Thesis for the degree of doctor of philosophy (PhD)

A NANNY MODEL FOR INTRINSICALLY DISORDERED PROTEINS

by Rashmi Sharma

Supervisor: Prof. Dr. Monika Fuxreiter

University of Debrecen

DOCTORAL SCHOOL OF MOLECULAR, CELLULAR AND IMMUNE BIOLOGY

Debrecen, 2019

ACKNOWLEDGEMENTS

I acknowledge the financial support of GINOP-2.3.2-15-2016-00044, HAS 11015, MTA-DE laboratory of Protein Dynamics by the Hungarian academy of Sciences and pre-doctoral scholarship program.

ABBREVIATIONS

AP-1	Activator Protein -1
AD	Activation domain
bZIP	Basic leucine zipper
bHLH	Basic helix-loop-helix
BCA	Bicinchoninic acid assay
CD	Circular dichroism
CDK	Cyclin-dependent kinase
DMEM	Dulbecco's Modified Eagle Medium
DTT	Dithiothreitol
DBDs	DNA binding domains
ELISA	Enzyme-linked immunosorbent assay
E2	Ubiquitin-carrier or conjugating enzyme
E3	Ubiquitin ligase
GDP	Guanosine diphosphate
GFP	Green fluorescent protein
GST	Glutathione S-transferase
GOR	Garnier-Osguthorpe-Robson
GnHCl	Guanidinium chloride or guanidine hydrochloride
HD	Huntington’s disease
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IDPs	Intrinsically disordered proteins
IDRs	Intrinsically disordered regions
ID	Intrinsically disordered
IPTG	Isopropyl β-D-thiogalactoside
LM	Linear motifs
LZ	Lucine zipper
LB	Luria Bertani
MS	Mass spectrometry
NES	Nuclear export signal sequence
NP40	Nonidet-P40
NLS	Nuclear localization signal sequence
NaCl	Sodium chloride

PBS	Phosphate-buffered saline
PyMOL	Python-enhanced molecular
PDB	Protein Data Bank
PMSF	Phenylmethylsulfonyl fluoride
RFU	Relative fluorescence units
RFP	Red Fluorescent Protein
RNAP II	RNA polymerase II
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH	Src homology
SLiMs	short linear motifs
SUMO	Small Ubiquitin-like Modifier
TA	Transactivating domains
TRE	12-O-tetradecanoylphorbol-13-acetate (TPA) response element
TR	Transrepressing domain
Tris-HCl	Tris hydrochloride
UD	Ubiquitin-dependent
UI	Ubiquitin-independent
VCA	Verprolin, cofilin, acidic
WASP	Wiskott–Aldrich syndrome protein

Table of Contents

1. INTRODUCTION

Proteins are biopolymers composed of 20 amino acids, which could be arranged in many ways to form a nearly infinite assortment of protein shapes. Researchers believed for a long part of the 20th century that protein is eventually functional in the cell when folded into a well-defined tertiary structure. The view that protein sequence defines the 3D structure, which determines function, has been widely accepted. This relationship connects function to the primary sequence. At the same time however biophysical techniques used to characterize protein folding also indicated heterogeneous ensembles of conformations for functional proteins. It turns out that for some proteins, termed as intrinsically disordered (ID) proteins the native state is composed of a multitude of structures. Experimental studies and genomic analysis have been initiated to assess the occurrence and functions of ID protein segments [1–5]. ID regions were found to be involved in intracellular signaling and regulatory processes, as well as interaction hubs for protein-protein interaction [6].

Although structural evidence proliferated to demonstrate the existence of ID proteins *in vitro,* their cellular organization and role have remained quite enigmatic. An urgent matter has been raised how exposed regions of proteins can survive under cellular conditions.

The proteasome system was well described for maintaining the protein turnover however; new mechanisms were predicted by which cells control intrinsically disordered proteins (IDPs) level within them. The process named “degradation by default” [7]. Also described as ubiquitin independent pathway where the protein degradation is mediated by the 20S proteasome core [8]. The ‘‘uncapped’ or free 20S particles are active in degradation of proteins that are either partly disordered or fully without prior modification, which makes this pathway a default or passive process which is unique to IDPs [9, 10]. In 2009 a possible model has been put forward for how proteins with ID regions escape 20S proteasome degradation. Termed as the 'nanny' model [7], the ID proteins ('clients') have associated, protective partners ('nannies'). The client-nanny interactions are weak, which - unlike chaperons - do not impose folding on the ID region, simply provides a shield from proteases or the proteasome. In this manner, nannies ensure the survival of newly synthesized IDPs.

In my thesis, we test the hypothesis that nannies control the levels of their ID clients using the AP-1 transcription complex as a model system. Here c-Fos is the ID client, which interacts with c-Jun serving as a nanny. We probed whether the interactions with c-Jun serve to regulate the turnover of c-Fos and therefore its function. We show c-Fos is a natural substrate of the degradation by default pathway, and complex formation with c-Fos has a protective role. Furthermore, we explore the nature of interactions between these two proteins and their coorelation in terms of protein turnover. Based on the availability of similar complexes, we propose that the nanny model could be a wide-spread phenomenon to rescue ID proteins.

1.1. Background of intrinsically disordered proteins

Intrinsically disordered proteins (IDPs) contain regions, which are devoid of a well-defined tertiary structure [4, 2]. These could be loop and linkers in a large folded protein or may affect a larger portion of a protein [11]. Intrinsically disordered regions (IDRs) usually defined by continuous stretches ≥ 30 amino acids.

A number of additional terms have been used to indicate the highly dynamical or polymorphic characteristics of these proteins, such as natively denatured [12], natively unfolded [13], intrinsically unstructured [14], intrinsically denatured [12], intrinsically unfolded [13], intrinsically disordered [4], floppy [15], flexible [16], mobile [17], partially folded [18], vulnerable [19], malleable [20], dancing protein [21].

A number of structural and functional aspects of IDPs are distinct from folded proteins. A strong amino acid compositional bias in IDPs, i.e. abundance of hydrophilic and charged residues and lack of hydrophobic core is responsible for their lack of structure [22–26]. Their existence *in vitro* has been reported with many state-of-the-art solution state-NMR, small-angle X-ray scattering and fluorescence spectroscopy, circular dichroism (CD), and single molecule fluorescence resonance energy transfer (smFRET) measurements [27-33].

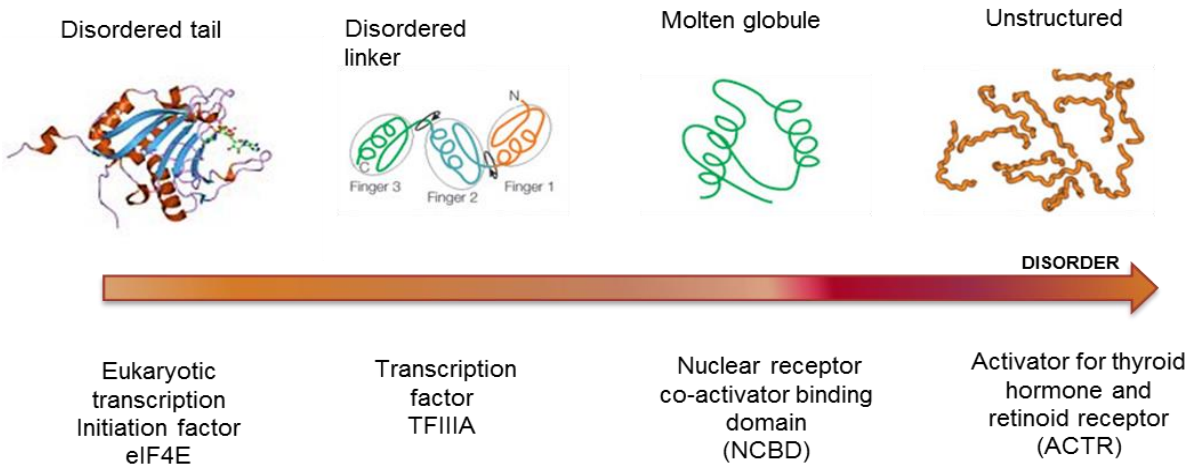

Figure 1. Example of intrinsically disordered proteins based on their degree of disorder. The figure is adopted and modified from [11]**.**

1.2 Emergence of a new paradigm

In general, evolution has produced increasingly complex organization and complexity is one of the fundamental indexes of evolution. Proteins are also the results of evolution. In the vastness of the protein sequence space available, Nature has extremely narrowed its evolution-driven search [34]. The potential amino acid sequence in the protein that would aggregate under generic physiological conditions is generally eliminated by evolution. As a result, only a few selected populations get its place in this large amino acid sequence space.

The first large fraction of selected amino acid sequences from globular protein where the amino acid sequence adopted a folded structure with the ability to perform function, which could thus be selected by evolution. Ultimately, these proteins were easy to purify and crystallize and get faster recognition in the field of structural biology. Up to experimental confirmation of structure–function relationship **(structure-function paradigm)** [35] with **Anfinsen's** Rnase A demonstration of folding from unfolded state to achieve its functionality [36] . Undeniably, it helped to dominance the idea a well-behaved protein was synonymous with well-structured, with one singular function, and one mechanism of action.

Figure 2. Classical structural biology is based on deterministic sequence-structure-function relationships.

Another evolutionary region of the amino acid sequence space is that of protein that lack constraints on maintaining packing interactions (neither aggregates nor collapse, still functional). These are known as intrinsically disordered proteins/regions. IDRs typically evolve faster than structural domains. These groups of protein residues display a wide range of evolutionary rates, depends on flexible, constrained and non-conserved disorder [37]. However, IDPs were initially regarded as a bit of an oddity. The failure to experimentally prove became a problem and eventually led to the ignorance of this protein class. Later, these anomalies accumulated over the years to the point where could no longer be neglect. The different definitions of IDPs emphasized not only the need of a **new structure-function paradigm,** but also new computational and experimental approaches to deal with them in depth.

More detailing of IDPs with respect to their different behaviors and functions doesn't place them against globular proteins, instead might place both of them at different points on a continuum. For example: like ordered proteins, the disordered proteins are common to all living organisms [5, 38, 39], are essentially important for basic cellular performance. Both states of protein in terms of their folding have movement at the atomic level and well fit their Ramachandran angles. In fact, the same building blocks - the 20 amino acids control structural and functional preferences-and the same laws of physics apply in both cases.

1.3 Prevalent features of IDPs

Different definitions of IDPs described them as having little or no ordered secondary or tertiary structure, in contrast to properties of ordered proteins. Their experimental properties are at odd with those of ordered proteins to the extent that these can be taken as indicators of disorder. Examples are enhanced sensitive behavior to proteolysis [40–42], residues

missing from electron density maps [43, 44], absence of secondary structure or negative values for 1H-15N heteronuclear NOEs in NMR spectroscopy, low intensity peaks from ~210 to ~240nm in circular dichroism and larger Stoke's radius than expected hydrodynamic dimensions.

There are several signs and scores to define disorder in a protein, however, not necessarily all features are mandatory to be present in each IDPs. They can be present to different extent in a system-dependent manner. Here, we will discuss very prevalent features of IDPs as an indicator of structural disorder.

1.3.1 Sequence

Compared to sequences from ordered proteins, the compositional trait of IDPs/IDRs is the low portion of bulky hydrophobic (Ile,Val and Leu,) and aromatic (Trp, Phe and Tyr,) groups, considered as 'ordering' amino acids [45]. By contrast, IDPs are constantly enriched in polar (Gly, Arg, Ser, Gln, Pro, Lys and Glu) and structure-breaking amino acid residues like Pro and Gly considered as a source of disorder [4, 46, 24, 47]. These biases in amino acid composition in IDPs enabled a coarse prediction of disorder from primary sequence information alone *via* a numerous developed disorder predictors [48-50]. It was moreover revealed that IDRs and low complexity sequences have similar compositional bias – less order-promoting residues (C,W,Y, I and V) in comparison to more disorder-promoting residues (R, K, E, P, and S) [24]. Groups of functionally related proteins were found to have similar disorder–complexity distributions within each of the group but reported to show notable differences between groups [51].

The long disordered regions in proteins exhibit a close relationship with low sequence complexity. The length of IDRs observed to be one of the structural related function indicator of disorder, emphasized that length of IDRs follows a power law distribution in human proteome [52]. Proteomic studies from eukaryotes and prokaryotes revealed to have similar disorder length profiles which were shown to be advantageous for cell functionality [53]. 44% of human protein coding genes have been reported to contain IDRs having >30 amino acids in length. However, the IDRs with different lengths have been shown to exhibit distinct types

of IDP functions. The frequently occurring short IDRs may be short linkers or linear motifs [54-56] conferring binding or post-translational modification. Longer ones may be longer linkers, a combination of motifs, or domains functioning in recognition or as entropic chains [56]. For example: transcription mediated functions are usually over-represented with long disordered regions which are typically more than 500 residues. Similarly, IDRs containing regions of intermediate length like 300-500 residues are enriched for kinase and phosphatase functions. Furthermore, short IDRs which are less than 50 residues tend to be linked to GTPase regulatory functions and metal ion binding, ion channel. Thus, the length of a IDR can be a determinant and provide a useful indication about the function of the protein containing it [53].

1.3.2 Conformational ensembles of IDPs

The adequate representation of single fixed conformation of a protein would fit to the "lock and key" theory [57] in which protein exists in a single well-defined state biased for the optimal ligand. Emil Fischer who laid this idea assumed that the interaction between a specific protein and its substrate take place only in case of a steric complementarity of protein crevice/binding surface and the substrate.

Later in the following years, ordered proteins were shown to exhibit a well-defined minimum energy states (one or few favored conformers) in a rugged energy landscape which means they are not fully rigid, they do possess some sort of flexibility and go through many unfavorable states that eventually collapse *via* multiple routes into possibly favorable folded state. Protein folding is therefore described as a conformational funnel [58]. Eventually, ordered protein achieves a unique equilibrium configuration.

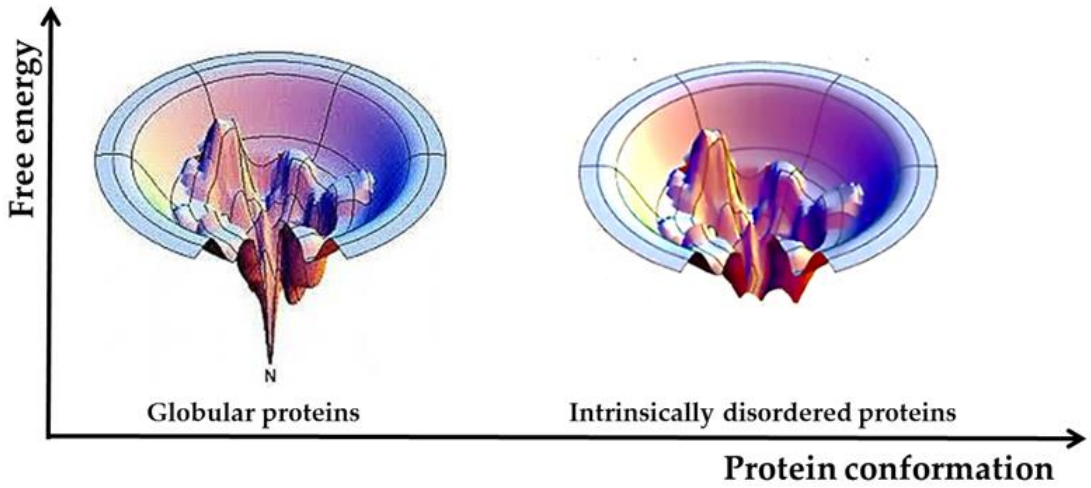

Figure 3. Free energy associated between intrinsically disordered proteins and globular proteins. The figure is adopted and modified from [59].

In contrast, IDPs are continuously fluctuating between multitudes of conformations states (collection of structurally similar and nearly energetically equivalent conformations of the protein) and rugged free energy landscape. Barriers between the wells of the landscape are comparable to k_BT (k_B is the Boltzmann constant and T is the temperature). Therefore, IDPs exhibit high conformational entropy in the free state, as compared to folded proteins. This raises a problem regarding IDP interactions: binding could be accompanied by a large entropy loss ('entropic penalty') [59].

1.3.3 Protein interactions by intrinsically disordered proteins

Another unique feature of intrinsically disordered proteins is their binding promiscuity which means the ability of one partner to bind to many partners [60]. In contrast to ordered proteins, IDPs are highly pliable and one IDP can form an array of unrelated structures being bound to different partner [61]. Many IDPs, being mostly disordered, tend to have transient elements containing preformed secondary structure which are highly interaction prone and used for binding to specific partners with high specificity while retaining low affinity with the help of IDRs [62]. Intrinsic disorder in a protein could allow one protein to bind with multiple partners (one-to-many signaling) or to enable multiple partners to bind to one protein (many-to-one signaling) [38]. For example, the kinase inhibitory domain

of the cyclin-dependent kinase (CDK) inhibitor $p21^{Cip1}$ which can bind to a diverse family of cyclin–CDK complexes [30]. Similarly, the GTPase-binding domain of the Wiskott–Aldrich syndrome protein (WASP) able to bind to its own C-terminal of VCA (Verprolin, cofilin, acidic) domain, which results in auto-inhibition, whereas in an altered conformation it bind to the GTPase Cdc42, results in initiation of actin polymerization through WASP activation [63].

There are several other structural properties of IDPs for example position of disordered regions, percentage of disorder present in a protein, tandem repeats required for normal cellular function and prone to cause aggregation. The general descriptions of each property will be discussed throughout the thesis in the sub-sections, so I will not make a detailed comparative discussion here. However, interested reader can find detailed information in these interesting articles [25, 26, 37, 53, 64, 65].

Last, but not least, unique features of IDPs include folding upon binding, conformational selection, formation of fuzzy complexes with their interacting partner. I will discuss these particularly interesting features of IDPs in the section Binding mechanisms of IDPs.

1.4. Binding mechanisms of IDPs

One of the most intriguing aspects of IDPs in order to perform their function is their ability to undergo disorder-to-order transitions upon binding [66, 67]. Adoption of the right conformation by the substrate taking place upon binding has garnered lot of attention in the field of structural biology. Although, the intermolecular interaction involving IDPs have divergent views on their binding mechanisms. Two main mechanisms for the interactions of disordered proteins gained popularity in the IDPs field: conformational selection and induced folding [68].

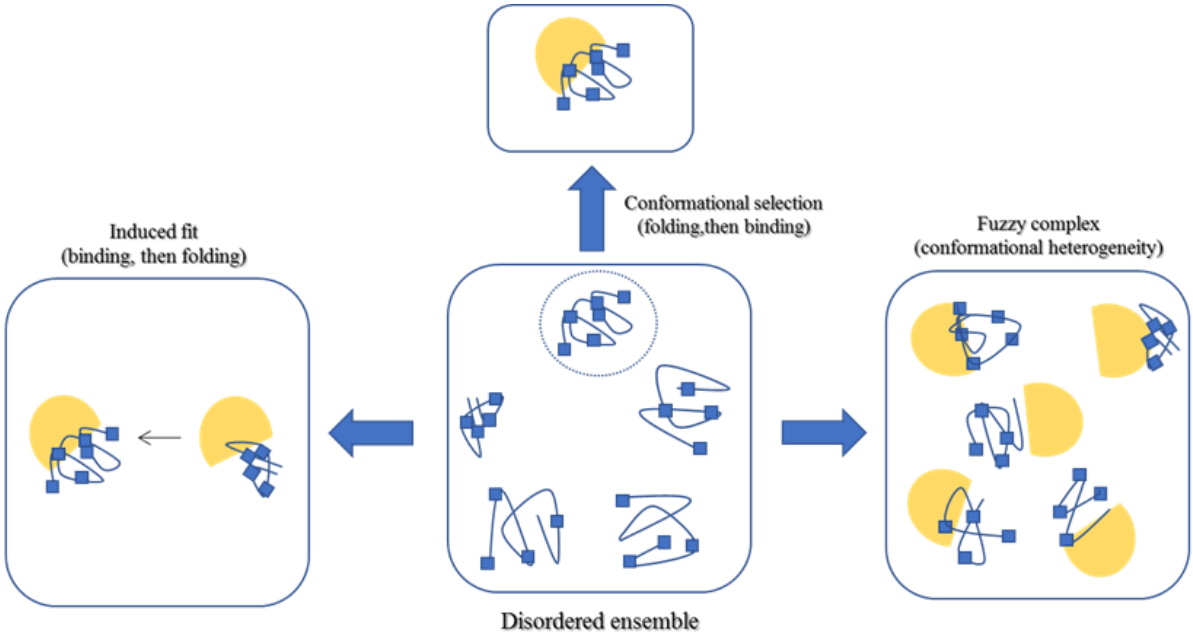

Figure 4. Schematic representation of the different binding modes observed for IDPs interactions (shown in blue) with folded proteins (yellow). In the "conformational selection" the folded binding partner bind to a specific conformation of the ensemble, which can also be a folded state. In the "induced-fit", the presence of the binding partner induces folding of the IDP. In the multivalent fuzzy complex, binding mode, remain disordered with alter native conformations reflecting different binding mechanisms. The figure is adopted and modified from [69].

The structural adaptability of ID proteins is studied extensively in numerous regulatory pathways, signaling, gene expression and cell-cycle regulation [70, 71]. ID proteins mechanisms are based on their interaction with DNA, RNA, small metabolites and their protein partners. As ID proteins lack stable tertiary structure in their free native forms, arise many thoughts regarding their interactions and specificity of these interactions. The first mechanism was termed coupled folding binding, where ID proteins lose its conformational heterogeneity in comparison to its native state [72]. This proposal has limitation as it couldn't differentiate the conformational diversity of free form/bound form of intrinsically disordered protein and assumed to achieve stable well-defined structure in the bound form [2, 38].

The factors like the low affinity, entropy loss governs the binding and folding of ID proteins interactions. To understand the broad range of IDPs' characteristics requires a comparison of ID proteins ensembles in their free native forms to their bound forms.

Emerging computational and experimental data suggested that ID may possess pre-organized secondary structures and may not be totally random in its free state [73, 74]. Following similar lines, the conformational selection mechanism was suggested, where the ligand (or target protein) selects among the conformations of the dynamically fluctuating protein and shifts the conformational ensemble which is compatible with binding. Hence, this model proposed the pre-existence of inter-converting conformations folds before binding to the partner shift the population to specific bound state. It was further noted that structural centric changes are not only dependent on protein itself, rather also on surrounding environment [72].

An alternate binding mechanism, which is also referred as 'binding–induced folding' also emerged. It proposed that the disordered protein undergoes a disorder-to-order transition process upon binding to their partner. This illuminates that folding does not occur until the IDP has interacted with the target protein. A shift can occur from one conformational state to another upon binding third partner or upon post-translational modifications such as phosphorylation or alternative splicing [72]. An important example of this is the increase in binding affinity by five-fold of the tumour suppressor p53 TRD to the KIX domain of the co-regulator CBP/p300 upon phosphorylation [75].

Finally, experiments and computational analysis revealed the commonness of both the events governing molecular recognition by IDPs. It turns out that the most often, folding coupled binding and induced folding work in synergy. Perhaps, their contribution to the full binding mechanism will be dependent on the binding rate required, IDP's concentration, the native local plasticity of IDPs and many more other disorder to order transition factors.

Eventually, increased appreciation of preformed structural elements proposal budded an interesting observation. In contrast to previous mechanisms, sequences of ID regions in their free state as well as in their bound state predicted to fall outside the allowed regions of the Ramachandran map, possibly suggesting disorder to disorder transition even in complexes. This phenomenon termed '*fuzziness*' represents the extension of the paradigm of structural disorder to the functional-bound-state [76, 77]. I will discuss this detail in the following section.

1.5 Physiological roles of IDPs

The intrinsically disordered protein field has been immensely expanded in the last 20 years; now considered as a new branch of unstructured biology. Since the late-90's, more experimental data's of intrinsic disorder in relevant proteins started accumulating with respect to their role in cellular signaling and human disease. Few prime examples includes studies of tau, ataxin-1, α-synuclein and huntingtin in several neurodegenerative disease [78], similarly the correlation of the expression of the IDPs Stathmin and p27 to various type of cancer [79, 80].

In general, IDPs seem to be more prevalent in signaling pathways, cell cycle regulation, transcriptional and translation regulation [71]. Target selectivity and fast association/dissociation rate in protein-protein interactions are highly specific in IDPs molecular recognition functions [81]. The most common functional interactors within IDR sequence are the linear motifs (LM) also known as short linear motifs (SLiMs) or molecular recognition features (MoRFs) [55, 82–85]. These short sequences ranging from 3 to 70 amino acids and more are the providers of both selective specificity and promiscuous multivalency in intrinsically disordered proteins [86]. Recently, it was proposed that linear motifs in IDR in synergy with folded domain regulate formation of higher order assemblies [87]. The role of interaction motif's functions and regulation of protein assemblies through post-translational modification brings important advantages to the signaling processes [71]. Sites of post-translational modifications are commonly present in IDRs [88] and important for modulating the binding activity of numerous proteins in signaling networks. The properties like flexibility and conformational dynamics of intrinsically disordered proteins increased their accessibility for post-translational modification [89]. Signaling can be regulated by sequential modification of multiple phosphorylation sites or by the removal (or addition) of single phosphoryl group, can give rise to range of signaling responses [89].

Therefore, multifunctionality, plasticity, dynamic tunability and modulations *via* post-translational modification are undoubtly advantageous in the context of complex cellular networking.

The precise control to maintain right endogenous amounts of IDPs and the integrity of the associated signaling and regulatory mechanisms is crucial for cell's normal cellular function [90]. Availability of IDPs is controlled by multiple mechanisms during transcription and translation, especially of those IDPs which are present at low levels or have shorter life-times [91]. These signify that IDPs are precisely and tightly regulated from transcript synthesis to degradation [90].

The regulation of the susceptibility of IDRs to proteasomal degradation is an important topic of the thesis. In the following sub-section I will thus discuss the mechanisms of protein degradation emphasizing on IDRs and their precise regulation mediated by proteasomal degradation.

1.6 IDPs in disease

Formerly, protein function deranging effects of missense mutations were interpreted in terms of destabilization and perturbation of structured elements. Later it was realized that mutations in IDRs can equally be deleterious by affecting dynamics, causing disorder-to-order transitions or altering MoRFs.

What is special about IDPs and makes them cardinal players in the development of pathological conditions is their unique, high specificity, low affinity interaction capabilities and interactions entailing transient context dependent many-to-one and one-to-many associations (binding plasticity and binding promiscuity) in signaling networks. Indeed, signaling networks in eukaryotes are highly enriched in ID proteins [11, 56]. 66% of cell-signaling related proteins were predicted to contain disordered regions of >30 residues, which is significantly higher than the prevalence of disorder in eukaryotic proteins in general (47%) [92]. It follows from the high prevalence of IDPs in homeostatic regulation that aberrant regulation of IDPs at different levels likely results in the development of various pathologies. These different levels include genetic alterations, such as mutations, chromosomal translocation, aberrant splicing, incorrect alternative splicing, and non-genetic ones, as levels of protein expression and availability, cleavage propensity, and post-translational modifications or disrupted trafficking.

It is directly pertinent to the subject of my dissertation, that for the purpose of physiological regulation of biological processes it is crucial that a given protein be available in appropriate amounts and not to be present longer than needed, IDR-containing proteins have to be tightly regulated and controlled. IDP encoding mRNAs have accelerated decay rates, lower rates of IDP protein synthesis, and shorter half-lives [91].

Misregulation of IDPs have been linked to many human diseases. Any such aberrations were shown to be directly responsible for the pathogenesis includes types of cancer (p53, BRCA1), neurodegenerative diseases (α-synuclein, tau, β-amyloid), cardiovascular disease (hirudin and thrombin, LDLR) [93], AIDS (HIV Rev protein) [94], cystic fibrosis (CFTR) [95], developmental disease and certain viral infections.

1.6.1 Cancer

It was predicted that 79% of cancer-associated proteins contain disordered regions which are >30 residues long [92]. For example:

p53

p53 function is inactivated in numerous tumors famous ones like colon, lung, breast, liver, brain, hemopoietic tumors and many more due to mutations or abberations in regulatory process [96]. p53 has promiscuous behaviour with multiple binding proteins to carry out various signal transduction function. Several of these are transcription factors, and activators/ inhibitors of its transactivating function. The N-terminal transactivation domain, the C-terminal regulatory domain, and the DNA binding domain (DBD) of p53 are the determinant for these effects on cellular level [97, 98]. However, the terminal domains contain disordered regions which are responsible to mediate and modulate interactions with other proteins. Roughly ∼70% of p53 interactions are mediated by IDRs and a multitude of PTMs in the disordered regions (acetylation, phosphorylation, and protein conjugation) are involved in their regulation [99] mutations of which can lead directly by disrupting interaction motifs or any perturbation of the dynamics of the disordered segments to aberrant function of p53 and cellular transformation.

As another example: BRCA-1

More than half of the BRCA1 residues are predicted to be disordered [99]. Surprisingly, BRCA1 binds to multiple proteins including in DNA damage sensors, DNA repair factors, transcription factors and signal transducers [100, 101]. The vast majority of these occur through dynamic interactions with the IDR, mutations of which are often linked to breast and ovarian epithelial cancer.

Besides constituting heterogenous interactions regions that can be perturbed by point mutations, structural disorder of connecting segments between structured domains confers viability on oncogenic fusion proteins, as translocation breakpoints need to be mostly located in the disordered regions or motifs, result in functional proteins. Consequently, IDRs are significantly enriched in oncoproteins arising from chromosomal translocations [102]. Translocation-generated fusions through flexible disordered peptide segments enable the long-range structural communication of binding and catalytic domains (BCR-ABL, BCR-RET) or kinase and dimerization domains [103].

1.6.2 Neurodegenerative disorder

Besides membrane bound organelles, cells harbor organelles that are 'membraneless' in the sense that they are not surrounded by a membrane. These supramolecular assemblies are usually composed of nucleic acids, proteins and other molecular components. Examples include the nucleolus, nuclear speckles, stress granules (SGs), processing bodies, and the centriole [104, 105]. Early studies already pointed out their dynamic nature and fluid-like properties as much as they fuse, separate and their components retain their movement and mix. We have recently come to understand the physical process by which is formed, which has been called phase separation or liquid-liquid demixing. The missense mutations in a number of low-complexity IDR containing stress granule proteins (hnRNPA1, FUS, TDP-43) cause neurodegenerative disorders such as ALS, where both mutant and wild type proteins are found to be aggregated in neurons [106, 107]. However SGs are assemblies which are dynamic in nature, have a fibrillar architecture aggregates [108, 109]. These disordered domains can form reversible hydrogels, and this is dependent on labile kinked β sheets [110, 111]. In liquid

droplets, the low-complexity domains retain their tendency to be disordered [114, 115]. However, with time an eventual maturation into fibrillar solid aggregates occurs, the rate of which is enhanced by ALS-causing mutations [112, 116]. Both the labile-to-stable gel [113] and liquid-to-solid [112, 116] transitions could explain the pathological conversion of SGs to aggregates in for ALS [117].

1.6.3 Aggregates

While IDPs contain regions that are functional in their unstructured state, many other proteins are known to unfold or misfold and they need to be either refolded by chaperones or to be removed by the proteasome or by autophagy. Often such misfolding occurs at such a scale that the protein degradation machinery is overwhelmed and the proteostasis of the cell is perturbed. The accumulating misfolded proteins tend to aggregate and/or sequester the protein quality control machinery exhausting the cell's capabilities to recover. This leads to several human diseases that originate from the deposition of protein aggregates formed from specific proteins or protein fragments in number of tissues [127–129].

For example*:* Missense mutations in human succinate dehydrogenase increases turnover rates and have been associated to play a crucial role in neuroendocrine tumors [130]. Another interesting example is of α-synuclein which constitutes ~1% of the total soluble protein in the brain and reported to present in the cells for longer period of time *via* multiple interactions or PTM modulations which promotes aggregation and lead to disease [131].

1.6.4 Viral assemblies

IDPs are important structural and functional components of paramyxoviruses, e.g. the measles virus (MeV), the Sindai virus and the newly emerged Nipah and Hendra (NiV and HeV) viruses, which cause [118] zoonoses associated with severe and often fatal encephalitis [119]. Paramyxoviruses are encapsidated by the N nucleoprotein. N proteins are over 500 residues long and consist of a globular N-terminal domain, N_{CORE}, which is crucial for the formation of self-assembly, RNA-binding, and of a C-terminal domain, $N_{TAIL.}$ The N_{TAIL} domain is disordered shows low sequence conservation with a compositional bias for polar

and charged residues and against hydrophobic residues. N_{TAIL} is exposed at the surface of the viral nucleocapsid and establishes a fuzzy interaction with the P phosphoprotein of the RNA-dependent RNA polymerase of the virus during viral infection once the viral ribonucleoprotein has been released into the cytoplasm of the infected cell and transcription of viral genes starts. P is an essential polymerase cofactor since it allows the large subunit of the polymerase to be recruited onto the nucleocapsid template triggering transcription and replication [120–123]. Hence, the N_{TAIL}–P interaction needs to be dynamically established and broken to ensure RNA synthesis. The affinity of the N_{TAIL}–P interaction is critical for viral transcription rates [125]. The N_{TAIL}–P interaction strength has to be kept into a precise window to ensure efficient transcription and replication. The dynamic binding of the complex provides the means to modulate the binding affinity [126].

1.7 Proteasomal degradation

The proteasome is a large assembly of proteins (2.5 MDa in size), it consists of two major entities: the 20S catalytic core and the 19S regulatory particle(s) (RP). The 19S RPs is composed of multiple subunits of proteins with molecular masses ranging from 10 kDa to 110 kDa, and is also designated as PA700. The 19S regulatory particle generally associates with one or both ends of 20S proteasome core in order to recognize ubiquitin tagged proteins and for their translocation into the interior of the 20S catalytic core. The 20S core, which on its own is also called 20S proteasome, is a cylinder like structure with a molecular mass of approximately 750 kDa, packed together from outer α rings and inner β rings. In other words, the catalytic core of the proteasome is a particle, which consists of seven α and β structurally similar subunits arranged in axial stacking style. Each inner ring which is consisting of three β-type subunits is associated with catalytically active threonine residue at their N-terminal and possesses N-terminal nucleophile hydrolase activity. These β subunits (β_1, β_2 and β_5) are linked with caspase-like/PGPH (peptidylglutamyl-peptide hydrolyzing), trypsin-like and chymotrypsin-like activities respectively. The enzymatic activities help the proteasome β-subunits to cleave peptide bonds present at the C-terminal or at the side of basic, acidic and hydrophobic amino-acid residues [132]. In the resolved crystal structure of the 20S proteasome, the center of the outer α ring is completely closed and it prevents internalization of proteins into the inner β subunit (proteolytic active site). In addition, N termini of the α subunits form

a physical barrier, which further restrict the access of proteins to the active sites. Thus, client proteins are only allowed after passing the narrow opening situated at the centre of the α rings [133].

Protein degradation plays role in almost every basic cellular regulation. Two alternative proteasomal degradation mechanisms have been described for IDPs: ubiquitin-dependent (UD) degradation and ubiquitin-independent (UI) degradation which are not mutually exclusive. These mechanisms are keys to maintain protein half-life in cells [135, 136].

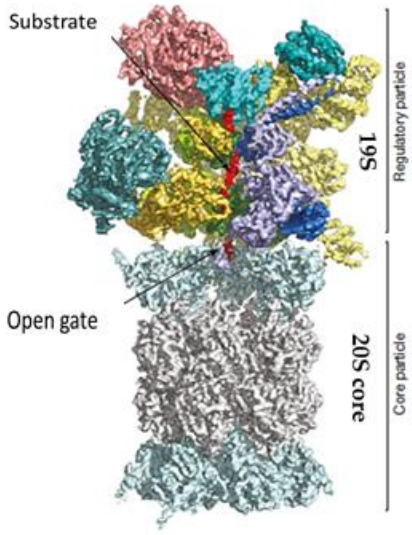

Figure 5. Structure of the Proteasome. The figure is adopted and modified from 134.

Many proteins in the cell are selectively degraded *via* the ubiquitin-26S proteasomal degradation pathway [137, 138]. The labelling step is an active process, performed on the substrate (IDPs or otherwise) in a highly regulated three step sequence during which ubiquitin, a 76 amino acid protein becomes covalently conjugated to the substrate protein making it for 26S proteasomal degradation. Structurally, ubiquitination is catalyzed by three proteins: the ubiquitin activating enzyme (E1), the ubiquitin-carrier or conjugating enzyme (E2), and the ubiquitin ligase (E3). In brief, the E1 enzyme stage 1 is ATP driven process and activation of Ub is the main goal which then transfer to another stage 2 E2 carrier enzymes, where formation of Ub-carrier enzymes E2 convert into an intermediate (E2-S-Ub). The stage 1 and stage 2, in combination with one of many E3 ubiquitin ligases, transfer the ubiquitin

monomer to the substrate. There are two main classes of E3 ligases which show difference in attaching Ub to the substrate. In the case of RING-domain-containing ligases, Ub is directly delivered to the subtrate however in some cases like HECT-domain containing ligases, dilvered after formation of intermediate (E3-s-Ub) [139].

First, the C-terminal of ubiquitin, forms an isopeptide bond with the amino-acid group of a Lys in the substrate [140]. The ubiquitin attachment can then be extended in a polyubiquitin chain through one of sever lysines in ubiquitin. For proteasomal recognition, as a minimal signal a chain of four Ub molecules need to be sequentially connected to through Lys48 linkages to the target protein [141, 142].

Recently, exciting discoveries in the field have revealed the susceptibility of proteins (mostly protein of inherent disorder) to ubiquitin-independent degradation which is mediated *via* the core 20S proteasome [10]. In contrast to UD, UI degradation is a passive process where the 'uncapped' or 'free' 20S particles are active in degradation of substrates that are either completely or regionally disordered in nature without prior modification of the substrates. It is also referred as 'degradation by default'. Thus, IDPs have different fate of degradation than the globular proteins in order to regulate them *in vivo*. The author also mentioned the possibilities why no significant correlation were observed between intrinsic protein disorder and shorter half-lives in the literature, by emphasizing on the existence of protective mechanisms of 'default degradation' [143].

Structurally, the design of the proteasome suggests that the entry gates of the 20S proteasome are too small for correctly folded globular proteins in their 'native' conformation and the substrate should be denatured by the regulatory 19S subunit prior to 20S degradation [144, 145]. However, IDPs override this requirement. For example, p21 and α-synuclein which are directly susceptible to uncapped 20S proteasomal degradation due to their inherent disorder behavior [146, 147]. However, it is not true in all the proteins, as certain motifs proved to be resistant to degradation as was reported for the polyQ chain [148].

Previously, it has been estimated that long IDRs contribute 36-63% of the eukaryotic proteomes *in vitro* studies [3]. Other findings confirmed that IDPs form functional complexes

in vivo and this might be the reason why IDPs to escape proteasomal system. Further, argued that properties of IDPs make them to form low affinity complexes with another protein *via* transient interactions [22].

A few years later, a study suggested that the proteasome cleaves a significant proportion of cellular proteins (approximately 20%) independent of the ubiquitination process. Furthermore, studies comparing IDPs with globular proteins *in vitro* demonstrated that only IDPs are susceptible to 20S proteasomal digestion. Based on these observations, it was predicted that protein-protein interactions, protein-nucleic acid interactions or any functional interaction can protect IDPs from 20S proteasomal degradation. In general, this may be achieved either by masking the unstructured domain or by folding it [149].

For clarification, several mechanisms have been proposed by which IDPs are prevented from default degradation: (i) IDPs form a functional complex either promoting order in or masking IDRs (ii) proteasome gatekeepers, which interact with both the 20S proteasome and the potential substrate can protect the latter (iii) nannies- where more specific binding to IDRs takes place (iv) the IDPs interact with nucleic acid where the IDPs gains more structure upon binding with DNA (v) Competitive binding to 'decoy' binding sites in DNA ,mainly applicable to transcriptional factors (TFs) (vi) intramolecular interactions which block the degradation signal and hence IDRs can survive in a proteases rich environment (vii) Local folding by interacting with the partner (viii) Interaction with ribosome or ribosome associated proteins [143].

1.8 The concept of fuzzy complexes

The term fuzzy logic was introduced by Lofti Zadeh with the proposal of fuzzy set theory in the year of 1965 [150]. In mathematical terms, fuzzy logic is a many valued logic with variables. It is in contrast to traditional binary logic, deviant logic employed to hold partial truth value, where truth range between completely true and completely false [151]. This theory can be easily exemplified as the definitions of its usage in the real world: the most used machine at one's home the television: uses sensed variables such as ambient lighting, time of

the day and user profile to adjust parameters such as screen brightness, color, contrast and sound.

In protein biology, the term fuzzy was borrowed by Tompa and Fuxreiter in 2008 in order to describe disorder in the bounded states in protein complexes. By definition, fuzzy protein complexes are usually composed of intrinsically disordered proteins, which retain their conformational heterogeneity in the bound form *via* transient interactions and required for the function. This structural multiplicity or dynamic disorder could contribute to the formation, function and regulation of the assembly [76, 77]. Proteins containing fuzzy regions not only preserve their conformational freedom in the bound state, but also impact the biological activity of the complexes [152]. In fuzzy complexes, the multiple conformations that IDPs adopt in the bound or complex state cover a continuum, similar to the structural spectrum of free, unbound IDPs [153], and ranges from static to dynamic, and from full to segmental disorder [77]. Undeniably, the introduction of fuzziness and fuzzy complexes as concepts has been tremendously widen the field of IDPs.

In broader term, two types of structural heterogeneity were considered: polymorphic and dynamic. In polymorphic complexes which represent static disorder where in the disordered regions have alternative conformations in the complex, whereas in dynamic disorder the IDRs exhibit a dynamical continuum of rapidly exchanging conformations [77, 76].

The static and dynamic disorder are considered situational as these properties totally depending on the spatial and temporal resolution of employed experiments.

In the context of mechanism of fuzzy interactions : four different mechanisms were defined to understand how distant ID segments in the bound form can affect cellular functions and biological activity [76] with proteins, RNA or DNA [154].

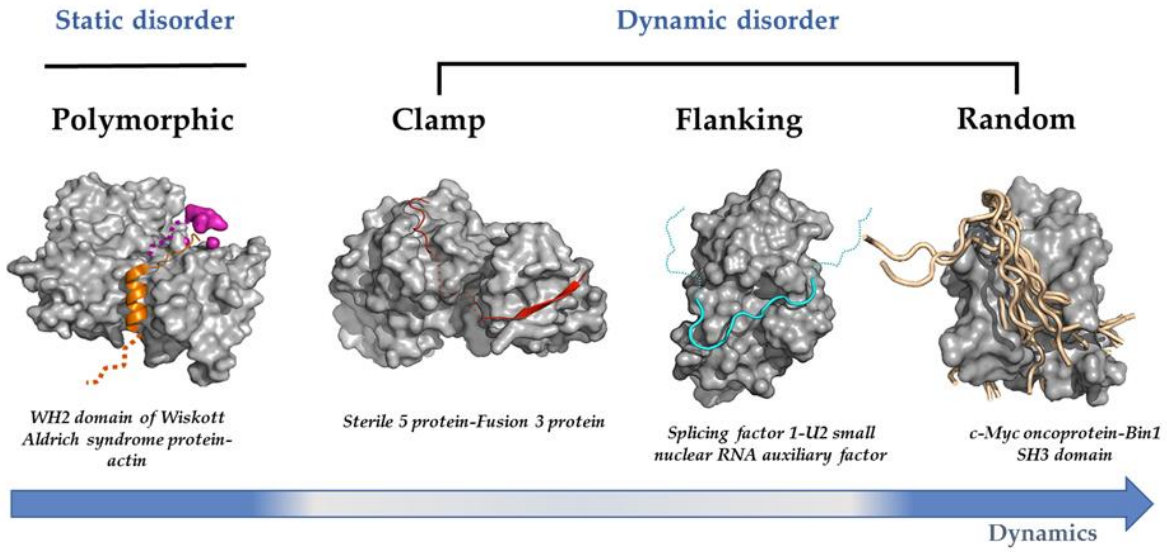

Figure 6. Topological classes of fuzzy complexes: A) polymorphic; B) clamp; C) flanking; D) random. Solid ribbons (well defined bound stretches), dotted ribbons (non-resolved but functional regions).The figure is adopted and modified from [77].

Both specific and non-specific interactions (mostly transient interactions) [154] were demonstrated to help IDRs to interact with the partner proteins. Some of the interactions were reported to increase the local concentration of the binding element leading to productive contacts [17] and some in contrast lead to auto inhibitory effects when fuzzy regions are involved in the binding interface [155]. All fuzzy interactions mode involves variety of mechanisms in relation to conformational heterogeneity of the complex [156].

(i) Conformational selection,

Selection of a thermodynamically populated conformer suitable for function by shift in the conformational equilibrium prior to binding.

For example interaction of the Max transcription factor with E-box (CACGTG) DNA sequences *via* (bHLH)/(LZ) segment to repress transcription of Myc target genes. The acidic N-terminus covers the destabilizing electrostatic potential present at the LZ region. The N- and C-terminal regions facilitate the formation of the recognition helices in the bHLH region, owing to a strong cooperativity between the bHLH and LZ domains, thus increasing the dissociation constant with the E-box by 10 to 100-fold [157].

(ii) Flexibility modulation,

Preservation of the conformational heterogeneity of the complex with DNA in order to improve the entropy of the binding interfaces.

The Ets-1 transcription factor is a well reported example which regulates various genes; stem cell development and tumorigenesis describing this mechanism. An auto inhibitory region in Ets-1-DNA binding regulates unfolding of the HI-1helix. The disordered Ser-rich region (SRR) further attenuate its interaction with DNA upon five phosphorylation sites within the SRR region and reduces DNA binding affinity by 100-1000 fold [158].

(iii) Competitive binding of the disordered region.

By competing with the binding partner via electrostatic interactions or steric hindrance.

The acidic fuzzy C-terminal tail of high mobility group protein B1 (HMGB1) negatively regulates interaction of the HMG DNA binding domains by occluding the basic DNA-binding surfaces [159].

Another example where similar competitive mechanism is observed is the human positive cofactor 4 (PC4), when it recruits commonly transcription factors and initiate RNA polymerase II (RNAP II). The disordered NTD of PC4 alone lacks considerable affinity for either ssDNA or dsDNA. However, the fuzzy region rich in Ser, acidic and Lys of PC4 interact transiently with the structured CTD and hence compete with ssDNA binding sites [160].

(iv) Tethering

Improving affinity by increasing the local concentration of the globular, weak affinity binding domain.

The human replication protein A (RPA) is a single stranded DNA binding protein that participates in nucleotide excision repair and combinatorial repair. RPA interacts with the

DNA through high affinity DNA binding domains (DBDs) which are connected *via* 78 amino acid long ID domain [161] and thus results in increase of the local concentration of DNA binding domain near the DNA, hence contribute to improve binding affinity [162].

All mechanisms of fuzzy complexes formation have different binding affinities to varying degrees and can be further fine-tuned by post-translational modifications, alternative splicing and asymmetric localization in the biological processes and assembly formations [76], [154].

The proposal of fuzzy region or regions in the protein complexes has definitely widened the scope of IDPs from protein-protein interactions to formation of higher order assemblies reviewed historical perspective of fuzziness in protein interactions in reference [163]. Dynamical ensembles could be shifted upon external signals or changes in the intracellular milieu, signifies context dependence. Over the years, it seems fuzziness is highly connected to context dependency in the cell however, presently handful of experimental data exist in literature. Also, dynamical behavior of higher order assemblies has been reviewed in recent study and described how intrinsic fuzziness regulates key feature in higher-order assemblies [87]. Taken together, fuzziness is an intrinsic property in protein-protein interactions, supramolecular protein organizations which provides mathematical framework for highly heterogeneous and diverse protein interactions. Indeed, fuzzy regions are often reported to be involved in protein half-life regulation, gene replication, cell division, immune response, signaling or biological activity of the assembly [152]. In further, our group have expanded the repertoire of available fuzzy complexes examples on the basis of experimental data: structural and biochemical evidences [152, 164].

Function	Protein name	Partner	Organism
Gene-expression			
Chromatin structure and dynamics			
	$H1^0$ linker histone	DNA	*Mus musculus*
	FACT	DNA	*Drosophila melanogaster*
	MeCP2	DNA	*Homo sapiens*
	MBD2 NurD	DNA	*Homo sapiens*
Transcription factors			
	Max	DNA	*Homo sapiens*
	NKX3.1	DNA	*Drosophila melanogaster*
	ApLLP	DNA	*Aplysia kurodai*
	Neurogenin 1	DNA	*Homo sapiens*
	Ultrabithorax	DNA, Exd	*Drosophila melanogaster*
	HMGB1	DNA	*Rattus norvegicus*
	Oct-1	DNA	*Homo sapiens*
	Ets-1	DNA	*Mus musculus*
	c-Myc	Bin1 SH3 domain	*Homo sapiens*
	Nrf2	Keap1	*Mus musculus*
	Prothymosin α	Keap1	*Homo sapiens*
Coactivator interactions			
	GCN4	Med15	*Saccharomyces cerevisiae*
	p65 (RelA)	CBP TAZ1	*Mus Musculus*
	p53 TAD	CBP NCBD	*Homo sapiens*
	KID	KIX	*Mus musculus*
Interactions with the basal machinery			
	EWS	PIC	*Homo sapiens*
	SP1	TFIID	*Homo sapiens*
	GCN4	PIC	*Saccharomyces cerevisiae*
	Gal4	PIC	*Saccharomyces cerevisiae*
	PC4	PIC	*Homo sapiens*
Nuclear receptors, transport			
	PPAR-γ	DNA	*Homo sapiens*
	NLS	Importin-α	*Xenopus laevis*
mRNA maturation, translation			
	Cup	eIF4E	*Drosophila melanogaster*
	UPF2	UPF1	*Homo sapiens*
	RNAP II CTD	mRNA maturation factors	*Saccharomyces cerevisiae*
	SF1	$U2AF^{65}$	*Homo sapiens*
	4E-BP2	eIF4E	*Homo sapiens*
	L7/L12	Ribosome	*Escherichia coli*
DNA repair			
	RPA	DNA	*Homo sapiens*
	UmuD′2	UmuD2	*Escherichia coli*
	UvrD	DNA	*Escherichia coli*
Signaling			
	Ste5	Fus3	*Saccharomyces cerevisiae*
	Tcf3, Tcf4	β-Catenin	*Homo sapiens*
	E-cadherin	β-Catenin	*Mus musculus*
	iκBα	NFκB	*Homo sapiens*
	Calmodulin	MBP	*Homo sapiens*
	RLP1	SH3	*Gallus gallus*
	I2	PP1	*Mus musculus*
Cell-cycle regulation			
	$p27^{Kip1}$	Cdk2/cyclin	*Homo sapiens*
	$p21^{WAF1/CIP1}$	Cdk2/cyclin	*Homo sapiens*
	Sic1	Cdc4	*Saccharomyces cerevisiae*
Cytoskeleton structure			
	Thymosin β4	Actin	*Bos taurus*
	Ciboulot	Actin	*Drosophila melanogaster*
	Myelin basic protein	Actin	*Mus musculus*
	Dynein IC	NudE	*Drosophila melanogaster*
Viral proteins			
	Nucleoprotein	Phosphoprotein	*Measles virus*
	Nucleoprotein	Phosphoprotein	*Henipah virus*
	Nucleoprotein	Phosphoprotein	*Hendra virus*
	E1A	CBP TAZ2	*Human adenovirus*
	NS5A	SH3	*Hepatitis C virus*
	preS1	γ2-Adaptin	*Hepatitis B virus*
	NS5B	VAPC	*Hepatitis C virus*
	Nucleoprotein	VP35	*Ebola virus*
Enzymes			
	Cellulase E	Cellulose	*Humicola insolens*
	Thymine–DNA glycosylase	DNA	*Homo sapiens*
	Anhydrin	DNA	*Aphelenchus avenae*
	RNase I	RNase inhibitor	*Homo sapiens*
Endocytosis/adhesion			
	LigB	Fibronectin	*Leptospira interrogans*
	SfbI	Fn3	*Streptococcus pyogenes*
	AP180	Clathrin	*Mus musculus*
Chaperones			
	Hsp90	Ppp5 TRP	*Homo sapiens*
	Hsp25	α-Lactalbumin	*Homo sapiens*
	αA-crystallin	HMM	*Bos taurus*
Self-assembly, aggregation			
	Sup35	Sup35 prion amyloid	*Saccharomyces cerevisiae*
	Ure2	Ure2 prion amyloid	*Saccharomyces cerevisiae*
	Elastin	Elastin	*Bos taurus*
	α-Synuclein	Membrane	*Homo sapiens*

Table 1: Illustration of fuzzy complexes examples involved in various cellular processes [152].

1.8.1 Nanny model for IDPs

The nanny model for IDPs is one of the plausible mechanisms to describe the turnover of inherently disordered proteins *in vivo*.

The main principle mechanism of the model is that when IDPs are in the monomer state, they are inherently unstable and tend to degrade by the 20S mediated UI proteasome pathway. However, this degradation can be prevented by forming a functional complex upon binding to a partner. Formation of a complex is to mask the disordered segment or reshape the IDPs structurally. The proteins that bind or mask the IDP and help escape their degradation by default termed as nannies [7].

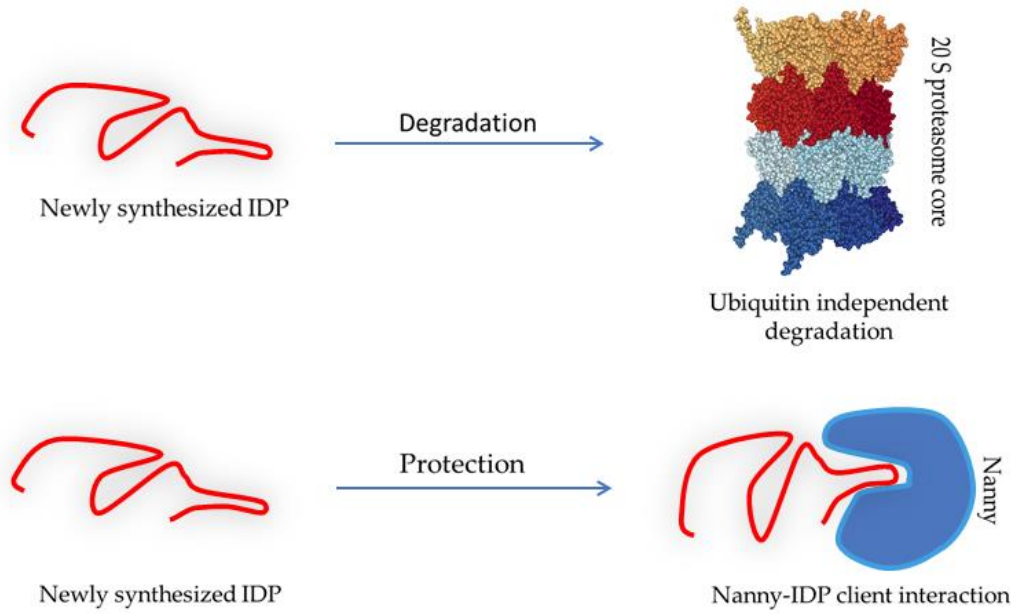

Figure 7. The nanny model for intrinsically disordered proteins. The figure is adopted and modified from [7].

The nanny model holds three principles: Firstly, a protein should have degradation signal attached to the IDR segment. Tumor suppressor p53 is an interesting example to fulfill this criterion. The p53 transactivation domain, as well as, regulatory domain is both unstructured at the N-terminal and C-terminal end of the protein. However, the N-terminal

has sensitivity to degradation by default. Second: a nanny protein should bind to the IDP and mask it from degradation. The classical example is again p53 where interactions through N-terminal are with protein partners block the degradation; however, C-terminal interactions fail to do so. Third: The interaction between the IDP and its nanny should be a transient association. The newly formed IDPs are at greater risk of proteolysis (20S specifically). Their susceptibility to the 20S proteasome and their protection from default degradation provides the ground for additional modes of IDP regulation [7].

The important fact is that the nannies are distinct from chaperones as they do not fold ID proteins. Their role is to eliminate IDPs destruction unlike chaperones, and further support IDPs maturation without affecting their dynamics. Therefore, fuzziness of ID protein-nanny interactions is well characterized by experimental data. An interesting and well-studied experimentally proven example is IkBα. Free IkBα has unstructured domains at the N- and C-termini [50]. Monomeric IkBα has been experimentally shown to rapidly degrade *via* the ubiquitin-independent pathway with a half-life of <10mins. Upon binding to NF-kB, IkBα escapes degradation by default, hence increases its half-life to many hours. Under this condition, IkBα can be destabilized by IκB Kinase (IKK)-mediated phosphorylation followed by ubiquitination and degradation [165, 166].

In vitro degradation studies with p21 illustrate that the protein is degraded by the 20S proteasome in an ubiquitin independent fashion. However, the degradation could be inhibit upon the binding to the proliferating cell nuclear antigen (PCNA) or in the presence of the cyclin E and Cdk2 complex [167, 168].

1.8.2 The AP-1 model system

According to the nanny model, a number of proteins with disordered regions are predicted to require specific nannies such as IκBα-NF-κB, Ku70 & Ku80, Cullin-3 complex and many more. The nanny-client interaction can be wide spread mechanism present in the cell to regulate intrinsically disordered protein turnover [7]. To test the plausible mechanism of nanny function, we hypothesize AP-1 (Activator Protein -1) complex formation is a nanny-client interaction where c-Jun acts as a nanny for c-Fos.

1.9 A historical overview of AP-1

Historically, AP-1 has been a most extensively studied topic, mostly because of its implication in a variety of pathologies, ranging from inflammation to tumorigenesis [169].

For the sake of clarity, I will selectively cover only a few aspects of AP-1 in detail relevant to my thesis work. However, overview of AP-1 in several diseases and therapeutic intervention in drug development are beautifully covered in several detailed studies [169–173].

AP-1 was first discovered as a TPA-activated transcription factor and found to be bound to a cis-regulatory element of the human metallothionein IIa (hMTIIa) promoter and SV40 [174]. The very same year, a study identified AP-1 binding site as the 12-O-tetradecanoylphorbol-13-acetate (TPA) response element (TRE) with the consensus sequence 5'-TGA G/C TCA-3' [175]. Further studies provided the information that AP-1 DNA binding activity is a dimeric transcription factor consisting of various members of the c-Fos, c-Jun and other families. c-Jun was identified as a novel oncoprotein of avian sarcoma virus [176], and c-Fos was first isolated as the cellular homologue of two viral v-fos oncogenes [177]. However, there is structurally a big difference between the v-Fos and that encoded by its cellular counterpart c-Fos. The difference is a 104-base-pair deletion in the carboxyl terminus of the v-Fos protein [178, 179]. Soon after, other family members of Fos and Jun were identified. These breakthrough findings ignited new interest to investigate the Fos and Jun gene as well as protein structures to gather more knowledge regarding the regulation and function of the family members. Eventually, over the years compiling of reviews provided much important insight into AP-1 transcriptional dysregulation and oncogenesis. Almost all field of biology from modern molecular genetics to structural biochemistry, embark on AP-1 and formulated European TMR network to proliferate regulation of AP-1 family members' function.

What makes AP-1 so special is its high degree of evolutionary conservation of gene structure among different species. Since its discovery, AP-1 has been connected with numerous regulatory and physiological functions, and new relationships are still under investigation. Not only the similarities but also the differences between species boundaries teach us how deeper is the underlying molecular pathways [180–183]. Furthermore, extended

studies in model organisms helped scientists to draw a general agreement that AP-1 is quite sensitive complex to its environmental information and accordingly translates it into expression of specific gene; found to be true across the species.

For clarity of thoughts, AP-1 seems quite fascinating as it has been reported to be extremely diverse to extracellular stimuli and also for catalyzing the very different biological functions of its individual components, however at the same time, it is frustrating to have unpredicted results from hypothesis driven research. There are numerous unanswered questions for example: how many AP-1 factors are there and do we really need to consider context dependency while studying AP-1 and its components diversity? Therefore, it is important to study diversity of AP-1 interaction factors together and individual regulation of these factors with more specificity.

1.9.1 Structure of the AP-1 complex; c-Fos and c-Jun

The best described and studied components of AP-1 are c-Fos and c-Jun. The Fos family consists of four proteins (FosB, c-Fos, Fra-2, Fra-1) and on the other hand the Jun family includes JunD ,c-Jun, JunB [184], reviewed in more detail [185]. Both the Fos and Jun families of eukaryotic transcription factors can homo or heterodimerize to form complexes that binds to the 20-nucleotide DNA consensus sequence TGAC/GTCA [186]. Fos-Jun dimerization occurs *via* the parallel interaction of leucine zipper motifs required for DNA binding. The crystallized structure of the bZIP domains of Fos and Jun can recognize the AP-1 site in both orientations related by a 180 rotational movement by the dyad axis of the complex. The structure of Fos and Jun bound to DNA was determined at a resolution of 3.05 Å [187].

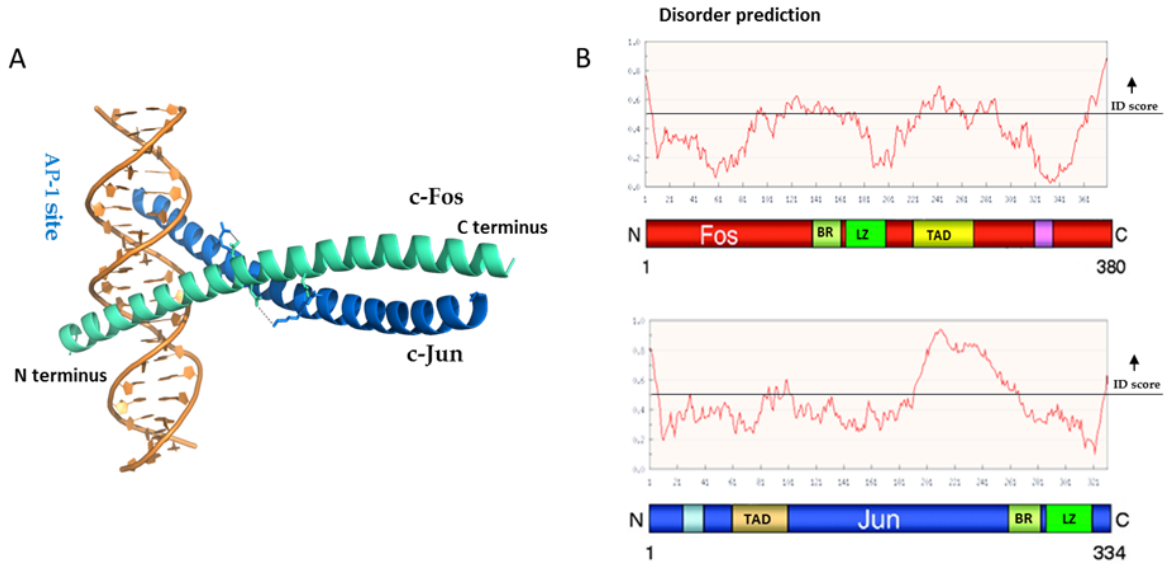

Figure 8. (A) The structure of the native bZIP domain of c-Fos/c-Jun bound to AP-1 site (PDB:1fos). c-Fos is shown in cyan color and c-Jun shown in blue color. The figure is adopted and modified from PDB. (B) The prediction of disorder in c-Fos and c-Jun protein based on IUPred.

Dimeric c-Fos/c-Jun bZIP domains form an X-shaped α-helical structure [187]. c-Jun is a 39 kDa protein consisting of an N-terminal transactivation domain (TA), the bZIP domain containing a basic DNA binding domain (DBD) followed by an alpha-helical lucine zipper dimerization domain (LZD) and a C-terminal region. On the other hand, c-Fos is a 65 kDa nuclear phosphoprotein containing several N and C terminal transactivating domains (TA), a transrepressing domain (TR) and a central bZIP region. The positively charged amino acids residues found within the DNA binding domain of bZIP proteins are required for DNA binding activity. The downstream leucine-zipper domain (LZD) contains a heptad repeat of leucine residues which mediates the dimerization of proteins in order to bring two DBDs into juxtaposition which in turn facilitates the protein dimer's interaction with DNA [188]. It has been found that amino (NH_2)- and carboxy (COOH)-terminal regions are quite divergent among all AP-1 proteins [189].

As clarification, Landschulz et al. in 1988 was the first scientist to propose a model of the structural basis of Fos-Jun interaction involving the so-called 'leucine zipper' motif (LZ) [190]. The importance of the leucine zipper in dimerization that was initially studied on the GCN4 transcription factor [191,192], eventually received extensive experimental support for

other similar classes of protein (reviewed in detail [185]). During the same time, the leucine repeats and the adjacent DBD were also shown to mediate heterodimer formation of c-Fos and c-Jun [193]. c-Fos forms stable heterodimers with c-Jun with higher affinity as compared to the Jun-Jun homodimer [194,195]. In contrast to c-Jun homodimerization, *in vitro* studies first inferred that the c-Fos homodimer is unstable therefore c-Fos can only interact with c-Jun. However, more recently the existence of c-Fos homodimerization has also been demonstrated *in vivo* [196].

1.9.2 Regulation of c-Fos protein turnover

c-Fos is an unstable protein, reported to have shorter half-life of approximately 60mins while c-Jun has a half-life of 90mins [197–199].

The instability of c-Fos explains its fate from synthesis to degradation. It also signifies why its accumulation is transient. c-Fos is a regulator of a diverse set of genes important for cell growth and differentiation [200, 201]. Regarding the expression of the *Fos* gene, it is induced in a variety of mammalian cell types rapidly and transiently within the few hours range by growth factors, phorbol esters, neurotransmitters, and membrane-depolarizing agents [202, 203]. The protein can be constitutively expressed in asynchronously growing cells [204] and transiently in other cell types upon various stimuli [205]. In order to define the regions based on the stability, some computational approaches were performed and found out c-Fos basically contain three PEST motifs at the C-terminus rich in proline, glutamic acid, serine, aspartic acid and threonine. These amino acids were proposed to provide instability to the protein, mainly PEST 3 motifs [206, 207]. Much later, Fos-AD was reported to be unstructured and highly mobile protein demonstrated with the help of advance methodologies: a lack of (1) H chemical shift dispersion, circular dichroism spectra indicative of unfolded proteins, and negative (1)H-(15)N heteronuclear nuclear Overhauser effects. Further, validated with the hydrodynamic factors of Fos-AD are found to be an extended structure. Thereby collectively results confirmed that the C-terminal domain of human c-Fos is an intrinsically disordered [208].

It was reported that c-Fos degradation is governed by a unique C-terminal which is bigger than the PEST3 motif in asynchronous cells; unlike in serum stimulated growing cells where N terminal together with C-terminal intrinsically disordered domain and the DBD/LZ region responsible for faster degradation of c-Fos. The author conclusively noted that multiple structural determinants are deciding factors for rapid c-Fos degradation or stabilization which thoroughly depends on the expressed conditions. Further, they indicated that both the mechanisms can be functional at the same time within the same c-Fos to determine its half-life [209].

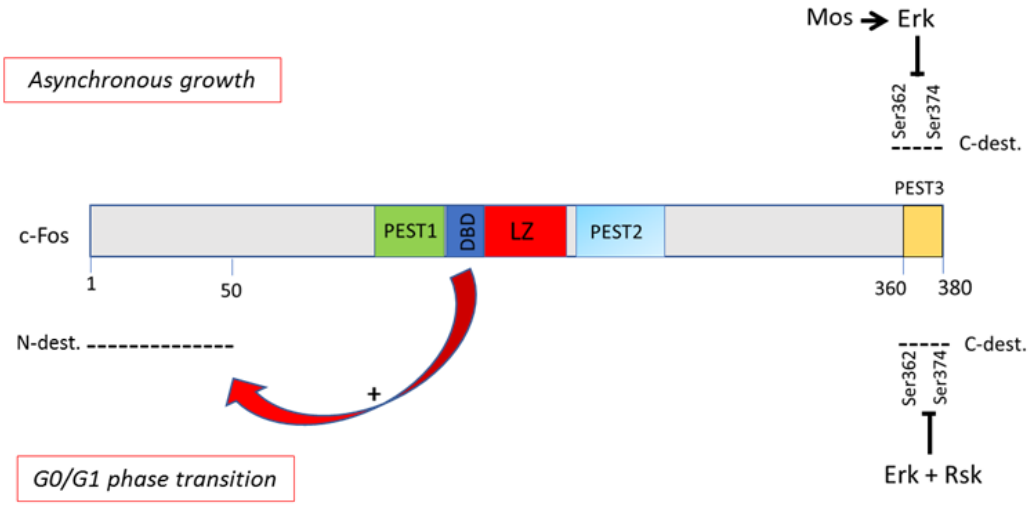

Figure 9. The different structural determinants responsible for c-Fos degradation in asynchronously growing cells and serum stimulated cells. PEST3 is necessary for the degradation of c-Fos compared to PEST 2 and PEST1. In exponentially growing cells, PEST3 upstream and full PEST3 length is responsible for c-Fos degradation. In cells undergoing the G0/G1 transition, both a C-terminal and N-terminal destabilizer are responsible for c-Fos destruction. The figure is adopted and modified from [210].

The proteolytic machinery found in cell irrespective of cytoplasmic and nuclear localization of proteins [211]. However, in case of c-Fos, although fractions of cytoplasmic c-Fos is mediated by 26 ubiquitin dependent proteolytic pathway, majority of nuclear c-Fos is subjected to ubiquitin independent degradation (UID) *in vivo* [212]. Recently, a study reported that *in vivo* c-Fos undergoes degradation by a default mechanism and regulated by NQO1

(20S proteasome gatekeeper) followed by c-Jun for the formation of function AP-1 complex. Furthermore, the author speculated that upon interaction, the functional complex adopts much lower conformational freedom, and the unstructured domain(s) of c-Fos substrate get masked by interacting protein partner c-Jun [213].

2. AIM OF THE STUDY

To probe the mechanisms of regulations of ID half-life *via* protein interactions

1. Relationship between nanny affinity and ID half-life.

- The interaction affinity between the nanny and ID client defines the ID client degradation by the 20S proteasome.

- Destabilizing the ID client –nanny interaction strength by affecting the hydrophobic contacts reduces the protecting role of the nanny (binding partner) and influence ID degradation.

- Stabilizing the ID client-nanny interaction by introducing a bulkier sidechain at the interface slows down the ID protein initial degradation rate and hence provides more protection to ID degradation.

2. What is the nature of the interaction of ID client with nannies?

- The interaction of ID client is both specific and transient, non-specific interactions with nannies which influence half-life of ID client and regulate protein degradation *via* fine-tuning the interaction affinities with a nanny.

3. How changes in ID regions contribute to changes in half-life?

- Deletion or truncation of ID region shortens the half-life as it weakens the complexation with nanny.

3. MATERIAL AND METHODS

3.1 Materials

3.1.1. Cell culture media

Dulbecco's Modified Eagle Medium - high glucose – (Sigma, St Louis, MO, United States)

Fetal bovine serum (FBS) – (Gibco, Waltham, MA, United States)

Gultamine – (Gibco, Waltham, MA, United States)

3.1.2. Antibodies/Reagents

GFP antibody (Biolegend, San Diego, California, United States)

60A8 Anti-Jun antibody (Cell signaling technologies)

9F6 Anti-Fos antibody (Cell signaling technologies)

His6- antibody (Sigma)

Anti-mouse (Advansta, Menlo Park, California, United States)

Anti-rabbit (Advansta, Menlo Park, California, United States)

Anti-rat (Advansta, Menlo Park, California, United States)

MG132 (Sigma, St Louis, MO, United States)

Protease inhibitor cocktail (Sigma)

Luria Bertani (Sigma)

Page blue staining solution (Thermo scientific)

Amicon Ultra 50 10000 MWCO ultrafilteration centrifuge tube

BCA reagent (Thermo scientific)

Snake memberane 10kDa (Thermo scientific)

All other materials were purchased from Sigma (St Louis, MO, USA) unless otherwise indicated.

3.2 Methods

3.2.1 Construction of plasmid based expression vector in different systems

The constructs pSV-c-Fos-EGFP, pSV-c-Jun-mRFP and pSV-c-Fos-215-EGFP (truncated version of c-Fos-EGFP lacking the last 164 amino acids) were kind gifts from Joerg Langowski, DKFZ, Heidelberg, containing coding sequence of human full length c-Fos and c-Jun. The sequences were PCR amplified and subcloned separately and together in the pET-Duet-1 vector between the EcoRI and XhoI sites to obtain expression of a fusion protein with an N-terminal His6-tag. The correct composition of these and all following plasmids was verified by DNA sequencing and restriction digestion. His6-c-Fos-EGFP and His6-c-Jun-mRFP were expressed in the Rosetta-2 *E.coli* cells (Novagen) and Rosetta™ *E.coli* cells (Promega) bacterial strain, which is engineered for optimal eukaryotic codon usage. The PCR amplified cDNAs of His6-c-Fos-EGFP and His6-c-Jun-mRFP were also subcloned in the pRSFDuet bacterial expression vector, which carries a different antibiotic resistance gene (kanamycin) to make simultaneous expression in the same cell from two separate plasmids (pET and pRSF) possible. Sequencing has been done using eurofins genomics facilities, Ebersberg, Germany.

We have also generated random mutant library of c-Fos and c-Jun with using GeneMorph II Random Mutagenesis Kit (Agilent technologies) and error-prone PCR method.

3.2.2 Expression and purification of recombinant plasmid fused with EGFP and mRFP in Rosetta bacterial system

Optimal expression condition of His6-c-Fos-EGFP and His6-c-Jun-mRFP were chosen by screening different temperatures (18°C, 25°C, 30°C and 37°C), expression times (24 hrs, 12 hrs, 6 hrs, 2 hrs) and IPTG concentrations (0.1 mM or 0.5 mM). High level of transgene expression was validated by detecting the proteins on 12% SDS-PAGE in bacterial lysates, by detecting GFP and RFP fluorescence by flow cytometry in intact bacteria and by verifying the GFP and RFP fluorescence spectra with a fluorimeter. Western blotting was

used to preclude the presence of prematurely terminated products using anti-His antibody (sigma) in (1:5000) dilution.

The cells were cultured in 300ml LB flask and grown at 37^{o}C until OD600 reached 0.5 followed by the IPTG induction at 16^{o}C with 0.5mM IPTG overnight. The cells were then collected by centrifugation at 4°C and lysed by sonication in 10 ml of lysis buffer (1% TritonX-100 in 1X PBS), 0.1 mM PMSF, protease inhibitor cocktail tablets (Sigma). The cells then centrifuged at 20,000 x g for 30 min. Clear supernatant was incubated with HisPur^{TM} Ni-NTA metal-affinity resin (Thermo Scientific) for overnight at 4 ^{o}C. Resins were washed several times with buffer A (lysis buffer + 20mM imidazole + 0.1 mM PMSF, protease inhibitor cocktail), buffer B (50mM Tris-HCl pH 7.9 + 500mM NaCl + 20 mM imidazole + 0.1 mM PMSF, protease inhibitor cocktail), buffer C (50mM Tris-HCl pH 7.9 + 100mM NaCl + 100 mM imidazole and 500mM Imidazole with 50mM EDTA + 0.1 mM PMSF, protease inhibitor cocktail). The purity of the protein and lacked any detectable partial product was determined by staining of 12% SDS polyacrylamide gels with Page blue protein staining solution (Thermo Scientific). Protein was eluted with buffer D (50mM Tris-HCl pH 7.9 + 100mM NaCl + 500mM Imidazole + 50mM EDTA + 0.1 mM PMSF, protease inhibitor cocktail). Recombinant His6-c-Fos-EGFP, His6-c-Jun-mRFP were dialyzed in “Mini dialysis tubes” (nominal molecular weight cutoff at 10kDa, Thermo Scientific) overnight in buffer D without imidazole. The purity of the protein was determined by staining of 12% SDS polyacrylamide gels with Page blue protein staining solution followed Western blotting using anti-His antibody (Sigma) in (1:5000) dilution. The bands were detected by chemiluminescence ECL Detection System (Millipore).

As His6-c-Fos-EGFP was poorly soluble and was present mainly in inclusion bodies. Thus, different non-denaturing extraction buffers were also tried. The cells were cultured in 300ml LB flask and grown at 37^{o}C until OD600 reached 0.5 followed by the IPTG induction at 37^{o}C with 0.5mM IPTG for three hours. The cells were then collected by centrifugation at 4°C and lysed by sonication in 10 ml of lysis containing 20mM Hepes, pH 7.9, at 4^{o}C. After centrifugation of the lysate, the pellet was used to purify insoluble His6-c-Fos-EGFP. The pellet was resuspended with 10 ml E-buffer (50 mM Hepes, pH 7.9, 5% (v/v) glycerol, 0.5 mM of 2-mercaptoethanol, 0.05% (w/v) Na-deoxycholate, and 1% (v/v) NP-40) and

homogenized with a Dounce homogenizer and a B-pestle. Extracted inclusion bodies were then recovered by centrifugation for 25 min, 14,000 rpm at 4 °C. The His6-c-Fos-EGFP in the pellet was solubilized with 10 ml S-buffer (10 mM Hepes, pH 7.9, 6 M guanidine-HCl, and 5 mM of 2-mercaptoethanol) on a shaker overnight at room temperature. Denatured His6-c-Fos-EGFP protein in S-buffer was adjusted to 15 mM imidazole and bound in batch to 1-2 ml HisPur^{TM} Ni-NTA resin on a rotator for 2-4 h at 4°C. The resin was washed 3 times with 5 ml each of S-buffer containing 15 mM imidazole, 3 times with 1 ml BC500 containing 7 M urea and 15 mM imidazole, once with 1 ml BC100 containing 7 M urea and 30 mM imidazole, and once with 1 ml BC100 containing 7 M urea and 45 mM imidazole. Recombinant His6-c-Fos-EGFP was eluted from the resin with 300mM or 500mM imidazole in BC100 containing 7 M urea. Although, very small quantity of protein was recovered from the protocol, the 6His-c-Fos-EGFP protein was >80% pure and lack any detectable partial product as judged by 12% SDS–PAGE and coomassie blue staining. Purified His6-c-Fos-EGFP protein in BC100 containing 7M urea and dialyzed in "Mini Dialysis Tubes" (nominal molecular weight cutoff at 10 kDa, Thermo scientific) in 7 successive steps of 2 h each against BC500 (as above, but with 0.1% NP-40) containing 4M urea (BC500- 4 M urea), BC500- 2 M urea, BC500- 1 M urea, BC500- 0.5 M urea, BC500, and twice with BC100. All dialysis steps were done at 4°C. After the last dialysis step, the dialyzed proteins were centrifuged at 13,000 rpm for 1 hour at 4 °C in an Eppendorf microfuge to remove insoluble aggregates. An aliquot (20 µl) was analyzed by 12% SDS–PAGE and coomassie staining and stored in small aliquots at -80°C.

3.2.3 Expression and purification of His6-c-Fos and His6-c-Jun without fusion proteins

Rosetta2(DE3) pLysS *E.coli* (Novagen) bacterial cultures transformed with the constructs containing His6-c-Fos and His6-c-Jun coding sequence were grown to OD600=0.6 and protein expression was induced with 0.1 mM IPTG for long 3 hours at 37 °C. The cells were harvested at 6,000g for 10mins at 4°C and sonicated in 6M guanidine-HCl; 25 mM Tris-HCl, pH 7.4. Buffers were supplemented with a protease inhibitor cocktail (Sigma). The proteins were purified by immobilized metal affinity chromatography on a Ni-Sepharose™ 6

Fast Flow resin (GE Healthcare) and washed with 50 volumes of lysis buffer. His6-c-Jun was eluted with lysis buffer containing 250 mM imidazole and dialyzed sequentially against 100 volumes of 4M urea, 2M urea and 1M urea and finally against 25 mM Tris-HCl, pH-7.4; 100 mM NaCl. For His6-c-Fos variants the buffer was changed to 25 mM Tris-HCl, pH-7.4; 100 mM NaCl while the proteins were still bound on the resin. They were eluted in the same buffer containing 250 mM imidazole and were dialysed against the same buffer without imidazole and concentrated with an Amicon-Ultra 50 10000 MWCO ultrafiltration device. Protein concentrations were measured by BCA protein assay.

3.2.4 Designing of c-Fos and c-Jun mutants

We designed mutants using two methods: Garnier-Osguthorpe-Robson (GOR) method and Python-enhanced molecular (PyMOL) graphics tool. The GOR method is an information theory-based method for the prediction of secondary structures in proteins. PyMOL excels at 3D visualization of proteins, small molecules, density, surfaces, and trajectories. We have designed targeted mutants at the c-Fos leucine zipper domain which is responsible for dimerization, also at the C-terminal of c-Jun using the GOR4 program and PyMoL tool. Single amino acid mutants were generated using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Point mutants were expressed in bacteria and purified from pellet as well as supernatant fraction (subsection-3.2.2) followed by 12% SDS-PAGE and Western blotting as mentioned above in expression and purification of recombinant protein with/without fusion proteins subsections (subsections- 3.2.2 & 3.2.3).

3.2.5 20S proteasomal degradation assay by ELISA

c-Fos variants and human 20S proteasome (Boston Biochem) were both diluted in 25 mM TRIS-HCl, pH7.4; 100 mM NaCl; 0.5 mM DTT to concentrations of 400 nM and 4 nM, respectively. The degradation assay was started by mixing 12.5-12.5 µl of the two. Degradation of the c-Fos substrates was followed at 37°C and 2 µl samples were collected at 0, 15, 40 and 80 minutes into 98 µl of PBS, 0.1% Tween-20, 1% BSA containing a protease inhibitor cocktail and 1 µM MG132 (Sigma). The c-Fos protein remaining in this solution

was bound to 96-well Pierce Nickel-coated plates (Thermo Scientific) for an hour at room temperature. The wells were washed three times with PBS-0.1% Tween-20 and they were reacted for an hour with 9F6 anti-c-Fos antibody (Cell Signaling Technologies) diluted 1 to 2000. After washing, the wells were reacted with HRP-conjugated anti-rabbit IgG and developed with tetramethyl-benzidine (Sigma). An equal volume of 1 N HCl was added to stop the reaction and develop the yellow color, read at 450 nm in a Synergy H1 microplate reader (BioTeK Instruments). Protein amounts were determined with a calibration curve relating the ELISA signal to protein concentration. The intensity values were first fitted by nonlinear fitting to an equation describing single exponential decay with a constant offset and the initial degradation rates were calculated as the value of the first derivative of the decay curve at time 0 using the software R program. Experiments were repeated three times with three independent parallels and the data reported are as mean ± SD from the experiment. Statistical significance was determined by unpaired two tailed t-test with Welch's correction. The $p \leq 0.05$ value was considered significant.

3.2.6 Co-Immunoprecipitation assay

Swollen protein G sepharose beads were taken with a larger excess of IP100 buffer by mixing slowly for an hour at room temperature (ca. 20 mg of dry beads will give 100 µl wet beads). Master mixture containing 50mg beads + 1 ml of IP100 buffer: 100mM KCl, 25 mM Tris-HCl, pH 7.9, 10 % glycerol, 0.1 mM NP40, 0.5 mM DTT. In 100 µl wet beads, 5µg of anti-Fos or anti-Jun antibody were added in a 0.5-1 ml of IP100 buffer and mixed by slow rotation for an hour at RT. For the antibody to bind for an hour followed by centrifugation at 1500rpm and rinse three times with excess of IP100. Diluted c-Fos or c-Jun or pre-complexed c-Fos/c-Jun into 1 ml of IP100 and add to 20ul antibody/protG beads. Incubate by rotational mixing at 4 °C for 2-4 hours. Beads were sediment by using centrifugation at 15000rpm for 1min followed 3-5 times washing with excess of buffer. For reference control, we used no antibodies on the beads and beads coupled with an unrelated antibody. Beads were centrifuged for a min to sediment and liquid phase was removed from the tube containing sedimented beads. Boiled in Laemmli loading buffer and Western blotting was performed to examine the interaction between c-Fos and c-Jun.

Dilution of antibodies used: anti-Fos antibody- (1:1000), anti-Jun antibody- (1:1000), anti-rabbit secondary antibody (1:10000).

3.2.7 Binding kinetics analysis

The binding kinetics of c-Fos variants to c-Jun were measured with a BLItz (PALL-ForteBio) biolayer interferometer. 1.5 μg of c-Jun was pre-complexed with 60A8 anti-Jun antibody (Cell Signaling Technologies) at 37°C for an hour. This resulted in saturation of the antibody with Jun. The complex was loaded onto Protein-A Dip and Read biosensors to a spectral shift of 3.5 nm. c-Fos variants were diluted into PBS containing 0.1% Tween-20 to various concentrations and their association to c-Jun was measured before the biosensor was dipped into the same buffer to record the dissociation of the proteins. The binding curves were fitted to a 1:1 binding model and k_a, k_d and K_D values were calculated with the BLItz Pro™ software.

3.2.8 Circular dichroism spectroscopy

Electronic circular dichroism (ECD or CD) far-UV spectra were recorded on a J-810 spectropolarimeter in 20 mM sodium phosphate, pH 7.4 in a 0.2 mm pathlength quartz cell at room temperature. Raw spectra were buffer-subtracted, converted to mean residue ellipticities and smoothed. Deconvolutions were made with the Bestsel program. CD spectroscopy has been performed by Dr. Sandor Balazs Kiraly and Dr. Tibor Kurtan, University of Debrecen.

3.2.9 Electrophorectic mobility shift assay

The AP-1 probe was hybridized together from the oligonucleotides 5'-GTCAGTCAG**TGACTCA**ATCGGTCA-3'and 5'-TGACCGAT**TGAGTCA**CTGACTGAC-3' by heating their 1:1 molar mixture in water to 95 °C and letting it cool down to room temperature. The purified His6-c-Fos and His6-c-Jun proteins were mixed in a 1:1 molar ratio in gel shift buffer (25 mM TrisHCl, pH 7.5; 50 mM KCl; 0.5 mM EDTA; 1 mM DTT),

incubated for 10 mins at 65 °C and then cooled to room temperature. 1 µl of the diluted oligo mixture containing 50 ng of DNA probe was added to 1µg of AP-1 complex in gel shift buffer and the binding complexes were allowed to form for 15 minutes at room temperature. After adding 2 µl of glycerol the complexes were resolved on an 8 % polyacrylamide gel in 0.5x TBE buffer at 50 V at room temperature. The bands were visualized by incubating the gel in 0.5 µg/ml ethidium bromide.

3.2.10 Mammalian cell culture and transfection

To check c-Fos and c-Jun co-expression in eukaryotic, pSV-c-Fos-EGFP and pSV-c-Jun-mRFP plasmids were used. HEK293AD cells were cultured in DMEM containing 10% FBS, 2mM glutamine, and penicillin/streptomycin in a 37 °C incubator in the presence of 5 % CO_2. Transfections were performed in Opti-MEM medium for 5 hours using Lipofectamine 2000 in a 2 µl to 1 µg DNA ratio and then the medium was changed back to the original culture medium. Images were taken under the fluorescence microscope after 24hours post-transfection.

4. RESULTS: PART-I – DATA USED IN THE STUDY

4.1 Mutant design

The structure of the AP-1 complex which is composed of c-Fos and c-Jun is a coiled-coil [214], which is held together tightly by interdigitating hydrophobic contacts ('leucine zipper'). These leucine zipper contacts are complemented by salt bridge at particular locations, which gives stability to heterodimer in comparison to homodimer like c-Jun-c-Jun.

We designed site directed mutants of c-Fos and c-Jun guided by their known crystal structure available in the PDB database (PDB: 1fos) using the GOR4 and Pymol softwares. We constructed five mutants harboring single point mutation at the interface of leucine zipper of c-Fos. Out of five mutants, two mutants (L165V, L172V) affect the hydrophobic network and three mutants (E175D, E189D, K190R) perturbing the hydrophilic contacts as shown in the Figure 10. Replacements of amino acid from L to V and E to D expected to maintain the polarity of the interaction while little deviations in the geometry due to the smaller sidechains. On the other hand, K to R substitution mutation provides a bulkier sidechain with pi-pi interactions and thus provides additional interactions. We also designed a C-terminal truncation c-Fos mutant containing 1-214 amino acid length, with the aim to probe the effect of the fuzzy C-terminal tail of c-Fos (215-380AA).

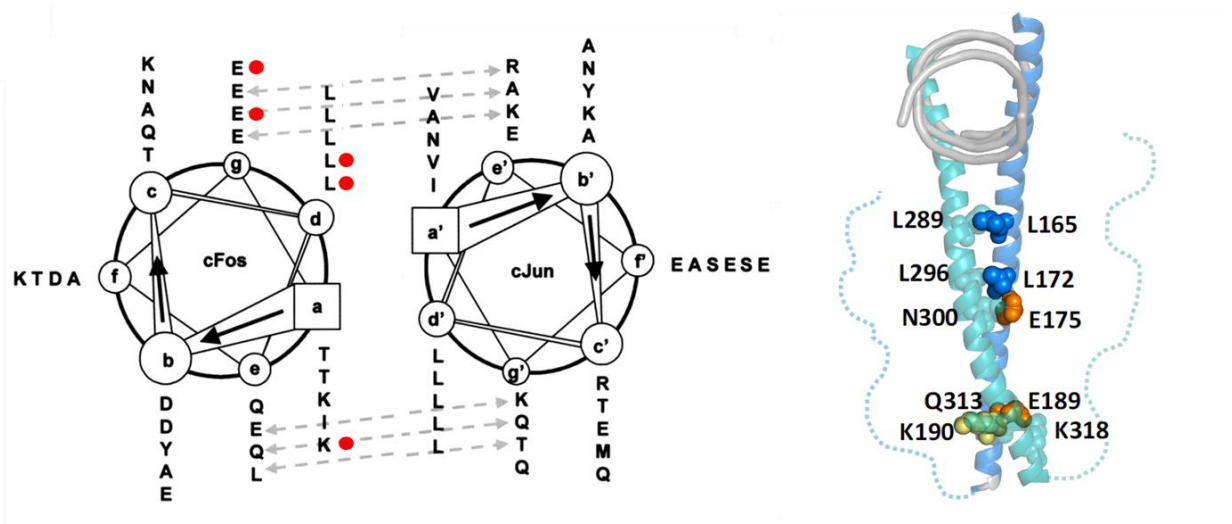

Figure 10. Designed mutations (red dots) in the cartwheel representation and the structure of the AP-1 complex (1fos.pdb). Mutations in c-Fos (marine, black labels) and the contacting residues in c-Jun (cyan, grey labels) are shown by spheres. Fuzzy C-terminal tails are displayed by dashed lines, the DNA is colored grey.

4.2 Expression and purification of recombinant His6-c-Fos and His6-c-Jun

Purification of c-Fos and c-Jun fused to larger globular protein tags EGFP/mRFP proved challenging (subsection 4.7). We expected that these fusion proteins would make the largely disordered c-Fos and c-Jun more soluble and stable. However, a single step immobilized metal affinity chromatography (IMAC) against the N-terminal 6xHis-tag was not sufficient to yield pure protein from the soluble fraction of the bacterial lysate homogenized with 1% Triton-X100. Denaturing purification with 6 M GnHCl yielded much purer proteins based on SDS-PAGE and Coomassie staining. However, during refolding by sequential dialysis these proteins mostly precipitated at the bottom of the tube. The globular protein part also never recovered its activity or fluorescence.

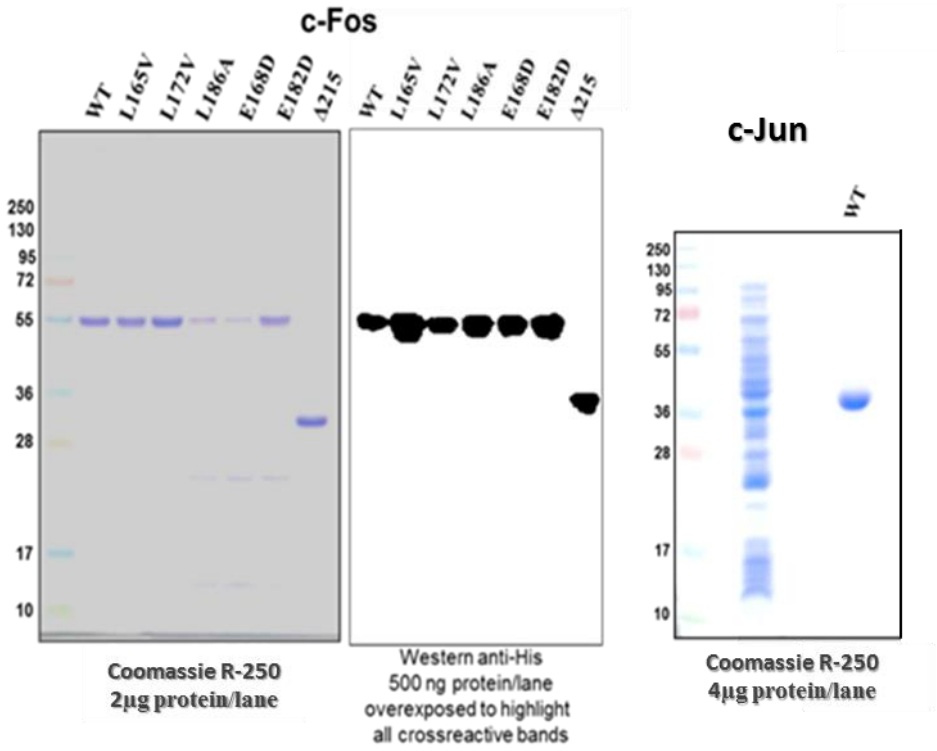

Figure 11. His6-tagged purified c-Fos and c-Jun proteins used in this study. (A) The purity and integrity of the His6-c-Fos variants is shown by Coomassie R-250 staining and western blot after SDS-PAGE. (B) The purity and integrity of the His6-c-Jun wildtype is shown by Coomassie R-250 staining after SDS-PAGE.

We therefore decided to remove the EGFP/mRFP fusion proteins from the recombinant constructs. Instead, these constructs tagged with His6 tag exclusively gave a high expression yield and could be purified under denaturing conditions in the presence of either 6 M GnHCl, 8 M urea or 1% lauroylsarcosine. In order to purify protein in functional form, His6-c-Fos, His6-c-Jun and His6-c-Fos mutants were transformed in Rosetta2 (DE3) pLysS *E. coli* (Novagen) cells by calcium chloride transformation method. The cells were grown to OD600=0.6 and protein expression was induced with 0.1 mM IPTG for 3 hours at 37°C. The cells were harvested and sonicated in 6M guanidine-HCl. The proteins were purified by immobilized metal affinity chromatography on a Ni^{2+}-NTA resin and washed extensively using 25 mM Tris-HCl, 100 mM NaCl at pH-7.5. The protein was dialyzed using

250 mM imidazole and concentrated with an Amicon-Ultra 50 10000 MWCO ultrafiltration device. Protein concentrations were determined by BCA protein assay. 12% SDS PAGE and Western blotting was used to preclude the presence of prematurely terminated products using anti-His antibody (1:5000) dilution (Figure 11). His6-c-Fos (refer as c-Fos) and His6-c-Jun (refer as c-Jun).

4.3 Binding kinetic analysis of c-Fos and variants with c-Jun

With the aim to determine the binding affinity and kinetics of heterodimers of c-Fos and c-Jun we took advantage of biolayer interferometry, which allows for kinetic analysis of macromolecular interactions using small amounts of material. In a typical experiment, a protein A biosensor (the tip of which is coated with protein A) was loaded with anti-Jun antibody/c-Jun complex. The 60A8 anti-c-Jun antibody was precomplexed with c-Jun as described in Materials and Methods section. The analyte, c-Fos or its mutants, were let bind to the c-Jun in known concentrations and the resulting shift in the interference profile was recorded as the association phase of the complex. Then, their dissociation into empty buffer was also followed in time. From the resulting two-phase sensorgram the association and dissociation rate constants (k_a and k_d) and the equilibrium binding constant (K_D) could be calculated after fitting the sensograms to either a Langmuir 1:1 or a 2-state conformational binding model (see under Material and Methods section). Most interactions could be fitted well with a 1:1 model; however, the L165V and L172V mutants better conformed to the 2-state model apparently reflecting more complex binding kinetics (Figure 12). The K_D determined for the wild-type complex matched well with literature data determined by other methods [215]. All substituted mutants hampered the stability of the complex when compared to the c-Fos complex with the exception of the K190R mutant, which showed a slightly increased affinity. The binding affinity was most drastically lowered in the case of the truncated c-$Fos_{\Delta 214}$, which had markedly decreased association and at least two times slower dissociation rates. As expected, changes in the hydrophobic interface residues, L165V and L172V lowered the affinity of dimerization arguably by perturbing the leucine zipper. Reducing the size of the negatively charged residues in E175D and E189D had a moderate effect on dimer formation. The mutant c-FosK190R showed tighter binding between c-Fos

and c-Jun compared to the c-Fos/c-Jun complex. As noted, interface mutations have larger impact on the association of the dimer compared to dissociation kinetics (Table 2).

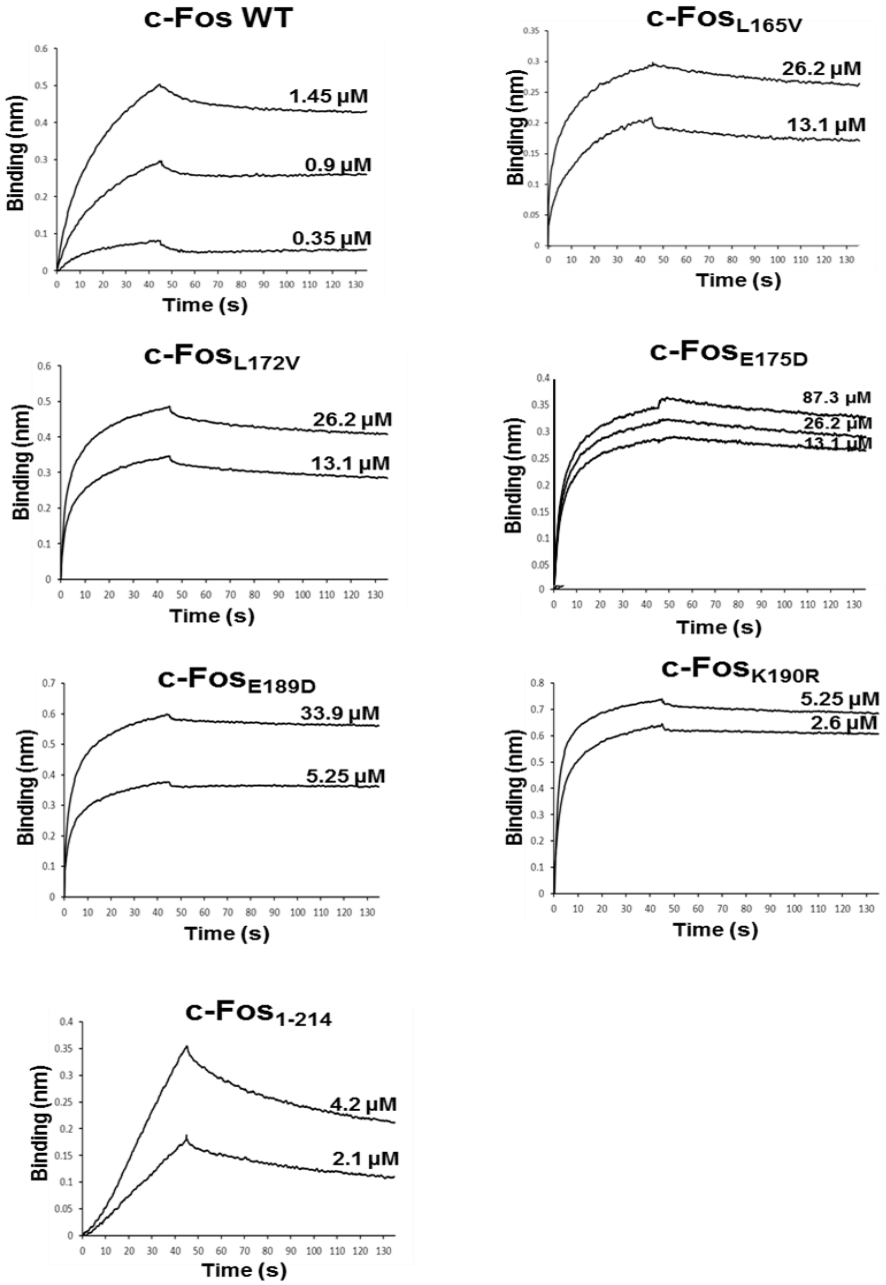

Figure 12. (A) Binding kinetics of c-Fos mutants to c-Jun. c-Jun was immobilized by attachment through an anti-c-Jun antibody to a protein A covered surface. Binding of free wild-type c-Fos, point mutants predicted to decrease, or increase the affinity of the complex, and of a C-terminal deletion mutant at the indicated concentrations to immobilized c-Jun were determined by biolayer interferometry.

	k_a (x10^3 $M^{-1}s^{-1}$)	k_d (x10^4 s^{-1})	K_D (nM)	affinity ratio to wild type
c-Fos	80.3 ± 18.5	32.4 ± 1.95	40.4 ± 22.6	–
c-Fos$_{L165V}$	1.82 ± 0.0337	23.8 ± 0.897	1310 ± 73.8	0.03
c-Fos$_{L172V}$	8.16 ± 0.197	55.1 ± 2.28	676 ± 79.7	0.06
c-Fos$_{E175D}$	7.76 ± 0.119	11.6 ± 0.738	149 ± 85.3	0.27
c-Fos$_{E189D}$	5.53 ± 0.0637	7.02 ± 0.474	127 ± 125	0.32
c-Fos$_{K190R}$	76.4 ± 0.876	7.75 ± 0.481	10.1 ± 5.79	4.0
c-Fos$_{\Delta 214}$	3.11 ± 0.421	77.8 ± 1.44	2500 ± 511	0.02

Table 2. Experimental binding affinities and computed interaction energies of c-Fos mutants with c-Jun. Association and dissociation curves were determined at 37℃ by BLItz analysis using a 1:1 binding model. c-Fos mutants were added to Jun labelled biosensors in different concentrations. Jun was marked by 60A8 Jun antibody (Cell Signaling Technologies). The equilibrium dissociation constant was determined as $K_D=k_d/k_a$, where k_d is the dissociation, k_a is the association rate.

4.4 Determination the degradation rates of c-Fos alone and in the complex with c-Jun

Proteins bearing intrinsically unstructured regions (IUPs) are prone to degradation *in vitro* by the 20S proteasome [10]. Also, *in vivo*, these proteins can undergo degradation in an ubiquitin-independent fashion by a "degradation-by-default" mechanism [216]. The C-terminal half of c-Fos is intrinsically unstructured [212] and involved in UID of c-Fos [213]. To examine the behavior of c-Fos as an IUP, we incubated c-Fos with purified 20S proteasomes and revealed that c-Fos was efficiently degraded alone. Next, we determined the sensitivity of c-Fos for 20S proteasomal degradation *in vitro* in the presence or absence of its binding partner, c-Jun. In the assay the amount of the remaining c-Fos was followed by ELISA over time. We fitted the decay curves of c-Fos alone or in the presence of c-Jun to a one phase decay model and determined the initial degradation rate (V_0) and the degradation half-life ($t_{1/2}$) (Figure 13 and Tables 3 and 4).

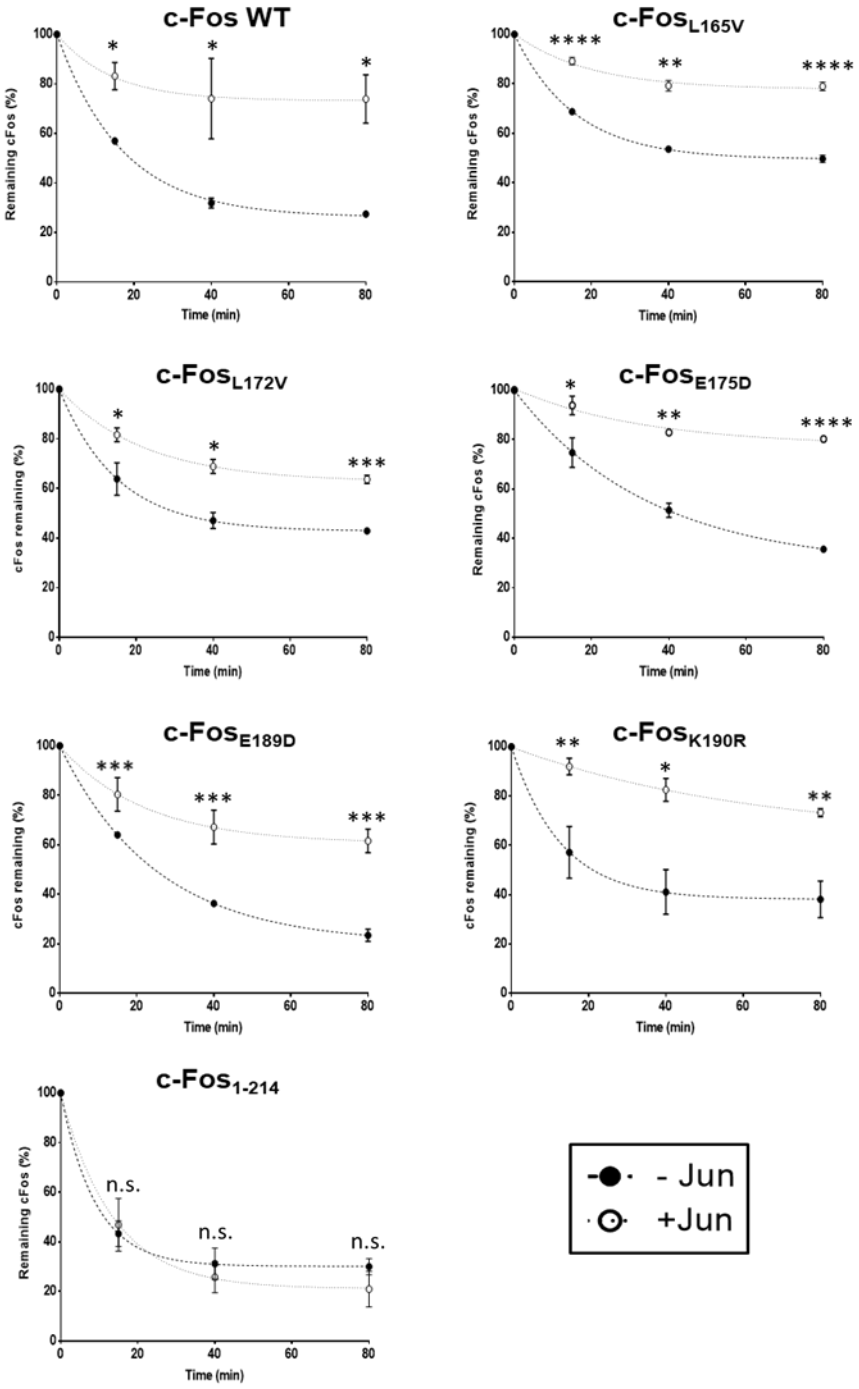

Figure 13. Degradation of various c-Fos mutants and their complexes formed with c-Jun by the 20S proteasome in vitro. (A) The residual amounts of wild-type c-Fos or its mutants were determined by anti-c-Fos ELISA after 0, 15, 40 and 80 minutes of incubation in the presence of purified 20S proteasome. The kinetics of the decay was also determined in the presence of equimolar c-Jun. Unpaired two tailed t-test.

	Initial degradation rate (O.D./min)		
	free (x 10^{-2})	+ c-Jun (x 10^{-2})	ratio
c-Fos	3.00 ± 0.0014	0.74 ± 0.0060	0.247
c-Fos$_{L165V}$	3.46 ± 0.0004	1.01 ± 0.0022	0.292
c-Fos$_{L172V}$	3.31 ± 0.0077	1.98 ± 0.0046	0.598
c-Fos$_{E175D}$	3.08 ± 0.0018	1.29 ± 0.0030	0.419
c-Fos$_{E189D}$	3.07 ± 0.0009	1.46 ± 0.0066	0.476
c-Fos$_{K190R}$	2.46 ± 0.0072	0.24 ± 0.0008	0.097
c-Fos$_{\Delta 214}$	2.86 ± 0.0048	2.69 ± 0.0048	0.941

Table 3. Initial degradation rates of c-Fos and its mutants with and without c-Jun during degradation by purified human 20S proteasomes. Degradation assays were performed using purified human 20S proteasome by measuring c-Fos concentration after 0, 15, 40 and 80 minutes. Nonlinear data fitting was carried out by the R program. All reactions were run in triplicates.

	Half-life (min)	
	free	+ c-Jun
c-Fos	11.64 ± 0.44	19.18 ± 9.00
c-Fos$_{L165V}$	10.84 ± 0.13	14.16 ± 3.74
c-Fos$_{L172V}$	10.93 ± 3.53	15.69 ± 2.76
c-Fos$_{E175D}$	23.95 ± 2.35	23.67 ± 7.90
c-Fos$_{E189D}$	17.36 ± 1.15	20.73 ± 2.32
c-Fos$_{K190R}$	9.28 ± 2.86	151.95 ± 33.87
c-Fos$_{\Delta 214}$	9.07 ± 1.25	13.85 ± 4.09

Table 4. Half-lifes of c-Fos and its mutants with and without c-Jun during degradation by purified human 20S proteasomes. Degradation assays were performed by measuring c-Fos concentration after 0, 15, 40 and 80 minutes. Nonlinear data fitting carried out by the R program. All reactions were run in triplicates.

E to D replacements stabilized the free c-Fos, whereas the K190R and c-Fos$_{\Delta 214}$ did not affect or very slightly shortened the half-life. As shown in Table 4 the interaction of c-Jun significantly slowed the degradation of c-Fos, increasing its half-life from 11.64 minutes to 19.18 minutes, which is a par excellence demonstration that it is the ‘nanny’ of its partner.

Since we found that a fraction of substrate escaped degradation at the end of the proteasomal degradation assay we also determined the initial degradation rates in order to characterize how the affinities manipulated by the site-specific mutations or C-terminal truncation influenced the ability of the nanny c-Jun to protect the c-Fos mutants from 20S proteasomal degradation.

We could show that the change in the initial degradation rate upon addition of c-Jun ($v_{0,+c\text{-}Jun}/v_{0,free}$) correlated positively with the K_D, that is, mutations decreasing the affinity between the two proteins to a larger degree prevented c-Jun from protecting its client more (Figure 14). We observed that all the designed mutations had a larger impact on the c-Fos degradation rate in the presence than in the absence of c-Jun, whereas the v_0 in the free forms were very comparable between mutants (Table 3). The c-Fos turnover is increased in all mutants, irrespective of the destabilizing or stabilizing behavior of the mutations. The impact on the initial degradation rates of c-Fos, however, is in accord with the interaction affinities observed by the BLItz measurements. As expected, destabilizing mutations affecting hydrophobic contacts (L165V, L172V) exhibited smaller decrease in the initial degradation rates ($\Delta V_{mut}/\Delta V_{wt}$) compared to those affecting hydrophilic interactions (E175D, E189D). Only mutant K190R slowed down the initial degradation rate as compared to wildtype c-Fos (Figure 13).

We included the C-terminal deletion mutant in the analysis together with the point mutants, because the deletion did not alter much the degradation rate of the mutant compared to the wild-type in the absence of c-Jun. The comparison of c-Fos and c-$Fos_{\Delta 214}$ revealed that the presence of the c-Fos tail within the complex also decreased the degradation rate almost by almost 4-fold (Table 3). Thus, removal of the fuzzy C-terminal segment considerably decreased c-Fos protection and complexation with c-Jun could barely influence the degradation rate (Table 4).

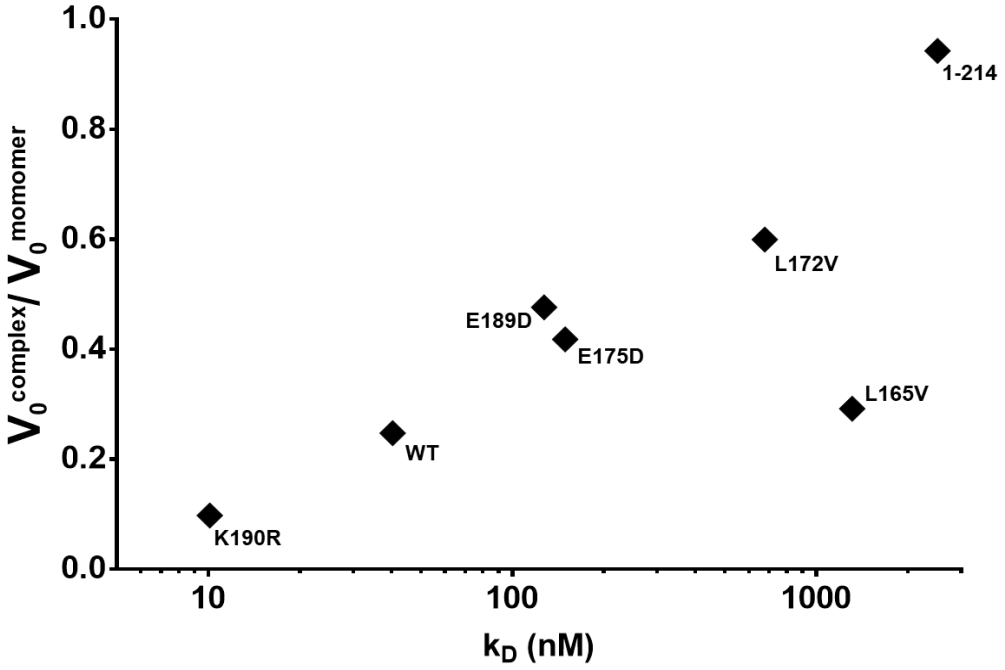

Figure 14. Change in c-Fos degradation rate via c-Jun interactions ($v_{0,\ +c\text{-}Jun}/v_{0,\ free}$) as a function of the binding affinity. The experimental K_D values are shown on a logarithmic scale.

4.5 Estimation of protein disorder in free c-Fos and c-Fos/c-Jun complex by CD-spectroscopy

To determine whether protein disorder is preserved when c-Fos and c-Jun interact with each other or one or both of the proteins fold up upon binding to its partner, we performed near-UV circular dichroism spectroscopy. We were also interested to assess the impact of the mutations on the structure of c-Fos in free form and in complex with c-Jun. Our observations are similar with the previous experimental results, almost half of the full length c-Fos (185 residues) does not possess regular secondary structures and remains extended form [208]. Our results show that the CD-spectrum of the complex is very similar to the average of the spectra of the individual proteins with slightly increased in helical population (Table 5). Secondary structure prediction estimated the disorder content to be around 60 % in all three cases (60.4 % for c-Fos, 61.3 % for c-Jun, 58.9 % for the complex), our results indicating that c-Fos and c-Jun retain as much disorder in the heterodimeric form as in their

monomeric forms, suggesting that fuzzy interactions between the disordered tails are responsible for the nanny effect.

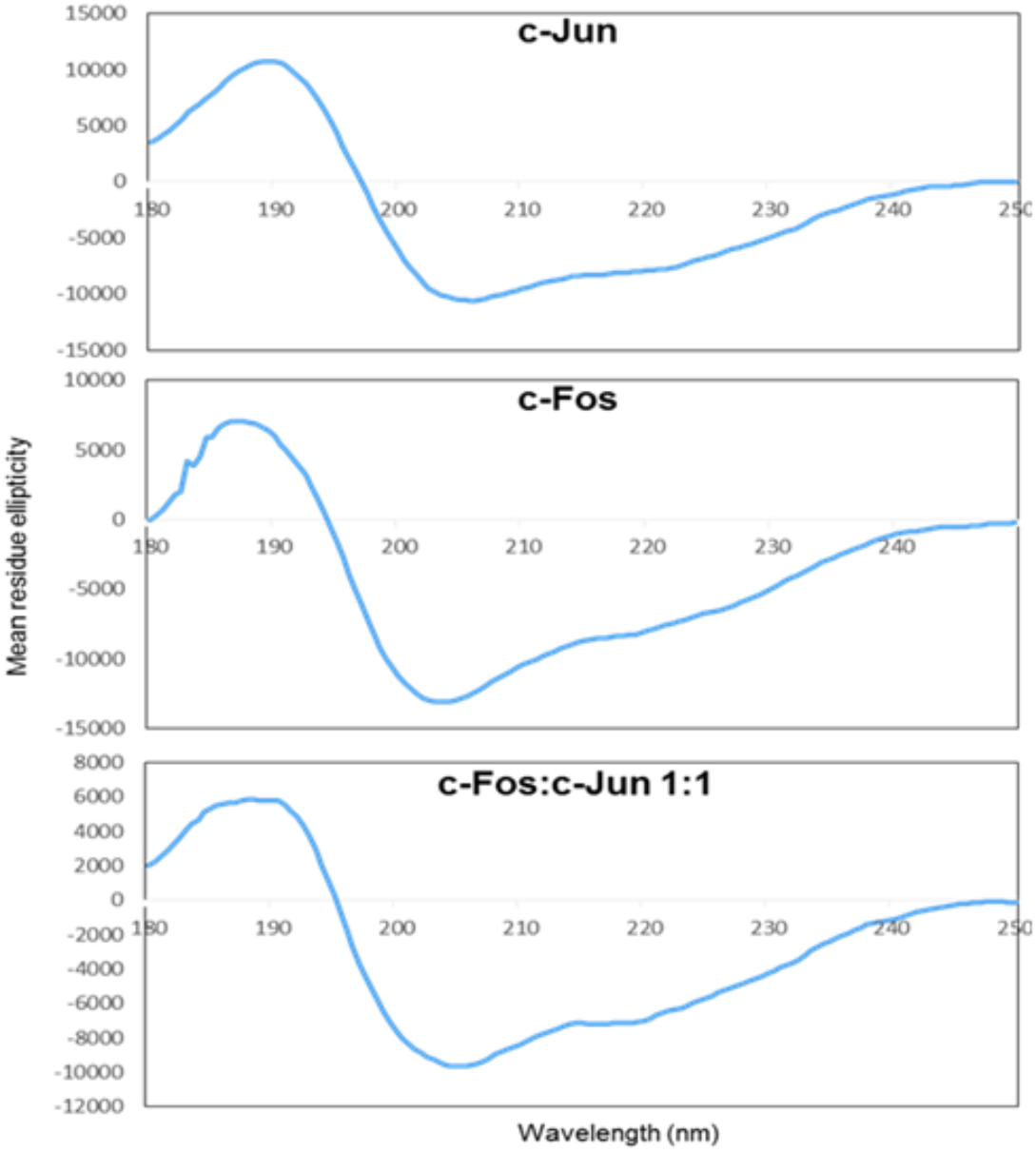

Figure 15. A) Near-UV ECD-spectra of c-Fos, c-Jun monomeric proteins and the c-Fos/c-Jun complex.

	α-Helix	β – strand (antiparallel)	β – strand (parallel)	Turn	Other
c-Fos	11.0	20.2	7.4	12.6	48.7
c-Jun	17.5	17.9	6.2	14.1	44.4
c-Fos/c-Jun	14.2	19.2	7.8	14.7	44.2
c-Fos$_{L172V}$	17.1	21.1	2.4	14.9	44.5
c-Fos$_{L172V}$/c-Jun	12.2	31.1	0.0	14.3	42.5
c-Fos$_{E175D}$	23.9	19.1	3.4	11.6	42.1
c-Fos$_{E175D}$/c-Jun	23.9	19.1	3.4	11.6	42.1
c-Fos$_{\Delta 214}$	19.2	18.5	1.1	15.0	46.2
c-Fos$_{\Delta 214}$/c-Jun	14.9	17.9	5.5	14.6	47.1

Table 5. Propensity of secondary strutcure elements (%) as determined by ECD spectroscopy. ECD spectra were recorded on a J-810 spectropolarimeter in 20 mM sodium phosphate, pH 7.4 buffer solution with a 0.2 mm quartz cell at room temperature. Raw spectra were corrected with the blank, then were converted to mean residue ellipticities and were smoothed. Deconvolution was performed by the BeStSel program.

Thus, we can highlight that c-Fos forms a fuzzy complex with c-Jun, where 45.8% of the residues remain to be disordered which is consistent with previous FRET and FCCS results [215].

On the other hand, L172V and E175D replacement cause only a minor increase in the secondary structure of c-Fos. As observed, structural disorder of the mutants is retained upon assembly with c-Jun, with minor impact on the secondary structure properties of the L172V and E175D mutant assemblies as compared to the wild-type protein. Although surprisingly, the removal of disordered C-terminal region from full-length c-Fos protein did not modify considerable ordering when compared with c-Fos, in the absence of c-Jun.

4. RESULTS: PART-II-UNPUBLISHED DATA

4.6 Generation of random mutant library

To design random mutant libraries of c-Fos and c-Jun, we made use of error-prone PCR method using a GeneMorph II Random Mutagenesis Kit. Error-prone PCR method is a polymerase chain reaction and was performed using Mutazyme (2.5 U/µl) enzyme to generate random mutations in the c-Fos and c-Jun wildtype sequences. PCR products of c-Fos and c-Jun were restriction digested followed their ligation with the plasmid vector containing EGFP and mRFP fusion proteins respectively. The ligated plasmids were transformed into DH5α competent cells in order to have isolated mutant colonies on the agar plate. To account the rate of mutations, sequencing was performed using forward and reverse sequencing primers. Simultaneously, to determine the size of each library, colonies were counted manually. In total, two random libraries with higher and lower mutation rates have been generated from full length cFos and cJun DNA sequence and will be used for future experiments (Table 6).

Library Name	Mutation Rate/Gene
c-Jun library A	4-9
c-Jun library B	1-4
c-Fos library A	1-5
LIBRARY SIZE – 10^5	

Table 6: Random mutant library of c-Fos and c-Jun DNA sequence using Gene Morph-II Random Mutagenesis.

4.7 Expression and purification of His6-c-Fos-EGFP and His6-c-Jun-mRFP.

The *in vitro* expression plasmid pETDuet-1 containing the His6-c-Fos or His6-c-Jun cDNA coding region was expressed in *E.coli* BL21 (DE3) competent cells. Cells were lysed by sonication and the insoluble fraction was dissolved in a guanidine-HCl-containing buffer and homogenized with a Dounce homogenizer and a B-pestle. Extracted inclusion bodies were then recovered overnight at room temperature. Denatured His6-cFos-EGFP protein bound in batch to 1-2 ml HisPur^{TM} Ni-NTA resin on a rotator for 2–4 hours at 4 °C. Following the elution, fractions purity was determined by 12% SDS-PAGE as shown below in the figure 16. Pure (>80 %) fractions were then pooled, and dialyzed. Western blotting was also used to verify the presence His6-c-Fos-EGFP and His6-c-Jun-mRFP proteins using anti-His antibody (1:5000) dilution in the eluted fractions. The expression and purification protocols were identical for both the wild-type and the mutants.

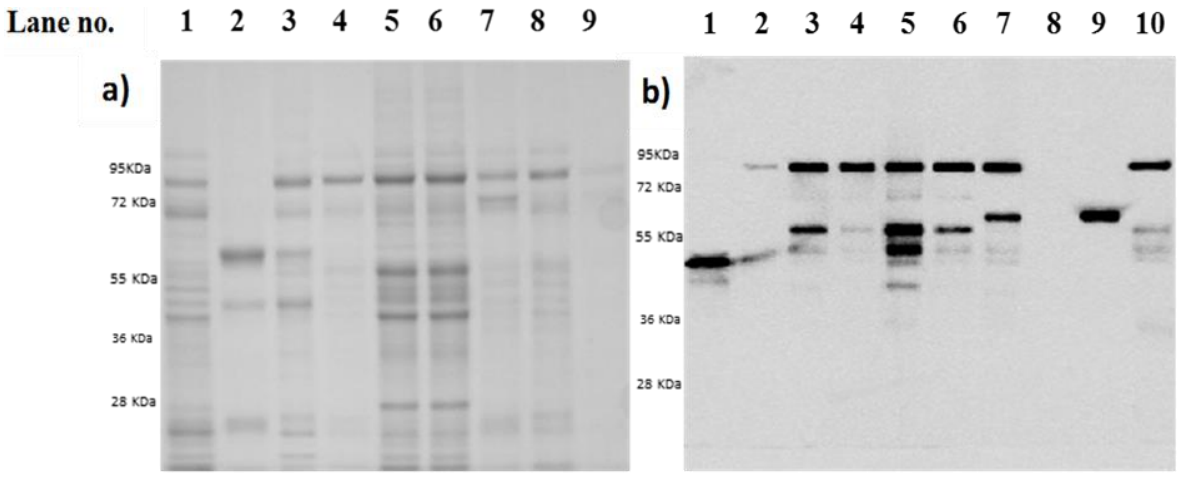

Figure 16. (A) Separation of purified His6-c-Fos-EGFP WT and mutant protein on SDS-PAGE gel. Lane1- WT, Lane2- His6-c-Fos$_{\Delta 214}$, Lane3- His6-c-Fos166-EGFP, Lane4- His6-c-Fos168-EGFP, Lane5- His6-c-Fos175-EGFP, Lane6- His6-c-Fos176-EGFP, Lane7- His6-c-Fos182-EGFP, Lane8- His6-c-Fos189-EGFP, Lane 9- His6-c-Fos190-EGFP. (B) Western Blot analysis after SDS-PAGE of purified His6-c-Fos-EGFP WT, His6-c-Jun and mutant proteins. Lane1- His6-c-Jun WT, Lane2- His6-c-Fos190-EGFP, Lane3- His6-c-Fos189-EGFP, Lane4- His6-c-Fos182, Lane5- His6-c-Fos176-EGFP, Lane6- His6-c-Fos175-EGFP, Lane7- His6-c-Fos168-EGFP, Lane8- His6-c-Fos166-EGFP, Lane9- His6-c-Fos$_{\Delta 214}$-EGFP, Lane10- His6-c-Fos-EGFP WT.

4.8 Expression and purification of recombinant His6-c-Jun and its mutants.

Purification c-Jun fused to larger globular protein tag mRFP proved challenging. We expected that these fusion proteins would make the largely disordered c-Jun more soluble and stable. However, a single step immobilized metal affinity chromatography (IMAC) against the N-terminal 6xHis-tag was not sufficient to yield pure protein from the soluble fraction of the bacterial lysate homogenized with 1% Triton-X100 (Figure 17). Denaturing purification with 6 M GnHCl yielded much purer proteins based on 12% SDS-PAGE and Coomassie staining. However, during refolding by sequential dialysis these proteins mostly precipitated at the bottom of the tube. The globular protein part also never recovered its activity or fluorescence.

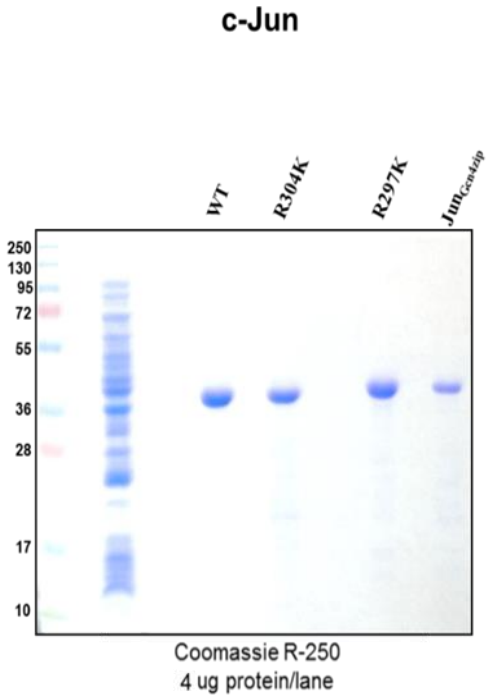

Figure 17. His6-tagged purification of c-Jun wildtype and mutants protein. The purity and integrity of the c-Jun variants is shown by Coomassie R-250 staining after SDS-PAGE.

We therefore decided to remove the mRFP fusion proteins from the recombinant constructs. Instead, these constructs tagged with His6 tag exclusively gave a high expression

yield and could be purified under denaturing conditions in the presence of either 6 M GnHCl, 8 M urea or 1% lauroylsarcosine. In order to purify protein in functional form His6-c-Jun and His6-c-Jun mutants were transformed in Rosetta2 (DE3) pLysS *E. coli* (Novagen) cells by calcium chloride transformation method. The cells were grown to OD600=0.6 and protein expression was induced with 0.1 mM IPTG for 3 hours at 37°C. The cells were harvested and sonicated in 6M guanidine-HCl. The proteins were purified by immobilized metal affinity chromatography on a Ni^{2+}-NTA resin and washed extensively using 25 mM Tris-HCl, 100 mM NaCl at pH-7.5. The protein was dialyzed using 250 mM imidazole and concentrated with an Amicon-Ultra 50 10000 MWCO ultrafiltration device. Protein concentrations were determined by BCA protein assay. 12% SDS PAGE and Western blotting was used to preclude the presence of prematurely terminated products using anti-His antibody (1:5000) dilution. We plan to use c-Jun mutants in future.

4.9 Verification of the interaction of the purified c-Fos and c-Jun proteins

We next investigated the physical interaction between the purified c-Fos and c-Jun proteins *in vitro*. To verify the interaction between His6-c-Fos (refer as c-Fos) and His6-c-Jun (refer as c-Jun) we co-immunoprecipitated them using 60A8 anti-c-Jun antibody. Controls included c-Fos added to protein A sepharose beads, c-Fos added to protein A sepharose beads carrying the antibody in the absence of c-Jun, and c-Fos added to protein A sepharose beads carrying an isotype matched unrelated rabbit mAb. As shown in Figure 18B, c-Fos co-immunoprecipitated with c-Jun.

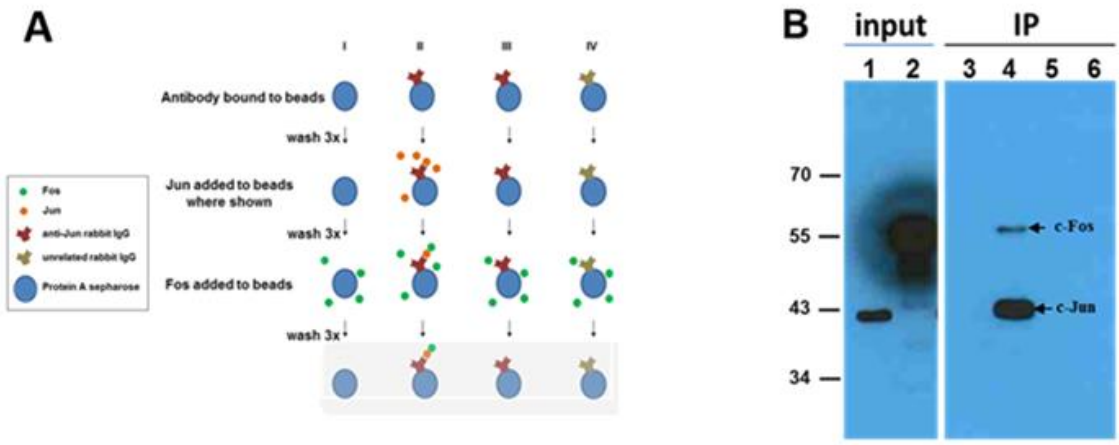

Figure 18. Interaction between c-Fos and c-Jun using in vitro technique. (A) Schematic representation of a standard co-immunoprecipitation protocol (B) lane 1; c-Jun alone, lane 2; c-Fos alone, lane 3; c-Fos bound to empty beads, lane 4; c- Fos bound to anti-Jun + c-Jun, lane 5; c- Fos bound to anti-Jun, lane 6; c-Fos bound to unrelated IgG.

c-Fos binding to c-Jun was specific, because it could not be detected in the absence of c-Jun. Importantly, binding of the anti-c-Jun antibody to c-Jun did not interfere with the c-Jun/c-Fos interaction, which will be crucial for the immobilization of c-Jun for binding kinetic measurements by biolayer interferometry.

4.10 Binding of purified c-Fos and c-Jun to the AP-1 element

To verify that the His6-tagged c-Fos and c-Jun form a functional heterodimer that is able to bind the consensus AP-1 response element a probe containing the TRE sequence TGACTCA was hybridized from two oligonucleotides and used for an electrophoretic mobility shift assay by incubation with the purified c-Fos and c-Jun. When both c-Fos and c-Jun were mixed with the oligonucleotide a shift could be detected in the mobility of the probe during acrylamide gel electrophoresis indicating that the c-Jun and c-Fos proteins were able to heterodimerize and the complex bound the cognate AP-1 response element in the DNA. In contrast, c-Jun alone failed to bind to the DNA (Figure 19).

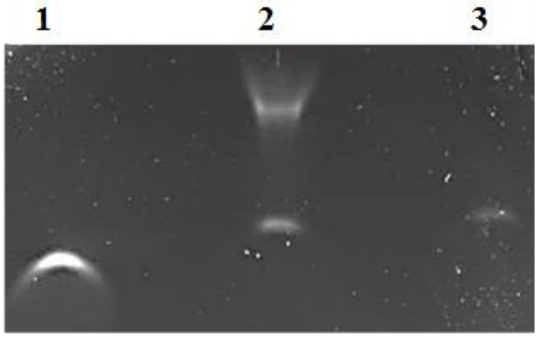

Figure 19. Electrophoretic mobility shift assay for His6-c-Fos and His6-c-Jun heterodimer binding to an AP-1 site containing fragment. Lane1; control (probeDNA), Lane2; c-Fos/c-Jun with probe DNA and Lane3; c-Jun/c-Jun with probe DNA.

4.11 Co-expression of c-Fos-EGFP and c-Jun-mRFP in mammalian system

In eukaryotic HEK293AD cells, co-expression of mRFP-tagged c-Jun with c-Fos-EGFP tremendously increased their expression (Figure 20). As the expression of both proteins was driven by strong CMV promoters, which would not be influenced by the overexpression of the factors themselves, the proteins post-translationally stabilized each other.

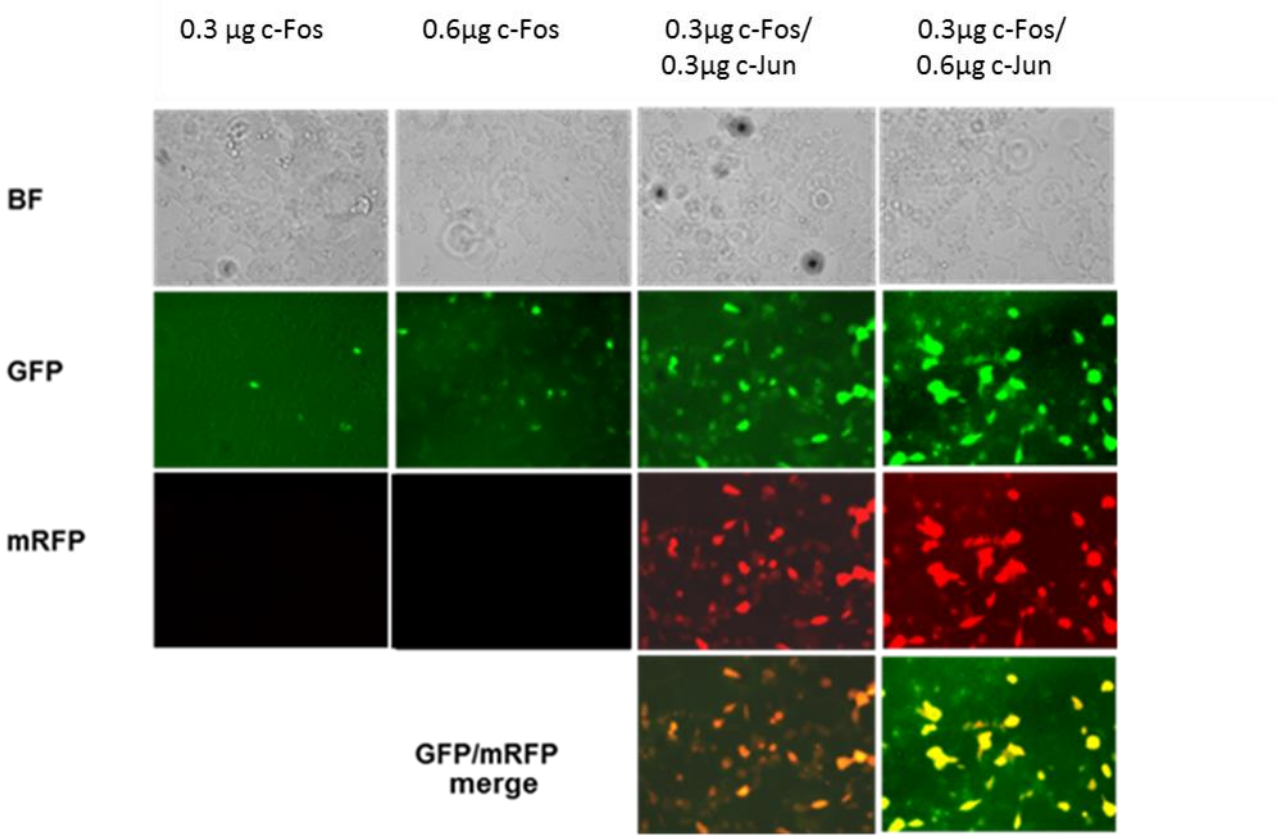

Figure 20. Co-expression of c-Jun increases the steady-state expression level of c-Fos-EGFP. c-Fos-EGFP was expressed either alone or together with c-Jun-mRFP in HEK293AD cells. Increasing amounts of co-expressed c-Jun-mRFP increased the level of the c-Fos-EGFP.

5. DISCUSSION

IDPs play indispensable functions in the life of the cell being involved in essentially every cellular process, from transcription regulation through DNA repair to signal transduction. Consequently, precise control of the ever required levels of IDPs is crucial for maintaining all cellular processes [139]. Besides being controlled at synthesis the termination of a protein's life-span also offers a point for regulation. The half-lives of proteins range from minutes to days and change in response to the biological context, attesting to the exquisite fine-tuning of protein turnover. Alterations of half-life affect the immune and stress responses, DNA repair, programmed cell death, and flaws of protein homeostasis have been linked to developmental problems, cancer and neurodegenerative disease [217–219]. For instance, a merely 8-minute increase in the half-life of the Hes7 transcription factor severely disrupts embryonic development in the mouse [220]. The proteasome is the major machinery that is responsible for the elimination of damaged, misfunctional or misfolded sequences and also for specific and controlled degradation of proteins in eukaryotic cells [145]. The eukaryotic proteasome consists of a 19S subcomplex for the recognition and active unfolding of ubiquitinated proteins, and a 20S core complex harboring the catalytic protease subunits. The assessment of protein half-life on a genomic scale in several organisms yielded results, that correlate surprisingly well with dynamic features of the protein sequences, suggesting that the decision to degrade a protein is strongly connected to the unstructured/unfolded state. The 20S particle is not necessarily complemented with the recognition particle and also exists and functions independently. Natively unstructured proteins can bypass the need for the 19S particle explaining that proteins which lack a tertiary structure in their native states are enriched among substrates of the 20S proteasome. This leads to the identification of specific biochemical mechanisms by which the amino acid sequence can determine the rate of a protein's degradation. The observations are consistent with a view in which dynamic polypeptide segments of lengths (>30 AA) sufficient to enter the proteasome's cavity (either at the termini of a protein or internally) are needed for efficient degradation and proteins containing such segments have shorter half-lives. [221]. The sequence composition of these unstructured initiation sites further modulates the affinity for the proteasome and determines the proteasome's processivity along a protein [132]. The other paramount element in deciding

the half-life of proteins is interaction with regulatory proteins. To distinguish non-functional segments and natively unstructured yet very functional sequences presents a challenge to regulated protein degradation [222]. The survival of disordered proteins, how they evade degradation by default by the 20S proteasome even in an ubiquitination independent manner needs to be accounted for. A plausible model that provides explanation for this proposes that protein interactions might shield exposed dynamic regions and make them inaccessible for the 20S core particle. This mechanism, which seems to have a wide-spread effect on intracellular levels of important disordered proteins, has been termed the 'nanny hypothesis', where the protective partner factor is referred to as 'nanny' [10, 7]. Although the nanny model has been applied to a number of examples, where complexation extends protein half-life (for example ikBa-NFkB [166], Ku70-Ku80 [223] the detailed molecular mechanisms remained to be elucidated.

In the present study we used the activator protein 1 (AP-1) as a model system to study the relationship between ID protein interactions and half-life. We aimed to explore how specificity and affinity of protein interactions between an IDP and a partner influences default 20S proteasomal degradation, that is, how c-Fos turnover is modulated by specific mutations affecting specific interactions by the structured elements at the contact sites with c-Jun and by the presence of the fuzzy tail.

We opted to use the full-length c-Fos (380 AA) and c-Jun (331 AA) and considerable effort had to be invested in the purification of the proteins. A single step IMAC against the N-terminal 6xHis-tag was not sufficient to yield pure protein from the bacterial lysates homogenized with Triton-X100 and a further reverse phase chromatography step also proved insufficient. Purification of c-Fos and c-Jun fused to larger globular protein tags, such as EGFP/mRFP proved challenging. We expected that these fusion proteins would make the largely disordered c-Fos and c-Jun more soluble and stable, surprisingly was insoluble in a broad spectrum of buffers. Denaturing purification with 6 M GnHCl yielded soluble and much purer proteins based on 12% SDS-PAGE and Coomassie staining, but during refolding by sequential dialysis these mostly fell out of solution and precipitated. The globular protein part also never recovered its activity. We finally removed the EGFP and mRFP fusion proteins, and constructs tagged with His6 exclusively gave a high expression yield and could

be purified under denaturing conditions in the presence of either 6 M GnHCl and could be successfully refolded.

We designed five mutants to affect either the hydrophilic interactions or the hydrophobic contact points within the leucine zipper. In the mutants E175D, E189D and L165V and L172V shorter side-chains were introduced in place of two glutamic acid residues or two leucines of the LZ to perturb the ideal geometry and to plausibly reduce binding affinity. The lysine to arginine (K190R) replacement introduces a longer charged side chain with additional π interactions. The C-terminus of c-Fos has been shown to behave dynamically both when free and as bound to c-Jun. We designed a truncation mutant lacking the C-terminus (c-Fos$_{\Delta 214}$) to investigate the effect of this C-terminal fuzzy tail and its interplay with the structured interacting regions.

The alteration of proteasomal removal can theoretically be the consequence of enhanced folding upon binding of two proteins to each other. To determine whether protein disorder is preserved when c-Fos and c-Jun interact with each other or one or both proteins fold up upon binding to their partners, we performed electronic circular dichroism (ECD) spectroscopy in the free and the bound states of c-Fos, a hydrophilic (E175D) and a leucine mutant (L172V), and the C-terminally truncated c-Fos$_{\Delta 214}$. The CD-spectrum of the c-Fos/c-Jun complex was very similar to the average of the spectra of the individual free proteins, and we noted that complex formation with c-Jun only slightly increased the helical population as compared to the c-Fos free state, meaning that extensive ordering did not occur. Consistent with previous FRET and FCCS results, we could confirm that c-Fos formed a fuzzy complex with c-Jun, where 45.8% of the residues remain disordered [215]. The mutants L172V and E175D showed only minor increases in their secondary structures with respect to c-Fos. The secondary structure content prediction estimated that the structural disorder of the c-FosL172V and c-FosL175D mutants is retained upon assembly with c-Jun, with small impact on the secondary structure properties as compared to the wild-type protein. On the other hand, we did not observe any substantial unfolding of the structured part in c-Fos$_{\Delta 214}$, as the number of residues with regular secondary conformations (> 200 AA in full-length c-Fos, 113 AA in c-Fos$_{\Delta 214}$) exceeded the size of the bZIP (62 AA). Taken together, our data

demonstrated consistently with previous observations by others that c-Fos and all the studied variants form fuzzy complexes with c-Jun and do not completely fold upon binding.

The biolayer interferometric binding kinetics measurements revealed that the binding affinities between c-Fos substitution mutants and c-Jun decreased by less than two orders of magnitude, except the one mutant K190R mutation when compared to full length c-Fos. We observed that inclusion of the fuzzy tail in full-length c-Fos improved binding affinity (40 nM) in contrast to interaction studies determined between the leucine zippers (63 aa, K_D=54 nM) [224]. Perturbing the hydrophobic interface (L165V, L172V) had the most considerable effect on the stability of the complex due arguably to looser zipper contacts. Reducing the size of the negatively charged residues (E175D, E189D) had a moderate impact on K_D, due to weaker electrostatic stabilizing interactions. The K190R mutation slightly stabilized the heterodimer *via* additional π-π interactions with Q313. The C-terminal truncation in c-$Fos_{\Delta 214}$ considerably destabilized dimer formation. The removal of the fuzzy C-terminal region in c-$Fos_{\Delta 214}$ considerably increased the dissociation rate from c-Jun as compared to the wild-type protein. The association of the bZIP domains is limited by electrostatic repulsion, which is usually masked by the flanking disordered regions. A similar trend was observed in c-Max and c-Myc assembly [157, 225]. Along these lines, truncating the fuzzy tail in c-$Fos_{\Delta 214}$ unfavorably affects formation of the coiled-coiled structure. We believe that the effect of the C-terminal fuzzy tail of c-Fos can also be interpreted within the framework of the fly-casting model. Disordered regions establishing nonspecific (e.g. electrostatic) interactions can increase the effective local concentration of a neighboring specific structured interacting region as this tethered segment is only free to sample a limited space [226, 227]. Along these lines fuzzy regions in protein complexes can serve as nonspecific anchors, which remain attached and decrease dissociation rates even in the absence of specific contacts [228] In accord with previous data, the association kinetics seems to be affected by mutations in full-length c-Fos [229, 230]. However, binding kinetics of full-length proteins without complete folding might exhibit complex kinetics [164, 231] (and references therein), as compared to truncated protein fragments, which fold upon binding [232].

Replacements of charged residues in c-Fos also slowed down the dissociation kinetics by 3-4 folds. We noted in case of the c-FosL165V and c-FosL172V mutants using a 2 state mode indicate more complex kinetics.

We evaluated how c-Fos turnover is modulated by the mutations at the specific contact sites with c-Jun and by the presence of the unstructured tail. We set up an ELISA based proteasomal degradation assay to quantitate *in vitro* the sensitivity of the different c-Fos mutants to 20S proteasomal degradation in the presence or absence of c-Jun. We used the change in the initial degradation rate (V_0) and the extension in the half-life ($t_{1/2}$) to characterize the stabilizing effect. An equimolar c-Jun slowed the rate of degradation of c-Fos ~5 times and increased its half-life from 11.64 minutes to 19.18 minutes, which was a par excellence demonstration that c-Jun is the 'nanny' of its partner, c-Fos. All the incorporated substitutions had a larger effect on c-Fos half-life in the presence rather than in the absence of c-Jun, and the V_0 values corresponding to the monomeric states did not change much in the site-directed mutants. This can be expected if the mutations do not influence the interaction with the proteasome per se, and if the degradation signal is conserved in the disordered tail. c-Fos turnover was prolonged in all cases irrespective of whether a mutation stabilized or destabilized the complex indicating that interactions with the binding partner - however weak or strong - influenced c-Fos degradation. We noted that the impact on the initial degradation rates correlated inversely with the interaction affinity of the mutant. The destabilizing L172V mutation, which affects hydrophobic contacts exhibited smaller decrease in the initial degradation rates than those affecting hydrophilic interactions (E175D, E189D). Destabilizing the complex reduced the protecting role of the binding partner as compared to the wild-type. On the other hand, the stabilizing K190R mutation slowed down the initial degradation as compared to the wild-type c-Fos. The c-FosL165V somewhat deviated from the trend most probably due to its considerably decreased helicity and decreased dissociation rates. The presence of the c-Fos fuzzy tail in the complex also decreased the degradation rate almost by 4-fold, consistent with its observed contribution to the stability of the AP-1. Thus, removal of the fuzzy C-terminal segment decreases c-Fos protection *via* weakening the complexation with c-Jun. Taken together, although interaction affinity is likely to be the major determinant in stabilizing c-Fos, the relationship appears to be rather complex. Overall, we showed that protection by c-Jun qualitatively correlates to the binding affinity; and

destabilizing the complex reduces the impact of the mutation on degradation rates. However, there are indeed additional factors, which can also influence stabilization, for example decreased dissociation rates or steric effects (most prominent may be in the larger system) indicating a rather complex protection mechanism. Taken together our results demonstrate that both specific and transient, nonspecific interactions influence half-life.

That c-Fos and c-Fos mutants retained as much disorder in the AP-1 complex as in their monomeric forms, suggests that fuzzy interactions between the disordered tails had to be responsible for the nanny effect. This is consistent with the fact that removal of the C-terminal of c-Fos dramatically increased the K_D of the interaction implying that the removed segment contributed substantially to binding. Also, when c-Fos is truncated in its C-terminal segment c-Jun was unable to exert nearly as significant a protection from the 20S proteasome. Our findings imply that protection of disordered regions can be achieved without inducing a proper stable structure and many binding configurations could be visited in the complex without disturbing the conformational entropy. The protective role of fuzzy interactions from the 20S proteasome could also provide a plausible explanation for how low-complexity sequence motifs might serve as selective inhibitors of proteolysis [233, 234] . Tandem repeats of short, linear sequences are frequently associated with proteins, which form higher-order protein structures or undergo liquid-liquid phase transition [235]. Interactions between these motifs are often not specific or well-defined, and multivalency [236] or fuzziness [237] is a ubiquitous feature of these associations. We may assume that such weak, heterogeneous contacts could also serve to protect disordered stretches with multivalent, low-complexity motifs [114] from the ubiquitin independent degradation pathway.

6. SUMMARY

- We demonstrated that structural order does not increase in the complex of full length c-Fos and c-Jun, therefore the two forms a fuzzy complex.

- We have shown that protein turnover can be modulated *via* fine-tuning the association with a binding partner.

- We demonstrate that protection of disordered regions does not require induced folding, but can be achieved within a fuzzy complex of nanny and client through weak heterogeneous interactions.

- The data suggest a plausible model of how post-translational modifications by changing the dynamics within such a complex may influence half-life.

- Our work predicts that low-complexity dynamic regions engaging in fuzzy interactions may serve as selective inhibitors of proteolysis.

Keywords: Intrinsically disordered proteins, folding to binding, fuzzy complexes, 20S proteasome, AP-1 complex, nanny model, protein regulation.

7. PERSONAL ACKNOWLEDGEMENTS

First and foremost I would like to express my gratitude to my supervisor Prof. Dr. Monika Fuxreiter for providing me with an opportunity to work in her research group and for the continuous financial support, motivation, numerous discussions and guidance. Her guidance helped me in all the time of research and writing of this thesis.

I would like to express my sincere gratitude to Professor Fésüs László and Professor József Tőzsér for giving me an opportunity to work at the Biochemistry and Molecular Biology department.

I am very grateful to Dr. Máté A. Demény for guiding me throughout my PhD and for helping me in experiments and for his valuable suggestions. A special thanks to our collaborator Dr. Pietro Gatti-Lafranconi from the University of Cambridge for his valuable inputs and feedback regarding my project.

Special thanks to Dr. Szentandrassyné Dr. Gönczi Mónika and Dr. Mahdi Mohamed for helping me in cell culture experiments and reversed-phase chromatography method and warm thanks goes to all my colleagues and friends at the department of Biochemistry and Molecular Biology for their consistent help and co-operation during my PhD work. I also thank them for helping me in any condition when I needed.

My gratitude also goes to Agota Veres, Iren Mezo, Szilvia Szalóki for their excellent technical assistance. I would also like to thank all the co-authors in the publications for their contributions to my project.

I would like to thank my friends especially Rita Elék, Károly Jambrovics, Jennifer Nagy, Fruzsina Zsólyomi , Pálma Szabó, Renuka Nair, Soniya Sharma and Varsha Jadhwani for their friendship, love and be there for me when I needed them the most. I am also grateful to my friends in Canada, United States and Germany for their encouragement and support during difficult times.

Especially, I am very grateful to my parents, sister, brother and bother in law for their love, prayers, sacrifice and emotional support throughout my PhD. Last but not the least; I am really thankful to my understanding husband Pramit Dutta for his unconditional love, patience and support during the PhD and my life in general.

8. REFERENCES

1. Romero P, Obradovic Z, Kissinger C, Villafranca JE, Dunker a. K. Identifying disordered regions in proteins from amino acid sequence. *Proc Int Conf Neural Networks*. 1997;(1):1-6.

2. Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J Mol Biol*. 1999;293(2):321-331.

3. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform*. 2000;11:161-171.

4. Dunker AK, Lawson JD, Brown CJ, et al. Intrinsically disordered protein. *J Mol Graph Model*. 2001;19(1):26-59.

5. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol*. 2004;337(3):635-645.

6. Gsponer J, Madan Babu M. The rules of disorder or why disorder rules. *Prog Biophys Mol Biol*. 2009;99(2-3):94-103.

7. Tsvetkov P, Reuven N, Shaul Y. The nanny model for IDPs. *Nat Chem Biol*. 2009;5(11):778-781.

8. Orlowski M, Wilk S. Ubiquitin-independent proteolytic functions of the proteasome. *Arch Biochem Biophys*. 2003;415(1):1-5.

9. Asher G, Shaul Y. p53 Proteasomal Degradation: Poly-Ubiquitination is not the whole story. *Cell Cycle*. 2005;4(8):1015-1018.

10. Asher G, Reuven N, Shaul Y. 20S proteasomes and protein degradation "by default". *Bioessays*. 2006;28(8):844-849.

11. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 2005;6(3):197-208.

12. Schweers O, Schönbrunn-Hanebeck E, Marx A, Mandelkow E. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J Biol Chem*. 1994;269(39):24290-24297.

13. Weinreb PH, Zhen W, Poon AW, Conway KA, Lansbury PT. NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*. 1996;35(43):13709-13715.

14. Tompa P. Intrinsically unstructured proteins. *Trends Biochem Sci*. 2002;27(10):527-533.

15. Lee J, O'Kane DJ, Visser AJ. Spectral properties and function of two lumazine proteins from Photobacterium. *Biochemistry*. 1985;24(6):1476-1483.

16. Pullen RA, Jenkins JA, Tickle IJ, Wood SP, Blundell TL. The relation of polypeptide hormone structure and flexibility to receptor binding: the relevance of X-ray studies on insulins, glucagon and human placental lactogen. *Mol Cell Biochem*. 1975;8(1):5-20.

17. Cary PD, Moss T, Bradbury EM. High-resolution proton-magnetic-resonance studies of chromatin core particles. *Eur J Biochem*. 1978;89(2):475-482.

18. Neurath H. Protein Structure and Enzyme Action. *Rev Mod Phys*. 1959;31(1):185-190.

19. Chen J, Liang H, Fernández A. Protein structure protection commits gene expression patterns. *Genome Biol*. 2008;9(7):R107.

20. Fuxreiter M, Tompa P, Simon I, Uversky VN, Hansen JC, Asturias FJ. Malleable machines take shape in eukaryotic transcriptional regulation. *Nat Chem Biol*. 2008;4(12):728-737.

21. Livesay DR. Protein Dynamics: Dancing on an Ever-Changing Free Energy Stage. *Curr Opin Pharmacol*. 2010;10(6):706-708.

22. Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P. Disorder and sequence repeats in hub proteins and their implications for network evolution. *J Proteome Res*. 2006;5(11):2985-2995.

23. Garner E, Cannon P, Romero P, Obradovic Z, Dunker AK. Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. *Genome Inform Ser Workshop Genome Inform*. 1998;9:201-213.

24. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. *Proteins*. 2001;42(1):38-48.

25. Weathers EA, Paulaitis ME, Woolf TB, Hoh JH. Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett*. 2004;576(3):348-352.

26. Williams RM, Obradovi Z, Mathura V, et al. The protein non-folding problem: amino acid determinants of intrinsic order and disorder. *Pac Symp Biocomput*. 2001;100:89-100.

27. Daughdrill GW, Chadsey MS, Karlinsey JE, Hughes KT, Dahlquist FW. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma σ28. *Nat Struct Biol*. 1997;4(4):285-291.

28. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chem Rev*. 2004;104(8):3607-3622.

29. Hecht O, Ridley H, Boetzel R, et al. Self-recognition by an intrinsically disordered protein. *FEBS Lett*. 2008;582(17):2673-2677.

30. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE. Structural studies of $p21^{Waf1/Cip1/Sdi1}$ in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc Natl Acad Sci U S A*. 1996;93(21):11504-11509.

31. Laskowski RA, Thornton JM. Understanding the molecular machinery of genetics through 3D structures. *Nat Rev Genet*. 2008;9(2):141-151.

32. Li M, Song J. The N- and C-termini of the human Nogo molecules are intrinsically unstructured: bioinformatics, CD, NMR characterization, and functional implications. *Proteins*. 2007;68(1):100-108.

33. McNulty BC, Young GB, Pielak GJ. Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. *J Mol Biol*. 2006;355(5):893-897.

34. Maynard Smith J. Natural Selection and the Concept of a Protein Space. *Nature*. 1970;225:563.

35. Edsall JT. Hsien Wu and the First Theory of Protein Denaturation (1931). *Adv Protein Chem*. 1995;46(C):1-5.

36. Anfinsen CB, Redfield RR, Choate WL, Page J, Carroll Wr. Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease. *J Biol Chem*. 1954;207(1):201-210.

37. Van der lee R, Buljan M, Lang B, et al. Classification of Intrinsically Disordered Regions and Proteins. Chem Rev.2014;114(13):6589-631.

38. Dunker AK, Garner E, Guilliot S, et al. Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pacific Symp Biocomput*. 1998:473-484.

39. Dosztányi Z, Chen J, Dunker AK, Simon I and Tompa P. Disorder and Sequence Repeats in Hub Proteins and Their Implications for Network Evolution. *Journal of proteome research*. 2006:5(11): 2985-2995.

40. Aviles FJ, Chapman GE, Kneale G, Crane-Robinson C, Bradbury EM. The Conformation of Histone H5: Isolation and Characterisation of the Globular Segment. *Eur J Biochem*. 1978;88(2):363-371.

41. Manalan AS, Klee CB. Activation of calcineurin by limited proteolysis. *Proc Natl Acad Sci U S A*. 1983;80(14):4291-4295.

42. Shaiu WL, Hu T, Hsieh TS. The hydrophilic, protease-sensitive terminal domains of eucaryotic DNA topoisomerases have essential intracellular functions. *Pac Symp Biocomput*. 1999;589:578-589.

43. Alber T, Gilbert WA, Ponzi DR, Petsko GA. The role of mobility in the substrate binding and catalytic machinery of enzymes. *Ciba Found Symp*. 1983;93:4-24.

44. Huber R. Conformational flexibility in protein molecules. *Nature*. 1979;280(5723):538-538.

45. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions?*Proteins*.2000;41(3):415-427.

46. Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. Intrinsic disorder and functional proteomics. *Biophys J*. 2007;92(5):1439-1456.

47. Kanter D, Mauzerall DL, Ravishankara AR, et al. Charge interactions can dominate the dimensions of intrinsically disordered proteins. PNAS. 2013;110(41).

48. Ferron F, Longhi S, Canard B, Karlin D. A practical overview of protein disorder prediction methods. *Proteins*. 2006;65(1):1-14.

49. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. Comparing and combining predictors of mostly disordered proteins. *Biochemistry*. 2005;44(6):1989-2000.

50. Sickmeier M, Hamilton JA, LeGall T, et al. DisProt: The database of disordered proteins. *Nucleic Acids Res*. 2007;35:786-793.

51. Weathers EA, Paulaitis ME, Woolf TB, Hoh JH. Insights into protein structure and function from disorder-complexity space. *Proteins Struct Funct Bioinforma*. 2006;66(1):16-28.

52. Tompa P, Kalmar L. Power Law Distribution Defines Structural Disorder as a Structural Element Directly Linked with Function. *J Mol Biol*. 2010;403(3):346-350.

53. Lobley A, Swindells MB, Orengo CA, Jones DT. Inferring function using patterns of native disorder in proteins. *PLoS Comput Biol*. 2007;3(8):1567-1579.

54. Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol*. 2004;338(5):1015-1026.

55. Fuxreiter M, Tompa P, Simon I. Local structural disorder imparts plasticity on linear motifs. *Bioinformatics*. 2007;23(8):950-956.

56. Tompa P, Fuxreiter M, Oldfield CJ, Simon I, Dunker AK, Uversky VN. Close encounters of the third kind: disordered domains and the interactions of proteins. *Bioessays*. 2009;31(3):328-335.

57. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. II. *Berichte der Dtsch Chem Gesellschaft*. 1894;27(3):3479-3483.

58. Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat Struct Biol*. 1997;4(1):10-19.

59. Flock T, Weatheritt RJ, Latysheva NS, Babu MM. Controlling entropy to tune the functions of intrinsically disordered regions. *Curr Opin Struct Biol*. 2014;26(1):62-72.

60. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit*. 2005;18(5):343-384.

61. Uversky VN, Dunker AK. Understanding protein non-folding. *Biochim Biophys Acta - Proteins Proteomics*. 2010;1804(6):1231-1264.

62. Zhou HX. Intrinsic disorder: Signaling *via* highly specific but short-lived association. *Trends Biochem Sci*. 2012;37(2):43-48.

63. Kim AS, Kakalis LT, Abdul-Manan N, Liu GA, Rosen MK. Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature*. 2000;404(6774):151-158.

64. Matsushima N, Tanaka T, Kretsinger RH. Non-globular structures of tandem repeats in proteins. *Protein Pept Lett*. 2009;16(11):1297-1322.

65. Simon M, Hancock JM. Tandem and cryptic amino acid repeats accumulate in disordered regions of proteins. *Genome Biol*. 2009;10(6):R59.

66. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol*. 2002;12(1):54-60.

67. Receveur-Brechot V, Bourhis J-M, Uversky VN, Canard B, Longhi S. Assessing protein disorder and induced folding. *Proteins*. 2006;62(1):24-45.

68. Espinoza-Fonseca LM. Reconciling binding mechanisms of intrinsically disordered proteins. *Biochem Biophys Res Commun*. 2009;382(3):479-482.

69. Lemke EA. The Multiple Faces of Disordered Nucleoporins. *J Mol Biol*. 2016;428(10):2011-2024.

70. Oldfield CJ DA. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem*. 2014;83:553-584.

71. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*. 2015;16(1):18-29.

72. Wright PE, Dyson HJ. Linking Folding and Binding. *Curr Opin Struct Biol*. 2009;19(1):31-38.

73. Tsai CJ, Ma B, Sham YY, Kumar S, Nussinov R. Structured disorder and conformational selection. *Proteins*. 2001;44(4):418-427.

74. Bienkiewicz EA, Adkins JN, Lumb KJ. Functional consequences of preorganized helical structure in the intrinsically disordered cell-cycle inhibitor p27(Kip1). *Biochemistry*. 2002;41:752-759.

75. Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. 2009;48(10):2115-24.

76. Fuxreiter M. Fuzziness: linking regulation to protein dynamics. *Mol Biosyst*. 2012;8(1):168-177.

77. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci*. 2008;33(1):2-8.

78. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem*. 2006;75:333-366.

79. Baldassarre G, Belletti B, Nicoloso MS, et al. p27(Kip1)-stathmin interaction influences sarcoma cell migration and invasion. *Cancer Cell*. 2005;7(1):51-63.

80. Grimmler M, Wang Y, Mund T, et al. Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic tyrosine kinases. *Cell*. 2007;128(2):269-280.

81. Umezawa K, Ohnuki J, Higo J and Takano M. Intrinsic disorder accelerates dissociation rather than association. *Proteins*. 2016;84(8):1124-1133.

82. Vacic V, Oldfield CJ, Mohan A, et al. Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res*. 2007;6(6):2351-2366.

83. Davey NE, Van Roey K, Weatheritt RJ, et al. Attributes of short linear motifs. *Mol Biosyst*. 2012;8(1):268-281.

84. Diella F, Haslam N, Chica C, et al. Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front Biosci*. 2008;13:6580-6603.

85. Dinkel H, Van Roey K, Michael S, et al. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res*. 2014;42:259-66.

86. Cumberworth A, Lamour G, Babu MM, Gsponer J. Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. *Biochem J*. 2013;454(3):361-369.

87. Wu H, Fuxreiter M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell*. 2016;165(5):1055-1066.

88. Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res*. 2004;32(3):1037-1049.

89. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci*. 2014;23(8):1077-1093.

90. Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: Regulation and disease. *Curr Opin Struct Biol*. 2011;21(3):432-440.

91. Gsponer J, Futschik ME, Teichmann SA and Babu MM. Tight regulation of unstructured proteins : from transcript synthesis to protein degradation. *Science*. 2008;322(5906):1365-1368.

92. Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol*. 2002;323(3):573-584.

93. Cheng Y, LeGall T, Oldfield CJ, Dunker AK, and Uversky V. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry*. 2006;45:10448–10460.

94. Casu F, Duggan BM and Hennig M. The arginine-rich RNA-binding motif of HIV-1 Revis intrinsically disordered and folds upon RRE binding. *Biophys J*. 2013;105:1004–1017.

95. Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, Forman-Kay JD. CFTR regulatory region interacts with NBD1 predominantly *via* multiple transient helices. *Nat Struct Mol Biol*. 2007;14:738-745.

96. Hollstein M, Sidransky D, Vogelstein B HC. p53 mutations in human cancers. *Science.*1991;253:49–53.

97. Dawson R, Muller L, Dehner A, Klein C, Kessler BJ. The N-terminal domain of p53 is natively unfolded. *J Mol Biol*. 2003;332:1131–41.

98. Lee H, Mok KH, Muhandiram R, Park KH, Suk JE, Kim DH, Chang J, Sung YC, Choi KY, Han KH. Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J Biol Chem*. 2000;275:29426.

99. Mark WY, Liao JC, Lu Y, Ayed A, Laister R et al. Characterization of segments from the central region of BRCA1:an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J Mol Biol*. 2005;(345):275–87.

100. Deny CX. BRCA1: cell cycle checkpoint, genetic instability, DNAdamage response and cancer evolution. *Nucleic Acids Res*. 2006;34:1416–26.

101. Venkitaraman A. Cancer susceptibility and the functions ofBRCA1 and BRCA2. *Cell*. 2002;108:171–82.

102. Hegyi L, Thway K, Newton R, Osin P, Nerurkar A, Hayes AJ, Fisher C. Malignant myoepithelioma arising in adenomyoepithelioma of the breast and coincident multiple gastrointestinal stromal tumours in a patient with neurofibromatosis type 1. *J Clin Pathol*. 2009;62:653–655.

103. Ballerini P, Struski S, Cresson C, Prade N, Toujani S, Deswarte C. et al. RET fusion genes are associated with chronic myelomonocytic leukemia and enhance monocytic differentiation. *Leukemia*. 2012;26:2384–2389.

104. Mitrea DM, Kriwacki RW. Phase separation in biology; functional organization of a higher order. *Cell Commun Signal*. 2016;14:1.

105. Uversky VN. Intrinsically disordered proteins in overcrowded milieu: membrane-less organelles, phase separation, and intrinsic disorder. *Curr Opin Struct Biol*. 2017:18-30.

106. Boeynaems S et al. Inside out: the role of nucleocytoplasmic transport in ALS and FTLD. *Acta Neuropathol*. 2016;132:159–173.

107. Neumann M et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 2006;314:130-133.

108. Lin WL, Dickson DW. Ultrastructural localization of TDP-43 in filamentous neuronal inclusions in various neurodegenerative diseases. *Acta Neuropathol*. 2008;116:2008; 116:205–213.

109. Kao PF, Chen YR, Liu XB, DeCarli C, Seeley WW, Jin LW. Detection of TDP-43 oligomers in frontotemporal lobar degeneration-TDP. *Ann Neurol*. 2015;78:211–221.

110. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, Grishin NV, Frantz DE, Schneider JW, Chen S, Li L, Sawaya MR, Eisenberg D, Tycko R, McKnight SL. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. 2012;149:753–767.

111. Murray DT et al. Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell*. 2017;171:615–627.

112. Molliex A et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*. 2015;163:123–133.

113. Murakami T et al. ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron*. 2015;88:678-690.

114. Burke KA, Janke AM, Rhine CL FN. Residue-by-residue view of *in vitro* FUS granules that bind the C-terminal domain of RNA polymerase II Kathleen. *Mol Cell*. 2015;60(2):231-241.

115. Brady JP et al. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc Natl Acad Sci U S A*. 2017;114:E8194– E8203.

116. Patel A et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell*. 2015;162:1066–1077.

117. Liu-Yesucevitz L et al. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One*. 2010;5:e13250.

118. Eaton BT, Mackenzie JS,Wang LF. Henipaviruses.*BN Fields, DM Knipe, PM Howley (Eds), Fields Virol*. 2007;5th ed.:1587–1600.

119. Eaton BT, Broder CC, Middleton D, Wang LF. Hendra and Nipah viruses: Different and dangerous. *Nat Rev Microbiol*. 2006;4(1):23–35.

120. Longhi S. Measles virus nucleoprotein. Hauppauge, NY: Nova Publishers Inc. 2007.

121. Longhi S. Nucleocapsid structure and function. *Curr Top Microbiol Immunol*. 2009;329:103-128.

122. Longhi S. Structural disorder within the measles virus nucleoprotein and phospho-protein: Functional implications for transcription and replication. *M Luo (Ed), Neg- ative strand RNA virus Singapore World Sci Publ*. 2011:95–125.

123. Longhi S, Bloyet LM, Gianni S, Gerlier D. How order and disorder within paramyxoviral nucleoproteins and phosphoproteins orchestrate the molecular interplay of transcription and replication. *Cellular and Molecular Life Sciences*. 2017;74(17), 3091–3118.

124. Kolakofsky D, Le Mercier P, Iseni F, Garcin D. Viral DNA polymerase scanning and the gymnastics of Sendai virus RNA synthesis. *Virology*. 2004;318(2):463-473.

125. Brunel J, Chopy D, Dosnon M, Bloyet LM, Devaux P, Urzua E.et al. Sequence of events in Measles virus replication: Role of phosphoprotein-nucleocapsid interactions. *J Virol*. 2014;88(18):10851–10863.

126. Troilo F, Bignon C, Gianni S, Fuxreiter M, Longhi S. Experimental Characterization of Fuzzy Protein Assemblies: Interactions of Paramyxoviral NTAIL Domains With Their Functional Partners. Methods Enzymol. 2018;611:137-192.

127. Kelly JW. The alternative conformations of amyloidogenic proteins and their multi step assembly pathways. *Curr Opin Struct Biol*. 1998;8:101-106.

128. Hipp MS, Park SH, Hartl FU. Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol*. 2014;24(9):506-14

129. Uversky VN. Wrecked regulation of intrinsically disordered proteins in diseases: pathogenicity of deregulated regulators. *Front Mol Biosci*. 2014;1:1-24.

130. Yang C, Matro JC, Huntoon KM, Ye DY, Huynh TT, Fliedner SM, Breza J, Zhuang Z, Pacak K. Missense mutations in the human SDHB gene increase protein degradation without altering intrinsic enzymatic function. *FASEB J*. 2012;26(11):4506-4516.

131. Iwai A, Masliah E, Yoshimoto M, Ge N, Flanagan L, de Silva HA, Kittel A, Saitoh T. The precursor protein of non-A beta component of Alzheimer's disease amyloidisa presynaptic protein of the central nervous system. *Neuron*. 1995;14:467–475.

132. Tanaka K. The proteasome: Overview of structure and functions. *Proc Jpn Acad, Ser*. 2009;(85):12-36.

133. Bochtler M, Ditzel L, Groll M, Hartmann C, Huber R. The proteasome. *Annu Rev Biophys Biomol Struct*. 1999;28:295-317.

134. Dong Y, Zhang S, Wu Z, et al. Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature*. 2018;565:49-55.

135. Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK. Quantification of protein half-lives in the budding yeast proteome. *Proc Natl Acad Sci U S A*. 2006;103(35):13004-13009.

136. Kristensen AR, Gsponer J, Foster LJ. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol Syst Biol*. 2014;9(1):689-689.

137. Glickman MH, Ciechanover A. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol Rev*. 2002;82(2):373-428.

138. Hershko A. Lessons from the discovery of the ubiquitin system. *Trends Biochem Sci*. 1996;21(11):445-449.

139. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem*. 1998;67:425-479.

140. Koegl M, Hoppe T, Schlenker S, Ulrich HD, Mayer TU, Jentsch S. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell*. 1999;96(5):635-644.

141. Breitschopf K, Bengal E, Ziv T, Admon A, Ciechanover A. A novel site for ubiquitination: the N-terminal residue, and not internal lysines of MyoD, is essential for conjugation and degradation of the protein. *EMBO J*. 1998;17(20):5964-5973.

142. Ben-Saadon R, Fajerman I, Ziv T, Hellman U, Schwartz AL, Ciechanover A. The tumor suppressor protein p16(INK4a) and the human papillomavirus oncoprotein-58 E7 are naturally occurring lysine-less proteins that are degraded by the ubiquitin system. Direct evidence for ubiquitination at the N-terminal residue. *J Biol Chem*. 2004;279(40):41414-41421.

143. Suskiewicz MJ, Sussman JL, Silman I, Shaul Y. Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci*. 2011;20(8):1285-1297.

144. Navon A, Goldberg AL. Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. *Mol Cell*. 2001;8(6):1339-1349.

145. Voges D, Zwickl P, Baumeister W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem*. 1999;68:1015-1068.

146. Bloom J, Pagano M. Experimental tests to definitively determine ubiquitylation of a substrate. *Methods Enzymol*. 2005;399(05):249-266.

147. Tofaris GK, Layfield R, Spillantini MG. alpha-synuclein metabolism and aggregation is linked to ubiquitin-independent degradation by the proteasome. *FEBS Lett*. 2001;509(1):22-26.

148. Venkatraman P, Wetzel R, Tanaka M, Nukina N, Goldberg AL. Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. *Mol Cell*. 2004;14(1):95-104.

149. Baugh JM, Viktorova EG, Pilipenko E V. Proteasomes Can Degrade a Significant Proportion of Cellular Proteins Independent of Ubiquitination. *J Mol Biol*. 2009 ;386(3):814-27

150. Zadeh L A. Fuzzy sets. *Inf Control*. 1965;8(3):338-353.

151. Tanaka K. An Introduction to Fuzzy Logic for Practical Applications. Springer; 1997.

152. Sharma R, Raduly Z, Miskei M, Fuxreiter M. Fuzzy complexes: Specific binding without complete folding. *FEBS Lett*. 2015;589(19):2533-2542.

153. Halfmann R, Alberti S, Krishnan R, et al. Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol Cell*. 2011;43(1):72-84.

154. Fuxreiter M, Simon I, Bondos S. Dynamic protein-DNA recognition: Beyond what can be seen. *Trends Biochem Sci*. 2011;36(8):415-423.

155.Watson M, Stott K, Thomas JO. Mapping Intramolecular Interactions between Domains in HMGB1 using a Tail-truncation Approach. *J Mol Biol*. 2007;374(5):1286-1297.

156. Welch GR. The 'fuzzy' interactome. *Trends Biochem Sci*. 2009;34(1):1-2.

157. Naud JF, McDuff FO, Sauve S, et al. Structural and thermodynamical characterization of the complete p21 gene product of Max. *Biochemistry*. 2005;44(38):12746-12758.

158. Pufall MA. Variable Control of Ets-1 DNA Binding by Multiple Phosphates in an Unstructured Region. *Science*. 2005;309(5731):142-145.

159. Stott K, Watson M, Howe FS, Grossmann JG, Thomas JO. Tail-mediated collapse of HMGB1 is dynamic and occurs *via* differential binding of the acidic tail to the A and B domains. *J Mol Biol*. 2010;403(5):706-722.

160. Jonker HRA, Wechselberger RW, Boelens R, Kaptein R, Folkers GE. The intrinsically unstructured domain of PC4 modulates the activity of the structured core through inter- and intramolecular interactions. *Biochemistry*. 2006;45(15):5067-5081.

161. Olson KE, Narayanaswami P, Vise PD, Lowry DF, Wold MS, Daughdrill GW. Secondary structure and dynamics of an intrinsically unstructured linker domain. *J Biomol Struct Dyn*. 2005;23(2):113-124.

162. Vise PD, Baral B, Latos AJ, Daughdrill GW. NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain. *Nucleic Acids Res*. 2005;33(7):2061-2077.

163. Fuxreiter M. Fuzziness in Protein Interactions—A Historical Perspective. *J Mol Biol*. 2018;430(16): 2278-2287

164. Miskei M, Antal C, Fuxreiter M. FuzDB: database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res*. 2017;45:D228-D235.

165. O'Dea EL, Kearns JD, Hoffmann A. UV as an amplifier rather than inducer of NF-kappaB activity.*Mol Cell*. 2008;30(5):632-641..

166. Mathes E, O'Dea EL, Hoffmann A, Ghosh G. NF-kappaB dictates the degradation pathway of IkappaBalpha. *Embo J*. 2008;27(9):1357-1367.

167. Bloom J, Amador V, Bartolini F, DeMartino G, Pagano M. Proteasome-mediated degradation of p21 *via* N-terminal ubiquitinylation. *Cell*. 2003;115(1):71-82.

168. Touitou R, Richardson J, Bose S, Nakanishi M, Rivett J, Allday MJ. A degradation signal located in the C-terminus of p21 WAF1 / CIP1 is a binding site for the C8 a-subunit of the 20S proteasome. *EMBO J*. 2001;20(10):2367-2375.

169. Garces de Los Fayos Alonso I, Liang HC, Turner SD, Lagger S, Merkel O, Kenner L. The role of activator protein-1 (AP-1) family members in CD30-positive lymphomas. *Cancers(Basel)*. 2018;10(4):1-22.

170. Muslin AJ. MAPK signalling in cardiovascular health and disease: molecular mechanisms and therapeutic targets. *Clin Sci (Lond)*. 2008;115(7):203-218.

171. Meijer CA, Le Haen PAA, van Dijk RA, et al. Activator protein-1 (AP-1) signalling in human atherosclerosis: results of a systematic evaluation and intervention study. *Clin Sci (Lond)*. 2012;122(9):421-428.

172. Giri RS, Thaker HM, Giordano T, et al. Design, synthesis and characterization of novel 2-(2,4-disubstituted-thiazole-5-yl)-3-aryl-3H-quinazoline-4-one derivatives as inhibitors of NF-kappaB and AP-1 mediated transcription activation and as potential anti-inflammatory agents. *Eur J Med Chem*. 2009;44(5):2184-2189.

173. Ball DP, Lewis AM, Williams D, Resetca D, Wilson DJ, Gunning PT. Signal transducer and activator of transcription 3 (STAT3) inhibitor, S3I-201, acts as a potent and non-selective alkylating agent. *Oncotarget*. 2016;7(15).

174. Lee W, Haslinger A, Karin M, Tjian R. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. *Nature*. 1987;325(6102):368-372.

175. Angel P, Imagawa M, Chiu R, et al. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell*. 1987;49(6):729-739.

176. Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK, Tjian R. Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science*. 1987;238(4832):1386-1392.

177. Curran T, Teich NM. Identification of a 39,000-dalton protein in cells transformed by the FBJ murine osteosarcoma virus.*Virology*. 1982;116(1):221-235.

178. Van Beveren C, van Straaten F, Curran T, Müller R, Verma IM. Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and cellular fos gene products have different carboxy termini. *Cell*. 1983;32(4):1241-1255.

179. Tom Curran A, Miller D, Zokas L, Verma IM. Viral and cellular fos proteins: A comparative analysis. *Cell*. 1984;36(2):259-268.

180. Hess J, Angel P, Schorpp-kistner M. AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci*. 2004;117(25):5965-5973.

181. Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer*. 2003;3(11):859-868.

182. Curran T, Franza BR. Fos and Jun: The AP-1 connection. *Cell*.1988;55:395-397.

183. Wagner EF. AP-1-Introductory remarks. *Oncogene*. 2001;20(19):2334-2335.

184. Angel P, Karin M. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochim Biophys Acta*. 1991;1072(2-3):129-157.

185. Piechaczyk M, Blanchard JM. C-Fos Proto-Oncogene Regulation and Function. *Crit Rev Oncol Hematol*. 1994;17(2):93-131.

186. Abate C, Curran T. Encounters with Fos and Jun on the road to AP-1. *Semin Cancer Biol*. 1990;1(1):19-26.

187. Glover JN, Harrison SC. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature*. 1995;373(6511):257-261.

188. Jochum W, Passegué E, Wagner EF. AP-1 in mouse development and tumorigenesis. *Oncogene*. 2001;20(19):2401-2412.

189. Shaulian E, Karin M. AP-1 in cell proliferation and survival. *Oncogene*. 2001;20(19):2390-2400.

190. Landschulz WH, Johnson PF, McKnight SL. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*. 1988;240(4860):1759-1764.

191. O'Shea E, Rutkowski R, Kim P. Evidence that the leucine zipper is a coiled coil. *Science*.1989;243(4890):538-542.

192. O'Shea E, Klemm J, Kim P, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science*. 1991;254(5031):539-544.

193. Turner R, Tjian R. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. *Science*. 1989;243(4899):1689-1694.

194. Nakabeppu Y, Ryder K, Nathans D. DNA binding activities of three murine Jun proteins: Stimulation by Fos. *Cell*. 1988;55(5):907-915.

195. John M, Leppik R, Busch SJ, Granger-Schnarr M, Schnarr M. DNA Binding of Jun and Fos bZip Domains: Homodimers and Heterodimers Induce a DNA Conformational Change in Solution. *Nucleic Acids Res*. 1996;24(22):4487-4494.

196. Szalóki N, Krieger JW, Komáromi I, Tóth K, Vámosi G. Evidence for Homodimerization of the c-Fos Transcription Factor in Live Cells Revealed by Fluorescence Microscopy and Computer Modeling. *Mol Cell Biol*. 2015;35:3785-3798.

197. Jariel-Encontre I, Pariat M, Martin F, Carillo S, Salvat C, Piechaczyk M. Ubiquitinylation is not an absolute requirement for degradation of c-Jun protein by the 26 S proteasome. *J Biol Chem*. 1995;270(19):11623-11627.

198. Lamph WW, Wamsley P, Sassone-Corsi P, Verma IM. Induction of proto-oncogene JUN/AP-1 by serum and TPA. *Nature*. 1988;334(6183):629-631.

199. Treier M, Staszewski LM, Bohmann D. Ubiquitin-dependent c-Jun degradation *in vivo* is mediated by the delta domain. *Cell*. 1994;78(5):787-798.

200. Ransone LJ, Visvader J, Sassone-Corsi P, Verma IM. Fos-Jun interaction: mutational analysis of the leucine zipper domain of both proteins. *Genes Dev*. 1989;3(6):770-781.

201. Corsi PA, Sisson JC, Verma IM. Transcriptional autoregulation of the proto-oncogene fos. *Nature*. 1988;314-319.

202. Greenberg EM, Edward ZB. Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. *Nature*. 1984;311:433–438.

203. Greenberg ME, Greene LA ZE. Nerve growth factor and epidermal growth factor induce rapid transient changes in proto-oncogene transcription in PC12 cells. *J Biol Chem*. 1985;260(26):14101-10.

204. Acquaviva C, Salvat C, Brockly F, et al. Cellular and viral Fos proteins are degraded by different proteolytic systems. *Oncogene*. 2001;20(8):942-950.

205. Acquaviva C, Brockly F, Ferrara P, et al. Identification of a C-terminal tripeptide motif involved in the control of rapid proteasomal degradation of c-Fos proto-oncoprotein during the G0-to-S phase transition. *Oncogene*. 2001;20(51):7563-7572.

206. Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*. 1986;234(4774):364-368.

207. Rogers SW, Rechsteiner MC. Microinjection studies on selective protein degradation: relationships between stability, structure, and location. *Biomed Biochim Acta*. 1986;45(11-12):1611-1618.

208. Campbell KM, Terrell AR, Laybourn PJ, Lumb KJ. Intrinsic structural disorder of the C-terminal activation domain from the bZIP transcription factor Fos. *Biochemistry*. 2000;39(10):2708-2713.

209. Ferrara P, Andermarcher E, Bossis G, et al. The structural determinants responsible for c-Fos protein proteasomal degradation differ according to the conditions of expression. *Oncogene*. 2003;22(10):1461-1474.

210. Ferrara P, Andermarcher E, Bossis G, et al. The structural determinants responsible for c-Fos protein proteasomal degradation differ according to the conditions of expression. *Oncogene*. 2003;22(10):1461-1474.

211. Farràs R, Bossis G, Andermarcher E, Jariel-Encontre I, Piechaczyk M. Mechanisms of delivery of ubiquitylated proteins to the proteasome: new target for anti-cancer therapy? *Crit Rev Oncol Hematol*. 2005;54(1):31-51.

212. Bossis G, Ferrara P, Acquaviva C, Jariel-Encontre I, Piechaczyk M. c-Fos Proto-Oncoprotein Is Degraded by the Proteasome Independently of Its Own Ubiquitinylation *in vivo*. *Mol Cell Biol*. 2003;23(20):7425-7436.

213. Adler J, Reuven N, Kahana C, Shaul Y. c-Fos Proteasomal Degradation Is Activated by a Default Mechanism, and Its Regulation by NAD(P)H:Quinone Oxidoreductase 1 Determines c-Fos Serum Response Kinetics. *Mol Cell Biol*. 2010;30(15):3767-3778.

214. Glover JNM, Harrison SC. Crystal structure of the heterodimeric bZIP transcription factor c-Fos–c-Jun bound to DNA. *Nature*. 1995;373(6511):257-261.

215. Vámosi G, Baudendistel N, Von Der Lieth CW, et al. Conformation of the c-Fos/c-Jun complex *in vivo*: A combined FRET, FCCS, and MD-modeling study. *Biophys J*. 2008;94(7):2859-2868.

216. Tsvetkov P, Asher G, Paz A, et al. Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins Struct Funct Bioinforma*. 2007;70(4):1357-1366.

217. Lakin ND, Jackson SP. Regulation of p53 in response to DNA damage. *Oncogene*. 1999;18:7644-7655.

218. Ciechanover A. Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Biochim Biophys Acta*. 2012;1824:3-13.

219. Van der Lee R, Lang B, Kruse K, et al. Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep*.2014;8(6):1832-1844.

220. Hirata H, Bessho Y, Kokubu H, et al. Instability of Hes7 protein is crucial for the somite segmentation clock. *Nat Genet*. 2004;36(7):750-754.

221. Van der Lee R, Lang B, Kruse K, et al. Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep*. 2014;8(6):1832-1844.

222. Basbous J, Jariel-Encontre I, Gomard T, Bossis G, Piechaczyk M. Ubiquitin-independent- versus ubiquitin-dependent proteasomal degradation of the c-Fos and Fra-1 transcription factors: Is there a unique answer? *Biochimie*. 2008;90(2):296-305.

223. Gu Y, Seidl KJ, Rathbun GA, et al. Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity*. 1997;7(5):653-665.

224. Kohler JJ, Schepartz A. Kinetic Studies of Fos , Jun , DNA Complex Formation : DNA Binding Prior to dimerization.*Biochemistry.* 2001;2:130-142.

225. Pursglove SE, Fladvad M, Bellanda M, et al. Biophysical properties of regions flanking the bHLH-Zip motif in the p22 Max protein. *Biochem Biophys Res Commun*. 2004;323(3):750-759.

226. Shoemaker BA, Portman JJ, Wolynes PG. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad Sci*. 2000;97(16):8868-8873.

227. Huang Y, Liu Z. Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding-Binding Process: A Critical Assessment of the "Fly-Casting" Mechanism. *J Mol Biol*. 2009;393(5):1143-1159.

228. Fuxreiter M. Fold or not to fold upon binding — does it really matter? *Curr Opin Struct Biol*. 2019;54:19-25.

229. Dogan J, Gianni S, Jemth P. The binding mechanisms of intrinsically disordered proteins. *Phys Chem Chem Phys*. 2014;16(14):6323-6331.

230. Gianni S, Dogan J JP. Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr Opin Struct Biol*. 2016;36:18-24.

231. Berlow RB, Dyson HJ, Wright PE. Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature*. 2017;543(7645):447-451.

232. Shammas SL, Crabtree MD, Dahal L, Wicky BIM, Clarke J. Insights into coupled folding and binding mechanisms from kinetic studies. *J Biol Chem*. 2016;291(13):6689-6695.

233. Sharipo A, Imreh M, Leonchiks A, Imreh S MM. A minimal glycine-alanine repeat prevents the interaction of ubiquitinated I kappaB alpha with the proteasome: a new mechanism for selective inhibition of proteolysis. *Nat Med*. 1998;4(8):839-844.

234. Yu H, Singh Gautam AK, Wilmington SR, et al. Conserved sequence preferences contribute to substrate recognition by the proteasome. *J Biol Chem*. 2016;291(28):14526-14539.

235. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, Tompa P, Fuxreiter M. Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol.* 2018;28(6):420-435.

236. Li P, Banjade S, Cheng H, et al. Phase Transitions in the Assembly of Multi-Valent Signaling Proteins. *Nature*. 2012;483(7389):336-340.

237. Wu H, Fuxreiter M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell*. 2016;165(5):1055-1066.

9. LIST OF PUBLICATIONS:

1. Rashmi Sharma, Zsolt Raduly, Marton Miskei, Monika Fuxreiter, Fuzzy complexes: Specific binding without complete folding; FEBS Letters 589 (2015) 2533–2542.

2. Miskei M, Gregus A, Sharma R, Duro N, Zsolyomi F, Fuxreiter M, Fuzziness enables context dependence of protein interactions; FEBS Letters 591(2017) 2682-2695.

3. Sharma M, Demeny M, Ambrus V, Kiraly S, Kurtan T, Lafranconi PG, Fuxreiter M, Specific and fuzzy interactions cooperate in modulating protein half-life; JMB (2019).

UNIVERSITY AND NATIONAL LIBRAY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/29/2019.PL
Subject: PhD Publikációs Lista

Candidate: Rashmi Sharma
Neptun ID: US03E0
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

List of publications related to the dissertation

1. **Sharma, R.**, Demény, M. Á., Ambrus, V. A., Király, S. B., Kurtán, T., Gatti-Lafranconi, P., Fuxreiter, M.: Specific and fuzzy interactions cooperate in modulating protein half-life. *J. Mol. Biol. "Accepted by Publisher"*, 1-18, 2019.
   IF: 4.894 (2017)
2. **Sharma, R.**, Ráduly, Z., Miskei, M., Fuxreiter, M.: Fuzzy complexes: specific binding without complete folding.
   FEBS Lett. 589 (19), 2533-2542, 2015.
   IF: 3.519

DEBRECENI EGYETEM * Egyetemi és Nemzeti Könyvtár *

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRAY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

3. Miskei, M., Gregus, A., **Sharma, R.**, Duró, N., Zsólyomi, F., Fuxreiter, M.: Fuzziness enables context dependence of protein interactions.
FEBS Lett. 591 (17), 2682-2695, 2017.
DOI: http://dx.doi.org/10.1002/1873-3468.12762

4. Noronha, V., Prabhash, K., Thavamani, A., Chougule, A., Purandare, N., Joshi, A., **Sharma, R.**, Desai, S., Jambekar, N., Dutt, A., Mulherkar, R.: EGFR Mutations in Indian Lung Cancer Patients: Clinical Correlation and Outcome to EGFR Targeted Therapy.
PLoS One. 8 (4), 1-5, 2013.
DOI: http://dx.doi.org/10.1371/journal.pone.0061561
IF: 3.534

5. **Sharma, R.**, Jadhav, U.: Antibacterial activity and brine shrimp lethality test of extracts of Ipomoea carnea and Clitoria ternatea.
Plant Archives. 10 (2), 803-806, 2010.

Total IF of journals (all publications): 11,947
Total IF of journals (publications related to the dissertation): 8,413

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

14 February, 2019