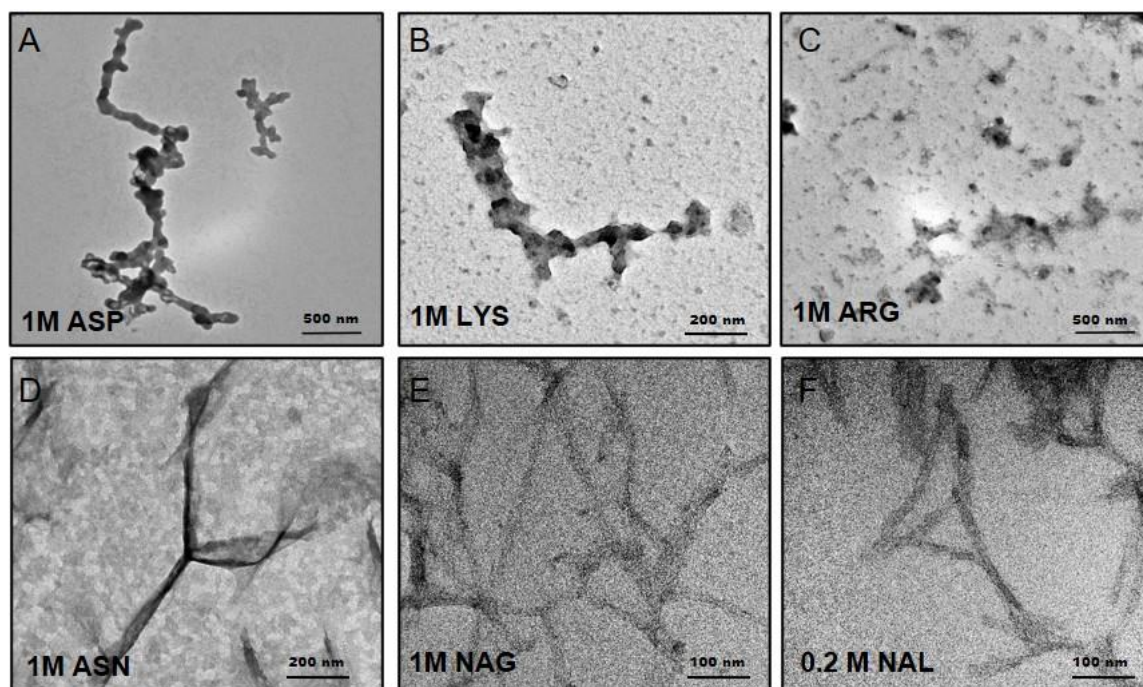


Fig. 3.6. Representative TEM images of α -syn fibrils in the presence of (A) 1 M Asp, (B) 1 M Lys, (C) 1 M Arg (no fibril is detectable), (D) 1 M Asn, (E) 1 M NAG and (F) 0.2 M NAL.

3.5 Discussion

The inherent characteristic of α -syn as a natively unfolded protein presents a significant challenge in understanding the interaction of small molecules that could potentially modulate its aggregation. Many direct therapeutic strategies targeting α -syn aggregation based on the sequence were developed. However, due to the lack of their environmental condition and conformational transition, these were not efficient.^[212] Therefore, it is foremost necessity to examine the effect of small compatible organic molecules on the fibrillar propensity of α -syn to design effective inhibitors against fibrillation.

3.5.1 Amino acid osmolytes

Fragniere *et.al.* reported that under hyperosmotic stress α -syn forms aggregates that are toxic to the cells.^[96] Though this feature of neural IDPs forming toxic aggregates is unique for α -syn, they could not understand how these aggregates were formed in specific cells (Neuro-2a (N2a) cells) and why. Since different molecules such as, sugar, salt, and alcohol were used to induce osmotic stress in cells the formation of toxic aggregates of α -syn is proposed to be cell-dependent and not due to the osmolytes.^[96] However, cell volume restoration mechanisms vary among different cells. The amino acids and their derivatives are found to be major organic osmolytes released to maintain the cell volume of neural cells.^[97,106,107,196] Therefore, the effect of natural amino acids, Lys, Arg, Glu, Asp, Gln and Asn on the fibrillation of α -syn is examined in this study.

A decrease in NAA (most abundant osmoregulatory amino acid derivative in neural cells) concentration is a neural biomarker for many neural diseases associated with neuron loss or dysfunction.^[107] Though the exact function of NAA is not known, it is found to be a precursor of NAAG (N-acetylaspartylglutamate) along with NAG^[213] and it can delay or inhibit the fibrillation of protein.^[107] Unfortunately, NAA was not soluble in water at high concentration, thus limiting our *in vitro* studies. In the present study, only NAG and NAL were chosen to understand the changes induced by the acetylated amino acids on α -syn fibrillation.

3.5.2 Fibrillation of α -syn in cosolvents

The fibrillation mechanism follows two pathways as described in section 3.2.2, lag-dependent and lag-independent. Lag-dependent fibrillation consists of protofibrils formation at the beginning, and these protofibrils are more toxic to the cells than the mature fibrils.^[214,215] Protofibrils can even induce aggregation of different other proteins present in the environment. In synucleinopathy, many heteropolymer aggregates were reported due to the high cross-seeding efficiency of α -syn.^[216–219] Therefore, it is important to understand how the selected osmolytes can modulate the α -syn fibrillation mechanism.

Various clinical and epidemiological reports suggest that occupational exposure to metal ions such as manganese, copper, lead, iron, mercury, zinc, and aluminum increase the risk of Parkinson's disease.^[220,221] It is shown by Uversky and co-workers that in the presence of all these ions α -syn followed faster fibrillation kinetics with and without lag phase.^[222,223] Even oxidized α -

Chapter 3

syn which does not form fibril, exhibits faster kinetics in presence of Pb^{2+} and Zn^{2+} without lag phase, whereas Al^{3+} induces shorter lag phase.^[222,223] Lower concentration of organic solvents favour partially folded or unfolded structures of α -syn leading to the fast fibrillation with and without lag phase. When concentration of these organic solvents is increased, it induces α -helical or β -strand structures of α -syn which inhibits fibrillation by forming amorphous aggregates.^[224] However, in 40% of TFE and 30% of HFIP α -syn doesn't form any aggregates due to its stable helical conformation.^[224] Osmolytes such as trehalose^[225], glycerol and lower concentrations of TMAO^[194] induce faster fibrillation of α -syn with shortened lag-phase, whereas higher concentrations of urea^[226] and TMAO^[194] inhibit α -syn fibrillation.

The charged amino acids, except Arg, reduce the lag time and increase the elongation rate. This combined effect results in faster fibrillation of the protein. This suggest that the amino acids might interact with the protein at monomeric and protofibril state as well. Higher concentration of polymers such as protein, polysaccharides and polyethylene glycol has also been reported to induce α -syn fibrillation with drastic decrease in the lag time.^[227–229] The concertation-dependent change on the effect of these amino acids is less (Fig. 3.5A and B). This suggests that the interaction between α -syn and these amino acids could be limited to specific sites such that any increased concentration does not affect the fibrillation pathway. A similar effect is reported by Ghosh et.al^[203] that arginine retards the fibrillation, glutamic acid accelerates fibrillation, lysine slightly accelerates α -syn fibrillation. However, their studies are limited to low concentrations that is up to 0.1 M.

Though polar amino acids, Gln and Asn also accelerate the fibril formation by altering both lag time and elongation rate, they show a concertation-dependent effect (Fig. 3.6B and D). The interaction sites of these amino acids may be more on the protein surface which might lead to the concentration-dependent effect on the fibrillation. Among the *N*-acetylated amino acids, NAG induces lag-independent fibril formation at all the concentrations whereas NAL shows an inhibitory effect at the concentrations above 0.4 M. Interestingly, the fibrillation time is faster in the presence of lower concentrations of NAL. This could be due to the change in interactions between the protein and NAL at higher concentrations. To further understand the residue-level interactions of α -syn with the amino acids, molecular dynamics simulation of the protein in the

presence of these amino acids were performed and the results are presented in the following Chapter.

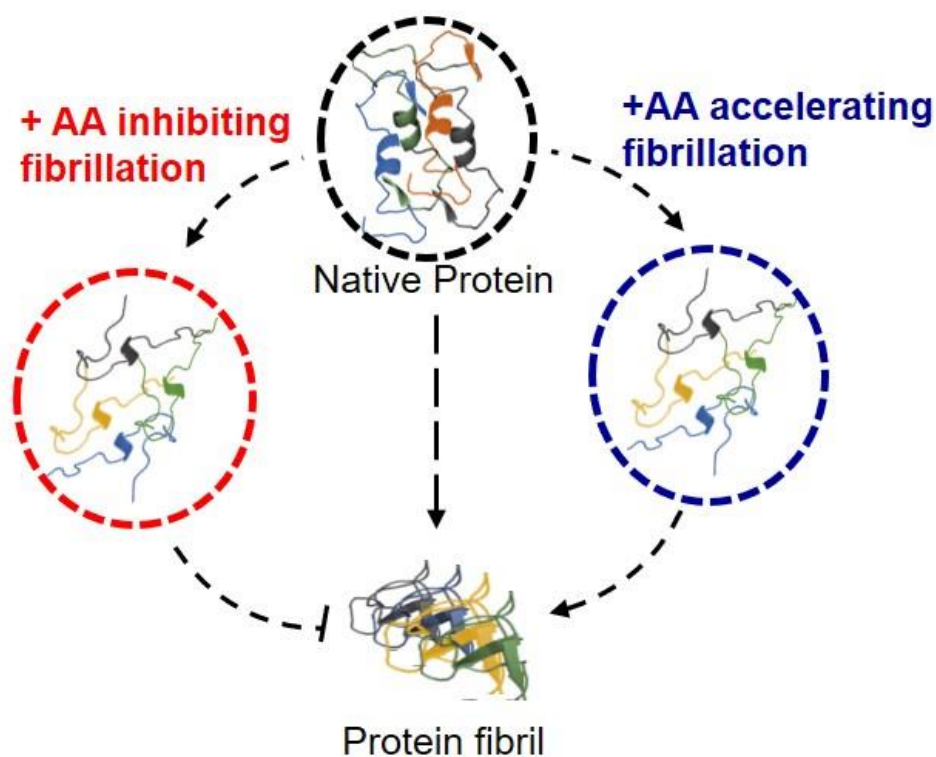
3.6 Conclusion

Osmolytes are the compatible organic molecules found in high concentration in the cells under different stress conditions and are released to protect proteins inside the cell. Therefore, it is important to investigate the concentration-based effects of naturally occurring osmolytes to understand their functional role. Effects of charged and polar amino acids, and their derivatives on the fibrillation kinetics of α -syn was probed with ThT-based fluorescence assay. It is observed that Gln, Asn, and NAG show a concentration-dependent effect on the fibrillation of α -syn. Arg and NAL promote the fibrillation at lower concentrations, whereas they inhibit at higher concentrations. Glu, Asp, and Lys accelerate the fibril formation and show little change with the concentration. The amino acids promoting the fibrillation might interact with the protein in its monomeric and protofibril form; thus, altering the lag time and the elongation rate. Further, the morphology of the fibrils formed in the presence of the amino acids are different from the fibrils formed in buffer.

Chapter 3

Chapter 4

Structural insights of the interaction between α -synuclein and amino acid osmolytes



Chapter 4

4.1 Abstract

Osmolytes are naturally occurring compounds that interact with proteins through preferential interaction leading to distinct effects on protein structure and function. α -Syn is an intrinsically disordered protein associated with synucleopathies like Parkinson's disease. Understanding the structure and fibril formation of α -syn in the presence of natural osmolytes is crucial for therapeutic development. The previous chapter examined the fibrillation process of α -syn in the presence of selected AAs. Here, we performed conventional molecular dynamic (MD) simulations of α -syn in the presence of 0.5 M of Glu (negatively charged), Lys and Arg (positively charged), Gln (amide side chain), and NAG and NAL (*N*-acetylated amino acid) representing different effects on the kinetics such as lag-dependent accelerated fibrillation, lag-independent accelerated fibrillation and inhibition of fibrillation. Analysis of the simulation trajectories revealed that AAs promote interdomain contacts between N-terminal (NTR) and fibril core (FC) regions, while the lack of intradomain interactions within the residues of FC accelerates fibrillation. The AAs interacting with the protein through non-coulombic interactions reduce the fibrillation time by shortening the lag phase. In contrast, AAs having preferential binding with α -syn through strong electrostatic interactions disrupt the intradomain contacts and favor the extended conformation of the proteins exclude water from the protein surface and inhibit fibril formation.

Chapter 4

4.2 Introduction

Monomeric α -synuclein (α -syn), a peptide consisting of 140 amino acids is a soluble protein and plays a crucial role in the transportation of vesicles within neurons^[230]. However, under certain stress conditions like hyperosmotic stress in neurons, it has the ability to aggregate into large and insoluble amyloid fibrils.^[96] These fibrils are linked to the degeneration of substantia nigra neurons and the development of Parkinson's disease.^[85,231] α -syn is divided into three distinct regions, N-terminal, fibrillar core and C-terminal. It also consists of nine hydrophobic regions (HPRs) and six conserved repeat domains (RTD) as described in Fig 1.4. The long-range intradomain interactions prevent the aggregation of α -syn through electrostatic and hydrophobic contacts.^[91,125,232] α -Syn exhibits a dynamic ensemble of structures due to the presence of multiple local minima separated by low energy barriers in its energy landscape. This unique characteristic of α -syn enables susceptibility to the cellular environment^[233]. By altering the surrounded solvent conditions such as temperature, pH, and ionic strength, the balance between favorable and unfavorable interactions within the protein can be disrupted.^[198,234,235] Further, the addition of small molecular osmolytes, crowding agents and metal ions can significantly impact the protein conformation. Osmolytes are small organic molecules that help in regulating osmotic pressure and protect proteins from denaturation under stress conditions.^[98,101] They modify the protein's stability by affecting the solvation properties or altering the strength of intermolecular interactions.^[236] Crowding agents simulate the crowded environment within cells, where proteins naturally reside, and can influence the protein stability by changing the effective concentration and molecular interactions.^[237] In presence of crowders and also in divalent and trivalent metals ions, α -syn α -syn shows faster fibrils formation.

4.2.1 Osmolytes

Osmolytes are a diverse group of naturally occurring compounds that interact with proteins and can exert distinct effects on their structure and function. The different classes of osmolytes are polyols, trimethylamine N-oxide (TMAO), sugars, and amino acids and their derivatives.^[99,193,238–240] Numerous computational approaches have been employed to investigate the impact of osmolytes on the folding and fibrillation of proteins at the atomic-level.^[102,241–245] Previous reports have proposed that certain osmolytes have the propensity to accumulate in the proximity of

proteins as a result of direct interactions with the unfolded protein states. This phenomenon, known as preferential interaction (Fig 4.1), is found to contribute to the enthalpic energy required for protein unfolding.^[246] Polyol molecules form clusters around proteins at the distance of 4 Å from the surface and encourages the formation of highly ordered water molecules within 2 Å from the protein surface resulting in preferential hydration (Fig 4.1). Polyols decrease water entropy in the first hydration shell and more significant effects are observed as the polyol size increases. These interactions indirectly protect the protein from thermal unfolding by offering a shielding effect.^[247] Certain osmolytes act as denaturants by promoting protein unfolding through van der Waals interactions.^[103,248–250] Conversely, specific osmolytes, such as trimethylamine N-oxide, act as protectants, aiding proteins to fold into a stable native conformations through electrostatic interactions.^[103,249–251] In case of α -syn, protectant osmolytes play a role in promoting the adoption of an intrinsically disordered conformation which is regarded as a stable native state for this protein.^[252,253]

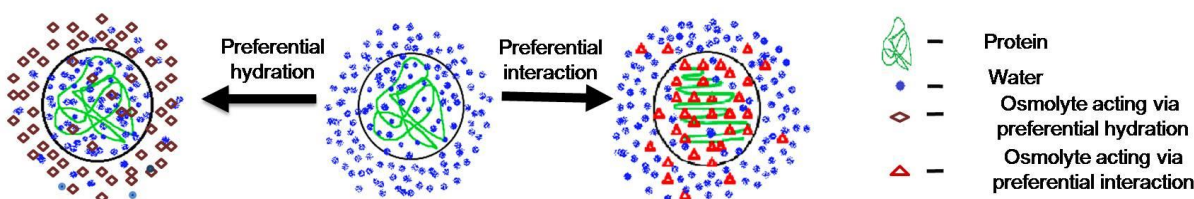


Fig. 4.1. Effect of osmolytes: Stabilizing osmolytes (brown diamonds) promote the close-packing of ordered water molecules around the protein (green wire) with minimal direct interaction between the osmolyte and the protein known as preferential hydration. Destabilizing osmolytes (red triangle) exclude water from the protein's surface and form clusters of osmolytes that have direct interactions with the protein referred as preferential interaction.

4.2.1 Osmolytes in protein aggregation

During the process of protein aggregation, initially misfolded or intrinsically disordered proteins come together to form oligomers. These oligomers can subsequently grow and organize into well-structured aggregates called amyloid fibrils, characterized by their abundance of β -sheet structures (as described in section 1.3).^[254,255] Procuring insights into the mechanisms of protein aggregation and the formation of amyloid fibrils is essential for comprehending the development of neurodegenerative diseases and devising therapeutic interventions. In the realm of protein aggregation, molecular dynamics (MD) simulations is a valuable tool for investigating the

Chapter 4

interactions and assembly of misfolded or intrinsically disordered proteins into oligomers and fibrils.^[256,257] Through the application of fundamental principles of physics and the utilization of force fields, molecular dynamics simulations offer valuable insights into the dynamic behavior, folding, and unfolding of proteins, along with their interactions with other molecules.

Notably, studies have reported strong interaction between methyl group of TMAO and the aromatic ring of Tyr residue in a short peptide sequence NNQQNY in its protofibril state. However, due to decrease in the accessibility of Tyr in the folded or aggregated states, TMAO indirectly interacts with the amide backbone Asn and Gln residues and destabilizes these states. Thus, addition of TMAO promotes the protofibril state of the peptide.^[258] Extensive MD simulation studies have been conducted on intrinsically disordered proteins, including A β protein, human amylin, tau protein, α -synuclein, and others.^[259] Previous investigations have examined the impact of urea and TMAO on the conformational behavior of tau protein.^[259–261] These studies revealed a population shift in the monomeric form of tau protein as a result of the influence of urea and TMAO.^[259–261] Preferential interaction coefficient calculations are used to examine the hydration properties of α -synuclein. The results demonstrate that urea causes a significant displacement of water molecules from the protein surface where the displacement of water is less in the presence of TMAO. This disparity in behavior led to the formation of distinct extended and compact conformational ensembles in the presence of urea and TMAO, respectively.^[253] By comparing the simulation results with experimental data, MD simulations can enhance our understanding of protein aggregation mechanisms at the atomic-level. This approach serves to complement and refine our knowledge. In the previous chapter, the revealed the inhibition of the fibrillation process in the presence of neural-osmoregulatory amino acids. To shed more light on the structural characteristics of α -syn and the potential regulatory role of specific amino acids and its derivatives in modulating the aggregation process of α -syn, MD simulation of the protein was carried out in the presence of Glu, Lys, and Gln in which the protein followed lag-dependent accelerated fibrillation kinetics, and in NAG in which the protein exhibited lag-independent accelerated fibrillation kinetics. Also, the MD simulations were performed in the presence of Arg and NAL which inhibited the fibril formation of the protein. The structures of these osmolytes are presented in Fig. 4.2. Interactions between the NTR and CTR domains and the loss of intradomain interaction in FC were identified as key factors accelerating the fibril formation of α -syn. Additionally, the preferential binding of AAs predominately via electrostatic interactions disrupted

interdomain interactions and lead to extended protein conformations which inhibited the fibrillation.

4.3 Methods

4.3.1 Conventional molecular dynamics (MD) simulation

The monomeric α -syn structure was obtained from the Protein Data Bank (<http://www.rcsb.org/>) with PDB id: 1XQ8.^[89] MD simulations were conducted using GROMACS package version 5.1.4^[158] and the CHARMM27 force field^[157]. To create the simulation environment, the protein was positioned within a dodecahedrane box and solvated with water molecules using the TIP3P water model. The system was subjected to energy minimization using the steepest descent algorithm. Subsequently, equilibration with NVT and NPT ensemble conditions for 1 ns each at 310 K and 1 atm pressure was performed. The temperature and pressure was maintained using V-rescale thermostat^[159] and Parrinello-Rahman barostat,^[160] respectively. A non-bonded interaction cut-off of 1.2 nm was applied. Following equilibration, production simulation was done for 100 ns. To study the effect of the selected osmolytes, Glu, Lys, Gln, Arg, NAG and NAL, a representative concentration of 0.5 M was used. To attain 0.5 M of concentration, required water molecules were replaced with cosolvents in the simulation box and the system was charge neutralized with appropriate counter ions (Table. 4.1.). The ternary systems of protein-water-cosolvent were simulated with the same conditions mentioned for the protein-water system. The parameters for amino acids and their derivatives were obtained from CGenFF.^[262,263]

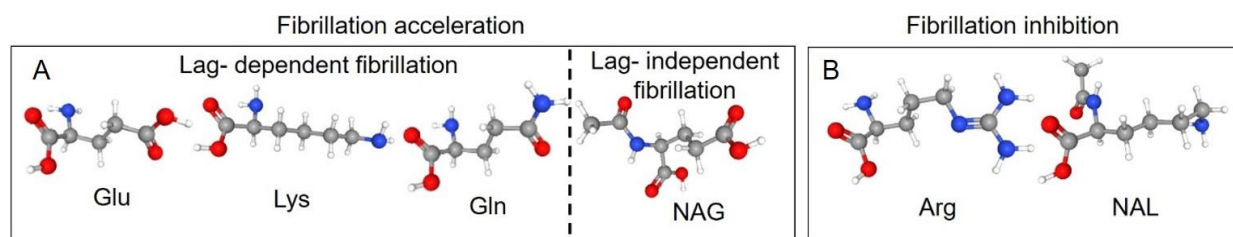


Fig. 4.2. Osmolytes selected for this study: (A) Fibrillation accelerating osmolytes: Glu (negatively charged), Lys (positively charged), and Gln (amide side chain) following lag-dependent kinetics and NAG (*N*-acetylated amino acid) following lag-independent kinetics. (B) Fibrillation inhibiting osmolytes: Arg (positively charged) and NAL (*N*-acetylated amino acid). The red, blue, gray, and white balls represent oxygen, nitrogen, carbon, and hydrogen atoms, respectively.

Chapter 4

Table 4.1 Details of the MD simulation systems.

System (with protein)	Box volume (nm ³)	Number of water molecules	Number of cosolvent molecules [†]	Number of ions
Water	3,968	1,29,436	--	9 Na ⁺
Glu	5,387	1,78,159	1,528	1,537 Na ⁺
Arg	5,387	1,77,753	1,525	1,516 Cl ⁻
Lys	5,387	1,78,277	1,525	1,516 Cl ⁻
Gln	5,387	1,80,580	1,527	9 Na ⁺
NAG	5,387	1,77,994	1,527	3,063 Na ⁺
NAL	5,387	1,81,205	1,524	9 Na ⁺

[†]-Corresponding to the concentration of 0.5 M.

4.3.2 Preliminary analysis

Root means square deviation (RMSD) of C α -atoms of the protein and solvent accessible surface area (SASA) were computed for full length (FL) and fibril core region (FC) of α -syn using gmx_rms and gmx_sasa modules in GROMACS, respectively. The RMSF (root mean square fluctuation) of residues and the radius of gyration (Rg) of FL protein were calculated using the gmx_rmsf and gmx_gyrate modules, respectively. Distances between different regions of α -syn (referred as H1, H2, and CTR in Fig 1.4) were determined using the gmx_distance module. The contact maps, showing the average distance between the residues during simulation, were obtained using the gmx_mdmat module with a distance cut-off of 1.5 nm. The radial distribution function (RDF) of water oxygen atoms around the protein's heavy atoms was evaluated using the gmx_rdf module. RDF for amino acids (AAs) around the protein were derived by considering the heavy atoms of the AA against the protein's heavy atoms. RDFs were calculated for three different regions of α -syn, NTR, FC, and CTR. The secondary structural contents were analyzed using the DSSP module in VMD.^[166] The non-covalent interaction energy between the AAs and the residues in α -syn was defined as the sum of L-J potential and coulombic interaction energy. These values were calculated using gmx_energy module after generating separate energy groups for AA and protein. All the analysis except RMSD were performed for last 60 ns of trajectories.

4.3.3 Preferential interaction

To examine the distribution of water and AAs around the protein, the number density analysis was performed using the last 60 ns of the trajectory. The gmx_select tool was utilized with distance cut-offs of 0.28 nm, 0.4 nm, and 0.6 nm from the protein surface defining them as the first, second and third hydration shells, respectively. This analysis aimed to gain insights into the spatial arrangement of cosolvents and water molecules around the protein. To investigate the preferential interaction of the cosolvent three distinct parameters were evaluated.

4.3.2.a. Hydration coefficient (χ_{hyd})

For each hydration shell, the hydration fraction was determined by calculating the ratio between the fraction of water molecules and AAs within the specified cut-off distance, which can be defined as

$$\chi_{hyd} = \frac{\langle n_w \rangle / N_w^{tot}}{\langle n_{AA} \rangle / N_{AA}^{tot}} \quad (4.1)$$

where n_w and n_{AA} are average number of water and AA molecules within the given cut-off during the last 60 ns of trajectory, while N_w^{tot} and N_{AA}^{tot} are the total number of water and AAs in the system.

4.3.2.b. Partition coefficient (K_p)

The local-bulk partition coefficient is inversely proportional to the hydration fraction^[264] which can be calculated as

$$K_p = \frac{\langle n_{AA} \rangle / N_{AA}^{tot}}{\langle n_w \rangle / N_w^{tot}} \quad (4.2)$$

where $K_p > 1$ would indicate that the accumulation of cosolvents around the protein is more compared to the bulk concentration. This analysis was conducted using the last 60 ns of the trajectory for all the three hydration shells of the protein.

Chapter 4

4.3.2.c. Preferential interaction coefficient (Γ_{P-AA})

Preferential interaction coefficient provides a quantitative measure for the selective binding or exclusion of a cosolvent relative to water molecules around a protein in a ternary system consisting of a solute (protein), solvent (water), and cosolvent. It explains the degree to which the cosolvent exhibits a preference for interacting with the protein, either by binding closely to it or by being excluded from its vicinity compared to water molecules.^[265] Preferential interaction coefficients of AAs with α -syn were calculated using following equation.^[104]

$$\Gamma_{P-AA} = \rho_{AA}(G_{P-w}(r) - G_{P-AA}(r)) \quad (4.3)$$

where ρ_{AA} is number density of AAs in bulk solution and was evaluated as number of AA molecules over the volume of simulation box. Here, $G_{P-w}(r)$ and $G_{P-AA}(r)$ are Kirkwood-buff integrals (KB integrals) for protein-water and protein-AA, respectively that was calculated at every 0.02 nm cutoff as,

$$G(r) = \int_0^r 4\pi r^2 (g(r) - 1) dr \quad (4.4)$$

where, $(g(r))$ refers to the integrated RDF values at the intervals of 0.02 nm up to the cut-off distance of 1 nm. A positive value of Γ_{P-AA} indicates the presence of favorable binding between AAs and the protein surface.

4.3.4 Cosolvent clustering

In order to gain insights into the differences in the distribution of AAs around the protein, the clusters of AAs formed on the protein surface was analyzed. The last 60 ns of the trajectories consisting of 3000 frames were selected for each α -syn-AA simulation. A distance matrix was generated by calculating the distance between C α -atom of each cosolvent in the simulation system. The distance matrix was converted into an adjacency matrix with a condition that the distance between two amino acids was less than or equal to 0.85 nm then the entry is one, otherwise zero. This was done using reshape2 and MASS libraries in R-studio. Based on this, interaction network was constructed such that each amino acid was considered as a node and the AAs within 0.85 nm distance from each other were connected with an edge. Using the igraph library, the edges were

categorized into different clusters of AAs. Subsequently, these clusters were categorized into four groups based on the number of AAs in the cluster. Small, medium, and large clusters were defined as having 4 to 10, 11 to 25, and 26 to 50 AAs, respectively. The largest clusters, which comprised of 50-150 AAs were referred as macro-clusters. To reduce the dimension of betweenness centrality of nodes in large and macro clusters, nodes with at least three edges were considered to be a part of a cluster.

4.4 Results

4.4.1. Global analysis

In order to gain a comprehensive understanding of the overall conformational changes in α -syn throughout the trajectory, RMSD of C α atoms were calculated using the respective initial conformations as reference (Fig. 4.3 A-D). The RMSD analysis revealed that α -syn in water exhibited an average deviation of 3.32 ± 0.76 nm for the full-length (FL) protein while the fibril core (FC) region alone showed an average RMSD value of 1.64 ± 0.37 nm. When 0.5 M of different cosolvents such as Glu, Arg, Lys, Gln, NAG, and NAL were present in the system, the RMSD values were reduced and the average RMSDs over the complete trajectory were observed to be 2.82 ± 0.60 nm, 3.09 ± 0.78 nm, 2.89 ± 0.43 nm, 2.81 ± 0.63 nm, 2.26 ± 0.41 nm, and 2.78 ± 0.66 nm, respectively. The RMSD values of FC region in these cosolvents were 0.75 ± 0.08 nm, 0.77 ± 0.13 nm, 0.79 ± 0.11 nm, 0.62 ± 0.11 nm, 0.71 ± 0.08 nm and 0.79 ± 0.15 nm, respectively. Further, to identify the regions of α -syn with significant fluctuations, we calculated the RMSF of C α atom of each residue for the FL protein (Fig. 4.3 E and F). It was observed that α -syn in the presence of Arg, Gln, and NAL exhibited the highest fluctuations in the N-terminal region (NTR). In contrast, fluctuations around FC region were higher in the presence of Arg, NAG, and NAL specifically at the hinge region between NTR-FC or FC-CTR. The C-terminal region (CTR) consistently showed the highest fluctuations in all systems, since it naturally lacks any ordered structure.^[266] Moreover, in the absence of any cosolvents, α -syn showed larger fluctuation in NTR and CTR. Also, FC of α -syn displayed more fluctuation in the HPR3 and HPR5 regions, as described in Fig. 1.4.

Chapter 4

The solvent accessible surface area of α -syn in all the systems was analyzed for the FL and FC regions of the protein (Fig. 4.4). The FL-SASA values revealed that NAG had the highest solvent accessibility, measured to be $143.24 \pm 2.51 \text{ nm}^2$. This was followed by NAL and Arg, with the values of $138.54 \pm 3.14 \text{ nm}^2$ and $137.51 \pm 3.85 \text{ nm}^2$, respectively. On the other hand, Glu, Lys, and Gln exhibited similar FL-SASA values as $111.20 \pm 3.67 \text{ nm}^2$, $119.43 \pm 4.67 \text{ nm}^2$, and $113.12 \pm 2.88 \text{ nm}^2$, respectively. The SASA of the protein in the absence of any cosolvent was $119.38 \pm 5.39 \text{ nm}^2$.

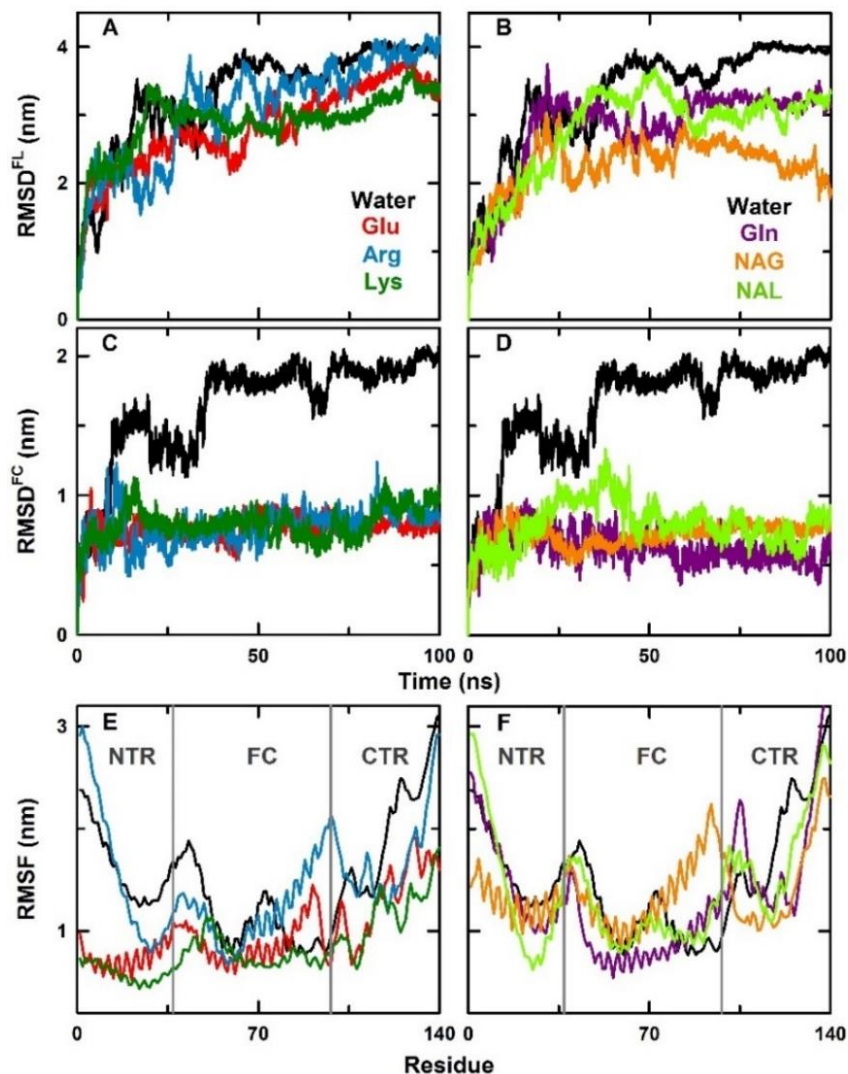


Fig. 4.3. Root mean square deviation (RMSD) of $\text{C}\alpha$ atoms of full length protein (A-B), only fibril core residues, V37-D98 (C-D). Root mean square fluctuation (RMSF) of FL (E-F) of α -syn in the absence (black) and in the presence of 0.5 M of cosolvent. Right side panels represent the presence of Glu (red), Arg (blue) and Lys (green), while left side panels represent Gln (purple), NAG (orange) and NAL (light green). The vertical lines in panels, E and F, separates α -syn into N-terminal region (NTR, residues 1-36), fibrillar core (FC, residues 37-98), and C-terminal region (CTR, residues 99-140).

Structural insights of the α -syn-AA interaction.

The FC-SASA values of α -syn in the presence of Glu, Lys, Gln, NAG, and NAL displayed similar values, ranging from $61.86 \pm 1.07 \text{ nm}^2$ to $63.06 \pm 1.29 \text{ nm}^2$. Arg exhibited slightly lower SASA of FC region with a value of $58.53 \pm 2.21 \text{ nm}^2$, whereas in the absence of cosolvents the FC region of α -syn showed the lowest SASA value, $54.70 \pm 1.94 \text{ nm}^2$.

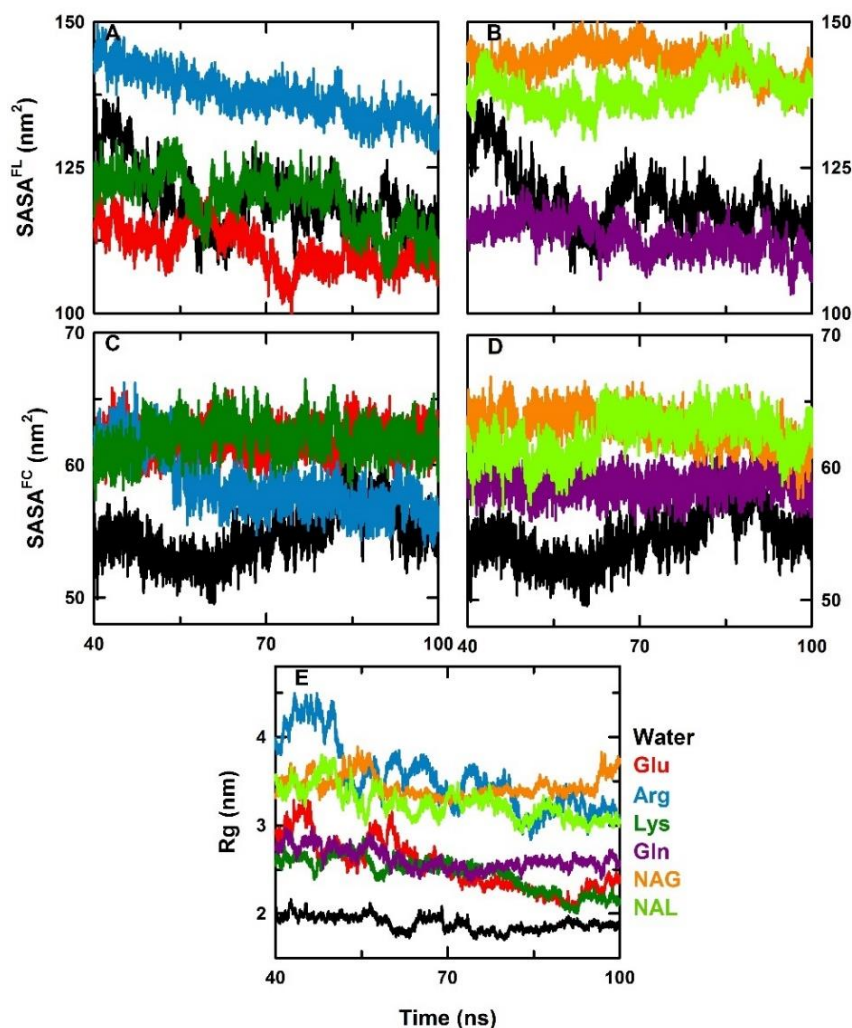


Fig. 4.4. Solvent accessible surface area (SASA) of FL (A-B) and only FC region (C-D), and radius of gyration, Rg (E) of α -syn in water (black) and in 0.5 M of amino acids, Glu (red), Arg (blue) and Lys (green), Gln (purple), NAG (orange) and NAL (light green).

The higher SASA of α -syn in the presence of Arg, NAG, and NAL complemented with the higher values of radius of gyration (Rg) of the protein (Fig. 4.4 E) in the presence of Arg ($3.52 \pm 0.38 \text{ nm}$), NAG ($3.44 \pm 0.12 \text{ nm}$), and NAL ($3.24 \pm 0.17 \text{ nm}$). In contrast, the absence of cosolvent led to the lowest radius of gyration (Rg) value, $1.89 \pm 0.081 \text{ nm}$. This indicated a more compact

Chapter 4

conformation of the protein due to the limited solvent accessibility of the FC region. When α -syn was exposed to 0.5 M concentrations of Glu, Lys, and Gln, it had slightly increased Rg values, ranging from 2.45 ± 0.19 nm to 2.62 ± 0.12 nm.

4.4.2. Surface solvation of α -syn

Intrinsically disordered proteins (IDPs) are sensitive to their surrounding solvent environment^[118,119]. This was evident in our initial analysis where significant differences in RMSD, Rg, and SASA were observed. These variations can be attributed to the differences in the interactions of solvent and cosolvents in the hydration shells surrounding α -syn, thus, leading to substantial alterations in the protein's conformation. Thus, the distribution of water and amino acids around the protein was analyzed using different parameters.

4.4.2.a Radial distribution function

RDF has been widely used by computational biologists to gain insights into atomic interactions between protein-water and protein-cosolvent systems^[267]. The RDF illustrates how the density of surrounding molecules changes as a function of distance. An increase in the RDF peak indicates a higher number density of neighboring molecules. Initially, the RDF of water molecules around the protein in the absence of any cosolvent was analyzed (Fig. 4.5 A). The RDF plot showed the first maximum at 0.28 nm. This is commonly associated with tetrahedrally oriented water molecules around the heavy atoms of proteins.^[268] In the presence of various AAs, the RDF values exhibited distinct peaks corresponding to the first, second, and third hydration spheres at 0.28 nm, 0.40 nm, and 0.6 nm, respectively that are referred as shell1, shell2, and shell3. The RDF values for water in all three hydration shells were highest in the protein-water system compared to systems containing AAs. The RDF plots for the cosolvents around the FL-protein, NTR, FC region and CTR were also evaluated (Fig. 4.5 C-J). The RDF values were particularly elevated in the presence of Arg, NAG, and NAL. The higher RDF values in the presence of Arg were mainly due to its increased presence around CTR region, whereas NAG and NAL exhibited similar distribution

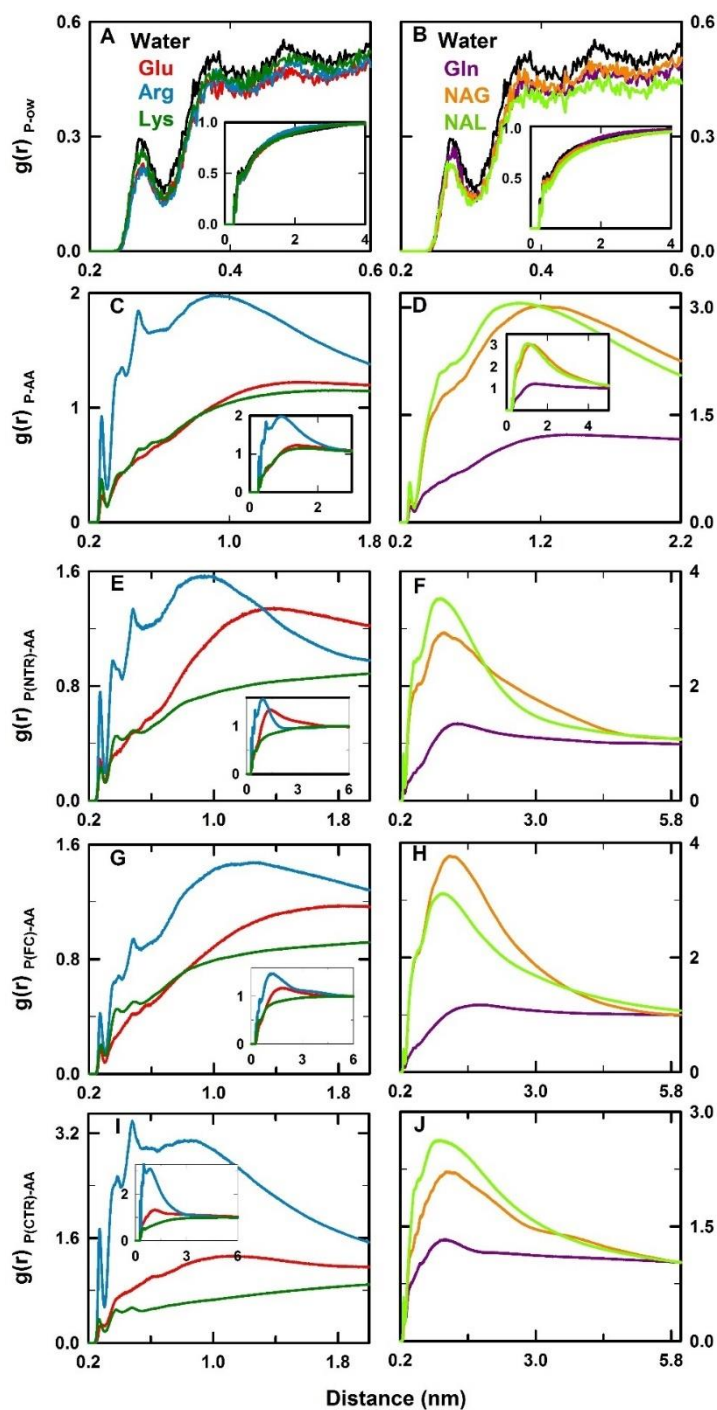


Fig. 4.5 Radial distribution function of (A-B) oxygen atoms of water around the heavy atoms of α -syn, (C-D) heavy atoms of cosolvent around the heavy atoms of the protein. The RDF of amino acids around NTR, FC, and CTR regions of α -syn are shown in (E-F), (G-H), and (I-J), respectively. Right side panels represent the presence of Glu (Red), Arg (blue) and Lys (green), whereas left side panels represent the presence of Gln (purple), NAG (orange) and NAL (light green). The insets show the distribution of water or cosolvent reaching the value 1 suggesting that the RDF is complete.

Chapter 4

across the protein. Lys demonstrated the least RDF value for FL α -syn, while Glu and Gln displayed similar behavior around the protein hydration shells.

4.4.2.b Distribution of water and AAs around α -syn

The RDF analysis of water and AAs revealed significant differences in the vicinity of α -syn which were evident from the changes in Rg and SASA values as well. The average number of water and AA around the protein were calculated (Fig. 4.6. A and B). Higher number of water molecules around α -syn in only water, Arg and NAG were observed. In presence of Lys, the number of water molecules were higher for the first hydration shell of the protein, but decreased with the increase in distance. However, number of AA around the protein were also significant in presence of Arg, NAG and NAL. α -Syn exhibited a range of compact to extended conformations (Fig. 4.6 E-G), which could influence the distribution of AA and water molecules around the protein.

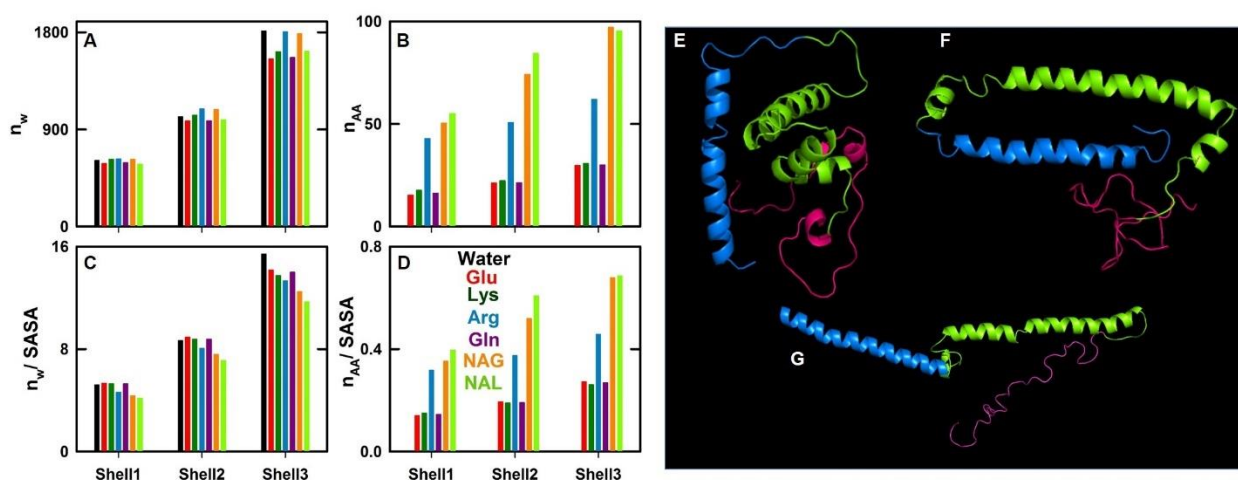


Fig. 4.6. The number of water (A) and amino acid (B) molecules in the first, second and third hydrations shell of α -syn and the number of water (C) and AA (D) molecules per solvent-exposed surface area of the protein in water (black) and in the presence of Glu (red), Arg (blue), Lys (green), Gln (purple), NAG (orange) and NAL (light green). Representative (E) compact, (F) closely-packed, and (G) fully-extended conformations of α -syn obtained from the simulation of the protein in the presence of different AAs. Blue, green and pink colours represent NTR, FC and CTR, respectively.

To evaluate the number of solvent and cosolvent molecules interacting with the surface exposed region of the protein, the molecules occupied per solvent-exposed surface area was calculated (Fig. 4.6 C and D). The number of water around the exposed surface of protein was higher for only water, Glu, Lys and Gln, whereas it decreases in presence of Arg, NAG and NAL.

Structural insights of the α -syn-AA interaction.

The decrease in the number of water molecules was complemented with the higher amount of Arg, NAG, and NAL occupied in all three solvation shells of the protein compared to Glu, Lys, and Gln molecules. This suggested that the density of cosolvents in the vicinity of α -syn varies for different amino acids and they might stabilize different conformational states in the protein.

4.4.2.c Hydration coefficient and partition coefficient

To investigate the solvation effects on α -syn, the hydration coefficient (χ_{hyd}) and partition coefficient (K_p) were calculated using Eqn. 4.1 and 4.2, respectively. These coefficients describe the proportional variation in density between the solvation shells and the bulk for water and cosolvent, respectively. They generally exhibit an inverse relationship for protein-cosolvent systems (Fig. 4.7). Hydration fraction of α -syn was found to be higher in the presence of Glu, Lys, and Gln compared to Arg which had larger χ_{hyd} values than NAG and NAL. On the other hand, the partition coefficients of the AAs showed an opposite trend with highest K_p values for Arg followed by NAG and NAL, whereas in the presence of Glu, Lys, and Gln the K_p value was reduced. As the distance from the protein increases the partition coefficient decreases indicating a stronger interaction of AAs in the first and second hydration shells (Fig. 4.7). Further, the solvent and cosolvent distribution around three distinct regions (NTR, FC, and CTR) of the protein was analyzed to understand the region-specific interactions of AAs (Fig 4.7 B-H). Glu, Lys and Gln showed higher χ_{hyd} for the amphipathic regions NTR and FC, and Arg had higher χ_{hyd} values in FC region than other regions, whereas the hydration coefficient of NAG and NAL was less across all the regions of the protein. The partition coefficient was highest for NAG and NAL in the NTR and FC regions whereas Arg showed higher K_p values for CTR. Glu, Lys and Gln showed less K_p values across the protein compared to other three AAs. In all the cases, the K_p values were found to be above one suggesting that the density of AAs near the protein surface was higher than the bulk. However, the extent of increased density around the protein significantly varied among the AAs.

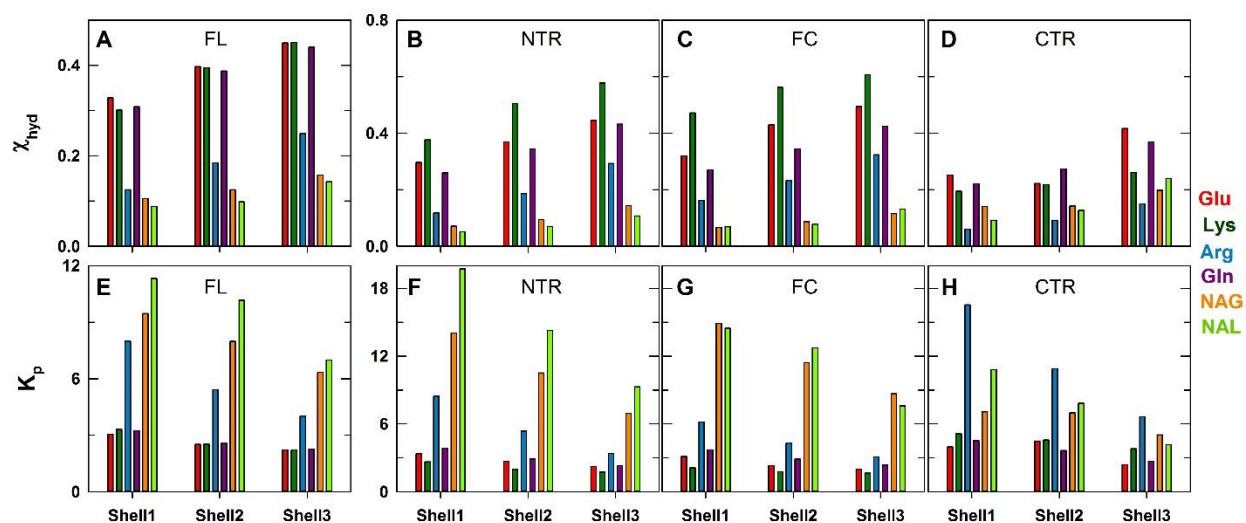


Fig.4.7 (A-D) Hydration fraction (χ_{hyd}) calculated using Eqn. 4.1 and (B) partition coefficient (K_p) calculated using Eqn. 4.2 for the first, second and third hydrations shells of the protein. The first, second and third hydration shells are considered to be at the distance of 0.28 nm, 0.4 nm and 0.6 nm from the protein surface, respectively based on the RDF plots shown in Fig. 4.5. (A and E) present χ_{hyd} and K_p for FL α -syn and (B-D and F-H) present χ_{hyd} and K_p for NTR, FC and CTR of α -syn.

4.4.2.d Preferential interaction coefficient

A low χ_{hyd} and high K_p values indicate a high affinity or direct interaction of AAs with the protein by excluding water from the surface of protein.^[104,269] This is known as preferential binding. To gain more quantitative understanding of the preferential interaction of AAs with α -syn, preferential interaction coefficient was calculated using Eqn. 4.3 (Fig 4.8 A-D). The plots demonstrated that all the AAs exhibited preferential binding to FL α -syn, with the order of preference as Arg > NAL > NAG > Lys > Glu > Gln. Lys showed preferential hydration for the NTR and FC regions, which corroborates with low K_p values and high χ_{hyd} values. However, it displayed high preferential binding with the CTR region. The AAs which inhibit the fibrillation of α -syn, Arg, and NAL, exhibited a strong preferential binding across all three regions of α -syn. For the NTR region, the order of preferential binding was NAL $\approx$ Arg > NAG > Glu > Gln, while for the FC region, it was NAG $\approx$ NAL > Arg > Gln > Glu.

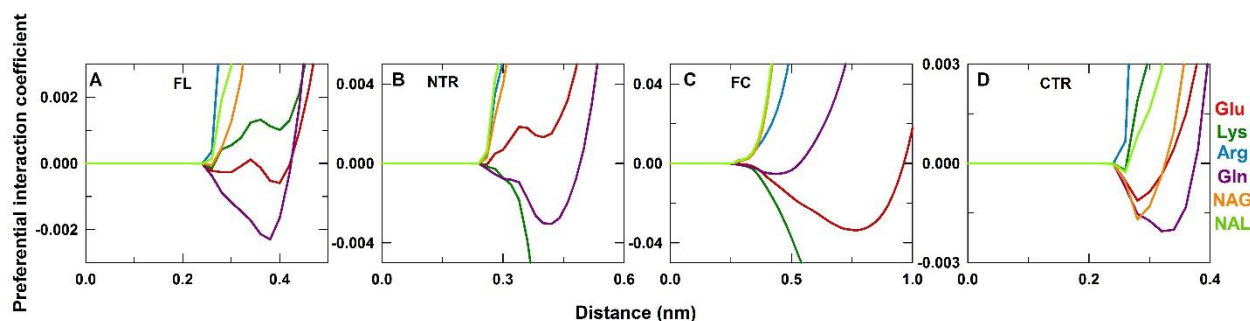


Fig.4.8. Preferential interaction coefficient for different AAs, Glu (red), Arg (blue), Lys (green), Gln (purple), NAG (orange) and NAL (light green) for (A) FL, (B) NTR, (C) FC and (D) CTR of α -syn derived from Eqn. 4.3

4.4.3 Dynamics of Cosolvent

Earlier studies on protein-osmolyte interactions have suggested formation of molecular assemblies of osmolytes on the surface of proteins.^[105,243,253,258] A significant amount of formation of clusters of AAs on the protein was noticed in the present study as well. The dynamics of these assemblies on the protein surface was examined. To distinguish the cluster formation of AAs on the protein surface from their inherent nature to form small aggregates,^[242,270] MD simulation of 0.5 M of each AA in the absence of the protein was also performed. The distribution of clusters and the number of each type of clusters, small, medium, large, and macro, were calculated (Fig. 4.9).

In the absence and presence of α -syn, the formation of small and medium-sized clusters was similar for Arg, Gln and NAL. Glu formed more number of medium clusters in the presence of α -syn by loss of small-sized clusters. The number of both small and medium clusters of NAG were reduced, but it formed a macro-sized cluster in the presence of the protein. Lys forms only small-sized clusters and the number of clusters were slightly reduced with the addition of protein. A small fraction of large clusters was observed only for NAG both in the absence and the presence of α -syn. NAL formed a macro-sized cluster only in the presence of α -syn without much changes in the number of small and medium-sized clusters. For all the amino acids, the number of small clusters are more than the medium or large clusters. An intriguing observation was the formation of a single macro-cluster by NAL comprising 65 AAs that completely covered the NTR region of α -syn (Fig. 4.9 B). Unlike NAL, NAG formed seven macro-clusters, one of which formed on the complete surface of α -syn and hinder the long-range interaction within the protein (Fig. 4.9 C).

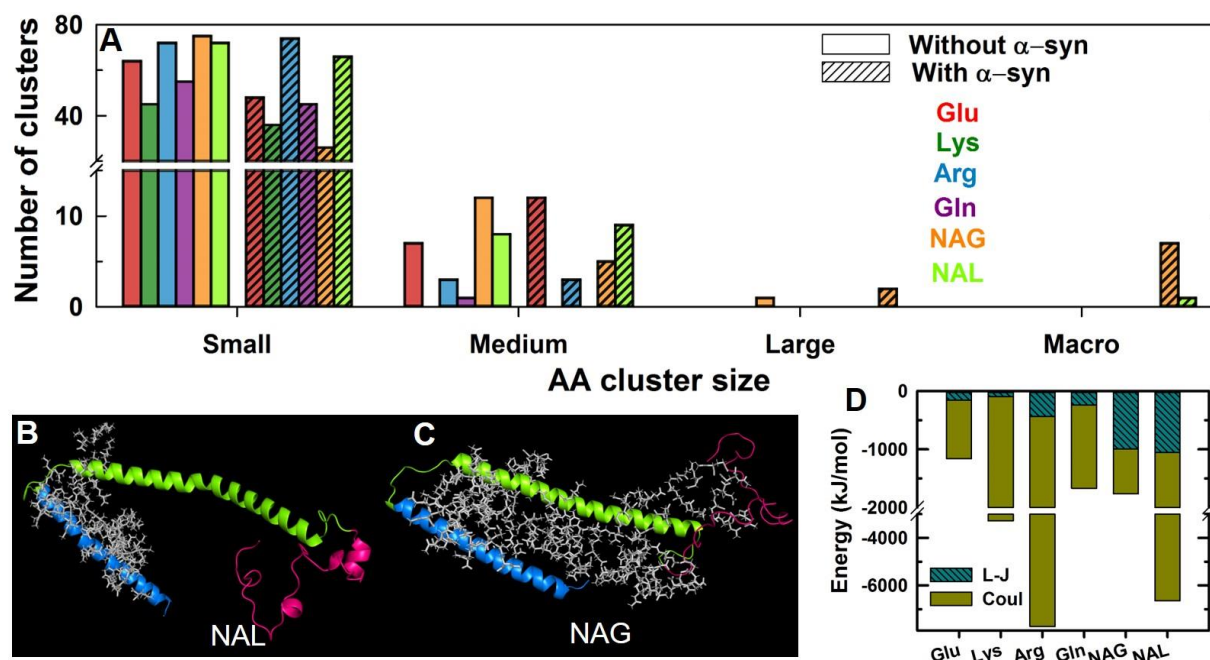


Fig.4.9. (A) Distribution of clusters formed by different AAs for system without α -syn (as empty bar) and with α -syn (as filled lined bar). Formation cluster and their distribution analysis is described in section 4.3.4. Representative macro-cluster of (B) NAL formed on NTR and (C) NAG on the FL α -syn. NTR, FC and CTR of α -syn are shown as blue, green and pink respectively and AAs are shown as grey coloured sticks. (D) Interaction energy of AA-FL α -syn where L-J potential (dark cyan) and coulombic interaction (Coul) energy (pink) contribution is presented of each system.

To analyze the energetics of interactions between the AAs and α -syn, the non-covalent interaction energies were calculated (Fig.4.9 D). Notably, the amino acids inhibiting the protein fibrillation, Arg and NAL, exhibited exceptionally high coulombic (Coul) interaction energy with α -syn compared to other AAs. Lys also displayed high coulombic energy; however, as shown in Fig. 4.8 (B-C), this can be attributed to the preferential interaction of Lys with the CTR region of α -syn rather than the FC region. In case of NAG and NAL, a significant contribution of L-J potential was observed indicating the presence of non-polar interactions. In the case of NAL, the interaction energies might be due to its high polarity and binding affinity on the FL the protein whereas NAG due to its negative charge exhibited less affinity towards CTR region (Fig. 4.8) NAG had significant interactions with NTR and FC regions that are amphipathic with KTKEGV repeat domains and formed a macro cluster holding the NTR and FC regions of α -syn.

4.4.4 Conformational characteristics of α -syn in presence of AA

4.4.4 a Contact index

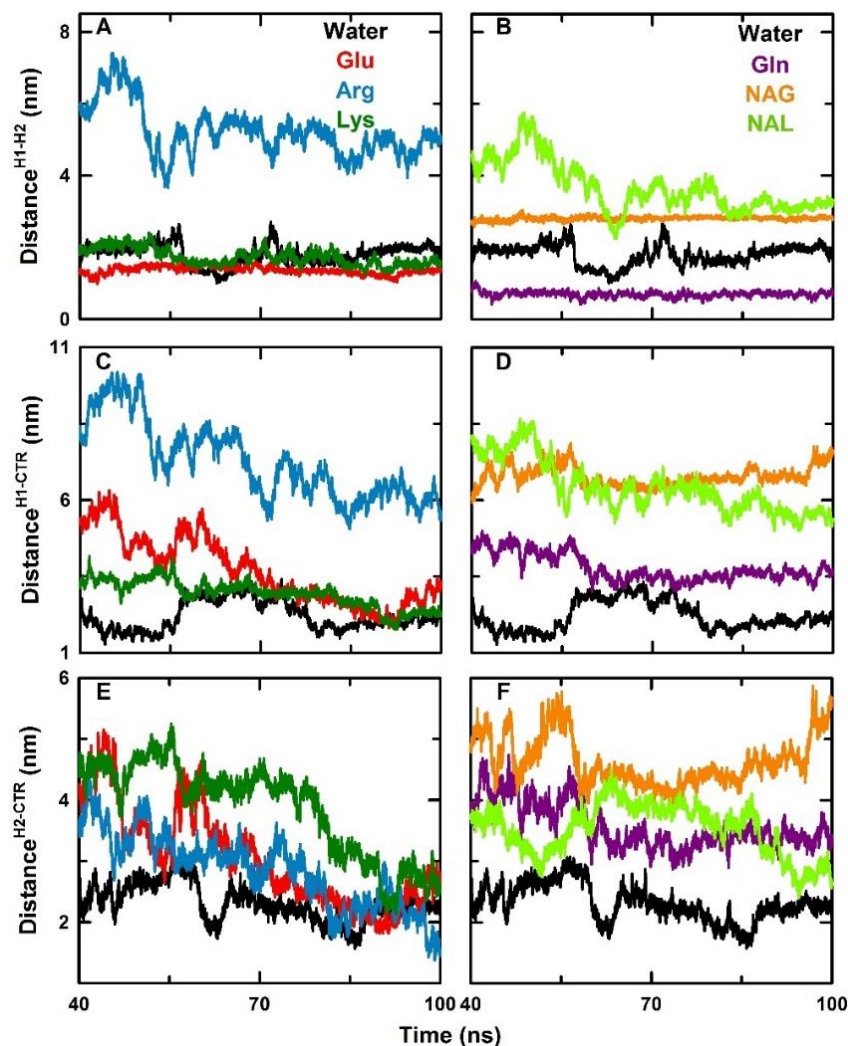


Fig. 4.10. The center of mass (COM) distance between (A-B) H1 and H2, (C-D) H1 and CTR and, (E-F) H2 and CTR calculated for the last 60 ns of the simulation trajectories. Right side panels (A, C and E) represent the presence of Glu (red), Arg (blue) and Lys (green) while left side panels (B, D and F) represent the presence of Gln (purple), NAG (orange) and NAL (light green). The black lines present the distance changes in the absence of any AA.

The interaction between the different regions of α -syn determines its fibrillation propensity. The formation of clusters by the AAs on the surface of α -syn had a significant influence on the contacts between different regions and leading to the adoption of various conformations. The interaction between helix-1 (H1: V3-G37), helix-2 (H2: T45-T92), and CTR (Q99-A140) was analyzed by measuring the distance between the center of mass (COM) of the residues in these regions. The

Chapter 4

analysis was done for the last 60 ns of the simulation trajectories (Fig. 4.10). The distance between H1 and H2 of α -syn increased in the order: Gln < Glu $\approx$ Lys $\approx$ in water < NAG < NAL < Arg. The distance between H1 and CTR of α -syn had larger fluctuations due to the disordered nature of CTR. The average distance, however, showed similar trend: Gln $\approx$ Glu $\approx$ Lys $\approx$ in water < NAL < NAG < Arg. The distance between H2 and CTR also had larger fluctuations. In the absence of AAs, α -syn exhibited compact conformations as evidenced from the minimal difference in distance between the regions. Conversely, in the presence of Arg, NAG, and NAL, the distance between the regions were higher suggesting the possibility of extended conformations. Notably, in the presence of NAG, the distance between H1 and H2 was considerably large and remained stable throughout the trajectory which can be attributed to the formation of a macro-cluster of NAG molecules between these helices (Fig. 4.9 C).

Further, the contact maps were constructed for FL protein (Fig. 4.11) that complemented the above observations. In the absence of AAs, α -syn exhibited a high number of contacts between FC and CTR, involving both intra-region and inter-region interactions. AAs that accelerated α -syn fibrillation in a lag-dependent manner, Glu, Lys, and Gln, showed stronger interactions between NTR and FC regions. However, the AAs that accelerated fibrillation in a lag-independent manner did not exhibit any significant contacts between NTR, FC and CTR of α -syn. NAL and Arg, that inhibit the fibrillation, did not show any interaction between NTR and FC, but showed possible interactions between RTD5 from FC and Q99 to Q109 of CTR.

4.4.4 b Secondary structural changes

The secondary structure analysis was conducted to investigate the impact of AA molecules on the secondary structure of α -syn. In the absence of any cosolvent, α -syn adopted a highly compact conformation with a loss of helicity in the HPR5 region of FC resulting in a separation of H2 into two helices and the disruption of the helical structures from HPR2 to HPR3 (Fig. 4.11A). A transient loss of helicity in HPR5 and HPR6 was observed in the presence of Glu and Lys whereas in the presence of Gln and NAG there were minimal changes observed in the secondary structure of α -syn. Arg and NAL induced an extended conformation of α -syn with a notable loss of helicity in the H2 region.

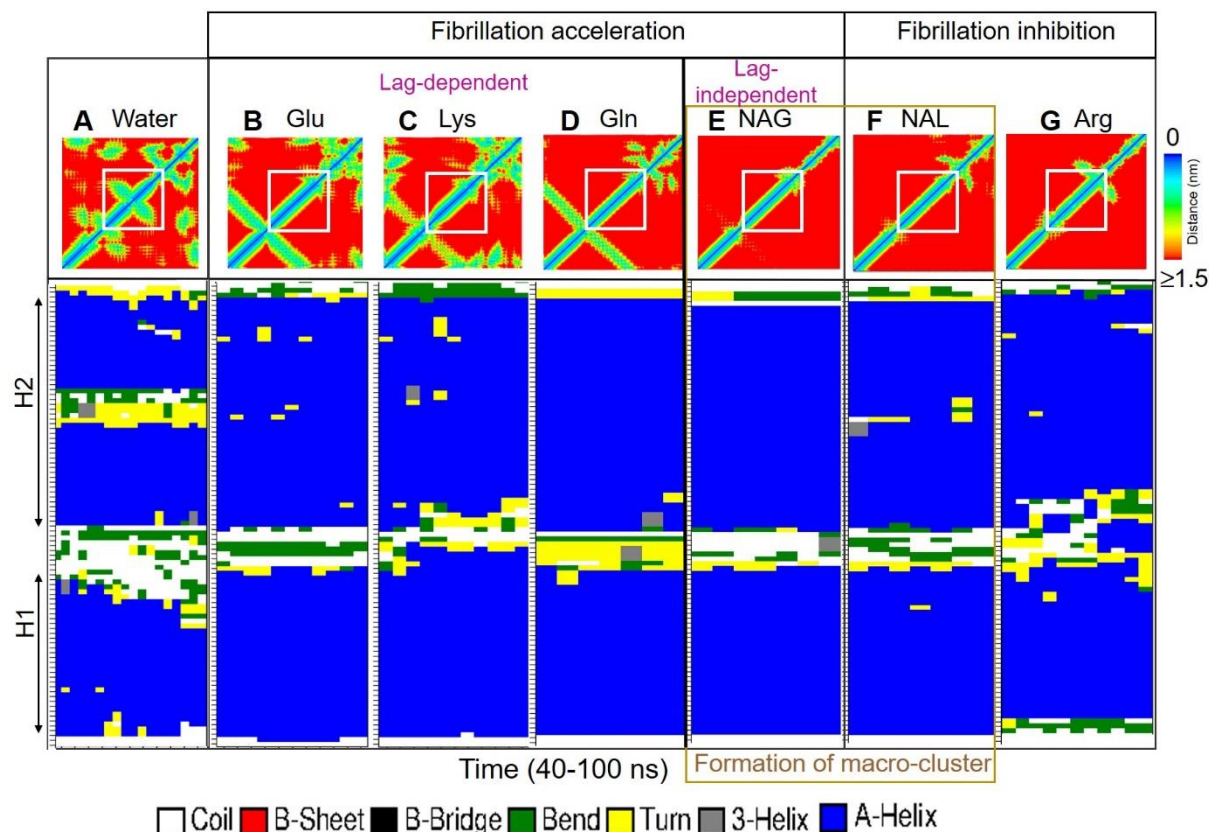


Fig. 4.11 The contact maps between the residues in α -syn (upper panels) and the secondary structural changes (lower panels) of the protein in the absence of AAs, (A) and in the presence of Glu (B), Lys (C), Gln (D), NAG (E), NAL (F) and Arg (G) during the last 60 ns of MD simulations. The color codes for distance map is given in the bar.

4.5 Discussion

Investigating the mechanism of α -syn fibrillation in presence of various osmolytes and crowding agents using different spectroscopic and computational methods aids in understanding the interactions dictating the structural transitions and in strategizing therapeutics for synucleinopathy. Here, residue-level interactions and dynamics of α -syn in the presence of L-amino acids involved in neural osmoregulatory is analyzed and their influence on the fibril formation of the protein, experimentally studied in the previous chapter, is examined. Three different effects of AAs on α -syn fibrillation is mainly observed: (i) acceleration of fibrillation with lag-dependent mechanism, (ii) acceleration of fibrillation with lag-independent mechanism and (iii) inhibition of fibrillation.

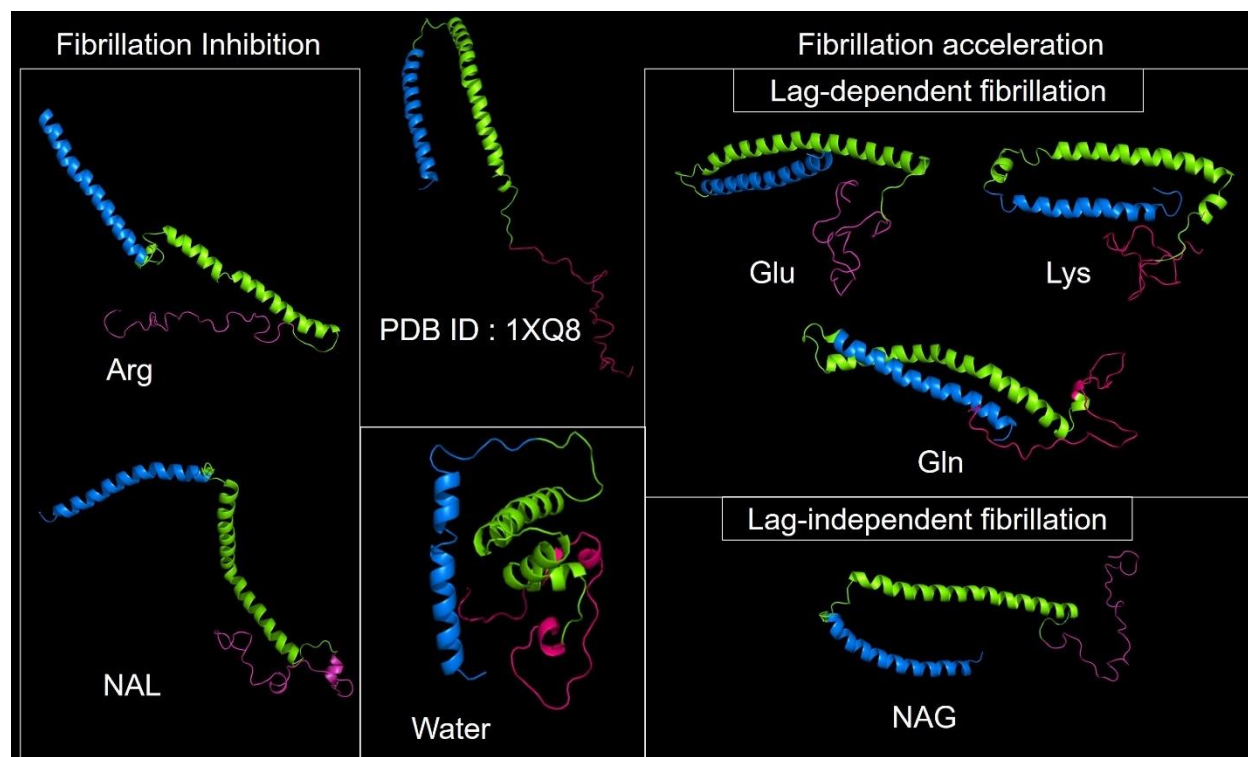


Fig.4.12 Representative structures of α -syn from molecular dynamics simulations: At the center NMR-resolved structure of α -syn from protein structure database (PDB ID:1XQ8) and α -syn in only water from MD simulation. Left side panels have α -syn structures in presence of AAs that inhibit fibrillation, Arg and NAL. Right side panels have AAs that accelerate fibrillation with lag-dependent (Glu, Lys and Gln) or lag-independent pathway (NAG). NTR, FC and CTR of α -syn are shown as blue, green and pink, respectively.

4.5.1 AAs accelerating fibrillation with lag-dependent mechanism

In the presence of Glu, Lys and Gln, the fibrillation of α -syn is accelerated, but without altering the lag-dependent pathway. In these AAs, α -syn exhibits significant interactions within NTR and FC (Fig. 4.10 and 4.11). A representative conformation of the protein in these AAs are presented in Fig. 4.12. Gln disrupts long-range interactions between the C-terminal region (CTR) and NTR or FC of α -syn (Fig. 4.11). Gln demonstrated preferential hydration around α -syn up to 0.42 nm, except for the FC region (Fig. 4.8), which is primarily composed of hydrophobic amino acids. However, the hydrophobic interactions between the two chains of α -syn in the HPR5 (V71-V83) region may compete with the weak interaction of Gln and might result in the early fibrillation of α -syn. On the other hand, Glu and Lys display preferential interactions with the NTR and CTR regions, respectively, owing to their charges. This led to a transient loss of helicity in the HPR5 and HPR6 regions of H2 enabling access of HPR5 between two chains and promoting fibril

formation. Notably, a compact conformation with significant intra-region interactions in the FC region is observed in the absence of any AAs indicating the necessity for additional structural transitions to access HPR5 between two chains of α -syn.

4.5.2 AAs accelerating fibrillation with lag-independent mechanism

NAG, as observed in Fig. 4.9C, forms a macro-cluster around H1 and H2 regions of α -syn. It exhibits preferential hydration around the CTR region up to 0.42 nm primarily due to the negative charge. Additionally, NAG interacts with both NTR and FC regions of α -syn (Fig. 4.8). Despite its negative charge, NAG's interaction with α -syn is mainly driven by non-coulombic interaction (Fig. 4.9) where water molecules might screen the charge-charge repulsion between NAG molecules and facilitate a tight packing to form macro-clusters (Fig. 4.9C). This conformational arrangement leads to the separation of H1 and H2 at the distance of ~ 2 nm (Fig. 4.10B). Although NAG and α -syn shows strong L-J potential, it might be easily replaced by stronger hydrophobic interactions formed between two HPR5 regions of two α -syn chains. Due to NAG's interactions with water molecules, it breaks water network around the protein and create hydrophobic environment which may guide the association of polypeptide chains. As the concentration of NAG increases, the charge-charge repulsion intensifies, requiring a greater number of water molecules to hold NAG in place. Consequently, the distance between NAG and α -syn might increase that leading to faster fibrillation.

4.5.3 Fibrillation inhibiting AA with α -syn

Arg and NAL inhibit the fibrillation of α -syn. Structural characteristics of α -syn in presence of inhibitory AAs is similar to the intermediate IC3 of helical conformational pathway demonstrated in the Chapter 2 except for the long-range interaction of NTR and CTR shown in Fig. 4.13. In the presence of these AAs, the distance between NTR and FC regions of α -syn become longer whereas significant interactions between RTD5 from FC and Q99 to Q109 region of CTR is notable (Fig. 4.11 and 4.12). Arg is known to interact with aromatic residues through cation- π interactions and form salt bridges with charged residues on proteins.^[242,270] The self-clustering behavior of Arg limits protein-protein interactions and inhibits fibrillation in many globular proteins.^[271,272] Similar effect observed for Arg in the case of IDP as well. A high Coulombic

Chapter 4

interaction energy raising due to a strong binding of these amino acids across all three regions of α -syn is observed. NAL has the highest affinity for NTR region, as indicated by its highest K_p value and lowest hydration fraction, and it forms clusters on this region (Fig. 4.7). Both Arg and NAL form small and medium-sized clusters that engage in high electrostatic interactions (Fig. 4.9D) with the protein, thereby inhibiting the hydrophobic interactions between the FC regions of two α -syn chains.

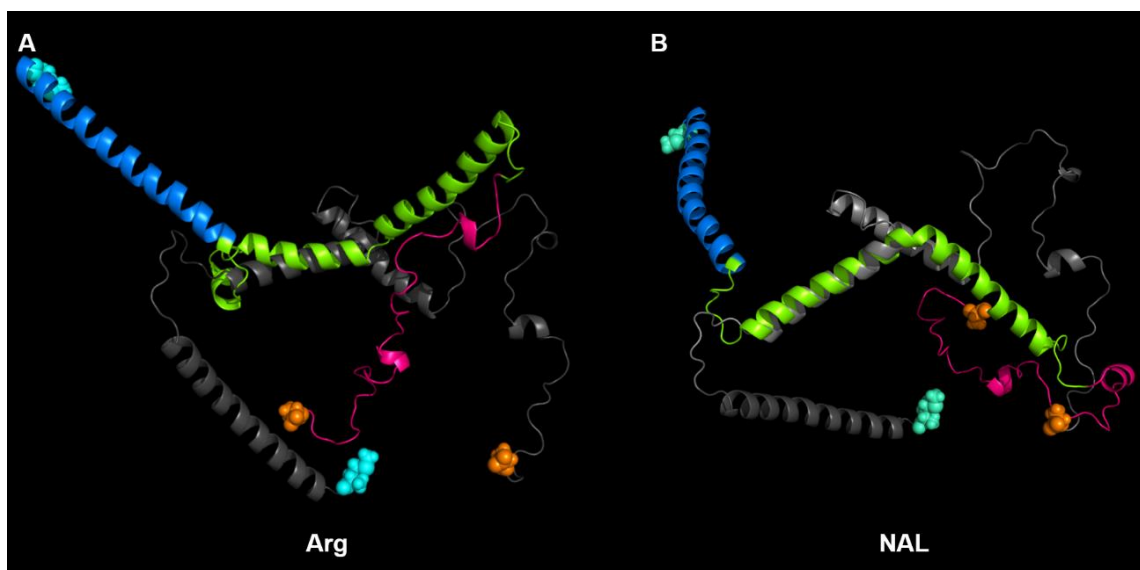


Fig.4.13 Representative structures of α -syn from MD simulation in presence of fibril-inhibiting AAs aligned with the intermediate IC3 (gray) from helical-conformation pathway (Chapter2). NTR, FC and CTR of α -syn in presence of (A) Arg and (B) NAL are shown as blue, green and pink respectively. Cyan and orange colour balls represents NTR and CTR respectively.

4.5.4 Molecular interactions and fibrillation propensity

Monomeric α -syn in a pure water system adopts a compact conformation characterized by a ternary contact between NTR, CTR, and FC regions. This compact structure has been previously identified as a fibril-prone conformation of α -syn in a pure solvent condition.^[125,127,134] In chapter 2, the transition of α -syn from its monomeric state to fibrillar structures was studied which also suggested the same. In this chapter, the factors that determine the acceleration or inhibition of α -syn fibrillation in presence of AAs was studied. It is observed that two key features play significant roles in the α -syn-AA system. Firstly, the conformational changes in α -syn induced by the AAs. The interdomain interactions between NTR and FC regions and the absence of intradomain interactions within FC region are associated with the acceleration of fibrillation. Conversely, the

Structural insights of the α -syn-AA interaction.

lack of interaction of NTR with other regions of the protein along with loss of helicity in the H2 region and the interaction between RTD5 from FC with the residues in CTR particularly Q99-Q109 resulted in inhibition of fibril formation. Previous studies have also indicated that the highly dynamic nature of RTD5 in oligomers leads to the formation of amorphous aggregates.^[273] Urea interacts with α -syn by forming H-bond and leads to extended conformation that inhibits fibrillation.^[253,260]

Secondly, the strength of interaction energy between AA and α -syn determines the competition between AA-protein and protein-protein interactions. If an AA exhibits preferential interaction with α -syn through strong Coulombic interaction energy, it inhibits α -syn fibrillation. In such cases, the AA effectively excludes water molecules from the protein surface due to the strong electrostatic interactions between the AA and α -syn, thereby inhibiting protein-protein interactions. Non-polar interaction of AAs with the protein surface induces a hydrophobic environment surrounding the protein leading to a fast phase separation and accelerated protein-protein interactions. Earlier studies have highlighted the significance of hydrophobicity for fibril stability, as amyloid fibril elongation is generally associated with a negative heat capacity, which correlates with the buried hydrophobic surface area upon fibril formation.^[274,275] Similarly in our observations, a stronger non-columbic interaction contribution (NAG > Gln > Glu > Lys) corresponds to shorter lag times (NAG < Gln < Glu < Lys, as shown in Chapter 3).

Chapter 4

4.6 Conclusion

Structural change in monomeric α -syn induced by various AAs were generated by CMD and the behavior of AAs were analyzed using the theory of preferential interaction. Major structural characteristics accelerating the fibril formation of α -syn is found to be interdomain interaction between NTR and CTR and loss of intradomain interaction in FC induced by strong non-columbic interactions. On the other hand, preferential interaction of AAs with α -syn through strong electrostatic potential resulting in disruption of interactions between the domains and leading to extended conformations of the protein. These findings shed light on the specific non-covalent interactions that play a crucial role in determining the fate of α -syn fibrillation and provide further understanding on the unique characteristics exhibited by different AAs in their interactions with α -syn.

Overall summary

α -Synuclein (α -syn) forms amyloid fibrils which are the major constituent of Lewy bodies, a hallmark of Parkinson's disease (PD) and many other neurodegenerative diseases. Due to experimental challenges in following the fibrillation mechanism of intrinsically disordered α -syn, concerted enhanced sampling methods (SMD, Umbrella Sampling and REMD) and conventional molecular simulations (CMD) were performed to characterize the intermediate structures in the fibril formation pathway(s). Nineteen distinct structures were obtained and arranged between monomer to fibril, based on their radius of gyration, solvent-accessible surface area, and secondary structure content to predict the fibril dissociation pathway. The analysis showed that the initial dissociation of the polypeptide chain from the fibril might follow either a compact-conformational pathway due to the presence of long-range interactions or an extended-conformational pathway that is stabilized by local interactions. These two pathways might converge after complete dissociation of the fibril and further follow a helical-conformational pathway to attain the monomeric structure. The overall analysis of the pathways suggests that the formation of β -turns, reorganization of salt bridges, and dihedral changes in the hydrophobic regions are the major driving forces for helix-fibril transition. Further, the structural features of some of the identified intermediates correlated with the earlier experimental and computational studies. This concerted approach helps to characterize the intermediate structures in the fibril formation pathway(s) of the intrinsically disordered protein that otherwise would be an impossible task with any single method. Also, understanding the interactions in these intermediates including solvation energies, and long-range contacts might aid in designing anti-fibrillation agents.

The high plasticity and lack of stable tertiary structure makes α -Syn highly susceptible to its surrounding environment. It is observed that α -syn aggregates in hyperosmotic stress caused by different osmolytes. It is also reported that charged amino acids and their derivatives (AA) are essential for neural cell volume in hyperosmotic stress. Therefore, the effect of charged amino acids (L-Glutamate, L-Aspartate, L-Lysine, and L-Arginine), amino acids with amide side chains (L- Glutamine and L-Asparagine), and *N*-acetylated amino acids (*N*-acetyl-L-glutamic acid and *N*-acetyl-L-lysine) on the fibrillation of recombinant human α -syn was examined. Arg and NAL inhibited the fibrillation of α -syn at the concentrations above 0.2 M and 0.4 M, respectively. In the presence of lysine, aspartate, and glutamate, due to the shorter lag time, the protein attained the

Chapter 5

fibrillar state earlier. NAG followed lag-independent α -syn fibrillation whereas Gln, and Asn exhibited concentration-dependent effects on α -syn fibrillation. Gln and Asn showed a decrease in lag time with an increase in their concentration resulting in acceleration of fibrillation kinetics. The elongation rate had an insignificant effect on the total fibrillation time. Therefore, it may be stated that the effect of AA on α -syn fibrillation is predominantly due to their role on altering the initial lag time.

Previous research has shown that osmolytes accumulate near proteins through preferential binding and contribute to enthalpic energy required for protein unfolding. In order to understand the molecular interaction between AA and α -syn, conventional MD simulations of α -syn were performed in the presence of 0.5 M of Glu, Lys, and Gln which showed lag-dependent accelerated fibrillation kinetics of α -syn. Additionally, MD simulation was performed in the presence of 0.5 M of NAG that induced lag-independent accelerated fibrillation kinetics. To analyze the fibrillation inhibition effect by AAs, simulations with 0.5 M Arg and NAL were also carried out. AAs that accelerate fibrillation enhance contacts between N-terminal region (NTR) and the fibril core (FC), while lacking interactions within the FC. These AAs exhibit stronger non-columbic interactions with surface residues of α -synuclein (NAG > Gln > Glu > Lys), resulting in reduced lag time for fibrillation. Inhibitory AAs preferentially interact with α -synuclein through strong electrostatic energy, causing no contact between NTR and the other two domains of α -synuclein and leading to the formation of an extended conformation.

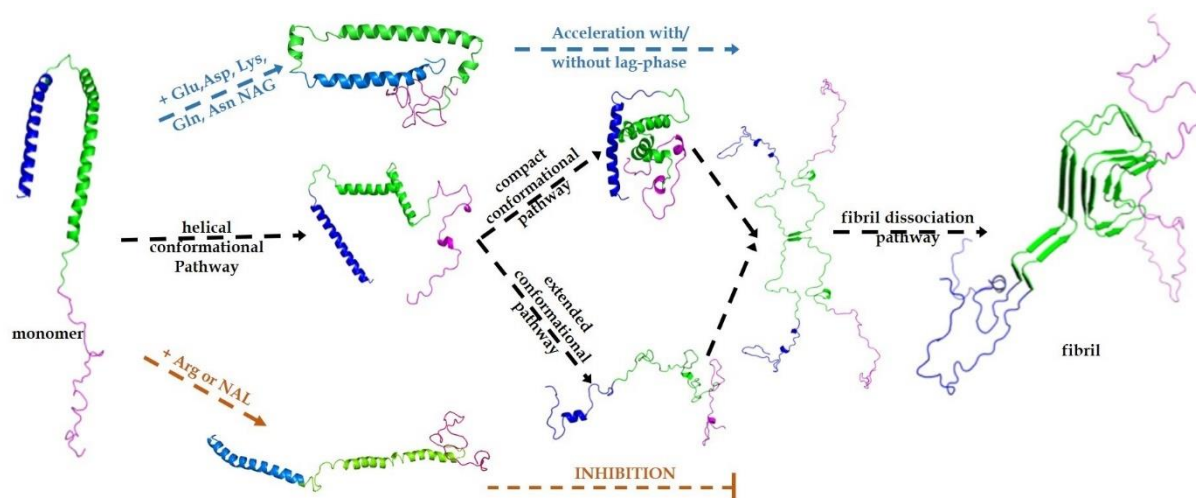


Figure. A schematic representation of monomer to fibril transition pathway(s) of α -syn demonstrated in Chapter 2 (black arrows). Structural changes in monomeric α -syn in presence of fibril accelerating osmolytes (blue arrows) and inhibitory osmolytes (orange arrows).

Bibliography

- [1] C. B. Anfinsen, E. Haber, M. Sela, F. H. J. White, *Proc. Natl. Acad. Sci. U. S. A.* **1961**, 47, 1309–1314.
- [2] I. Langmuir, *Science (80-.)*. **1938**, 87, 119–123.
- [3] J. Traube, *Berichte der Dtsch. Chem. Gesellschaft* **1891**, 24, 737–745.
- [4] A. C. Clark, *Arch. Biochem. Biophys.* **2008**, 469, 1–3.
- [5] *J. Chem. Phys.* **1933**, 1, 515–548.
- [6] W. M. Latimer, W. H. Rodebush, *J. Am. Chem. Soc.* **1920**, 42, 1419–1433.
- [7] W. Kauzmann, *Adv. Protein Chem.* **1959**, 14, 1–63.
- [8] E. Fischer, *J. Am. Chem. Soc.* **1914**, 36, 1170–1201.
- [9] L. Pauling, R. B. Corey, H. R. Branson, *Proc. Natl. Acad. Sci. U. S. A.* **1951**, 37, 205–211.
- [10] J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, *Nature* **1960**, 185, 422–427.
- [11] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, 181, 662–666.
- [12] J. F. Brandts, *J. Am. Chem. Soc.* **1964**, 86, 4291–4301.
- [13] R. Lumry, R. Biltonen, *Biopolymers* **1966**, 4, 917–944.
- [14] C. Tanford, *Adv. Protein Chem.* **1968**, 23, 121–282.
- [15] Levinthal, Cyrus, *J. Chim. Phys.* **1968**, 65, 44–45.
- [16] C. R. Matthews, *Annu. Rev. Biochem.* **1993**, 62, 653–683.
- [17] P. S. Kim, R. L. Baldwin, *Annu. Rev. Biochem.* **1990**, 59, 631–660.
- [18] D. A. Dolgikh, R. I. Gilmanishin, E. V Brazhnikov, V. E. Bychkova, G. V Semisotnov, V. SYu, O. B. Ptitsyn, *FEBS Lett.* **1981**, 136, 311–315.
- [19] T. E. Creighton, *J. Mol. Biol.* **1974**, 87, 579–602.

- [20] K. P. Wong, C. Tanford, *J. Biol. Chem.* **1973**, *248*, 8518–8523.
- [21] M. Ohgushi, A. Wada, *FEBS Lett.* **1983**, *164*, 21–24.
- [22] J. Brange, L. Andersen, E. D. Laursen, G. Meyn, E. Rasmussen, *J. Pharm. Sci.* **1997**, *86*, 517–525.
- [23] E. Dupont, Y.-S. Ko, S. Rothery, S. R. Coppen, M. Baghai, M. Haw, N. J. Severs, *Circulation* **2001**, *103*, 842–849.
- [24] S. M. C. Hardman, K. P. Pfeiffer, T. Kenner, M. I. M. Noble, W. A. Seed, *Basic Res. Cardiol.* **1994**, *89*, 438–455.
- [25] J. D. Sipe, A. S. Cohen, *J. Struct. Biol.* **2000**, *130*, 88–98.
- [26] B. H. Toyama, J. S. Weissman, *Annu. Rev. Biochem.* **2011**, *80*, 557–585.
- [27] R. Nelson, D. Eisenberg, *Curr. Opin. Struct. Biol.* **2006**, *16*, 260–265.
- [28] L. C. Serpell, *Biochim. Biophys. Acta* **2000**, *1502*, 16–30.
- [29] M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, C. C. Blake, *J. Mol. Biol.* **1997**, *273*, 729–739.
- [30] A. L. Fink, *Acc. Chem. Res.* **2006**, *39*, 628–634.
- [31] S. Ghosh, A. Saurabh, N. P. Prabhu, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *264*, 120307.
- [32] E. K. Kumar, N. Haque, N. P. Prabhu, *Int. J. Biol. Macromol.* **2017**, *100*, 3–10.
- [33] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, *Proteins* **1995**, *21*, 167–195.
- [34] K. A. Dill, *Biochemistry* **1985**, *24*, 1501–1509.
- [35] J. D. Bryngelson, P. G. Wolynes, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7524–7528.
- [36] P. E. Leopold, M. Montal, J. N. Onuchic, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 8721–8725.
- [37] K. A. Dill, H. S. Chan, *Nat. Struct. Biol.* **1997**, *4*, 10–19.

- [38] J. D. Dunitz, G. F. Joyce, L. E. Orgel, *Natl. Acad. Sci.* **2013**.
- [39] D. Berichte, **1898**, 3479–3483.
- [40] A. Borgia, M. B. Borgia, K. Bugge, V. M. Kissling, P. O. Heidarsson, C. B. Fernandes, A. Sottini, A. Soranno, K. J. Buholzer, D. Nettels, B. B. Kragelund, R. B. Best, B. Schuler, *Nature* **2018**, 555, 61–66.
- [41] D. E. Koshland, *Proc. Natl. Acad. Sci. U. S. A.* **1958**, 44, 98–104.
- [42] J.-R. Yang, S.-M. Zhuang, J. Zhang, *Mol. Syst. Biol.* **2010**, 6, 421.
- [43] D. A. Drummond, C. O. Wilke, *Cell* **2008**, 134, 341–352.
- [44] K. B. Zeldovich, P. Chen, E. I. Shakhnovich, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, 104, 16152–16157.
- [45] D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 14338–14343.
- [46] J.-R. Yang, B.-Y. Liao, S.-M. Zhuang, J. Zhang, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109, E831-40.
- [47] A. W. R. Serohijos, Z. Rimas, E. I. Shakhnovich, *Cell Rep.* **2012**, 2, 249–256.
- [48] A. W. R. Serohijos, S. Y. R. Lee, E. I. Shakhnovich, *Biophys. J.* **2013**, 104, L1-3.
- [49] E. D. Levy, S. De, S. A. Teichmann, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109, 20461–20466.
- [50] N. Tokuriki, D. S. Tawfik, *Nature* **2009**, 459, 668–673.
- [51] S. L. Rutherford, *Nat. Rev. Genet.* **2003**, 4, 263–274.
- [52] S. L. Rutherford, S. Lindquist, *Nature* **1998**, 396, 336–342.
- [53] A. N. Gupta, K. Neupane, N. Rezajooei, L. M. Cortez, V. L. Sim, M. T. Woodside, *Nat. Commun.* **2016**, 7, 12058.
- [54] A. S. Kadibalban, D. Bogumil, G. Landan, T. Dagan, *Genome Biol. Evol.* **2016**, 8, 1590–1599.

- [55] V. N. Uversky, A. K. Dunker, *Biochim. Biophys. Acta* **2010**, 1804, 1231–1264.
- [56] V. N. Uversky, *J. Biomed. Biotechnol.* **2010**, 2010, 568068.
- [57] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, D. T. Jones, *J. Mol. Biol.* **2004**, 337, 635–645.
- [58] A. K. Dunker, Z. Obradovic, P. Romero, E. C. Garner, C. J. Brown, *Genome Inform. Ser. Workshop Genome Inform.* **2000**, 11, 161–171.
- [59] A. K. Dunker, J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, Z. Obradovic, *J. Mol. Graph. Model.* **2001**, 19, 26–59.
- [60] B. Xue, A. K. Dunker, V. N. Uversky, *J. Biomol. Struct. Dyn.* **2012**, 30, 137–149.
- [61] Z. Peng, J. Yan, X. Fan, M. J. Mizianty, B. Xue, K. Wang, G. Hu, V. N. Uversky, L. Kurgan, *Cell. Mol. Life Sci.* **2015**, 72, 137–151.
- [62] B. Kahali, T. C. Ghosh, *J. Biomol. Struct. Dyn.* **2013**, 31, 472–476.
- [63] R. Pancsa, P. Tompa, *PLoS One* **2012**, 7, e34687.
- [64] I. Korneta, J. M. Bujnicki, *PLoS Comput. Biol.* **2012**, 8, e1002641.
- [65] U. Midic, Z. Obradovic, *Proteome Sci.* **2012**, 10 Suppl 1, S19.
- [66] H. Hegyi, P. Tompa, *Mol. Biosyst.* **2012**, 8, 229–236.
- [67] P. Tompa, Z. Dosztanyi, I. Simon, *J. Proteome Res.* **2006**, 5, 1996–2000.
- [68] C. J. Oldfield, Y. Cheng, M. S. Cortese, C. J. Brown, V. N. Uversky, A. K. Dunker, *Biochemistry* **2005**, 44, 1989–2000.
- [69] V. N. Uversky, *J. Biol. Chem.* **2016**, 291, 6681–6688.
- [70] A. K. Dunker, M. S. Cortese, P. Romero, L. M. Iakoucheva, V. N. Uversky, *FEBS J.* **2005**, 272, 5129–5148.
- [71] V. N. Uversky, *Protein Sci.* **2002**, 11, 739–756.

- [72] H. J. Dyson, P. E. Wright, *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.
- [73] V. N. Uversky, *Eur. J. Biochem.* **2002**, *269*, 2–12.
- [74] A. K. Dunker, C. J. Brown, J. D. Lawson, L. M. Iakoucheva, Z. Obradović, *Biochemistry* **2002**, *41*, 6573–6582.
- [75] P. E. Wright, H. J. Dyson, *J. Mol. Biol.* **1999**, *293*, 321–331.
- [76] V. N. Uversky, *Methods Mol. Biol.* **2020**, *2141*, 895–945.
- [77] H. J. Dyson, P. E. Wright, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.
- [78] B. Strodel, *J. Mol. Biol.* **2021**, *433*, 167182.
- [79] A. Toto, F. Malagrinò, L. Visconti, F. Troilo, L. Pagano, M. Brunori, P. Jemth, S. Gianni, *J. Biol. Chem.* **2020**, *295*, 6586–6593.
- [80] Y. Chebaro, A. J. Ballard, D. Chakraborty, D. J. Wales, *Sci. Rep.* **2015**, *5*, 10386.
- [81] V. N. Uversky, C. J. Oldfield, A. K. Dunker, *Annu. Rev. Biophys.* **2008**, *37*, 215–246.
- [82] J. Q. Trojanowski, V. M.-Y. Lee, *Ann. N. Y. Acad. Sci.* **2003**, *991*, 107–110.
- [83] J. E. Galvin, V. M. Lee, J. Q. Trojanowski, *Arch. Neurol.* **2001**, *58*, 186–190.
- [84] M. Goedert, *Philos. Trans. R. Soc. London. Ser. B, Biol. Sci.* **1999**, *354*, 1101–1118.
- [85] M. G. Spillantini, M. Goedert, *Ann. N. Y. Acad. Sci.* **2000**, *920*, 16–27.
- [86] C. B. Lücking, A. Brice, *Cell. Mol. Life Sci.* **2000**, *57*, 1894–1908.
- [87] H. R. Lucas, R. D. Fernández, *Neural Regen. Res.* **2020**, *15*, 407–415.
- [88] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.
- [89] T. S. Ulmer, A. Bax, N. B. Cole, R. L. Nussbaum, *J. Biol. Chem.* **2005**, *280*, 9595–9603.
- [90] M. D. Tuttle, G. Comellas, A. J. Nieuwkoop, D. J. Covell, D. A. Berthold, K. D. Kloepper, J. M. Courtney, J. K. Kim, A. M. Barclay, A. Kendall, W. Wan, G. Stubbs, C. D. Schwieters, V. M. Y. Lee, J. M. George, C. M. Rienstra, *Nat. Struct. Mol. Biol.* **2016**,

23, 409–415.

- [91] A. Saurabh, N. P. Prabhu, *Int. J. Biol. Macromol.* **2022**, *223*, 1024–1041.
- [92] R. Guerrero-Ferreira, L. Kovacik, D. Ni, H. Stahlberg, *Curr. Opin. Neurobiol.* **2020**, *61*, 89–95.
- [93] A. Khammari, S. S. Arab, M. R. Ejtehadi, *Sci. Rep.* **2020**, *10*, 12175.
- [94] G. Bianchi, S. Longhi, R. Grandori, S. Brocca, *Int. J. Mol. Sci.* **2020**, *21*, 6208.
- [95] J. R. Allison, P. Varnai, C. M. Dobson, M. Vendruscolo, *J. Am. Chem. Soc.* **2009**, *131*, 18314–18326.
- [96] A. M. C. Fragniere, S. R. W. Stott, S. V Fazal, M. Andreassen, K. Scott, R. A. Barker, *Sci. Rep.* **2019**, *9*, 2288.
- [97] R. Franco, O. Quesada, H. Pasantes-Morales, *J. Neurosci. Res.* **2000**, *61*, 701–711.
- [98] D. Harries, J. Rösger, *Methods Cell Biol.* **2008**, *84*, 679–735.
- [99] S. H. Khan, N. Ahmad, F. Ahmad, R. Kumar, *IUBMB Life* **2010**, *62*, 891–895.
- [100] A. Ghahghaei, S. Mohammadian, *Appl. Biochem. Biotechnol.* **2014**, *174*, 739–750.
- [101] S. Mojtavavi, N. Samadi, M. A. Faramarzi, *Iran. J. Pharm. Res. IJPR* **2019**, *18*, 13–30.
- [102] L. A. Ferreira, X. Fan, P. P. Madeira, L. Kurgan, V. N. Uversky, B. Y. Zaslavsky, *RSC Adv.* **2015**, *5*, 59780–59791.
- [103] M. Mukherjee, J. Mondal, *J. Phys. Chem. B* **2018**, *122*, 6922–6930.
- [104] B. Anumalla, N. P. Prabhu, *J. Mol. Graph. Model.* **2020**, *98*, 107602.
- [105] N. Kushwah, V. Jain, D. Yadav, *Biomolecules* **2020**, *10*, 132.
- [106] D. L. Taylor, S. E. Davies, T. P. Obrenovitch, M. H. Doheny, P. N. Patsalos, J. B. Clark, L. Symon, *J. Neurochem.* **1995**, *65*, 275–281.
- [107] M. Warepam, A. K. Mishra, G. S. Sharma, K. Kumari, S. Krishna, M. S. A. Khan, H. Rahman, L. R. Singh, *Front. Cell. Neurosci.* **2021**, *15*, 617308.

- [108] D. Sarfaraz, C. L. Fraser, *Am. J. Physiol.* **1999**, 276, E596-601.
- [109] W. Hoyer, T. Antony, D. Cherny, G. Heim, T. M. Jovin, V. Subramaniam, *J. Mol. Biol.* **2002**, 322, 383–393.
- [110] A. J. Lees, J. Hardy, T. Revesz, *Lancet (London, England)* **2009**, 373, 2055–2066.
- [111] N. M. Bonini, B. I. Giasson, *Cell* **2005**, 123, 359–361.
- [112] A. Abeliovich, Y. Schmitz, I. Fariñas, D. Choi-Lundberg, W. H. Ho, P. E. Castillo, N. Shinsky, J. M. Verdugo, M. Armanini, A. Ryan, M. Hynes, H. Phillips, D. Sulzer, A. Rosenthal, *Neuron* **2000**, 25, 239–252.
- [113] M. G. Spillantini, R. A. Crowther, R. Jakes, M. Hasegawa, M. Goedert, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 6469–6473.
- [114] M. G. Spillantini, M. L. Schmidt, V. M. Lee, J. Q. Trojanowski, R. Jakes, M. Goedert, *Nature* **1997**, 388, 839–840.
- [115] R. Jakes, M. G. Spillantini, M. Goedert, *FEBS Lett.* **1994**, 345, 27–32.
- [116] M. Goedert, *Nature* **1997**, 388, 232–233.
- [117] L. Pontoriero, M. Schiavina, M. G. Murrall, R. Pierattelli, I. C. Felli, *Angew. Chem. Int. Ed. Engl.* **2020**, 59, 18537–18545.
- [118] L. Geist, M. A. Henen, S. Haiderer, T. C. Schwarz, D. Kurzbach, A. Zawadzka-Kazimierzczuk, S. Saxena, S. Zerko, W. Koźmiński, D. Hinderberger, R. Konrat, *Protein Sci.* **2013**, 22, 1196–1205.
- [119] F.-X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, *Chem. Rev.* **2014**, 114, 6661–6714.
- [120] G. Comellas, L. R. Lemkau, D. H. Zhou, J. M. George, C. M. Rienstra, *J. Am. Chem. Soc.* **2012**, 134, 5090–5099.
- [121] C. C. Jao, B. G. Hegde, J. Chen, I. S. Haworth, R. Langen, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, 105, 19666–19671.

- [122] H. L. Roberts, D. R. Brown, *Biomolecules* **2015**, *5*, 282–305.
- [123] V. N. Uversky, J. Li, A. L. Fink, *J. Biol. Chem.* **2001**, *276*, 10737–10744.
- [124] W. Hoyer, D. Cherny, V. Subramaniam, T. M. Jovin, *Biochemistry* **2004**, *43*, 16233–16242.
- [125] C. W. Bertoncini, Y.-S. Jung, C. O. Fernandez, W. Hoyer, C. Griesinger, T. M. Jovin, M. Zweckstetter, *Proc. Natl. Acad. Sci.* **2005**, *102*, 1430–1435.
- [126] D. P. Hong, W. Xiong, J. Y. Chang, C. Jiang, *FEBS Lett.* **2011**, *585*, 561–566.
- [127] S. Esteban-Martín, J. Silvestre-Ryan, C. W. Bertoncini, X. Salvatella, *Biophys. J.* **2013**, *105*, 1192–1198.
- [128] J. A. Rodriguez, M. I. Ivanova, M. R. Sawaya, D. Cascio, F. E. Reyes, D. Shi, S. Sangwan, E. L. Guenther, L. M. Johnson, M. Zhang, L. Jiang, M. A. Arbing, B. L. Nannenga, J. Hattne, J. Whitelegge, A. S. Brewster, M. Messerschmidt, S. Boutet, N. K. Sauter, T. Gonen, D. S. Eisenberg, *Nature* **2015**, *525*, 486–490.
- [129] P. J. Salveson, R. K. Spencer, J. S. Nowick, *J. Am. Chem. Soc.* **2016**, *138*, 4458–4467.
- [130] S. W. Chen, S. Drakulic, E. Deas, M. Oubrai, F. A. Aprile, R. Arranz, S. Ness, C. Roodveldt, T. Williams, E. J. De-Genst, D. Klenerman, N. W. Wood, T. P. J. Knowles, C. Alfonso, G. Rivas, A. Y. Abramov, J. M. Valpuesta, C. M. Dobson, N. Cremades, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E1994-2003.
- [131] E. A. Mirecka, H. Shaykhalishahi, A. Gauhar, Ş. Akgül, J. Lecher, D. Willbold, M. Stoldt, W. Hoyer, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 4227–4230.
- [132] H. Shaykhalishahi, A. Gauhar, M. M. Wördehoff, C. S. R. Grüning, A. N. Klein, O. Bannach, M. Stoldt, D. Willbold, T. Härd, W. Hoyer, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 8837–8840.
- [133] R. Guerrero-Ferreira, L. Kovacic, D. Ni, H. Stahlberg, *Curr. Opin. Neurobiol.* **2020**, *61*, 89–95.
- [134] W. Paslawski, S. Mysling, K. Thomsen, T. J. D. Jørgensen, D. E. Otzen, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7560–7563.

- [135] H.-Y. Kim, M.-K. Cho, A. Kumar, E. Maier, C. Siebenhaar, S. Becker, C. O. Fernandez, H. A. Lashuel, R. Benz, A. Lange, M. Zweckstetter, *J. Am. Chem. Soc.* **2009**, *131*, 17482–17489.
- [136] T. Bartels, J. G. Choi, D. J. Selkoe, *Nature* **2011**, *477*, 107–110.
- [137] S. Nuber, M. Rajsombath, G. Minakaki, J. Winkler, C. P. Müller, M. Ericsson, B. Caldarone, U. Dettmer, D. J. Selkoe, *Neuron* **2018**, *100*, 75-90.e5.
- [138] B. Fauvet, M. K. Mbefo, M.-B. Fares, C. Desobry, S. Michael, M. T. Ardah, E. Tsika, P. Coune, M. Prudent, N. Lion, D. Eliezer, D. J. Moore, B. Schneider, P. Aebischer, O. M. El-Agnaf, E. Masliah, H. A. Lashuel, *J. Biol. Chem.* **2012**, *287*, 15345–15364.
- [139] T. Gurry, O. Ullman, C. K. Fisher, I. Perovic, T. Pochapsky, C. M. Stultz, *J. Am. Chem. Soc.* **2013**, *135*, 3865–3872.
- [140] M. M. Apetri, N. C. Maiti, M. G. Zagorski, P. R. Carey, V. E. Anderson, *J. Mol. Biol.* **2006**, *355*, 63–71.
- [141] A. K. Dunker, M. M. Babu, E. Barbar, M. Blackledge, S. E. Bondos, Z. Dosztányi, H. J. Dyson, J. Forman-Kay, M. Fuxreiter, J. Gsponer, K.-H. Han, D. T. Jones, S. Longhi, S. J. Metallo, K. Nishikawa, R. Nussinov, Z. Obradovic, R. V Pappu, B. Rost, P. Selenko, V. Subramaniam, J. L. Sussman, P. Tompa, V. N. Uversky, *Intrinsically Disord. proteins* **2013**, *1*, e24157.
- [142] R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V Pappu, P. Tompa, V. N. Uversky, P. E. Wright, M. M. Babu, *Chem. Rev.* **2014**, *114*, 6589–6631.
- [143] J. Habchi, P. Tompa, S. Longhi, V. N. Uversky, *Chem. Rev.* **2014**, *114*, 6561–6588.
- [144] M. Arbesú, M. Pons, *Arch. Biochem. Biophys.* **2019**, *677*, 108161.
- [145] N. Rezaei-Ghaleh, G. Parigi, A. Soranno, A. Holla, S. Becker, B. Schuler, C. Luchinat, M. Zweckstetter, *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 15262–15266.
- [146] T. Zhang, Y. Tian, Z. Li, S. Liu, X. Hu, Z. Yang, X. Ling, S. Liu, J. Zhang, *J. Chem. Inf.*

- Model.* **2017**, *57*, 2281–2293.
- [147] J. Yoon, S. Jang, K. Lee, S. Shin, *J. Biomol. Struct. Dyn.* **2009**, *27*, 259–269.
- [148] S. Park, J. Yoon, S. Jang, K. Lee, S. Shin, *J. Biomol. Struct. Dyn.* **2016**, *34*, 376–383.
- [149] H. Heise, W. Hoyer, S. Becker, O. C. Andronesi, D. Riedel, M. Baldus, *Proc. Natl. Acad. Sci.* **2005**, *102*, 15871–15876.
- [150] M. M. Dedmon, K. Lindorff-Larsen, J. Christodoulou, M. Vendruscolo, C. M. Dobson, *J. Am. Chem. Soc.* **2005**, *127*, 476–477.
- [151] A. Iyer, S. J. Roeters, V. Kogan, S. Woutersen, M. M. A. E. Claessens, V. Subramaniam, *J. Am. Chem. Soc.* **2017**, *139*, 15392–15400.
- [152] R. A. Crowther, R. Jakes, M. G. Spillantini, M. Goedert, *FEBS Lett.* **1998**, *436*, 309–312.
- [153] J. V Vermaas, E. Tajkhorshid, *Biochim. Biophys. Acta* **2014**, *1838*, 3107–3117.
- [154] A. Balupuri, K.-E. Choi, N. S. Kang, *Sci. Rep.* **2019**, *9*, 59.
- [155] I. F. Tsigelny, Y. Sharikov, M. A. Miller, E. Masliah, *Nanomedicine* **2008**, *4*, 350–357.
- [156] K. K. Sahu, M. T. Woodside, J. A. Tuszynski, *Biochimie* **2015**, *116*, 133–140.
- [157] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, E. Lindahl, *J. Chem. Theory Comput.* **2010**, *6*, 459–466.
- [158] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19–25.
- [159] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 14101.
- [160] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [161] J. A. Lemkul, D. R. Bevan, *J. Phys. Chem. B* **2010**, *114*, 1652–1660.
- [162] J. S. Hub, B. L. de Groot, D. van der Spoel, *J. Chem. Theory Comput.* **2010**, *6*, 3713–3720.
- [163] A. Patriksson, D. van der Spoel, *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073–2077.

- [164] H. Liu, D. Song, Y. Zhang, S. Yang, R. Luo, H.-F. Chen, *Phys. Chem. Chem. Phys.* **2019**, *21*, 21918–21931.
- [165] H. Liu, D. Song, H. Lu, R. Luo, H.-F. Chen, *Chem. Biol. Drug Des.* **2018**, *92*, 1722–1735.
- [166] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 27-28,33-38.
- [167] K. Masutani, Y. Yamamori, K. Kim, N. Matubayasi, *J. Chem. Phys.* **2019**, *150*, 145101.
- [168] N. Matubayasi, *Curr. Opin. Struct. Biol.* **2017**, *43*, 45–54.
- [169] O. Ullman, C. K. Fisher, C. M. Stultz, *J. Am. Chem. Soc.* **2011**, *133*, 19536–19546.
- [170] S. Chandra, X. Chen, J. Rizo, R. Jahn, T. C. Südhof, *J. Biol. Chem.* **2003**, *278*, 15313–15318.
- [171] X. Chen, I. Weber, R. W. Harrison, *J. Phys. Chem. B* **2008**, *112*, 12073–12080.
- [172] K.-P. Wu, D. S. Weinstock, C. Narayanan, R. M. Levy, J. Baum, *J. Mol. Biol.* **2009**, *391*, 784–796.
- [173] A. G. de Brevern, *Sci. Rep.* **2016**, *6*, 33191.
- [174] Y.-H. Sung, D. Eliezer, *Protein Sci.* **2018**, *27*, 1314–1324.
- [175] H. Yu, W. Han, W. Ma, K. Schulten, *J. Chem. Phys.* **2015**, *143*, 243142.
- [176] M. Ziaunys, A. Sakalauskas, K. Mikalauskaite, V. Smirnovas, *Int. J. Mol. Sci.* **2021**, *22*, 12382.
- [177] B. Bryson, *A Short History of Nearly Everything Special Illustrated Edition.*, Broadway Books, **2003**.
- [178] A. D. Stephens, M. Zacharopoulou, G. S. Kaminski Schierle, *Trends Biochem. Sci.* **2019**, *44*, 453–466.
- [179] L. Stefanis, *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a009399.
- [180] J. G. Verbalis, M. D. Drutarosky, *Kidney Int.* **1988**, *34*, 351–360.
- [181] H. Pasantes-Morales, R. Franco, B. Ordaz, L. D. Ochoa, *Arch. Med. Res.* **2002**, *33*, 237–244.

- [182] E. Syková, *Neuroscience* **2004**, *129*, 861–876.
- [183] R. G. Thorne, C. Nicholson, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5567–5572.
- [184] S. K. Fisher, A. M. Heacock, R. F. Keep, D. J. Foster, *J. Physiol.* **2010**, *588*, 3355–3364.
- [185] K. Tuz, C. Peña-Segura, R. Franco, H. Pasantes-Morales, *Eur. J. Neurosci.* **2004**, *19*, 916–924.
- [186] A. M. Heacock, M. S. Dodd, S. K. Fisher, *J. Pharmacol. Exp. Ther.* **2006**, *317*, 685–693.
- [187] S. K. Fisher, T. A. Cheema, D. J. Foster, A. M. Heacock, *J. Neurochem.* **2008**, *106*, 1998–2014.
- [188] J. E. Olson, E. J. Martinho, *J. Neurochem.* **2006**, *96*, 1375–1389.
- [189] L. Rosso, B. Peteri-Brunbäck, P. Poujeol, N. Hussy, J.-M. Mienville, *J. Physiol.* **2004**, *554*, 731–742.
- [190] A. Wang, D. W. Bolen, *Biochemistry* **1997**, *36*, 9101–9108.
- [191] E. S. Courtenay, M. W. Capp, C. F. Anderson, M. T. J. Record, *Biochemistry* **2000**, *39*, 4455–4471.
- [192] P. H. Yancey, M. E. Clark, S. C. Hand, R. D. Bowlus, G. N. Somero, *Science* **1982**, *217*, 1214–1222.
- [193] S. N. Timasheff, *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67–97.
- [194] V. N. Uversky, J. Li, A. L. Fink, *FEBS Lett.* **2001**, *509*, 31–35.
- [195] V. Dall'Asta, O. Bussolati, R. Sala, A. Parolari, F. Alamanni, P. Biglioli, G. C. Gazzola, *Am. J. Physiol.* **1999**, *276*, C865-72.
- [196] E. Bajorunaite, J. Sereikaite, V.-A. Bumelis, *Protein J.* **2007**, *26*, 547–555.
- [197] U. Das, G. Hariprasad, A. S. Ethayathulla, P. Manral, T. K. Das, S. Pasha, A. Mann, M. Ganguli, A. K. Verma, R. Bhat, S. K. Chandrayan, S. Ahmed, S. Sharma, P. Kaur, T. P. Singh, A. Srinivasan, *PLoS One* **2007**, *2*, e1176.
- [198] W. Ariesandi, C.-F. Chang, T.-E. Chen, Y.-R. Chen, *PLoS One* **2013**, *8*, e53487.

- [199] P. Kushwaha, N. P. Prabhu, *New J. Chem.* **2022**, 46, 11082–11094.
- [200] J. Burré, S. Vivona, J. Diao, M. Sharma, A. T. Brunger, T. C. Südhof, *Nature* **2013**, 498, E4-6; discussion E6-7.
- [201] J. Kim, *Mol. Cells* **1997**, 7, 78–83.
- [202] P. H. Weinreb, W. Zhen, A. W. Poon, K. A. Conway, P. T. Lansbury, *Biochemistry* **1996**, 35, 13709–13715.
- [203] S. Ghosh, A. Kundu, K. Chattopadhyay, *Sci. Rep.* **2018**, 8, 5481.
- [204] G. Chiellini, *Int. J. Mol. Sci.* **2021**, 22, 6620.
- [205] M. R. Ajmal, *Dis. (Basel, Switzerland)* **2023**, 11, 30.
- [206] T. C. T. Michaels, L. X. Liu, G. Meisl, T. P. J. Knowles, *J. Phys. Condens. Matter* **2017**, 29, 153002.
- [207] C. Xue, T. Y. Lin, D. Chang, Z. Guo, *R. Soc. open Sci.* **2017**, 4, 160696.
- [208] J. M. Yoo, Y. Lin, Y. Heo, Y.-H. Lee, *Front. Mol. Biosci.* **2022**, 9, 959425.
- [209] S. Mehra, L. Gadhe, R. Bera, A. S. Sawner, S. K. Maji, *Biomolecules* **2021**, 11, 1419.
- [210] M. Schweighauser, Y. Shi, A. Tarutani, F. Kametani, A. G. Murzin, B. Ghetti, T. Matsubara, T. Tomita, T. Ando, K. Hasegawa, S. Murayama, M. Yoshida, M. Hasegawa, S. H. W. Scheres, M. Goedert, *Nature* **2020**, 585, 464–469.
- [211] D. R. Boyer, B. Li, C. Sun, W. Fan, K. Zhou, M. P. Hughes, M. R. Sawaya, L. Jiang, D. S. Eisenberg, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, 117, 3592–3602.
- [212] F. Longhena, G. Faustini, V. Brembati, M. Pizzi, A. Bellucci, *IUBMB Life* **2020**, 72, 590–600.
- [213] E. Alonso, M. A. García-Pérez, J. Bueso, V. Rubio, *Neurochem. Res.* **1991**, 16, 787–794.
- [214] R. Cascella, A. Bigi, N. Cremades, C. Cecchi, *Cell. Mol. Life Sci.* **2022**, 79, 174.
- [215] M. Verma, A. Vats, V. Taneja, *Ann. Indian Acad. Neurol.* **2015**, 18, 138–145.
- [216] J. Vaneyck, I. Segers-Nolten, K. Broersen, M. M. A. E. Claessens, *J. Biol. Chem.* **2021**,

296, 100358.

- [217] S. Subedi, S. Sasidharan, N. Nag, P. Saudagar, T. Tripathi, *Molecules* **2022**, *27*, 1776.
- [218] T. Watanabe-Nakayama, M. Nawa, H. Konno, N. Kodera, T. Ando, D. B. Teplow, K. Ono, *ACS Nano* **2020**, *14*, 9979–9989.
- [219] E. Katorcha, N. Makarava, Y. J. Lee, I. Lindberg, M. J. Monteiro, G. G. Kovacs, I. V Baskakov, *PLOS Pathog.* **2017**, *13*, e1006563.
- [220] J. M. Gorell, B. A. Rybicki, C. Cole Johnson, E. L. Peterson, *Neuroepidemiology* **1999**, *18*, 303–308.
- [221] J. M. Gorell, C. C. Johnson, B. A. Rybicki, E. L. Peterson, G. X. Kortsha, G. G. Brown, R. J. Richardson, *Neurotoxicology* **1999**, *20*, 239–247.
- [222] G. Yamin, C. B. Glaser, V. N. Uversky, A. L. Fink, *J. Biol. Chem.* **2003**, *278*, 27630–27635.
- [223] V. N. Uversky, J. Li, A. L. Fink, *J. Biol. Chem.* **2001**, *276*, 44284–44296.
- [224] L. A. Munishkina, C. Phelan, V. N. Uversky, A. L. Fink, *Biochemistry* **2003**, *42*, 2720–2730.
- [225] V. Naik, J. Kardani, I. Roy, *Mol. Biotechnol.* **2016**, *58*, 251–255.
- [226] S. Maity, K. Shimanaka, L. N. Rivet, M. O’Dell, A. M. Rashid, N. B. M. Isa, R. S. Kepczynski, U. Dettmer, B. Borhan, J. S. Fortin, *J. Mol. Struct.* **2022**, *1249*, 131569.
- [227] M. D. Shtilerman, T. T. Ding, P. T. J. Lansbury, *Biochem. (American Chem. Soc.* **2002**, *55–60*.
- [228] V. N. Uversky, E. M Cooper, K. S. Bower, J. Li, A. L. Fink, *FEBS Lett.* **2002**, *515*, 99–103.
- [229] S. Biswas, A. Bhadra, S. Lakhera, M. Soni, V. Panuganti, S. Jain, I. Roy, *Eur. Biophys. J. Ebj* **2021**, *50*, 59–67.
- [230] C. Del Mar, E. A. Greenbaum, L. Mayne, S. W. Englander, V. L. J. Woods, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15477–15482.

- [231] M. G. Spillantini, R. A. Crowther, R. Jakes, N. J. Cairns, P. L. Lantos, M. Goedert, *Neurosci. Lett.* **1998**, *251*, 205–208.
- [232] P. Radivojac, L. M. Iakoucheva, C. J. Oldfield, Z. Obradovic, V. N. Uversky, A. K. Dunker, *Biophys. J.* **2007**, *92*, 1439–1456.
- [233] K. K. Turoverov, I. M. Kuznetsova, V. N. Uversky, *Prog. Biophys. Mol. Biol.* **2010**, *102*, 73–84.
- [234] Z. Lv, A. V Krasnoslobodtsev, Y. Zhang, D. Ysselstein, J. C. Rochet, S. C. Blanchard, Y. L. Lyubchenko, *Biopolymers* **2016**, *105*, 715–724.
- [235] B. Ahmad, Y. Chen, L. J. Lapidus, *Proc. Natl. Acad. Sci.* **2012**, *109*, 2336–2341.
- [236] D. R. Canchi, A. E. García, *Annu. Rev. Phys. Chem.* **2013**, *64*, 273–293.
- [237] H.-X. Zhou, G. Rivas, A. P. Minton, *Annu. Rev. Biophys.* **2008**, *37*, 375–397.
- [238] T. O. Street, D. W. Bolen, G. D. Rose, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13997–14002.
- [239] M. Gao, C. Held, S. Patra, L. Arns, G. Sadowski, R. Winter, *Chemphyschem* **2017**, *18*, 2951–2972.
- [240] S. N. Timasheff, *Biochemistry* **2002**, *41*, 13473–13482.
- [241] G. König, S. Bruckner, S. Boresch, *Biophys. J.* **2013**, *104*, 453–462.
- [242] D. Shukla, B. L. Trout, *J. Phys. Chem. B* **2011**, *115*, 1243–1253.
- [243] T. O. Street, D. W. Bolen, G. D. Rose, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13997–14002.
- [244] S. Qin, H.-X. Zhou, *Biophys. J.* **2009**, *97*, 12–19.
- [245] S. Dwivedi, *Open Nutraceuticals J.* **2010**, *3*, 213–219.
- [246] D. R. Canchi, D. Paschek, A. E. García, *J. Am. Chem. Soc.* **2010**, *132*, 2338–2344.
- [247] F.-F. Liu, L. Ji, L. Zhang, X.-Y. Dong, Y. Sun, *J. Chem. Phys.* **2010**, *132*, 225103.
- [248] J. A. GORDON, W. P. JENCKS, *Biochemistry* **1963**, *2*, 47–57.

- [249] C. Y. Hu, H. Kokubo, G. C. Lynch, D. W. Bolen, B. M. Pettitt, *Protein Sci.* **2010**, *19*, 1011–1022.
- [250] G. Borgohain, S. Paul, *J. Mol. Liq.* **2017**, *231*, 174–184.
- [251] Q. Zou, B. J. Bennion, V. Daggett, K. P. Murphy, *J. Am. Chem. Soc.* **2002**, *124*, 1192–1202.
- [252] A. Rani, P. Venkatesu, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23151.
- [253] I. Jahan, S. M. Nayeem, *Biophys. J.* **2019**, *117*, 1922–1934.
- [254] F. Chiti, C. M. Dobson, *Annu. Rev. Biochem.* **2017**, *86*, 27–68.
- [255] M. C. Owen, D. Gnutt, M. Gao, S. K. T. S. Wärmländer, J. Jarvet, A. Gräslund, R. Winter, S. Ebbinghaus, B. Strodel, *Chem. Soc. Rev.* **2019**, *48*, 3946–3996.
- [256] R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, D. E. Shaw, *Annu. Rev. Biophys.* **2012**, *41*, 429–452.
- [257] M. Carballo-Pacheco, B. Strodel, *J. Phys. Chem. B* **2016**, *120*, 2991–2999.
- [258] A. N. Muttathukattil, G. Reddy, *J. Phys. Chem. B* **2016**, *120*, 10979–10989.
- [259] V. T. Duong, Z. Chen, M. T. Thapa, R. Luo, *J. Phys. Chem. B* **2018**, *122*, 10455–10469.
- [260] S. Jamal, A. Kumari, A. Singh, S. Goyal, A. Grover, *Front. Neurosci.* **2017**, *11*, 684.
- [261] Z. A. Levine, L. Larini, N. E. LaPointe, S. C. Feinstein, J.-E. Shea, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 2758–2763.
- [262] K. Vanommeslaeghe, E. P. Raman, A. D. J. MacKerell, *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.
- [263] K. Vanommeslaeghe, A. D. J. MacKerell, *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.
- [264] H. Zhang, Y. Jiang, Z. Cui, C. Yin, *J. Chem. Inf. Model.* **2018**, *58*, 1669–1681.
- [265] J. G. Kirkwood, F. P. Buff, *jcp* **1951**, *19*, 774–777.
- [266] D. Eliezer, *J. Mol. Biol.* **2013**, *425*, 2393–2396.

- [267] M.-C. Bellissent-Funel, K. F. Bradley, S. H. Chen, J. Lal, J. Teixeira, *Phys. A Stat. Mech. its Appl.* **1993**, *201*, 277–285.
- [268] H. Hofmann, D. Nettels, B. Schuler, *J. Chem. Phys.* **2013**, *139*, 121930.
- [269] K. T. Naidu, N. P. Prabhu, *J. Biomol. Struct. Dyn.* **2022**, *40*, 820–832.
- [270] V. Vagenende, A. X. Han, M. Mueller, B. L. Trout, *ACS Chem. Biol.* **2013**, *8*, 416–422.
- [271] D. Shukla, B. L. Trout, *J. Phys. Chem. B* **2010**, *114*, 13426–13438.
- [272] C. P. Schneider, B. L. Trout, *J. Phys. Chem. B* **2009**, *113*, 2050–2058.
- [273] W. Paslawski, S. Mysling, K. Thomsen, T. J. D. Jørgensen, D. E. Otzen, *Angew. Chemie* **2014**, *126*, 7690–7693.
- [274] J. H. M. van Gils, E. van Dijk, A. Peduzzo, A. Hofmann, N. Vettore, M. P. Schützmann, G. Groth, H. Mouhib, D. E. Otzen, A. K. Buell, S. Abeln, *PLoS Comput. Biol.* **2020**, *16*, e1007767.
- [275] T. S. Choi, J. Y. Han, C. E. Heo, S. W. Lee, H. I. Kim, *Biochim. Biophys. Acta - Biomembr.* **2018**, *1860*, 1854–1862.

List of Publications

1. **Saurabh, A.** and Prabhu, N. P. (2022). Concerted enhanced-sampling simulations to elucidate the helix-fibril transition pathway of intrinsically disordered α -Synuclein. *International Journal of Biological Macromolecules*, 223, 1024-1041.
<https://doi.org/10.1016/j.ijbiomac.2022.11.079>
2. Ghosh, S.*, **Saurabh, A.*** and Prabhu, N. P. (2022). Spectroscopic studies on the stability and nucleation-independent fibrillation of partially-unfolded proteins in crowded environment. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 264, 120307. (*equal contribution)
<https://doi.org/10.1016/j.saa.2021.120307>
3. Singh, D., Patri, S., Narahari, V., Segireddy, R. R., Dey, S., **Saurabh, A.**, Vijay, M., Prabhu, N.P., Srivastava, A., Kolli, S.K. and Kumar, K. A. (2022). A conserved Plasmodium structural integrity maintenance protein (SIMP) is associated with sporozoite membrane and is essential for maintaining shape and infectivity. *Molecular Microbiology*, 117, 1324-1339.
<https://doi.org/10.1111/mmi.14894>
4. **Saurabh, A.**, Chauhan, S., Jogadhenu, P.S.S. and Prabhu, N. P., Mechanistic insight of the human α -synuclein fibrils formation in the presence of charges amino acid and its derivatives: spectroscopic and computational analysis. (Manuscript under preparation).

Unraveling the Intricate Structural Transitions in α -Synuclein: Conformations Governing Fibril Formation and the Influence of Osmolytes

by Archi Saurabh



Librarian
Indira Gandhi Memorial Library
UNIVERSITY OF HYDERABAD
Central University P.O.
HYDERABAD-500 046.

Submission date: 03-Jul-2023 10:35AM (UTC+0530)

Submission ID: 2125860507

File name: Archi_Saurabh.pdf (7.99M)

Word count: 28612

Character count: 151574

Unraveling the Intricate Structural Transitions in α -Synuclein: Conformations Governing Fibril Formation and the Influence of Osmolytes

ORIGINALITY REPORT

46%

SIMILARITY INDEX

4%

INTERNET SOURCES

45%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1

Archi Saurabh, N. Prakash Prabhu. "Concerted enhanced-sampling simulations to elucidate the helix-fibril transition pathway of intrinsically disordered α -Synuclein", International Journal of Biological Macromolecules, 2022

Publication

40%

N. Prakash

N. PRAKASH PRABHU PhD
Associate Professor,
Dept. Biotechnology & Bioinformatics
School of Life Sciences
University of Hyderabad
Hyderabad-500 046

2

Submitted to University of Hyderabad,
Hyderabad

Student Paper

1%

3

Bramhini Anumalla, N. Prakash Prabhu. "Surface hydration and preferential interaction directs the charged amino acids-induced changes in protein stability", Journal of Molecular Graphics and Modelling, 2020

Publication

1%

4

www.researchgate.net
Internet Source

<1%

5

90thsbciauh.live

<1 %

6

Subhasree Ghosh, Archi Saurabh, N. Prakash Prabhu. "Spectroscopic studies on the stability and nucleation-independent fibrillation of partially-unfolded proteins in crowded environment", Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022
Publication

<1 %

7

www.pnas.org
Internet Source

<1 %

8

www.veteranstoday.com
Internet Source

<1 %

9

Pratibha Kushwaha, N. Prakash Prabhu. "Imidazolium-based ionic liquids with increasing alkyl chain length of cations decrease the stability and fibrillation propensity of lysozyme", New Journal of Chemistry, 2022
Publication

<1 %

10

www.papercamp.com
Internet Source

<1 %

11

PillaSankara Krishna, BalagaRadha Rani, M. Karthik Mohan, Iwane Suzuki, Sisinthy Shivaji, JogadhenuS. S. Prakash. " A novel transcriptional regulator, SII1130, negatively

<1 %

regulates heat-responsive genes in sp.
PCC6803 ", Biochemical Journal, 2013

Publication

12

Pratibha Kushwaha, N. Prakash Prabhu.
"Imidazolium-based ionic liquids with
increasing alkyl chain length of cation
decrease the stability and fibrillation
propensity of lysozyme", New Journal of
Chemistry, 2022

Publication

<1 %

13

hdl.handle.net

Internet Source

<1 %

14

A. Clay Clark. "Protein folding: Are we there
yet?", Archives of Biochemistry and
Biophysics, 2008

Publication

<1 %

15

www.science.gov

Internet Source

<1 %

16

Orly Ullman, Charles K. Fisher, Collin M. Stultz.
"Explaining the Structural Plasticity of α -
Synuclein", Journal of the American Chemical
Society, 2011

Publication

<1 %

17

Naghmeh Saadati-Eskandari, Latifeh
Navidpour, Parichehreh Yaghmaei, Azadeh
Ebrahim-Habibi. "Amino Acids as Additives
against Amorphous Aggregation: In Vitro and

<1 %

In Silico Study on Human Lysozyme", Applied Biochemistry and Biotechnology, 2019

Publication

18

www.mdpi.com

Internet Source

<1 %

19

en.wikipedia.org

Internet Source

<1 %

20

jcp.edpsciences.org

Internet Source

<1 %

21

Submitted to University of Nottingham

Student Paper

<1 %

22

Lingaswamy Bantu, Suraj Chauhan, Afshan Srikumar, Yoshihisa Hirakawa, Iwane Suzuki, Martin Hagemann, Jogadhenu S.S. Prakash. "A membrane-bound cAMP receptor protein, SyCRP1 mediates inorganic carbon response in Synechocystis sp. PCC 6803", Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2022

Publication

<1 %

23

Submitted to Spring Branch Independent School District

Student Paper

<1 %

24

Manjeet Kumar, Magdalena I Ivanova, Ayyalusamy Ramamoorthy. " Non-micellar ganglioside GM1 induces an instantaneous conformational change in A β leading to the

<1 %

modulation of the peptide amyloid-fibril
pathway ", Cold Spring Harbor Laboratory,
2023

Publication

25

Shayon Bhattacharya, Liang Xu, Damien
Thompson. "Molecular Simulations Reveal
Terminal Group Mediated Stabilization of
Helical Conformers in Both Amyloid- β 42 and
 α -Synuclein", ACS Chemical Neuroscience,
2019

Publication

<1 %

26

epub.uni-regensburg.de
Internet Source

<1 %

27

"Computer Simulations of Aggregation of
Proteins and Peptides", Springer Science and
Business Media LLC, 2022

Publication

<1 %

28

Submitted to W. P. Davidson High School
Student Paper

<1 %

29

atrium.lib.uoguelph.ca
Internet Source

<1 %

30

www.tandfonline.com
Internet Source

<1 %

31

Izawa, Yasutaka, Hironobu Tateno, Hiroshi
Kameda, Kazuya Hirakawa, Keiko Hato,
Hisashi Yagi, Kunihiro Hongo, Tomohiro
Mizobata, and Yasushi Kawata. "Role of C-

<1 %

terminal negative charges and tyrosine residues in fibril formation of α -synuclein", Brain and Behavior, 2012.

Publication

32

Tingting Zhang, Yuanxin Tian, Zhonghuang Li, Siming Liu, Xiang Hu, Zichao Yang, Xiaotong Ling, Shuwen Liu, Jiajie Zhang. "Molecular Dynamics Study to Investigate the Dimeric Structure of the Full-Length α -Synuclein in Aqueous Solution", Journal of Chemical Information and Modeling, 2017

Publication

<1 %

33

aip.scitation.org

Internet Source

<1 %

34

theses-test.ncl.ac.uk

Internet Source

<1 %

35

Submitted to University of Nebraska, Lincoln

Student Paper

<1 %

36

patents.google.com

Internet Source

<1 %

37

rcastoragev2.blob.core.windows.net

Internet Source

<1 %

38

scholars.cityu.edu.hk

Internet Source

<1 %

39

Ioana M. Ilie, Amedeo Caflisch. "Simulation Studies of Amyloidogenic Polypeptides and

<1 %

Their Aggregates", Chemical Reviews, 2019

Publication

40

Submitted to Monash University

Student Paper

<1 %

Exclude quotes On

Exclude matches < 14 words

Exclude bibliography On