Intrinsically unstructured proteins and neurodegenerative diseases: Conformational promiscuity at its best

ARTICLE *in* INTERNATIONAL UNION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY LIFE · JULY 2011

Impact Factor: 3.14 · DOI: 10.1002/iub.498 · Source: PubMed

2 AUTHORS:

Samir Das
University of British Columbia - Vancouver
10 PUBLICATIONS **38** CITATIONS

SEE PROFILE

Debashis Mukhopadhyay
Saha Institute of Nuclear Physics
45 PUBLICATIONS **407** CITATIONS

SEE PROFILE

IUBMB *Life*, 63(7): 478–488, July 2011

Critical Review

Intrinsically Unstructured Proteins and Neurodegenerative Diseases: Conformational Promiscuity at its Best

Samir Das and Debashis Mukhopadhyay
Structural Genomics Division, Saha Institute of Nuclear Physics, 1/AF Bidhan Nagar, Kolkata, WB, India

Summary

Neurodegenerative diseases are complex, multifactorial disorders where misfolding of proteins cause aberrant protein–protein interactions. They are not usually characterized by specific mutations especially for nonfamilial disease types. Most of the causative proteins, however, are intrinsically unstructured (IUP), loss of whose fine balance could play pivotal role in these processes. Very fast conformational switch of these IUPs between different functional forms, so as to choose different interaction partners and different functional niches within the cell, is the basic premise on which these proteins maintain their interaction network. We are working on the hypothesis that even small perturbations in conformation leads to disruption of the network and to the disease phenotype. Based on a comprehensive data search, the evidence was obtained for the role of IUPs in neurodegenerative disorders, and their mode of action through conformational promiscuity is elaborated through three case studies.

IUBMB *Life*, 63(7): 478–488, 2011

Keywords IUP; neurodegenerative diseases; conformational switch; Parkinson's disease; Alzheimer's disease; Huntington's disease.

INTRODUCTION

Neurodegeneration is characterized by slow progression, which is barely recognizable before the patient develops any symptoms. For example, in Alzheimer's disease (AD) patients, any recognizable symptoms are evident only after 10–20 yr of brain damage, whereas the Parkinson's disease (PD) is manifested after the loss of more than 70% dopaminergic neurons in the *substantia nigra*. Despite a wide range of clinical symptoms, different neurodegenerative diseases (NDDs) commonly originate from the dysfunction of different regions of central and peripheral nervous system (*1*). Additionally, although there are differences in components and machineries, specific protein aggregation and misfolding triggers a cascade of events including oxidative damage, mitochondrial dysfunction, impaired bioenergetics, and disruption of Golgi apparatus and transport eventually leading to NDDs. It is well established, therefore, that formation of aggregates and plaques of various proteins, in association with failure of neuronal networks, are hallmarks of different neurodegenerative disorders [NDDs; (*2*)]. Occasionally, the aggregates form definitive structures. For AD itself, existence of globular or oligomeric intermediates of different Aβ fibrils have been shown independently by different groups. β-Sheet-like structures connected by β-turns have been found both in polyglutamine (Gln) tracks, characteristic of Huntington disease (HD), as well as in Aβ peptides indicating the commonality between the amyloid forms (*3*). Structural studies have revealed that the N-termini of Prion proteins are highly unstructured while they undergo conformational transformations in plaques where β-sheets mostly prevail (*4*). However, a number of subtypes of NDDs are characterized by the absences of such specific aggregates such as Gerstmann–Straussler syndrome. Moreover, amyloid plaques have been found throughout the cortex of 70-years old without any neurodegenerative consequences (*2*). Therefore, it is relevant to hypothesize those additional structural features, or the mere lack of it might be responsible for different NDDs.

Intrinsically unstructured proteins (IUPs) are evolutionarily selected in more complex eukaryotic systems where multiple interactions with distinctly related binding partners play essential roles in cellular signaling events. IUPs are extremely flexible, devoid of any compact globular fold and have little or no ordered secondary structures under physiological conditions (*5*). The disordered proteins or flexible regions of a given "folded protein" often undergo totally different folding paradigm while binding to different partners (*6*). Conformational switching or structural reorganization on binding to different interactors has

Received 11 March 2011; accepted 19 April 2011
Address correspondence to: Debashis Mukhopadhyay, Structural Genomics Division, Saha Institute of Nuclear Physics, 1/AF, Bidhan Nagar, West Bengal, Kolkata 700 064, India. Tel: +9133-2337-5345-49. Fax: +9133-2337-4637. E-mail: debashis.mukhopadhyay@saha.ac.in

ISSN 1521-6543 print/ISSN 1521-6551 online
DOI: 10.1002/iub.498

been characterized for IUPs (*7*). Thirty percent of eukaryotic proteins are composed of either completely disordered polypeptides or partially unfolded polypeptides, and their prevalence increases with the complexity of organisms (*8*). In the case of globular protein, "hydrophobic collapse" is the guiding principle to create the hydrophobic core of the protein. However, IUPs lack these "order promoting" hydrophobic amino acids and are characterized by prevalence of hydrophilic and charged residues resulting in high net charge and low mean hydrophobicity (*9*). Despite the thermodynamic cost of its existence, IUPs are evolutionarily selected and prevail mostly in higher organisms, especially in proteins, which evolved early, and, therefore, having greater functional significance (*10*). It was pointed out that expansions of internal repeats such as minisatellites or microsatellites might be responsible for the origin of IUPs. As an example, Prion proteins, associated with different encephalopathies, have been found to contain unstructured N-terminal regions with an octapeptide repeat (*11*).

The main advantage of IUPs comes from structural flexibility, which enables them to bind to an ensemble of partners (*12*). One of the key facts of IUPs is that their "unstructuredness" increases the size of intermolecular surfaces (*13*). Monomeric proteins with similar intermolecular surfaces such as IUPs would have two to three times larger size and that would increase molecular crowding or volume of the cell. Additional advantages come from exposed recognition elements that fold on binding, accessible post translational modification sites, and presence of short linear interaction motifs (*14*). As a result, IUPs evolve to perform more than one functions and inactivation of one of them would not generally affect its other functions (*15*) providing some amount of "robustness" of action. Further evidence for this comes from the fact that of the 30% eukaryotic proteins, which are intrinsically disordered, a majority of them are acting as "hubs" to play a key role in protein–protein interaction (PPI) networks. As the removal of a "hub" protein from the PPI network causes disruption of the cellular system, one can speculate that the "unstructuredness" of these "hub" proteins are also evolutionarily selected, increasing the complexity of the PPI networks. In human PPI networks, the occurrences of IUPs are higher when compared with compactly folded proteins (*16*).

These functional benefits entail further pressure on the cell in terms of abundance and half life of the IUPs (*14*). Higher abundance may simply lead to mass action driven interaction promiscuity, resulting in artifactual interactions of no or unscrupulous biological relevance (*17*). Cells, therefore, evolved robust mechanisms to control transcription, translation, post-translational modifications, and degradation of IUPs. Cells take best possible ways to reduce the number of IUPs. The length of polyadenylate [poly(A)] tail as well as half lives of mRNA transcripts, those encode IUPs, are much smaller compared to the mRNA transcripts encoding more structured proteins. The abundance of IUPs in cellular system is also very low when compared with the structured proteins. Protein synthesis and protein half-lives are shorter for unstructured proteins than that of the properly folded ones (*14*). IUPs are also selectively degraded by default through a ubiquitin-independent 20S proteosome machinery (*18*). It is unlikely, therefore, that loss of this fine regulation at any level could influence the disease mechanism. In this review, we are going to estimate the effects of protein misfolding or aggregation in NDDs from the perspective of IUPs.

IUPs AND NEURODEGENERATIVE DISEASES

NDDs such as AD, HD, PD, and various Prion diseases such as Creutzfeld–Jacob disease are characterized by accumulation of misfolded protein aggregates in brain and other tissues. Previously, we have shown that PPI network of three most common forms of NDDs (HD, PD, and AD) are composed of proteins that are predominantly unstructured (*19*). To enhance the scope of this review, we did a search of all the human NDDs in Online Mendelian Inheritance in Man, OMIM (http://www.ncbi.nlm.nih.gov/omim), with the search code "neurodegenerative disease," which gave 251 NDDs. After removal of redundancies (*e.g.*, multiple familial forms or variants) "OMIM" search identified 52 unique hits. Manual checking of all these diseases with available literature revealed that although all of these diseases were "neurological" in nature, many of them failed to satisfy the stringent criteria of neurodegeneration at the protein level. Our literature search identified 29 unique NDDs among all the OMIM hits and annotated 98 unique proteins involved in these disease pathways. We checked the unstructuredness of these 98 proteins using "FoldIndex" (*20*). Classification was made in regard to "FoldIndex" score as well as by the occurrence of 30 unstructured amino acids at a stretch (*19*). About 83% of these proteins turned out to be IUPs (Table 1) as opposed to about half of the human proteome as reported elsewhere (*21*). Neurofilament heavy (NEFH) polypeptide, a protein involved in amyotrophic lateral sclerosis, consists of a stretch of as long as 585 disorder promoting amino acids and microtubule-associated protein tau, involved in PD, Pick's disease, and supranuclear palsy contains a 171-residue long-unstructured region, which represents 80% of all residues in the protein being the highest percentage in our finding.

At the backdrop of statistically significant prevalence of IUPs in NDDs, protein conformational switching emerges as a significant molecular event in higher eukaryotes. Correct switching is likely to be important for cellular homeostasis, whereas failures in appropriate switching are likely to lead to pathogenesis. This is illustrated in the following three cases taken from commonest NDDs.

Case Study I: Htt in Huntington Disease

HD, an autosomal dominant neurodegenerative disease, is characterized by coding for glutamine repeats of more than 36 in exon 1 of Huntington. The Huntingtin protein (Htt) with elongated polyGln tracts in excess of 36 repeats, misfolds and aggregates as antiparallel β strands (*22*), and the occurrence of

Table 1
Intrinsically unstructured proteins associated with 29 different neurodegenerative diseases, retrieved from OMIM database, are presented here

S. No.	Disease name	Protein name	Length of longest disordered region (no. of amino acids)	Position	% Disordered
1.	Alzheimer disease (AD)	Amyloid Precursor's Protein (APP)	112	D177-R288	55.1
		Presenilin-1 (PSEN1)	71	L295-S365	29.6
		Apolipoprotein E (APOE)	57	A217-E273	54.6
		Amyloid beta A4 precursor protein-binding family B member 2 (APBB2)	201	R108-W308	48.9
		Nitric oxide synthase, endothelial (NOS3)	41	S594-S634	26.1
		Presenilin-2 (PSEN2)	74	M1-G74	29.5
2.	Creutzfeldt–Jakob disease (CJD)	Major prion protein (PRNP)	81	P26-K106	58.9
3.	Huntington disease (HD)	Huntingtin (HTT)	129	P499-I627	14.9
		Junctophilin-3 (JPH3)	144	G538-L681	70.1
		Major prion protein (PRNP)	81	P26-K106	58.9
		Apolipoprotein L1 (APOL1)	49	D109-R157	26.1
4.	Menkes disease	Copper-transporting ATPase 1 (ATP7A)	23	K243-G265	2.6
5.	DEMENTIA, Lewy body (DLB)	Alpha-synuclein (SNCA)	44	K97-A140	31.4
6.	Ceroid lipofuscinosis, neuronal	Cathepsin D (CTSD)	–	–	–
		Major facilitator superfamily domain-containing protein 8 (MFSD8)	29	M1-E29	7.5
		Palmitoyl-protein thioesterase 1 (PPT1)	12	H187-H198	9.8
		Tripeptidyl-peptidase 1 (TPP1)	15	S37-F51	7.1
		Battenin (CLN3)	13	S247-P259	3
		Ceroid-lipofuscinosis neuronal protein 6 (CLN6)	–	–	–
7.	Parkinson disease (PD)	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCHL1)	35	L118-C152	18.8
		Alcohol dehydrogenase 1C (ADH1C)	25	R102-D126	6.7
		Microtubule-associated protein tau (MAPT)	171	P439-G606	80.1
		TATA-box-binding protein (TBP)	73	T39-T111	24.8
8.	Adrenoleukodystrophy (ALD)	Alpha-synuclein (SNCA)	44	K97-A140	31.4
		ATP-binding cassette sub-family D member 1 (ABCD1)	24	M566-L589	5.8
		Peroxisome biogenesis factor 1 (PEX1)	59	T1225-A1283	20.3
		Peroxisome biogenesis factor 10 (PEX10)	41	S231-P237	19.9
		Peroxisomal membrane protein PEX13 (PEX13)	39	M1-L39	35.3
		Peroxisome assembly protein 26 (PEX26)	25	R192-A216	8.2
9.	Machado-Joseph disease (MJD)	Ataxin-3 (ATXN3)	122	A197-R318	52.8
10.	Myoclonic epilepsy	Gamma-aminobutyric acid receptor subunit gamma-2 (GABRG2)	55	S366-C420	21.5
		EF-hand domain-containing protein 1 (EFHC1)	60	Y523-A582	29
11.	Gerstmann–Straussler disease (GSD)	Major prion protein (PRNP)	81	P26-K106	58.9

Table 1
Contitnued

S. No.	Disease name	Protein name	Length of longest disordered region (no. of amino acids)	Position	% Disordered
12.	Charcot–Marie-tooth disease, AXONAL	Heat shock protein beta-1 (HSPB1)	32	V101-D132	23.4
		Ganglioside-induced differentiation-associated protein 1 (GDAP1)	71	T167-P237	36.3
		Heat shock protein beta-8 (HSPB8)	29	M1-M29	31.1
		Dynamin-2 (DNM2)	57	Q283-D339	27.8
		Alanyl-tRNA synthetase, cytoplasmic (AARS)	52	K153-V204	11.1
13.	Frontotemporal dementia (FTD)	Transitional endoplasmic reticulum ATPase (VCP)	110	L697-G806	25.3
14.	Pick disease	Microtubule-associated protein tau (MAPT)	171	P439-G606	80.1
		Presenilin-1 (PSEN1)	71	L295-S365	29.6
15.	Amyotrophic lateral sclerosis-Parkinsonism/dementia Complex 1	Transient receptor potential cation channel subfamily M member 7 (TRPM7)	91	T523-P613	26.4
16.	Neurodegeneration with brain iron accumulation	Ferritin light chain (FTL)	33	E57-E89	21.7
		Pantothenate kinase 2, mitochondrial (PANK2)	63	N56-Q118	27
		85 kDa calcium-independent phospholipase A2 (PLA2G6)	87	R550-S636	11.4
17.	GM2-Gangliosidosis, AB VARIANT	Ganglioside GM2 activator (GM2A)	–	–	–
18.	Agenesis of the corpus callosum with peripheral neuropathy (ACCPN)	Solute carrier family 12 member 6 (SLC12A6)	110	M60-A169	21
19.	Amyotrophic lateral sclerosis (ALS)	TAR DNA-binding protein 43 (TARDBP)	50	Y155-E204	35.3
		Polyphosphoinositide phosphatase (FIG4)	63	N731-N793	25.1
		Probable helicase senataxin (SETX)	284	E1136-S1419	52.1
		RNA-binding protein FUS (FUS)	303	M1-D303	90.3
		Vesicle-associated membrane protein-associated protein B/C (VAPB)	42	H86-A127	51.4
		Angiopoietin-4 (ANGPT4)	68	K317-D384	41
		Charged multivesicular body protein 2b (CHMP2B)	92	K81-I172	83.6
		Superoxide dismutase [Cu-Zn] (SOD1)	37	T55-D91	24
		ALS2 C-terminal-like protein (ALS2CL)	61	K433-H493	27.9
		Dynactin subunit 1 (DCTN1)	344	L220-K563	43.7
		Neurofilament heavy polypeptide (NEFH)	585	E442-K1026	74.9
		Peripherin-2 (PRPH2)	43	E304-G346	20.5
20.	Pontocerebellar hypoplasia	Serine/threonine-protein kinase VRK1 (VRK1)	53	F86-L138	53.3
		tRNA-splicing endonuclease subunit Sen54 (TSEN54)	83	L323-V405	41.6
		tRNA-splicing endonuclease subunit Sen2 (TSEN2)	113	K90-T202	41.5

Table 1
Contitnued

S. No.	Disease name	Protein name	Length of longest disordered region (no. of amino acids)	Position	% Disordered
		tRNA-splicing endonuclease subunit Sen34 (TSEN34)	85	R88-S172	40.3
		Probable arginyl-tRNA synthetase, mitochondrial (RARS2)	8	N54-I61	2.4
21.	Leukodystrophy, hypomyelinating	Gap junction gamma-2 protein (GJC2)	96	G96-T191	34.2
		60 kDa heat shock protein, mitochondrial (HSPD1)	14	L358-I371, L390-V403	5.8
		Hyccin (FAM126A)	109	I323-S431	24.8
22.	Frontotemporal lobar degeneration	Granulins (GRN)	53	W541-L593	12.3
		TAR DNA-binding protein 43 (TARDBP)	50	Y155-E204	35.3
23.	Spinocerebellar ataxia	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B beta isoform (PPP2R2B)	53	W97-R149	38.4
		Protein kinase C gamma type (PRKCG)	30	M1-V30	23.5
		Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)	90	V1867-A1956	30.1
		TATA-box-binding protein (TBP)	73	T39-T111	24.8
		AFG3-like protein 2 (AFG3L2)	82	S53-G134	18.1
		Protein BEAN1 (BEAN1)	146	I59-G204	58.7
		Ataxin-8 (ATXN8)	80	M1-Q80	100
		Putative Wiskott-Aldrich syndrome protein family member 4 (WASF4)	78	D269-K346	54.6
		Wiskott-Aldrich syndrome protein family member 3 (WASF3)	194	N115-P308	55
		Zinc finger protein 592 (ZNF592)	120	S936-F1055	49.3
		Nesprin-1 (SYNE1)	160	L6165-V6324	49.9
		Chaperone activity of bc1 complex-like, mitochondrial (CABC1)	51	M156-A206	19.3
		Tyrosyl-DNA phosphodiesterase 1 (TDP1)	166	M1-F166	36.2
		Ataxin-1 (ATXN1)	75	N741-K815	34.7
		Ataxin-10 (ATXN10)	31	A445-P475	8.6
		Tau-tubulin kinase 2 (TTBK2)	194	S1051-R1244	53.7
		Potassium voltage-gated channel subfamily C member 3 (KCNC3)	65	M1-R65	29.5
		Ataxin-2 (ATXN2)	302	A352-N653	57.8

Table 1
Contitnued

S. No.	Disease name	Protein name	Length of longest disordered region (no. of amino acids)	Position	% Disordered
		Fibroblast growth factor 14 (FGF14)	84	M1-Q84	52.6
		Spectrin beta chain, brain 2 (SPTBN2)	164	K2060-G2223	45.8
		Voltage-dependent P/Q-type calcium channel subunit alpha-1A (CACNA1A)	540	Y1966-C2505	48.9
		Ataxin-7 (ATXN7)	152	P303-R454	63.8
24.	Friedreich ataxia	Frataxin, mitochondrial (FXN)	47	A34-R80	43.3
25.	Supranuclear palsy	Microtubule-associated protein tau (MAPT)	171	P439-G606	80.1
26.	Leigh syndrome	Cytochrome c oxidase assembly protein COX15 homolog (COX15)	29	109K-137M	8.3
		NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2 (NDUFA2)	–	–	–
		Mitochondrial chaperone BCS1 (BCS1L)	24	P187-D210	12.4
		Dihydrolipoyl dehydrogenase, mitochondrial (DLD)	–	–	–
		NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial (NDUFS3)	36	E229-K264	24.6
		NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial (NDUFS4)	35	N141-K175	53.7
		NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial (NDUFS7)	–	–	–
		NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial (NDUFS8)	46	I165-R210	27.6
		NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial (NDUFV1)	45	F32-I76	22
		Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial (SDHA)	47	E587-G633	16.1
		Leucine-rich PPR motif-containing protein, mitochondrial (LRPPRC)	56	I987-A1042	10.4
		Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial (PDHA1)	59	V332-S390	29
		Surfeit locus protein 1 (SURF1)	27	E198-R224	12.3

Table 1
Contitnued

S. No.	Disease name	Protein name	Length of longest disordered region (no. of amino acids)	Position	% Disordered
27.	Wolfram syndrome	CDGSH iron-sulfur domain-containing protein 2 (CISD2)	17	K77-L93	12.6
		Wolframin (WFS1)	107	M1-G107	23.4
28.	Fragile x tremor/ataxia syndrome (FXTAS)	Fragile X mental retardation 1 protein (FMR1)	203	L430-P632	49.2
29.	Striatonigral degeneration	Nuclear pore glycoprotein p62 (NUP62)	110	L344-D453	27.8

Longest disordered region with disordered percentage as predicted using "FoldIndex" are tabulated. The initial and final amino acids of longest disordered regions with their positions are also mentioned with single letter code.

this repeat number solely determines the age at onset. Though the monomeric Gln is unstructured, aggregated ones contain β-sheet-like conformation that occurs through nucleation-dependent polymerization. Molecular simulation of the first 17 residues of Htt ($N17^{Htt}$) show that these regions can exist in two states: a single extended helix or a two-helix bundle. Both of these states exhibit a large hydrophobic surface, which compliments a region that contains an amphipathic surface (*23*). The crystal structure of exon1 of Htt-17Gln has identified many small secondary "structural modules" such as an amino terminal α-helix, followed by a poly 17Gln regions and a poly proline helix. This poly 17Gln region can switch among different conformations-like random coil, α-helix, and extended loop, depending on the conformations of the neighboring residues of the protein (*24*). Recent simulation and circular dichroism studies has shown that $N17^{Htt}$ and polyGln regions become unstructured depending on the length of polyGln. In monomeric conditions, the amphipathic $N17^{Htt}$ region interacts with the random coil of polyGln tracts and inhibits its affinity for aggregation (*25*). It has been hypothesized that in pathogenic conditions, the increased polyGln repeats, in tandem, also influence the length of the random coil, which finally promotes Htt aggregation and nonspecific interactions with other binding partners leading to aggregation, sequestration, and pathologic conditions (*24*).

A quick look at the functional distribution of Htt interactors, which are IUPs as well (*19*), points toward possible involvement of Htt in apparently unrelated functions-like cell proliferation, apoptosis, electron transport, carbohydrate metabolism, and so forth. Exposed linear motifs such as $N17^{Htt}$ or polyproline helix are reportedly involved in many of these interactions. But, above all, it indicates gain of function through dosage sensitivity under pathological conditions (*17*) and gross disruption of cellular homeostasis, as a result. Nonspecific structural features, resulting from genetic aberration, therefore, lie at the heart of this disease.

Case Study II: α-Synuclein in Parkinson's Disease

PD, the second most common form of NDDs, is clinically recognized as a consequence of intracellular proteinaceous inclusions, known as Lewy bodies, and Lewy neuritis formed due to aggregation of the presynaptic protein, α-Synuclein (α-Syn) (*26*). Moreover, patients with multiple system atrophy contain α-Syn fibrils aggregates, deposited in glial cytoplasmic inclusions. Interestingly in addition to PD, α-Syn is also involved in the pathogenesis of AD and is reported to cause aberrant synaptogenesis during AD development (*27*).

This 140 amino acids long molecule is devoid of any well-defined rigid structure but may either stay substantially unfolded, or adopt an amyloidogenic partially folded conformation, or fold into α-helical or β-structural species depending on the environment, a characteristic feature of IUPs. Several morphologically different types of aggregates, including oligomers, amorphous aggregates and amyloid-like fibrils are also formed by α-Syn (*28*). NMR studies have revealed that α-Syn can adopt a bipartite structure after it interacts with membranes: first N-terminal 95 residues form helical structure while the highly negatively charged C-terminal portion remains unfolded, enabling it to interact with other protein molecules (*29*). Within a membrane mimetic environment the helical 1–95 region of α-Syn forms two helical stretches, whereas residues 42–44 break the helical continuity (*30*). The negatively charged C-terminus of α-Syn could attract different cations like Ca^{2+} in cellular conditions and induce the shielding of negatively charged residues and promote aggregation of α-Syn (*31*). It has been found that the amino and carboxy termini of α-Syn were more

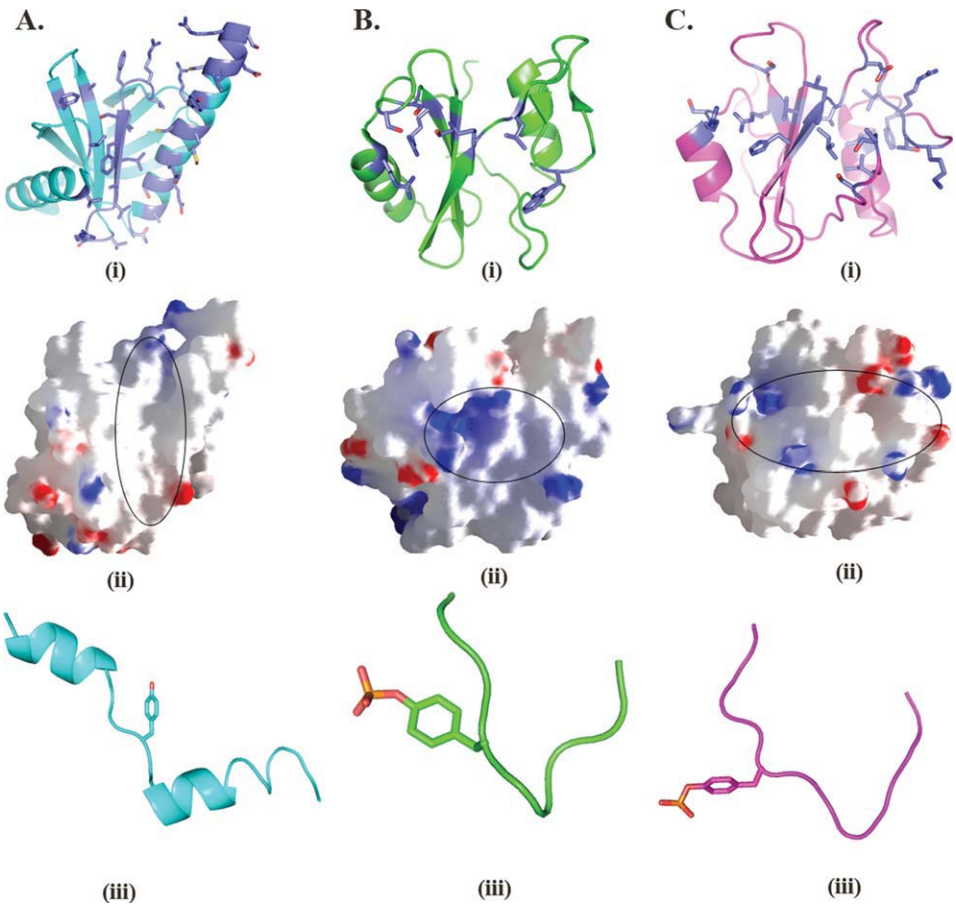

Figure 1. Structural organizations of different binding partners of AICD showing the need for its conformational switch. Cartoon diagrams (*45*) and surface representation (*46*) of three protein domains (A-ii) Fe65-PTB (PDB code: 3DXC.pdb), (B-ii) Grb2-SH2 (PDB code: 1JYR.pdb), and (C-ii) Shc-SH2 (PDB code: 1TCE.pdb) are shown. Positively and negatively charged residues are shown in blue and red, respectively. The binding clefts of these domains with their cognate peptides are marked. Residues that interact with their cognate binding partners are presented here as "stick" models. Larger binding cleft for Shc-SH2 domain compared to Grb2-SH2 domain is evident. The structures of cognate binding partners of these domains are shown in A-iii, B-iii, and C-iii. Specific tyrosines involved in this recognition are also shown. Distinct conformational variations of these peptides are apparent.

solvent exposed rather than the central part, which formed the core of α-Syn filaments, whereas the negatively charged carboxyl end inhibited the filament formation (*32*). Circular dichroism spectroscopy of α-Syn 108-140 fragment has revealed that though this portion was unfolded under the native conditions, it could form α-helical conformation in presence of 10 mM SDS. Limited proteolysis experiment has also revealed the existence of this conformational transition in the presence of SDS molecules (*31*). The conformation of α-Syn is also sensitive to pH and temperature as the high negative charge of α-Syn (pI 4.7) is neutralized in acidic pH leading to an increase in partially folded conformations with significant proportions of β sheet. The transition from unfolded to partially folded conformation take place between pH 5.5 and 3.0 indicating that one or more carboxylate group is responsible for this structural change, and this process is highly reversible. Low pH and elevated temperature can also cause partially folded intermediates to form that finally facilitate the fibril formation mostly driven by hydrophobic collapse. The low pH reduces the net negative charge with concomitantly minimal charge–charge intermolecular repulsion, whereas higher temperature favors additional hydrophobic interactions (*33*). Aggregation of α-Syn is also facilitated when Ser^{129} is phosphorylated, and this phosphorylation may have a role in the Lewy body formation in PD (*3*). These collectively are suggestive of the constant conformational transitions of highly unstructured α-Syn, influenced by the ambience to a large extent, in the pathways of PD.

Most of the IUPs in PD interactome (*19*) are involved in signal transduction in addition to cell cycle and other neuronal activities. Fast conformational switching of α-Syn would be definitely helpful for these processes. However, the way this protein responds to minute changes in the cellular

environment also contribute to its instability and disease condition in consequence.

Case Study III: AICD as an IUP in Alzheimer's Disease

Over the years, Aβ aggregation and the resulting extracellular plaques are considered as the primary causative agent of AD pathogenesis influenced to a large extent by environmental factors (*34*). The proteolytic machinery also generates APP intracellular C-terminal domain (AICD), which is liberated into the cytosol (*35*). In recent years, much of insights have been gained about this intrinsically unstructured proteolytic fragment AICD and its "adaptor" molecules (*36*). Transgenic mice overexpressing AICD develop an AD phenotype could originate from synergistic effect of both AICD and Aβ-deposits (*37*). Structural studies using multidimensional solution NMR spectroscopy have shown that AICD is devoid of any folded three-dimesional conformation over a broad range of pH while having only secondary structural elements such as N-terminal helix capping box, a type I β-turn and nascent helix (*38*).

AICD possesses several Ser, Thr, and Tyr residues with high potential of phosphorylation under different cellular circumstances. It is known that potential phosphorylation sites in eukaryotes belong to either fully or partially disordered regions (*39*). Phosphorylated Tyr-682 of AICD (as of APP-695 isoform) recruits SH2 domain containing different proteins such as Grb2 and Shc, which in turn trigger different signalling cascades (*40–42*). ShcA, ShcB, ShcC, Grb2, Grb7, Crk, and Nck bind to AICD when Tyr-682 is in the phosphorylated state and it is facilitated when Thr-668 is also phosphorylated. For different tyrosine kinases such as Abl, Lyn, and Src, this interaction occurs antagonistically, such that binding occurs only when Thr-668 is in phosphorylated state and facilitated in presence of Tyr-682 phosphorylation (*43*).

There is also minimal structural similarity between different domains that bind to AICD. For example, the N-terminal phosphotyrosine-binding (PTB) domain of Fe65 consists of seven antiparallel β-strands with two orthogonal β-sheets and three α helices [(*44*); Fig. 1A-i]. Another AICD interactor, Grb2-SH2 domain contains three antiparallel β-sheets, which are flanked by two α-helices and two short parallel β-strands [(*47*); Fig. 1B-i]. On the other hand, the Shc-SH2 domain contain larger connecting loops between β-strands resulting in a wider cleft for binding [(*48*); Fig. 1C-i]. The difference of structural organizations of these domains should also force the cognate binding partner AICD to take different conformations in the bound state [Fig. 1A–1C-iii]. For AICD itself, ^{680}NGYE^{683} motif is in a nascent helical conformation when it is in free solution; however, when bound to the X11-PTB domain, this motif forms β-stands and is sandwiched between the β-strands of the PTB domain (*49*). So, a helix to β-stands transition is apparent for specific regions of AICD. Surprisingly, phosphorylation of Thr^{668} forces a cis-isomerization of Pro^{669}, destabilizing the "helix capping box" of AICD (*50*). Though structures of phosphopeptides bound to the Shc-SH2 domain adopt an extended conformation as happens in other SH2 domains, significant differences have been found for Grb2-SH2 domain (*47, 48*). The presence of a bulky indoyl moiety in Trp^{121} of the EF loop, which determines the specificity of Grb2-SH2 interactions, forces the phosphopeptide to take a type-1 β-turn conformation. This β-turn conformation is stabilized by a hydrogen bonding between the carbonyl oxygen of phosphotyrosine residue and the main chain nitrogen of a Val residue (*51*). From this structural information, it can be hypothesized that AICD, when phosphorylated at Tyr^{682}, will form extended conformation and a type-1 β turn while binding to Shc-SH2 and Grb2-SH2 domains, respectively.

AICD is a classic case of conformational switching regulated by post-translational modifications. Phosphorylation at a single tyrosine residue (Tyr^{682}) determines the conformation. The choice of interaction partner depends on the conformation leading to a specific function as a result. Different downstream events can be initiated based on the initial trigger and the functional divergence of its "adaptors", starting from oncogenesis to apoptosis and cellular traffic originates out of this IUP.

DISCUSSION

Though the initiation of misfolding in a particular cell type is a stochastic event, various factors contribute to this process, such as increased protein concentration as in the case of familial PD where a triplicated locus of α-Syn has been found and covalent modifications such as phosphorylation and sumoylation of Ataxin-1 and Huntingtin, respectively, which destabilize protein conformations and initiate protein aggregation under different circumstances (*3*). It has now become clear that intrinsic disorder also contributes immensely to NDDs. There are two primary mechanisms by which disorder is utilized in PPI networks: one disordered regions binding to many partners and many disordered regions binding to one partner. AICD, like p53, for example (*52*), acts as a disordered hub protein, where it can interact with numerous binding partners (*13*). We have supported our hypothesis with data from proteins involved in different NDDs and shown the preponderance of unfolded or partially folded conformations, which occur in nature resulting in a tendency to form aggregates under normal physiological conditions (*53*). As specific cases discussed in this article, the essential proteins involved in three different NDDs such AD, HD, and PD are also highly unstructured with several distinct conformational possibilities driven by minimal chemical alterations. All these case studies elaborate the influence of dosage sensitivity, post-translational modifications and structural flexibility as typical characteristics of IUPs in determining cell fate. On IUPs, minor fluctuations in cellular environment would impart conformational changes on a much larger scale when compared with folded molecules. Whether or not these perturbations are a result of stress, lifestyle, or pathological conditions, protein disorderedness has emerged to play a decisive role in the disease process.

ACKNOWLEDGEMENTS

Udayaditya Sen is acknowledged for his help to gather knowledge of protein structure.

REFERENCES

1. Uversky, V. N. (2011) Flexible nets of malleable guardians: intrinsically disordered chaperones in neurodegenerative diseases. *Chem. Rev.* **111**, 1134–1166.
2. Lansbury, P. T. and Lashuel, H. A. (2006) A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature* **443**, 774–779.
3. Ross, C. A. and Poirier, M. A. (2004) Protein aggregation and neurodegenerative disease. *Nat. Med.* **10**, S10–S17.
4. Wright, P. E. and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331.
5. Uversky, V. N. (2002) What does it mean to be natively unfolded? *Eur. J. Biochem.* **269**, 2–12.
6. He, B., Wang, K., Liu, Y., Xue, B., Uversky, V. N., et al. (2009) Predicting intrinsic disorder in proteins: an overview. *Cell. Res.* **19**, 929–949.
7. Tompa, P., Szasz, C., and Buday, L. (2005) Structural disorder throws new light on moonlighting. *Trends Biochem. Sci.* **30**, 484–489.
8. Fink, A. L. (2005) Natively unfolded proteins. *Curr. Opin. Struct. Biol.* **15**, 35–41.
9. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., et al. (2001) Intrinsically disordered protein. *J. Mol. Graph Model.* **19**, 26–59.
10. Jeffery, C. J. (1999) Moonlighting proteins. *Trends Biochem. Sci.* **24**, 8–11.
11. Tompa, P. (2003) Intrinsically unstructured proteins evolve by repeat expansion. *Bioessays* **25**, 847–855.
12. Tompa, P. (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* **579**, 3346–3354.
13. Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) Extended disordered proteins: targeting function with less scaffold. *Trends Biochem. Sci.* **28**, 81–85.
14. Gsponer, J., Futschik, M. E., Teichmann, S. A., and Babu, M. M. (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **322**, 1365–1368.
15. Huberts, D. H. and van der Klei, I. J. (2010) Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta* **1803**, 520–525.
16. Shimizu, K. and Toh, H. (2009) Interaction between intrinsically disordered proteins frequently occurs in a human protein-protein interaction network. *J. Mol. Biol.* **392**, 1253–1265.
17. Vavouri, T., Semple, J. I., Garcia-Verdugo, R., and Lehner, B. (2009) Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell* **138**, 198–208.
18. Tsvetkov, P., Reuven, N., and Shaul, Y. Ubiquitin-independent p53 proteasomal degradation. *Cell Death Differ.* **17**, 103–108.
19. Raychaudhuri, S., Dey, S., Bhattacharyya, N. P., and Mukhopadhyay, D. (2009) The role of intrinsically unstructured proteins in neurodegenerative diseases. *PLoS One* **4**, e5566.
20. Prilusky, J., Felder, C., Zeev-Ben-Mordehai, T., Rydberg, E., Man, O., et al. (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435–3438.
21. Agorogiannis, E. I., Agorogiannis, G. I., Papadimitriou, A., and Hadjigeorgiou, G. M. (2004) Protein misfolding in neurodegenerative diseases. *Neuropathol. Appl. Neurobiol.* **30**, 215–224.
22. Hague, S. M., Klaffke, S., and Bandmann, O. (2005) Neurodegenerative disorders: Parkinson's disease and Huntington's disease. *J. Neurol. Neurosurg. Psychiatr.* **76**, 1058–1063.
23. Williamson, T. E., Vitalis, A., Crick, S. L., and Pappu, R. V. (2010) Modulation of polyglutamine conformations and dimer formation by the N-terminus of huntingtin. *J. Mol. Biol.* **396**, 1295–1309.
24. Kim, M. W., Chelliah, Y., Kim, S. W., Otwinowski, Z., and Bezprozvanny, I. (2009) Secondary structure of Huntingtin amino-terminal region. *Structure* **17**, 1205–1212.
25. Ekman, D. and Elofsson, A. (2010) Identifying and quantifying orphan protein sequences in fungi. *J. Mol. Biol.* **396**, 396–405.
26. Munishkina, L. A., Phelan, C., Uversky, V. N., and Fink, A. L. (2003) Conformational behavior and aggregation of alpha-synuclein in organic solvents: modeling the effects of membranes. *Biochemistry* **42**, 2720–2730.
27. Kim, S., Seo, J. H., and Suh, Y. H. (2004) Alpha-synuclein, Parkinson's disease, and Alzheimer's disease. *Parkinsonism Relat. Disord.* **10**(Suppl 1), S9–S13.
28. Uversky, V. N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J. Biomol. Struct. Dyn.* **21**, 211–234.
29. Eliezer, D., Kutluay, E., Bussell, R., Jr., and Browne, G. (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. *J. Mol. Biol.* **307**, 1061–1073.
30. Bussell, R., Jr. and Eliezer, D. (2003) A structural and functional role for 11-mer repeats in alpha-synuclein and other exchangeable lipid binding proteins. *J. Mol. Biol.* **329**, 763–778.
31. de Laureto, P. P., Tosatto, L., Frare, E., Marin, O., Uversky, V. N., et al. (2006) Conformational properties of the SDS-bound state of alpha-synuclein probed by limited proteolysis: unexpected rigidity of the acidic C-terminal tail. *Biochemistry* **45**, 11523–11531.
32. Murray, I. V., Giasson, B. I., Quinn, S. M., Koppaka, V., Axelsen, P. H., et al. (2003) Role of alpha-synuclein carboxy-terminus on fibril formation in vitro. *Biochemistry* **42**, 8530–8540.
33. Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.* **276**, 10737–10744.
34. Hardy, J. and Selkoe, D. J. (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356.
35. Mattson, M. P. (2004) Pathways towards and away from Alzheimer's disease. *Nature* **430**, 631–639.
36. Raychaudhuri, M. and Mukhopadhyay, D. (2007) AICD and its adaptors—in search of new players. *J. Alzheimers Dis.* **11**, 343–358.
37. Ghosal, K., Vogt, D. L., Liang, M., Shen, Y., Lamb, B. T., et al. (2009) Alzheimer's disease-like pathological features in transgenic mice expressing the APP intracellular domain. *Proc. Natl. Acad. Sci. USA* **106**, 18367–18372.
38. Ramelot, T. A., Gentile, L. N., and Nicholson, L. K. (2000) Transient structure of the amyloid precursor protein cytoplasmic tail indicates preordering of structure for binding to cytosolic factors. *Biochemistry* **39**, 2714–2725.
39. Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., et al. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049.
40. Russo, C., Dolcini, V., Salis, S., Venezia, V., Zambrano, N., et al. (2002) Signal transduction through tyrosine-phosphorylated C-terminal fragments of amyloid precursor protein via an enhanced interaction with Shc/Grb2 adaptor proteins in reactive astrocytes of Alzheimer's disease brain. *J. Biol. Chem.* **277**, 35282–35288.
41. Zhou, D., Noviello, C., D'Ambrosio, C., Scaloni, A., and D'Adamio, L. (2004) Growth factor receptor-bound protein 2 interaction with the tyrosine-phosphorylated tail of amyloid beta precursor protein is mediated by its Src homology 2 domain. *J. Biol. Chem.* **279**, 25374–25380.
42. Raychaudhuri, M. and Mukhopadhyay, D. (2010) Grb2-mediated alteration in the trafficking of AbetaPP: insights from Grb2-AICD interaction. *J. Alzheimers Dis.* **20**, 275–292.

43. Tamayev, R., Zhou, D., and D'Adamio, L. (2009) The interactome of the amyloid beta precursor protein family members is shaped by phosphorylation of their intracellular domains. *Mol. Neurodegener.* **4**, 28.
44. Radzimanowski, J., Ravaud, S., Schlesinger, S., Koch, J., Beyreuther, K., et al. (2008) Crystal structure of the human Fe65-PTB1 domain. *J. Biol. Chem.* **283**, 23113–23120.
45. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*. Available at: http://www.pymol.org. Accessed on February 2011.
46. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
47. Ogura, K., Tsuchiya, S., Terasawa, H., Yusawa, S., Hatanaka, H., et al. (1999) Solution structure of the SH2 domain of Grb2 complexed with the Shc-derived phosphotyrosine-containing peptide. *J. Mol. Biol.* **289**, 439–445.
48. Zhou, M. M., Meadows, R. P., Logan, T. M., Yoon, H. S., Wade, W. S., et al. (1995) Solution structure of the Shc SH2 domain complexed with a tyrosine-phosphorylated peptide from the T-cell receptor. *Proc. Natl. Acad. Sci. USA* **92**, 7784–7788.
49. Zhang, Z., Lee, C. H., Mandiyan, V., Borg, J. P., Margolis, B., et al. (1997) Sequence-specific recognition of the internalization motif of the Alzheimer's amyloid precursor protein by the X11 PTB domain. *EMBO J.* **16**, 6141–6150.
50. Ramelot, T. A. and Nicholson, L. K. (2001) Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR. *J. Mol. Biol.* **307**, 871–884.
51. Rahuel, J., Gay, B., Erdmann, D., Strauss, A., Garcia-Echeverría, C., et al. (1996) Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. *Nat. Struct. Biol.* **3**, 586–589.
52. Liu, Y. and Kulesz-Martin, M. (2001) p53 protein at the hub of cellular DNA damage response pathways through sequence-specific and non-sequence-specific DNA binding. *Carcinogenesis* **22**, 851–860.
53. Chi, E. Y., Krishnan, S., Randolph, T. W., and Carpenter, J. F. (2003) Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm. Res.* **20**, 1325–1336.