SHORT 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

A nanny model for intrinsically disordered proteins

By Rashmi Sharma, MSc

Supervisor: Prof. Dr. Monika Fuxreiter, PhD

Doctoral School of Molecular, Cellular and Immune Biology, University of Debrecen

Head of the **Examination Committee:**	Prof. Dr. Fésüs László, Member of the Hungarian Academy of Sciences
Members of the Examination Committee:	Dr. Balázs Papp, PhD
	Dr. András Penyige, PhD

The Examination takes place at the Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen Debrecen; at 1 PM; 28th of Jan 2019

Head of the **Defense Committee:**	Prof. Dr. Fésüs László, Member of the Hungarian Academy of Sciences
Reviewers:	Dr. Balázs Papp, PhD
	Dr. Krisztina Tar, PhD
Members of the Defense Committee:	Dr. László Homolya, DSc
	Dr. András Penyige, PhD

The PhD defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen Debrecen; at 11 AM; 26th of March 2019

1.INTRODUCTION

1.1 Intrinsically disordered proteins (IDPs)

Intrinsically disordered proteins contain regions, which are devoid of a well-defined tertiary structure. These could be loop and linkers in a large folded protein or may affect a larger portion of a protein. Intrinsically disordered regions (IDRs) usually defined by continuous stretches ≥ 30 amino acids. 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 functions. The frequently occurring short IDRs may be short linkers or linear motifs 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. 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. Their existence *in vitro* has been studied with many state-of-the-art methods: solution state-NMR, small-angle X-ray scattering and fluorescence spectroscopy, circular dichroism (CD), and single molecule fluorescence resonance energy transfer (smFRET) measurements. 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, residues missing from electron density maps, 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.

1.2 Prevalent features of IDPs

1.2.1 Sequence

Compared to sequences from ordered proteins, sequence 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. 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. These biases in amino acid composition in IDPs enabled a coarse prediction of disorder from primary sequence information alone via numerous developed disorder predictors.

1.2.2 Conformational ensembles of IDPs

IDRs are continuously fluctuating between multitudes of conformations states (collection of structurally similar and nearly energetically equivalent conformations of the protein) and have a rugged free energy landscape. Barriers between the wells of the landscape are comparable to kBT (kB 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').

1.2.3 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. In contrast to ordered proteins, IDPs are highly pliable and one IDP can form an array of unrelated structures being bound to different partner. 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. 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).

1.3. 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. Two main mechanisms

for the interactions of disordered proteins gained popularity in the IDP field: conformational selection and induced folding. The first mechanism was termed folding coupled to binding, where ID proteins lose their conformational heterogeneity in comparison to their native state. The factors like the low affinity, entropy loss governs the binding and folding of ID proteins interactions. Emerging computational and experimental data suggested that IDs may possess pre-organized secondary structures and may not be totally random in their free state. 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 proposes the pre-existence of inter-converting conformations/folds before binding to the partner shifts the population to a specific bound state. 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. A shift can occur from one conformational state to another upon binding a third partner or upon post-translational modifications such as phosphorylation or alternative splicing. 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.

1.4 The concept of fuzzy complexes

In contrast to the previous mechanisms, sequences of some ID regions in their free state as well as in their bound state were 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. The term was borrowed from the field of mathematics, where fuzzy logic is a many valued logic with variables. By definition, fuzzy protein complexes are 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. Proteins containing fuzzy regions not only preserve their conformational freedom in the bound state, but also impact the biological activity of the complexes. 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, and ranges from static to dynamic, and from full to segmental disorder. In polymorphic complexes which represent static disorder, the disordered regions have alternative conformations in the complex, whereas in dynamic disorder the IDRs exhibit a dynamical continuum of rapidly exchanging conformations. Dynamic ensembles can be shifted upon external signals or changes in the intracellular milieu, this signifies context dependence. Over the years fuzziness has been closely connected to context dependency in the cell. Fuzziness is also a key feature of the dynamical behavior of higher-order assemblies. 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. 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.

1.5 Proteasomal degradation

The proteasome is a large protein complex (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 seven α and seven β structurally similar subunits arranged in axial stacking style. The center of the outer α ring is completely closed and it prevents internalization of proteins into the inner β subunit (proteolytic active

site). Substrate proteins are only allowed in after passing the narrow opening situated at the centre of α ring. Two alternative proteasomal degradation mechanisms have been described for IDPs: ubiquitin-dependent (UD) degradation and ubiquitin-independent (UI) degradation which are not mutually exclusive. Many proteins in the cell are selectively degraded via the ubiquitin-26S proteasomal degradation pathway. The ubiquitin 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). For proteasomal recognition, as a minimal signal a chain of four Ub molecules need to be sequentially connected through Lys48 linkages to the target protein.

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'.

1.6 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. Importantly, nannies are distinct from chaperones as they do not fold ID proteins. Their role is to eliminate IDPs destruction unlike chaperones. 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.

1.7 The AP-1 model system

Activator protein 1 (AP-1) has been a most extensively studied topic, mostly because of its implication in a variety of pathologies, ranging from inflammation to tumorigenesis.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. 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'. 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, and c-Fos was first isolated as the cellular homologue of two viral v-fos oncogenes. 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. 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. Fos-Jun dimerization occurs via the parallel interaction of leucine zipper motifs required for DNA binding. Dimeric c-Fos/c-Jun bZIP domains form an X-shaped α-helical structure. 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. It has been found that amino (NH_2)- and carboxy (COOH)-terminal regions are quite divergent among all AP-1 proteins. c-Fos forms stable heterodimers with c-Jun with higher affinity as compared to the c-Jun-c-Jun homodimer. 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*.

1.7.3 Regulation of c-Fos protein turnover

c-Fos is an unstable protein, reported to have short half-life of approximately 60 minutes while c-Jun has a half-life of 90 minutes. 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. 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. The protein can be constitutively expressed in asynchronously growing cells and transiently in other cell types upon various stimuli. c-Fos contains 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 the PEST3 motif. The 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 intrinsically disordered.

c-Fos degradation is governed in asynchronous cells by a unique C-terminus which is bigger than the PEST3 motif; unlike in serum stimulated growing cells where N

terminus together with C-terminal intrinsically disordered domain and the DBD/LZ region responsible for faster degradation of c-Fos.

The degradation of cytoplasmic c-Fos is also mediated by 26 ubiquitin dependent proteolytic pathway, but the majority of nuclear c-Fos is subjected to ubiquitin independent degradation (UID) *in vivo*. *In vivo* c-Fos undergoes degradation by a default mechanism, which is regulated by NQO1 (a 20S proteasome gatekeeper) followed by c-Jun for the formation of functional AP-1 complex.

A pressing question has been raised how exposed regions of proteins can survive under cellular conditions.

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 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. 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).

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

Proteins were expressed in *E.coli* bacterial cultures (Rosetta2(DE3)pLysS, Novagen) using IPTG induction for 3 hours. Proteins purification was performed by immobilized metal affinity chromatography using a Ni-Sepharose™ 6 Fast Flow resin (GE Healthcare) with a lysis buffer containing 250 mM imidazole. Proteins were dialyzed against 100 volumes of 4M, 2M and 1M urea consecutively, then the buffer was changed to 25 mM Tris-HCl, pH7.4, 100 mM NaCl. The eluted proteins were concentrated with an Amicon- Ultra 50 10000 MWCO ultrafiltration device. Protein concentrations were determined by BCA protein assay.

3.3 Binding kinetics measurements

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 at 37 °C for an hour. 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.

3.4 20S proteasomal degradation assay by ELISA

c-Fos variants and human 20S proteasome (Boston Biochem) were both diluted to concentrations of 400 nM and 4 nM, respectively. The degradation was performed by mixing 12.5 µl of each solution. Degradation of the c-Fos was followed at 37 °C and 2 µl samples were collected at different time points into 98 µl of PBS, 0.1% Tween-20, 1% BSA containing a protease inhibitor cocktail and 1 µM MG132 (Sigma).The amount of c-Fos at each time point was determined using an ELISA assay. The c-Fos solution was loaded onto 96-well Pierce Nickel-coated plates for an hour at room temperature. The wells were washed three times with PBS-0.1% Tween-20 and incubated 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 R software.

3.5 Circular dichroism spectroscopy

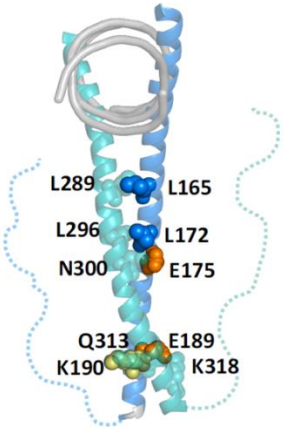
Electronic circular dichroism (ECD or CD) far-UV spectra were recorded on a J-810 spectropolarimeter in 20 mM sodium phosphate, pH 7.4 in a 0.2 mm pathlength quartz cell at room temperature. Raw spectra were buffer-subtracted, converted to mean residue ellipticities and smoothed. Deconvolutions were made with the Bestsel program.

4. RESULTS

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. 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).



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 c-Fos and c-Jun tagged with 6xHis

In order to purify protein in functional form, c-Fos, c-Jun and 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.

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. 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. The k_D determined for the wild-type complex matched well with literature data determined. 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.

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

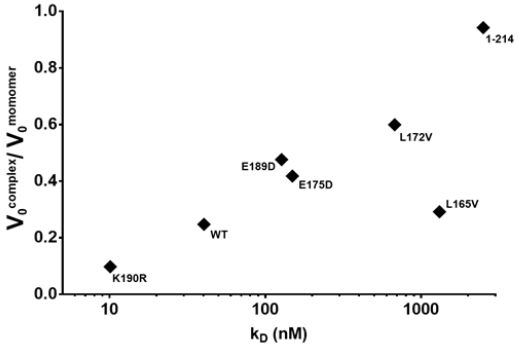
To examine the behavior of c-Fos as an IUP, we incubated c-Fos with purified 20S proteasomes and revealed that c-Fos was efficiently degraded alone. Next, we determined the sensitivity of c-Fos for 20S proteasomal degradation *in vitro* in the presence or absence of its binding partner, c-Jun. In the 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 ($t_{1/2}$).

The 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. 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-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. 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. 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 (ΔVmut/ΔVwt) compared to those affecting hydrophilic interactions (E175D, E189D). Only mutant K190R slowed down the initial degradation rate as compared to wildtype c-Fos.



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.

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. Thus, removal of the fuzzy C-terminal segment considerably decreased c-Fos protection and complexation with c-Jun could barely influence the degradation rate.

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. 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. 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.

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.

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.

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. 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.

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 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. We also 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 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. 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, kD=54 nM). 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 kD, 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. Along these lines, truncating the fuzzy tail in c-Fos$_{\Delta 214}$ unfavorably affects formation of the coiled-coiled structure. In accord with previous data, the association kinetics seems to be affected by mutations in full-length c-Fos. However, binding kinetics of full-length proteins without complete folding might

exhibit complex kinetics, as compared to truncated protein fragments, which fold upon binding.

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. Tandem repeats of short, linear sequences are frequently associated with proteins, which form higher-order protein structures or undergo liquid-liquid phase transition. Interactions between these motifs are often not specific or well-defined, and multivalency or fuzziness 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 from the ubiquitin independent degradation pathway.

6. SUMMARY

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. Our work also highlight a novel aspect of protein fuzziness in regulating of protein half-life. First, we demonstrate that protection of disordered regions from degradation could be achieved without inducing a stable structure. Binding to a partner generates a fuzzy complex, where significant conformational entropy is retained. Second, we show that both specific contacts and fuzzy interactions can impair protein degradation, proportionally to their contribution to binding affinity; suggesting that protein turnover can be regulated via fine tuning protein assembly. Third, low-complexity motifs may selectively inhibit proteolysis by generating higher-order structures via fuzzy interactions, suggesting the role of membraneless cellular compartments in protecting disordered regions from degradation.

7. LIST OF PUBLICATIONS

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

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

DEBRECENI EGYETEM Egyetemi és Nemzeti Könyvtár