Chapter 7

Intrinsically Disordered Proteins in Replication Process

Apolonija Bedina Zavec

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51673

1. Introduction

Intrinsically disordered proteins (IDPs) are proteins that lack stable tertiary conformation (3D structure) under physiological conditions and are biologically active in their unstructured form. IDPs are disordered either along their entire lengths, but more often they are disordered only in localized regions, intrinsically disordered regions (IDRs).

IDRs often undergo transitions to more ordered states after binding to their targets and adopt a fixed three dimensional structures. Folding transition enables specificity without excessive binding strength. Important characteristic of IDRs is multispecificity. One IDR is able to bind multiple targets (multispecific recognition) because it can adopt different conformations upon interaction with different binding partners [1]. IDPs are able to simultaneously bind their partners, which enable the assembly of large complexes. An additional functional advantage of IDPs is increased speed of the interaction due to greater capture radius and larger interaction surfaces.

The level of IDPs is tightly regulated in a cell and diverse post-translational modifications facilitate regulation of their function [2].

IDRs with multispecific recognition capabilities are especially important for the complex recognition processes. Therefore, IDRs are particularly enriched in proteins implicated in cell signalling. It is known that the majority of transcription factors and proteins involved in signal transduction contain long disordered segments [3]. How about IDPs in replication process? The analysis of the yeast proteome showed that IDPs are often located in the cell nucleus [4]. In addition, IDRs are abundant in DNA-binding proteins and many replication and recombination proteins are DNA-binding proteins. Many IDPs are involved in recognition and regulation pathways, because interactions with multiple partner molecules and high-specificity/low-affinity interactions are extremely important in these pathways. Additional interesting feature of IDRs is that they are very sensitive to the environment (Subchapter 4.2.). Summarizing these

findings, a high level of protein disorder is to be expected in processes that take place in the cell nucleus and the highest level of disorder is expected in processes involved in responses to environmental changes. Therefore it is expected that in the nucleus, transcription is a process with the highest level of IDPs. Recombination and repair processes are also expected to have many IDPs; however, these processes are tightly linked to DNA replication and many proteins are used by all three processes. DNA replication is a process that proceeds by a precise program with a defined temporal order. The structural and functional properties of IDPs indicate that a disordered structure is likely present to a lesser extent in DNA replication process. Because of the need for responsiveness to the environment, the initiation of DNA replication should engage more IDPs than the elongation of DNA replication. It is expected that the majority of IDPs in these processes are regulatory proteins.

In this chapter, the binding mechanism of IDRs, the level of IDRs in replication and recombination proteins, and the role of IDPs in replication and recombination processes are discussed.

2. Intrinsically disordered proteins

IDPs contain one or more long intrinsically disordered regions (IDRs) or they are disordered along their entire lengths. Structural disorder can span from short stretches, through long regions, to entire proteins [5]. The majority of IDRs is not fully disordered, but contains some secondary structure and sometimes even partial tertiary structure. IDRs are dynamic fluctuating systems that exist as structural ensembles of rapidly interconverting alternative conformations and perform their biological functions in a highly dynamic disordered state; however, they often have more compact configurations than simply a random coil and contain sites of molecular recognition [6]. The structure of IDPs is similar to a molten globule or pre-molten globule, which preserve the main elements of the native secondary structure and the approximate positions of the folded state, while the loops and ends are flexible. Structural flexibility is a major feature and a major functional advantage of these proteins. IDRs are rich in binding sites for various partners and these binding sites mean that many IDPs with flexible structure are polyfunctional proteins.

The disordered structure gives IDPs specific properties. They need no stable conformation to remain functional; therefore, they are more robust to different changes. Contrary to globular proteins, IDPs are stable at extreme temperatures and extreme pHs [7]. Increases and decreases in temperature or pH can even induce partial folding of IDPs. It has been shown that IDPs partially fold at extreme pH due to minimization of their large net charge present at neutral pH. An increase in temperature can also induce the partial folding of IDPs; in addition, they are resistant to freeze-thaw treatment [8].

2.1. Amino acid (AA) composition of IDPs

IDPs have a specific AA composition that differs from the AA composition of ordered proteins. In particular, IDPs are depleted in hydrophobic (Ile, Leu, Val) and aromatic AA (Trp, Tyr, Phe) that stabilize the structure of folded proteins, while they are enriched in

hydrophilic and charged AAs. The charge/hydropathy (C/H) ratio has been suggested to govern the degree of compaction in IDPs [9]. The combination of low hydrophobicity and high net charge represents an important prerequisite for the disordered structure under physiological conditions.

2.2. Evolution

IDPs are more abundant in eukaryotes than in archaea and prokaryotes, while multicellular eukaryotes have much more predicted disorder than unicellular eukaryotes [4,10]. IDPs play an important role in complex organisms by participating in recognition and in various signalling and regulatory pathways.

IDRs show higher robustness against mutations [11], presumably because changes in protein sequence do not affect protein stability and function as severely. IDRs are more tolerant of mutations than structured proteins. It was found that flexible proteins exhibiting functional promiscuity are the foundation stones of protein evolvability [12]. They are able to accumulate a large number of mutations and thereby facilitate adaptation. Structural disorder seems to enable the rapid appearance of novel, 'less-evolved' proteins [13]. It has been shown that in alternative splicing both alternative proteins have high disorders, because the chance is very low that dual coding would result in two sequences that are both capable of folding into well-defined, functional, 3D structures [14].

3. The binding mechanism of IDRs

IDPs bind to their molecular partners and perform their biological functions by regulation of the function of their binding partners or by promotion the assembly of multi-molecular complexes. One IDR is able to bind many different partners because of its flexible structure; on the other hand, some IDRs do not bind to any partner, but they provide flexible linkers between domains that maintain constant motion during functioning or they provide flexible tails that regulate the structured domains [7,15].

IDPs have functionally relevant characteristics:

- They frequently fold up upon binding to their biological targets [16]. The interaction of a disordered protein with a structured partner, very often induces a disorder-to-order transition thereby forming stable structures, enabling high-specificity-low-affinity interactions [17,18].
- They have possibility of overlapping binding sites (binding diversity) due to extended linear conformation [19]. Structural flexibility of IDPs enables their interactions with numerous biological targets.
- IDRs enable a very large accessible surface area [20]. Greater capture radius and larger interaction surfaces enable increased speed of interactions [15].

- IDPs undergo tighter regulation by post-translational modification as compared to structured proteins [2].

3.1. Complexes with IDPs

Molecular complexes with IDPs are diverse: the IDR may bind on the surface of the binding partner (Figure 1), by wrapping around the binding partner, or by penetrating deep inside the binding partner [21].

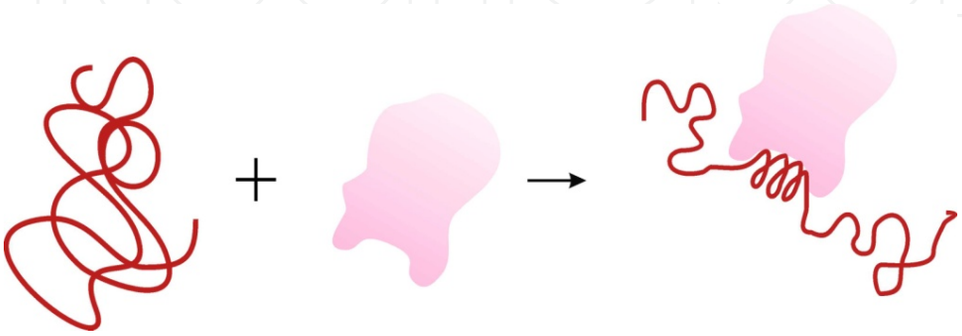

Figure 1. Intrinsically disordered protein forms complex with structured protein.

IDRs in complexes may control the degree of motion between domains, mask binding sites, enable transient binding of different binding partners, and be targets of post-translational modifications. IDPs are often involved in the binding of large partners or they are proteins involved in the binding of large number of small partners. In the latter case, they often function as scaffold proteins that enable the assembly of the relevant proteins into specific multi-molecular complexes and increase the efficiency of the interaction between partner molecules (Figure 2).

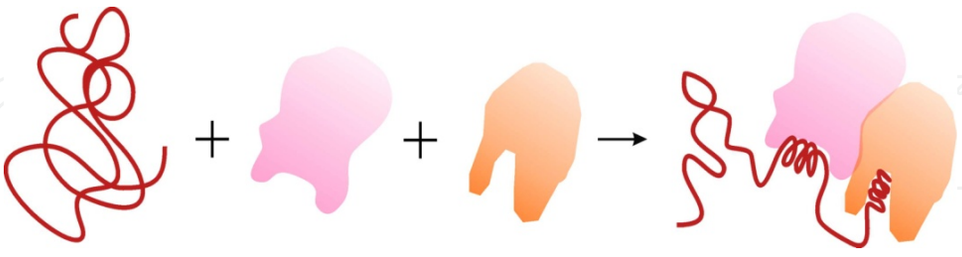

Figure 2. Intrinsically disordered proteins often function as scaffold proteins that enable assembling the relevant proteins into multi-molecular complexes.

The majority of intrinsically disorder-based complexes are ordered and relatively static due to disorder-to-order transitions; however, there are also dynamic complexes where IDRs go through an ensemble of rapidly interconverting conformations. Dynamic complexes do not involve significant ordering of the interacting protein segments but rely exclusively on transient contacts.

3.2. Mechanisms of formation of the complexes with IDPs

The primary mechanism by which disorder is utilized in molecular interactions is that the same IDR may fold differently and bind to several structurally diverse partners. On the other hand, different IDRs with different AA sequences may use their flexibility to bind to the same protein partner [6,22]. Their associations are dynamic. The lack of structure of highly flexible IDRs enables more diverse functionality [23]. IDRs are ensembles of conformations and each individual conformation has a dynamic structure. The binding partner selects the most binding-compatible conformation from this ensemble to form a complex [21,15]. The equilibrium is thus shifted towards this interaction-prone member of the conformational ensemble.

Models of IDRs interaction processes:

- The 'binding and folding' mechanism with disorder-to-order transition is the most accepted model for the binding of IDR, where a highly structured conformation is formed by binding to the partner molecule. A structured conformation is formed on binding IDR (the local disorder-to-order transition) or on the entire molecule of IDP (the global disorder-to-order transition) [25,26]. An IDR binds weakly at a relatively large distance followed by folding when the protein comes close to the binding site. One model utilizes a prediction that an IDR with an open structure has a larger binding surface and a greater capture radius for a specific binding site than the ordered protein and therefore the binding rate is significantly enhanced over the binding rate of the ordered proteins [21]. The binding induced disorder-to-order transition is accompanied by a dramatic decrease in accessible surface area and by the release of a large number of water molecules [6]. A large decrease in conformation entropy during this process enables highly specific but easily reversible interactions.

- The 'polyelectrostatic' model describes the interaction of highly charged IDR with several similar binding motifs and a folded partner with one binding site [27,28]. Multiple disordered binding motifs interact with the partner's folded binding site in a dynamic equilibrium. The flexibility of the IDR makes all binding motifs equally accessible. Weak affinities of the individual interactions permit their efficient exchange. In this model, the IDR generates an electrostatic field representing the cumulative electrostatic interaction of all charges in the IDR.

- The 'multi-step interaction' model describes the binding of an IDR that depends on the conformational selection of the structural ensemble via the pre-formed elements that dominate the ensemble [29]. When the IDP comes close to the binding site of the partner molecule, an encounter complex is formed that either proceeds towards the final complex or dissociates again. Electrostatic forces are the most important for encounter complex formation [30]. Interacting partners in the encounter complex affect the conformational landscapes of each other. Consecutive steps depend on the preceding steps and cooperation between protein partners. This process is called an interdependent protein dance [31,32]. The structural variability of complexes with IDPs can be considered a reflection of interdependent protein dynamics, where the structure of the complex is a result of coordinated mutual co-folding [21]. In such encounters 'pre-

organized' complexes, mainly non-specific electrostatic interactions are involved and multiple conformations and orientations are employed. In the 'multi-step interaction' model, IDPs interacts with their partners by a biphasic process with a fast Phase I leading to the formation of disordered complexes and slower Phase II leading to the formation of ordered complexes. Phase II includes the 'Binding and folding' model that may or may not (binding without folding) follow a Phase I [33]. 'Polyelectrostatic' complexes are probably the stopped stages of encounter complexes [21].

It is the most likely that the IDR contains a conformational preference for the structure it will take upon binding.

3.3. Levels of IDPs in the cell and modulation of their activity

The level of IDPs inside the cell is precisely controlled. IDPs are more tightly regulated as compared to structured proteins. Obviously it is very important that they are available at the appropriate time and in the appropriate amount. The level of IDPs is controlled at the synthesis and clearance levels and their activity is further modulated via interaction with specific binding partners and post-translational modifications. IDRs are more solvent-accessible then folded regions and therefore suitable for diverse post-translational modifications, such as phosphorylation, sumoylation, ubiquitination, acetylation, etc. Such modification can change the electrostatic properties of IDRs and affect their affinity for charged molecules like DNA.

The predicted intrinsic disorder is the strongest determinant of dosage sensitivity - proteins become harmful when they are overexpressed [34]. The likely cause of dosage sensitivity is the binding promiscuity of IDPs [11]. IDRs are prone to make promiscuous interactions when their concentration is increased; it has been demonstrated that this is a likely cause of pathology when genes are overexpressed [34].

4. The role of IDPs in replication processes

This chapter refers to the proteins of budding yeast *Saccharomyces cerevisiae*, because *S.cerevisiae* is the best studied eukaryotic model organism that providing the most integrated view of replication and recombination processes.

4.1. DNA replication process is tightly linked to recombination process

DNA replication and DNA recombination are central characteristics of life that cooperate to maintain biological inheritance and genomic integrity. Replication enables the formation of two identical DNA molecules from a single double-stranded DNA, while recombination enables accurate repair of errors that occur on both strands of DNA, as well as the formation of new combinations of genes. Both processes are tightly intertwined [35]. The recombination system plays a crucial role in DNA replication ensuring that the replication machines can complete their task of genome duplication. DNA replication forks stall or collapse at DNA

lesions or problematic genomic regions. When replication forks collapse, recombination is the most important rescue mechanism. The recombination mechanism forms substrates for the assembly of a new replication fork thus allowing continued DNA replication. On the other hand, DNA synthesis is a crucial step during the recombination process. After Rad51-mediated DNA strand invasion, DNA synthesis is the next step in recombination to restore the integrity of the chromosome. Repair DNA synthesis during the recombination process is similar to normal S-phase replication, but has specific properties. Thus recombination is part of DNA replication and, vice versa, DNA synthesis is part of the recombination process.

Clearly then, the replication process requires both, replication and recombination proteins, but then again so does the recombination process. This is why replication and recombination proteins are discussed within the same functional group.

4.2. Predicted level of IDPs in replication and recombination processes

There are some facts to consider when predicting the level of IDPs in replication and recombination processes:

- An analysis of the yeast proteome showed that IDPs are often located in the cell nucleus [4]. IDRs are abundant in DNA-binding proteins, while many replication and recombination proteins are DNA-binding proteins. IDRs play a crucial role in DNA-binding proteins by increasing the affinity and specificity of DNA binding [36]. The ability of IDRs to interact with DNA is tightly linked to the high content of charged residues in IDRs; IDRs that bind to DNA are rich in positively charged residues and their positive charges are highly clustered.
- Many IDPs are involved in recognition and regulation pathways, because interactions with multiple partner molecules and high-specificity/low-affinity interactions are extremely important in these pathways [2].
- Interesting feature of IDRs is that they are very sensitive to the environment. Flexible IDPs more readily undergo conformational change in response to environmental perturbations than rigid proteins [37,38]. Due to flexible structure, their local and global structures can easily be shaped by their environment. High-specificity/low-affinity interactions with their partners enable extremely sensitive functioning of IDPs, which is favourable for responses to the environmental changes. In addition, the level of IDPs inside the cell is precisely controlled (Subchapter 3.4.) allowing rapid and accurate responses of the cell to changing environmental conditions. Higher and more regulated synthesis, higher degradation rates, and tightly regulated activity make the levels of IDPs very sensitive to the environment.

Summarizing these findings, a high level of protein disorder is to be expected in processes that take place in the cell nucleus, especially within regulatory proteins. The highest level of disorder is expected in processes involved in responses to environmental changes. According to those findings, in the nucleus, transcription should be the process with the highest level of IDPs. Recombination and repair processes are also expected to have many IDPs; however, these processes are tightly linked to DNA replication and many proteins are used by all these processes. DNA replication is a process that proceeds by a precise program with

a defined temporal order. The structural and functional properties of IDPs indicate that a disordered structure is likely present to a lesser extent in DNA replication process. However, the initiation of DNA replication would be expected to engage more IDPs than the elongation of DNA replication due to the need for responsiveness to the environment (Figure 3). It is expected that the majority of IDPs in these processes are regulatory proteins.

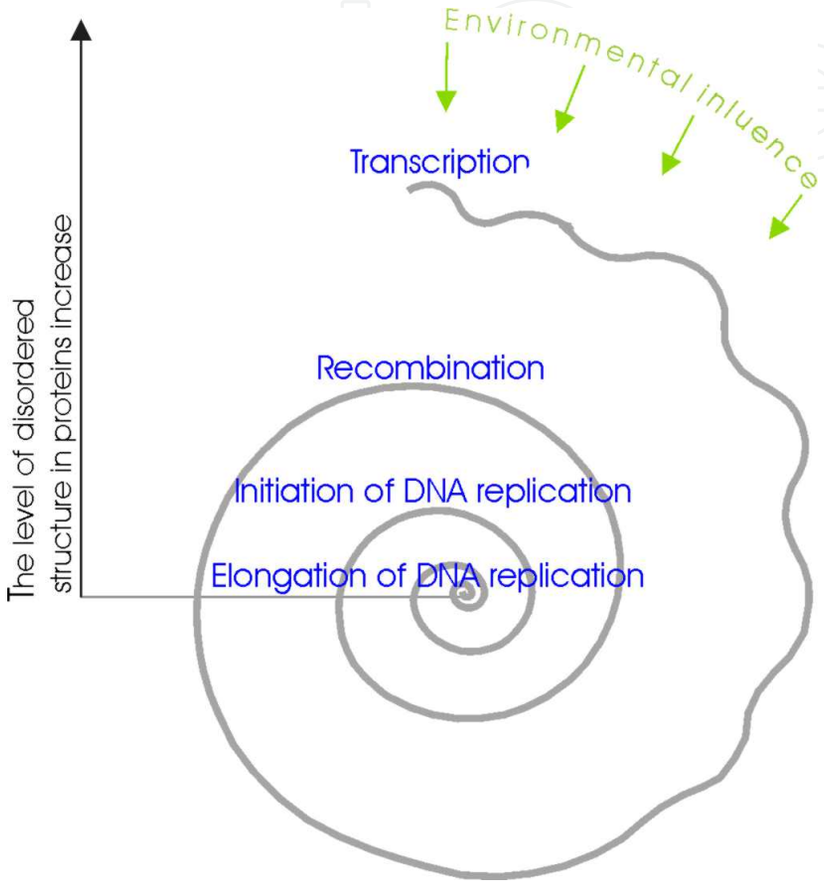

Figure 3. Processes in the nucleus: global prediction of protein disordered structure in the processes linked to DNA, considering responsiveness to changes in the environment.

4.3. IDRs in replication proteins

Analysis of predicted IDRs within proteins that have a role in DNA replication was done (Table 1). The majority of them functions also in recombination and repair processes.

It was found that proteins with the role in initiation of DNA replication have more predicted disordered structure (26%) than proteins with the role in elongation of DNA replication (20%). Difference is significant among proteins with very short IDRs; there is 35% proteins with the role in initiation of DNA replication that contain less than 10% disordered structure, while there is as much as 60% such proteins with the role in elongation of DNA replication. However, difference is tiny among proteins with very large IDRs; 22% proteins with the role in initiation of DNA replication contain more than 50% disordered structure and there is 20% such proteins with the role in elongation of DNA replication.

Protein	Protein Length (AA)	*Location of IDR (AA)	*IDR Length (AA)	% of disorder	Recombination and Repair
		Elongation of DNA replication			
Pol2	2222	0-44, 1186-1263	44, 77	5.4	R
Dpb2	689	92-159	67	9.7	R
Dpb3	201	96-201	105	**52.2**	R
Dpb4	196	0-26, 119-196	26, 77	**52.6**	R
Pol3	1097	0-100	100	9.1	R
Pol31	487	0-32	32	6.6	R
Pol32	350	118-350	232	**66.3**	R
PCNA	258	-	0	0	R
Cdc9	755	0-144	144	19.1	R
Rfa1	621	126-183	57	9.2	R
Rfa2	273	0-39, 177-235	39, 58	35.5	R
Rfa3	122	-	0	0	R
Rfc1	861	0-155, 230-296, 780-861	155, 66, 81	35.1	R
Rfc2	353	0-25	25	7.1	
Rfc3	340	-	0	0	
		Initiation of DNA replication			
Pol1	1468	82-341	259	17.6	R
Pol12	705	70-203	133	18.8	R
Pri1	409	-	0	0	R
Pri2	528	0-43	43	8.1	R
Orc1	914	0-42, 195-428	42, 233	30.1	
Orc2	620	0-267	267	43.1	
Orc3	616	0-40	40	6.5	
Mcm2	868	0-68, 112-189	68, 77	16.7	R
Mcm3	971	742-897	155	16.0	R
Mcm4	933	0-177	177	19.0	R
Mcm10	571	41-154, 338-571	113, 233	**60.6**	
Sld2	453	0-453	453	**100**	
Sld3	668	87-155, 292-335, 417-668	68, 43, 271	**57.2**	
Sld5	294	17-45	28	9.5	
Psf1	208	-	0	0	
Psf2	213	195-213	18	7.8	
Psf3	194	-	0	0	
Dpb11	764	226-321, 567-764	95, 197	38.2	R
Cdt1	604	427-507	80	13.2	R
Cdc6	513	0-61	61	11.9	
Cdc7	507	-	0	0	R
Dbf4	704	0-110, 317-658	110, 341	**64.1**	R
Ecm11	302	0-183	183	**60.6**	R

Data about proteins were obtained from *Saccharomyces genome* database [39] and references therein.
Server Disopred2 [40] was used for protein disorder prediction.
*Predicted disorder regions more than 30 AA long (more than 20 AA long for very small proteins).
R - proteins involved in recombination and repair processes;
Disordered regions at N- or C- terminus are underlined.

Table 1. Predicted disorder regions of replication proteins.

The majority (63%) of IDRs are at N- or C- terminus in proteins that have a role in DNA replication. Most often IDRs are at the N-terminus, in 44% of replication proteins; 19% of replication proteins have IDRs at the C-terminus.

4.4. IDRs in DNA polymerases

4.4.1. DNA polymerase α

Polymerase α is the only enzyme that can synthesize DNA de novo. It is required for initiation of chromosomal DNA replication during mitosis and meiosis, intragenic recombination, and repair of double stranded DNA breaks. Pol1, the catalytic subunit of polymerase α complex, has a lot of disordered structure in comparison to Pol3 and Pol2 (Table 1). This fact is consistent with the hypothesis concerning the expected average level of disordered structures, since Pol1 is required for the initiation of DNA replication and Pol3 and Pol2 are required for the elongation of DNA replication.

It was shown that IDR of Pol1 interacts with Cdc13: Pol1 residues 13-392 [41] or Pol1 residues 47-560 [42]. Actually, a short fragment of Pol1 consisting only of residues 215-250 is necessary and sufficient for binding with Cdc13. This disordered region of Pol1 becomes well ordered, folded into a single amphipathic α-helix, when it is in complex with Cdc13, as evidenced by good electron density in the crystals [43]. The interaction between IDR of Pol1 and Cdc13 is primarily mediated by a highly positively charged groove of Cdc13 and a negatively charged acidic convex surface of Pol1. These two surfaces are not only opposite in charge distribution but also complementary in shape.

4.4.2. DNA polymerase δ

DNA polymerase δ is a major replicative DNA polymerase and is primarily required for the lagging strand synthesis. It is a heterotrimeric complex composed of the catalytic subunit Pol3, the structural subunit Pol31, and an additional auxiliary subunit Pol32. Pol32 is highly disordered protein (Figure 4). While structured Pol3 and Pol31 are essential for viability, the disordered Pol32 is not essential. Pol3 and Pol31 are highly conserved in eukaryotes; on the other hand, the disordered Pol32 shows an extreme divergence in its AA sequence [44]. Hydrodynamic studies of polymerase δ have shown an unusually high Stokes radius [45]. This deviation from globularity may be due to the disordered structure of Pol32.

Pol32 is bound to Pol3 through Pol31. The C-terminus of Pol3 interacts with the conserved region of Pol31 [46]. Deletion of the last four C-terminal AAs of Pol3, which are required for the interaction between the Pol3 and Pol31, does not affect DNA replication but leads to defects in homologous recombination and in break-induced replication repair pathways. Deletion of Pol32 leads to signs of DNA replication defects and DNA repair defects, with increased sensitivity to ultraviolet (UV) radiation and methylation damage [47].

Pol32 binds to Pol31 by the N-terminus (92 AA) and to PCNA by the C-terminus [48]. The structured N-terminus of Pol32, which enables binding to Pol3 through Pol31, is essential for damage-induced mutagenesis. Highly disordered C-terminus of Pol32 interacts with the

http://dx.doi.org/10.5772/51673

C-terminus of PCNA during DNA synthesis. Although the C-terminus of Pol32 is highly disordered, there is one motif that is highly conserved in this region: the consensus PCNA-binding motif 338-QGTLESFFKRKAK-350 (conserved amino acids in bold).

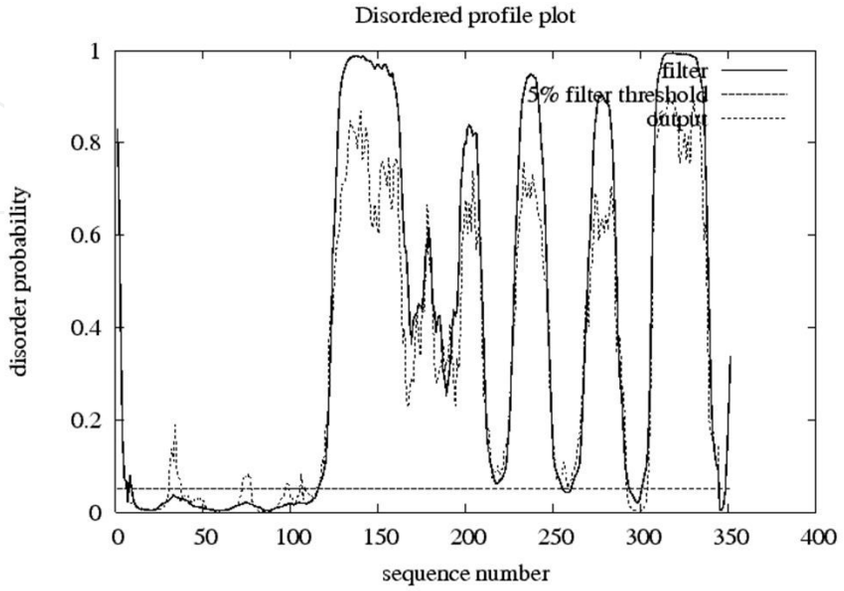

Figure 4. Predicted disordered regions of Pol32 (Disopred2).

It has also been shown that Pol32 interacts also with Pol1 that is a part of polymerase α, suggesting that Pol δ and Pol α interact *via* the Pol32 subunit [48]. These findings show diverse role of Pol32 as typical IDP.

For the replication of the lagging strand where the polymerase must dissociate from the DNA after extension of each Okazaki fragment, Polδ utilizes a collision-release mechanism where the Polδ is released from PCNA. Polδ exhibits a very high processivity in synthesizing DNA with the PCNA sliding clamp. It has been shown that the N-terminal region of Pol3 interacts with PCNA, and that this interaction increases Pol3 processivity [49]. The N-terminal of Pol3 is predicted to be highly disordered (Table 1). Pol31 and Pol32 also have binding sites for PCNA and all three subunits contribute to PCNA-stimulated DNA synthesis by Polδ [50].

4.4.3. DNA polymerase ε

DNA polymerase ε is primarily required for the leading strand synthesizes. Pol2 is the catalytic subunit of the polymerase ε complex. It has been shown that the highly structured C-terminus of Pol2 is essential for DNA replication [51].

Dpb3 and Dpb4 are nonessential small subunits of the DNA polymerase ε complex, which have a histone fold. Both Dpb3 and Dpb4 are highly disordered. They form a subassembly that interacts with histones and functions in transcriptional silencing caused by chromatin structures [52].

4.5. Dbf4, IDP with the role in replication and recombination

The complex Cdc7–Dbf4, also known as Dbf4-dependent kinase (DDK), has a role at eukaryotic origins of replication. DDK is required for origin firing and replication fork progression in mitotic S phase, for pre-meiotic DNA replication, meiotic double strand break formation, recruitment of the monopolin complex to kinetochores during meiosis I and as a gene-specific regulator of the meiosis-specific transcription factor Ndt80p. DDK is a Ser/Thr kinase whose activity depends on the association of the Cdc7 catalytic subunit with a regulatory subunit Dbf4. The level of Dbf4 is changes during the cell cycle and is the highest during metaphase I [53]. Both subunits Cdc7 and Dbf4 are essential for growth.

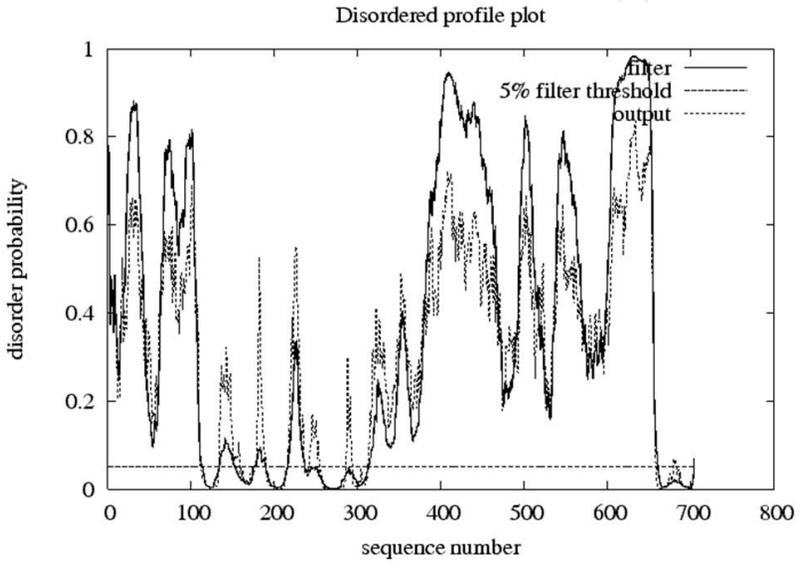

Figure 5. Predicted disordered regions of Dbf4 (Disopred2).

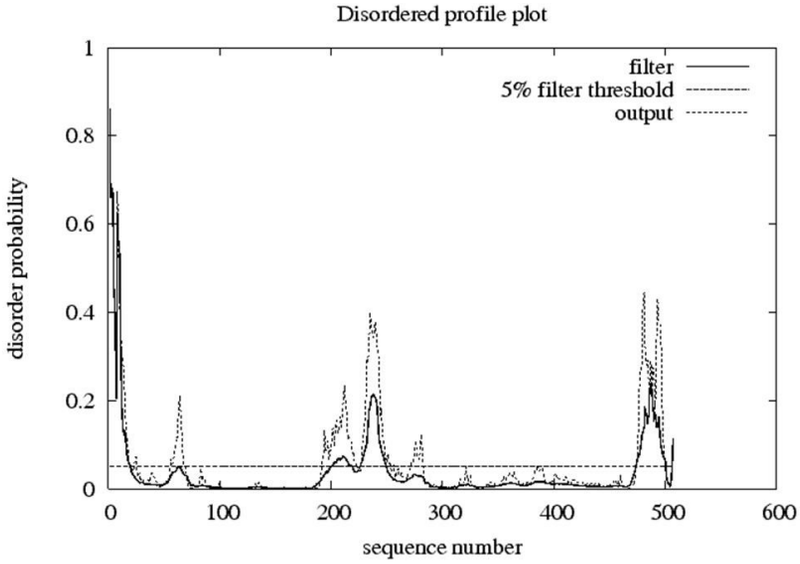

Figure 6. Predicted disordered regions of Cdc7 (Disopred2).

As expected, the regulatory subunit Dbf4 is highly disordered (Figure 5), while catalytic subunit Cdc7 is a highly structured protein (Figure 6).

Dbf4 is a highly disordered protein with a disordered N-terminus (110 AA) and an IDR that is half the length of the protein at the C-terminus; it has only a 200 AA long structural region between both IDRs (Figure 5). It was shown that highly disordered C-terminus of Dbf4 has a role in response to mutation by HU and that it is required in meiosis [54]. Superfamily assignments [55] show no confident structure prediction for Dbf4, while Cdc7 has a predicted cyclin-dependent protein kinase function at 1-304 AA with protein, ATP, and DNA binding activity. Cdc7 has well conserved subdomains (30-195, 275-348, 438-465) found in the eukaryotic protein kinase superfamily, while Dbf4 contains only three short conserved regions, termed N (135-179), M (260-309), and C (659-696) [56]. Two of the three conserved regions (N and M) are found in the the structural region of Dbf4.

4.5.1. Dbf4 in initiation of replication during mitosis

DDK phosphorylates the Mcm2-7 helicase, and is probably required for helicase activation or for recruitment of pre-IC factors. DDK preferentially phosphorylates the MCM complexes that are most tightly linked to the DNA [57]. Dbf4 associates with origins in an ORC-dependent manner [58]. The pre-RC components Mcm2, Mcm4, Orc2, and Orc3 have each been identified as binding partners for Dbf4 [59,60,61]. The N-terminal half of Dbf4 is critical for recruitment of DDK to the origin. The highly disordered C-terminal half of Dbf4 is required to bind the Cdc7 kinase [58]; more precisely, region 573-695 is required for interaction with Cdc7, while the structured region of Dbf4 (110-296) is required for binding the Mcm2–7 complex [60].

4.5.2. Dbf4 in checkpoint control

During the replication checkpoint response, Dbf4 is phosphorylated by checkpoint kinase Rad53 allowing inhibition of initiation of replication at late origins. Checkpoint control during S-phase slows the rate of DNA replication in response to DNA damage and blocks the replication fork. This regulation is achieved through the Rad53 kinase-dependent block of late origins of replication [62]. Dbf4 has been shown to be phosphorylated in a Rad53-dependent manner in response to replication stress, which correlates with a reduced DDK activity [63]. It was shown that mutations at predicted Rad53 phosphorylation sites (Ser84, Ser235, Ser377, Thr467, Thr506, Ser507, and Thr551) contribute to bypassing such control [64].

The conserved region N of Dbf4 (66-221) is necessary for the interaction of Cdc7-Dbf4 with the checkpoint kinase Rad53. The core of this binding region folds as a BRCT domain; in addition, it includes an additional N-terminal helix unique to Dbf4 that is essential for the interaction with Rad53 [65]. This unique N-terminal part of the conserved region N is predicted to be an IDR (Figure 6) and probably becomes helix-structured after binding with Rad53.

4.5.3. Dbf4 in meiosis

DDK is required for replication, recombination and segregation events during meiosis in yeast. It has been shown that in addition to the initiation of DNA replication, DDK has an

important role in the initiation of meiotic recombination [66]. DDK phosphorylates the double strand break protein Mer2 and facilitates meiotic recombination [67]. CDK-S and DDK function sequentially phosphorylate Mer2 on adjacent serines, Ser30 and Ser29, allowing formation of meiotic double strand breaks.

DDK plays a role in meiotic segregation. DDK allows expression of *NDT80*, a global transcription factor in meiosis, required for the induction of genes required for meiotic progression and spore formation. DDK promotes *NDT80* transcription by relieving repression mediated by a complex of Sum1, Rfm1, and histone deacetylase Hst1. Sum1 exhibits meiosis-specific Cdc7-dependent phosphorylation. By this function, DDK links DNA replication to the segregation of homologous chromosomes in meiosis I [68,69].

DDK is also necessary for recruitment of monopolin Mam1 to sister kinetochores, which is required for mono-orientation of sister kinetochores in the reductional segregation occurring during meiosis I.

The use of the same Cdc7-Dbf4 complex to regulate many distinct meiosis-specific processes could be important for the coordination of these processes during meiosis [68]. DDK is a link between DNA replication, recombination and mono-orientation during meiosis I in budding yeast [70]. In addition to the unifying role in meiosis, DDK has a role in initiation of replication during mitosis and in checkpoint control. Highly flexible structure of Dbf4 is very likely crucial for such a complex role of DDK.

4.6. Ecm11, IDP with the role in replication and recombination

Ecm11 is a protein with a strong meiotic phenotype; it affects meiotic DNA synthesis and recombination [71]. Homozygous deletion of the *ECM11* gene causes delay in a process of meiosis, lower efficiency of ascii formation and lower spore viability.

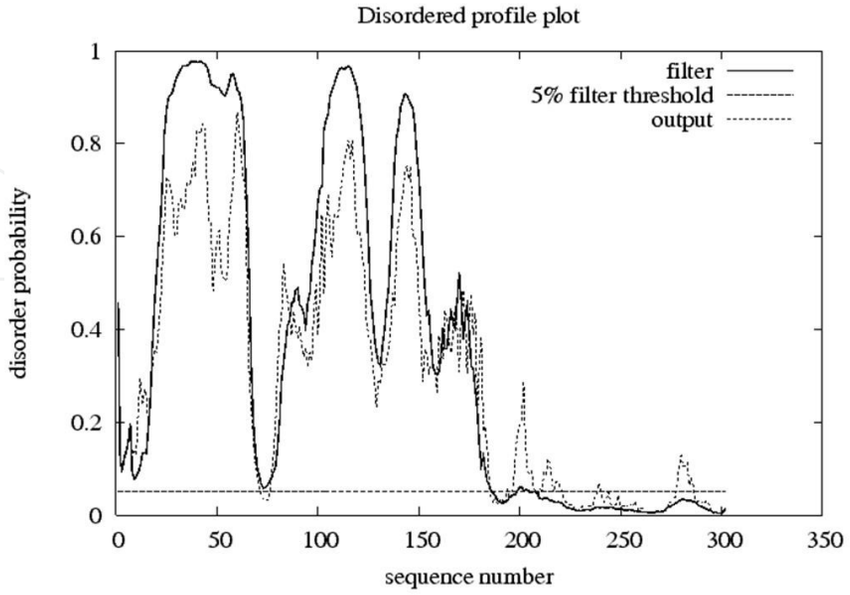

Figure 7. Predicted disordered regions of Ecm11 (Disopred2).

Ecm11 is highly disordered protein; 2/3 of Ecm11 is unstructured (Table 1). Ecm11 has 302 mostly hydrophilic AA as expected for IDP (Subchapter 2.1.). IDR of 183 AA is located at the N-terminal end of Ecm11 (Figure 7). The C-terminal end is predicted to be mostly helical and contain coiled-coil motif at the very C-terminus. Superfamiliy assignments [55] show no confident structure prediction for Ecm11.

4.6.1. The ecm11 mutation affects sporulation efficiency

It was showed that *ecm11* homozygous diploid strains sporulate more slowly and less efficiently than the wild type strains [71]. Wild type strains carrying additional *ECM11* on the centromeric plasmid also showed reduced sporulation efficiency comparing to wild types. Obviously, sporulation efficiency depends on the copy number of Ecm11 protein in the cell during meiosis. As more Ecm11 than usual in the cell make lower sporulation efficiency, Ecm11 is probably a part of heterologous protein complex, demanding exactly correct balance among those proteins.

4.6.2. Ecm11 has a role in meiotic recombination

It was showed that *ecm11* homozygous spores have reduced viability for 50% [71]. The majority of *ecm11* ascii (56%) produced only two viable spores, while only 1% of such ascii were observed in the parental strain. This result shows non-disjunction of homologous chromosomes at the first meiotic division. By recombination tests was demonstrated that *ECM11* is required for crossing over, but not for gene conversion. This result raises the possibility that *ecm11* mutation impairs the crossover process at an early step of recombination, at the differentiation of intermediates into crossovers or non-crossovers.

4.6.3. Ecm11 is required for meiotic DNA replication

Deletion of the *ECM11* gene cause diminished DNA replication in meiosis [71].

In the two-hybrid screen it was found out that Ecm11 strongly interacts with Cdc6 that has a pivotal role in the initiation of DNA replication [72]. Genetic interactions between Cdc6 and Ecm11 were also observed. Moderate supression of *cdc6-1* mutation by overexpression of *ECM11* was detected [72] and deletion of *ECM11* in *cdc6-1* genetic background enhances thermo-sensitivity of *cdc6-1* mutation (Zavec AB, unpublished result). These data suggest direct involving of Ecm11 in initiation of DNA replication process.

4.6.4. Ecm11 is modified by SUMO during meiosis

IDPs are tightly regulated in a cell and diverse post-translational modifications (such as ubiquitination, sumoylation, and phosphorylation) facilitate regulation of their function (Subchapter 3.4.). It was shown that the majority of Ecm11 protein in the cell is sumoylated during meiosis [73]. Lys5 at the highly disordered N-terminus of Ecm11 is modified by SUMO. It was shown that sumoylation is essential for biological role of Ecm11 in meiosis and that sumoylation directly regulates Ecm11 function in meiosis.

5. Conclusion

Cell nuclei contain high levels of IDPs. In this work, a hypothesis has been made, that in the nucleus, transcription is a process with the highest level of IDPs and that a disordered structure is likely present to a lesser extent in DNA replication process. However, the initiation of DNA replication would be expected to engage more IDPs than the elongation of DNA replication due to the need for responsiveness to the environment. By analysis of predicted disordered structure in replication proteins, it was confirmed that proteins with the role in initiation of DNA replication have more disordered structure than proteins with the role in elongation of DNA replication. The majority of IDRs in these proteins are at N- or C- terminus, most often IDRs are at the N-terminus of IDPs.

Acknowledgements

Drago Cerjan provided help with figures.

Financial support was received from the Slovenian Research Agency, Grant P1-0104.

Author details

Apolonija Bedina Zavec*

Address all correspondence to: polona.bedina@ki.si

National Institute of Chemistry, Department for Molecular biology and Nanobiotechnology, Slovenia

References

[1] Uversky, N., Oldfield, C. J., & Dunker, A. K. (2005). Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit*, 18, 343-384.

[2] Babu, M. M., Lee, R., Groot, S. N., & Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Bio*, 21(3), 432-40.

[3] Lakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z., & Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol.*, 323(3), 573-84.

[4] Ward, J. J., Sodhi, J. S., Mc Guffin, L. J., Buxton, B. F., & Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol*, 337(3), 635-45.

[5] Tompa, P., Fuxreiter, M., Oldfield, C. J., Simon, I., Dunker, A. K., & Uversky, V. N. (2009). Close encounters of the third kind: disordered domains and the interactions of proteins. *Bioessays*, 31(3), 328-35.

[6] Uversky, V. N., & Dunker, A. K. (2010). Understanding protein non-folding. *Biochim Biophys Acta*, 1804(6), 1231-64.

[7] Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*, 11(4), 739-56.

[8] Tantos, P., Friedrich, P., & Tompa, P. (2009). Cold stability of intrinsically disordered proteins. *FEBS Lett*, 583(2), 465-9.

[9] Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L., & Pappu, R. V. (2010). Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci*, USA, 107(18), 8183-8.

[10] Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., & Dunker, A. K. (2005). Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, 44(6), 1989-2000.

[11] Bellay, J., Han, S., Michaut, M., Kim, T., Constanzo, M., Andrews, B. J., Boone, C., Bader, G. D., Myers, C. L., & Kim, P. M. (2011). Bringing order to protein disorder through comparative genomics and genetic interactions. *Genome Biol*, 12(2), R14.

[12] Tokuriki, N., & Tawfik, D. S. (2009). Protein Dynamism and Evolvability. *Science*, 324, 203-207.

[13] Tompa, P. (2011). Unstructural biology coming of age. *Current Opinionin Structural Biology*, 21, 1-7.

[14] Kovacs, E., Tompa, P., Liliom, K., & Kalmar, L. (2010). Dual coding in alternative reading frames correlates with intrinsic protein disorder. *Proc Natl Acad Sci*, USA, 107(12), 5429-5434.

[15] Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem Sci*, 27(10), 527-33.

[16] Dyson, H. J., & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol*, 12, 54-60.

[17] Demchenko, A. P. (2001). Recognition between flexible protein molecules: induced and assisted folding. *J Mol Recognit*, 14, 42-61.

[18] Tompa, P., Szasz, C., & Buday, L. (2005). Structural disorder throws new light on moonlighting. Trends Biochem Sci. , 30, 484-489.

[19] Dyson, H. J., & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.*, 6(3), 197-208.

[20] Marsh, J. A., & Teichmann, S. A. (2011). Relative Solvent Accessible Surface Area Predicts Protein Conformational Changes upon Binding Structure. 19(6), 859-867.

[21] Uversky, V. N. (2011). Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. *Chem Soc Rev.*, 40(3), 1623-34.

[22] Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., & Dunker, A. K. (2008). Flexible nets: disorder and induced fit in the associations of 53 and 14-3-3 with their partners. *BMC Genomics*, 9(1), S1.

[23] Cortese, M. S., Uversky, V. N., & Dunker, A. K. (2008). Intrinsic disorder in scaffold proteins: getting more from less. *Prog Biophys Mol Biol*, 98(1), 85-106.

[24] Tobi, D., & Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc Natl Acad Sci*, USA, 102(52), 18908-13.

[25] Shoemaker, B. A., Portman, J. J., & Wolynes, P. G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci*, USA, 97(16), 8868-73.

[26] Pontius, B. W. (1993). Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends Biochem Sci*, 18(5), 181-6.

[27] Mittag, T., Orlicky, S., Choy, W. Y., Tang, X., Lin, H., Sicheri, F., Kay, L. E., Tyers, M., & Forman-Kay, J. D. (2008). Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci*, USA, 105(46), 17772-7.

[28] Mittag, T., Marsh, J., Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M., & Forman-Kay, J. D. (2010). Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase. *Structure*, 18(4), 494-506.

[29] Fuxreiter, M., Simon, I., Friedrich, P., & Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol*, 338(5), 1015-26.

[30] Ubbink, M. (2009). The courtship of proteins: understanding the encounter complex. *FEBS Lett*, 583(7), 1060-6.

[31] Kovács, I.A. , Szalay, M. S., & Csermely, P. (2005). Water and molecular chaperones act as weak links of protein folding networks: energy landscape and punctuated equilibrium changes point towards a game theory of proteins. *FEBS Lett*, 579(11), 2254-60.

[32] Antal, M. A., Böde, C., & Csermely, P. (2009). Perturbation waves in proteins and protein networks: applications of percolation and game theories in signaling and drug design. *Curr Protein Pept Sci*, 10(2), 161-72.

[33] Sigalov , A. B. (2010). Protein intrinsic disorder and oligomericity in cell signaling. *Mol Biosyst*, 6(3), 451-61.

[34] Vavouri, T., Semple, J. I., Garcia-Verdugo, R., & Lehner, B. (2009). Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell*, 138(1), 198-208.

[35] Kusic-Tisma, J. (2011). *DNA Replication and Related Cellular Processes*, Rijeka, InTech.

[36] Vuzman, D., & Levy, Y. (2012). Intrinsically disordered regions as affinity tuners in protein-DNA interactions. *Mol Biosyst*, 8(1), 47-57.

[37] Wright, P. E., & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J of Mol Bio.*, 293(2), 321-331.

[38] Romero, P., Obradovic, Z., & Dunker, A. K. (2004). Natively disordered proteins: functions and predictions. *Appl Bioinformatics*, 3(2-3), 105-13.

[39] Saccharomyces genome database. (2012). http://www.yeastgenome.org/, accessed 20 June.

[40] The DISOPRED2 Prediction of Protein Disorder Server. (2012). http://bioinf.cs.ucl.ac.uk/disopred/, accessed 20 June.

[41] Qi, H., & Zakian, V. A. (2000). The Saccharomyces telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes Dev*, 14(14), 1777-88.

[42] Hsu, C. L., Chen, Y. S., Tsai, S. Y., Tu, P. J., Wang, M. J., & Lin, J. J. (2004). Interaction of Saccharomyces Cdc13p with Pol1Imp4pSir4p and Zds2p is involved in telomere replication, telomere maintenance and cell growth control. *Nucleic Acids Res*, 32(2), 511-21.

[43] Sun, J., Yang, Y., Wan, K., Mao, N., Yu, T. Y., Lin, Y. C., De Zwaan, D. C., Freeman, B. C., Lin, J. J., Lue, N. F., & Lei, M. (2011). Structural bases of dimerization of yeast telomere protein Cdc13 and its interaction with the catalytic subunit of DNA polymerase α. *Cell Res*, 21(2), 258-74.

[44] Cliften, P. F., Hillier, L. W., Fulton, L., Graves, T., Miner, T., Gish, W. R., Waterston, R. H., & Johnston, M. (2001). Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis. *Genome Res*, 11(7), 1175-86.

[45] Johansson, E., Majka, J., & Burgers, P. M. (2001). Structure of DNA polymerase delta from Saccharomyces cerevisiae. *J Biol Chem*, 276(47), 43824-8.

[46] Brocas, C., Charbonnier, J. B., Dhérin, C., Gangloff, S., & Maloisel, L. (2010). Stable interactions between DNA polymerase δ catalytic and structural subunits are essential for efficient DNA repair. *DNA Repair*, 9(10), 1098-111.

[47] Haracska, L., Prakash, S., & Prakash, L. (2000). Replication past O(6)-methylguanine by yeast and human DNA polymerase eta. *Mol Cell Biol*, 20(21), 8001-7.

[48] Johansson, E., Garg, P., & Burgers, P. M. (2004). . The Pol32 subunit of DNA polymerase delta contains separable domains for processive replication and proliferating cell nuclear antigen (PCNA) binding. J Biol Chem , 279(3), 1907-15.

[49] Brown, W. C., & Campbell, J. L. (1993). Interaction of proliferating cell nuclear antigen with yeast DNA polymerase delta. *J Biol Chem*, 268(29), 21706-10.

[50] Acharya, N., Klassen, R., Johnson, R. E., Prakash, L., & Prakash, S. (2011). PCNA binding domains in all three subunits of yeast DNA polymerase δ modulate its function in DNA replication. *Proc Natl Acad Sci*, USA, 108(44), 17927-32.

[51] Dua, R., Levy, D. L., & Campbell, J. L. (1999). Analysis of the essential functions of the C-terminal protein/protein interaction domain of Saccharomyces cerevisiae pol epsilon and its unexpected ability to support growth in the absence of the DNA polymerase domain. *J Biol Chem*, 274(32), 22283-8.

[52] Tsubota, T., Tajima, R., Ode, K., Kubota, H., Fukuhara, N., Kawabata, T., Maki, S., & Maki, H. (2006). Double-stranded DNA binding, an unusual property of DNA polymerase epsilon, promotes epigenetic silencing in Saccharomyces cerevisiae. *J Biol Chem*, 281(43), 32898-908.

[53] Matos, J., Lipp, J. J., Bogdanova, A., Guillot, S., Okaz, E., Junqueira, M., Shevchenko, A., & Zachariae, W. (2008). Dbf4-dependent CDC7 kinase links DNA replication to the segregation of homologous chromosomes in meiosis I. *Cell*, 135(4), 662-78.

[54] Davey, M. J., Andrighetti, H. J., Ma, C. J., & Brandl, X. (2011). A synthetic human kinase can control cell cycle progression in budding yeast. *G3*, 1(4), 317-25.

[55] Malmström, L., Riffle, M., Strauss, C. E., Chivian, D., Davis, T. N., Bonneau, R., & Baker, D. (2007). Superfamily assignments for the yeast proteome through integration of structure prediction with the gene ontology. *PLoS Biol*, e76.

[56] Masai, H., & Arai, K. (2000). Dbf4 motifs: conserved motifs in activation subunits for Cdc7 kinases essential for S-phase. *Biochem Biophys Res Commun.*, 275(1), 228-32.

[57] Francis, L. I., Randell, J. C., Takara, T. J., Uchima, L., & Bell, S. P. (2009). Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev.*, 23(5), 643-54.

[58] Dowell, S. J., Romanowski, P., & Diffley, J. F. (1994). Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. *Science*, 265(5176), 1243-6.

[59] Duncker, B. P., Shimada, K., Tsai-Pflugfelder, M., Pasero, P., & Gasser, S. M. (2002). An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing. *Proc Natl Acad Sci*, USA, 99(25), 16087-92.

[60] Varrin, A. E., Prasad, A. A., Scholz, R. P., Ramer, M. D., & Duncker, B. P. (2005). A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. *Mol Cell Biol.*, 25(17), 7494-504.

[61] Sheu, Y.J. , & Stillman, B. (1999). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell 2006*, 24(1), 101-13.

[62] Zegerman, P. , & Diffley, J. F. (2010). Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature*, 467(7314), 474-8.

[63] Weinreich, M. , & Stillman, B. Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. *EMBO J.*, 18(19), 5334-46.

[64] Duch, A. , Palou, G., Jonsson, Z. O., Palou, R., Calvo, E. , Wohlschlegel, J. , & Quintana, D. G. (2011). A Dbf4 mutant contributes to bypassing the Rad53-mediated block of origins of replication in response to genotoxic stress. *J Biol Chem*, 286(4), 2486-91.

[65] Matthews, L. A., Jones, D. R., Prasad, A. A., Duncker, B. P., & Guarné, A. (2012). Saccharomyces cerevisiae Dbf4 has unique fold necessary for interaction with Rad53 kinase. *J Biol Chem*, 287(4), 2378-87.

[66] Wan, L., Niu, H., Futcher, B., Zhang, C., Shokat, K. M., Boulton, S. J., & Hollingsworth, N. M. (2008). Cdc28-Clb5 (CDK-S) and Cdc7-Dbf4 (DDK) collaborate to initiate meiotic recombination in yeast. *Genes Dev.*, 22(3), 386-97.

[67] Sasanuma, H., Hirota, K., Fukuda, T., Kakusho, N., Kugou, K., Kawasaki, Y., Shibata, T., Masai, H., & Ohta, K. (2008). Cdc7-dependent phosphorylation of Mer2 facilitates initiation of yeast meiotic recombination. *Genes Dev.*, 22(3), 398-410.

[68] Lo, H. C., Wan, L., Rosebrock, A., Futcher, B., & Hollingsworth, N. M. (2008). Cdc7-Dbf4 regulates NDT80 transcription as well as reductional segregation during budding yeast meiosis. *Mol Biol Cell*, 19(11), 4956-67.

[69] Matos, J., Lipp, J. J., Bogdanova, A., Guillot, S., Okaz, E., Junqueira, M., Shevchenko, A., & Zachariae, W. (2008). Dbf4-dependent CDC7 kinase links DNA replication to the segregation of homologous chromosomes in meiosis I. *Cell*, 135(4), 662-78.

[70] Marston, A. L. (2009). Meiosis: DDK is not just for replication. *Curr Biol*, 19(2), 74-6.

[71] Zavec, A. B., Lesnik, U., Komel, R., & Comino, A. (2004). The Saccharomyces cerevisiae gene ECM11 is a positive effector of meiosis. *FEMS Microbiol Lett*, 241(2), 193-9.

[72] Zavec, P. B., Comino, A., Watt, P., & Komel, R. (2000). Interaction trap experiment with CDC6. *Pflugers Arch*, 439(3), R 94-6.

[73] Zavec, A. B., Comino, A., Lenassi, M., & Komel, R. (2008). Ecm11 protein of yeast Saccharomyces cerevisiae is regulated by sumoylation during meiosis. *FEMS Yeast Res*, 8(1), 64-70.

IntechOpen