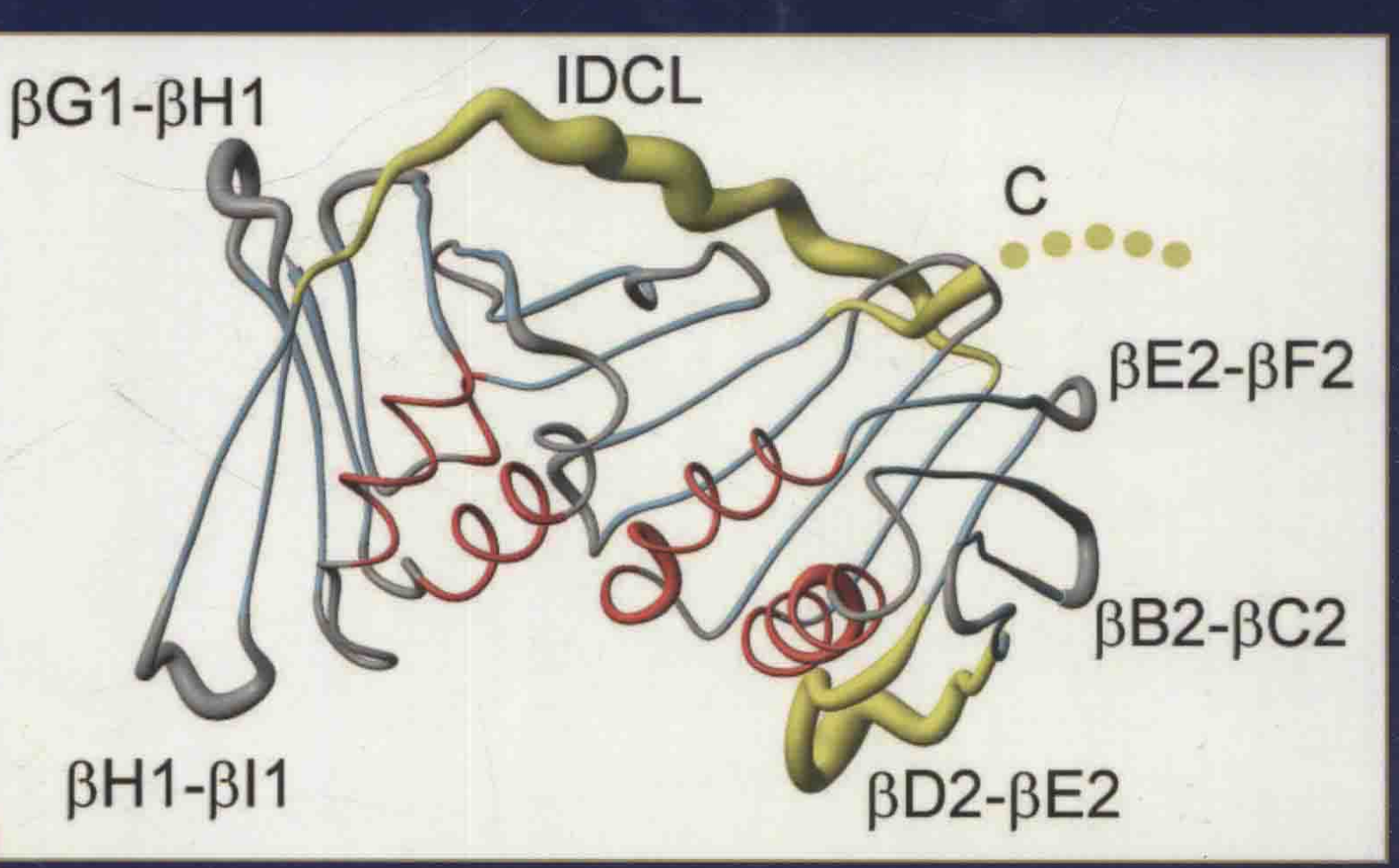
Advances in

PROTEIN CHEMISTRY and STRUCTURAL BIOLOGY

VOLUME 91

Protein-Nucleic Acids Interactions



Edited by
Rossen Donev

VOLUME NINETY ONE

ADVANCES IN PROTEIN CHEMISTRY AND STRUCTURAL BIOLOGY

Protein-Nucleic Acids Interactions

Edited by

ROSSEN DONEV

Institute of Life Sciences, College of Medicine
Swansea University, United Kingdom

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK
32 Jamestown Road, London NW1 7BY, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
225 Wyman Street, Waltham, MA 02451, USA
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA

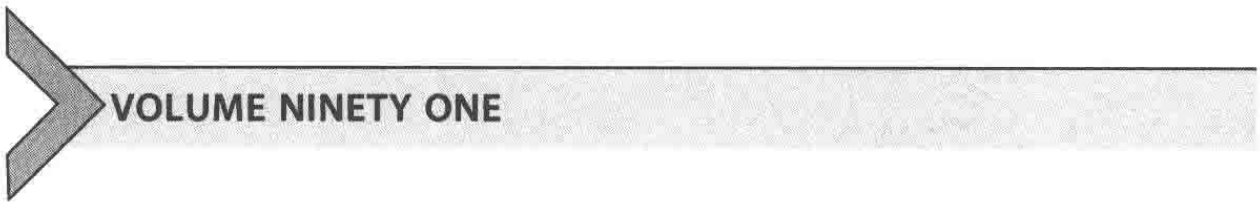
First edition 2013

ISBN: 978-0-12-411637-5
ISSN: 1876-1623

For information on all Academic Press publications visit our website at store.elsevier.com

Printed and bound in USA

13 14 15 16 10 9 8 7 6 5 4 3 2 1



ADVANCES IN PROTEIN CHEMISTRY AND STRUCTURAL BIOLOGY

Protein-Nucleic Acids Interactions

PREFACE

At the postgenomic era, when the primary structure of our genome has been determined, a significant challenge for investigators is now to determine not only the role of polymorphisms appearing in gene coding sequence for the function of the corresponding proteins but also the protein–nucleic acid interactions that regulate numerous vital processes in our cells such as gene expression, replication, cell proliferation and apoptosis, RNA stability, etc. Therefore, in this volume of the *Advances in Protein Chemistry and Structural Biology* dedicated to protein–nucleic acid interactions, we discuss some of these issues.

The first chapter in this book is focused on the current knowledge on PCNA interactions, the DNA sliding clamp found in eukaryotes and archaebacteria, from a structural point of view and highlights the questions that remain still open for these interactions. The second chapter brings insights into aggregation, nuclear location, and nucleic acid interaction as common features shared by a number of proteins related to neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, transmissible spongiform encephalopathy, Huntington's disease, etc. This chapter reviews those common features suggesting that neurological diseases may share also a transcriptional disorder as an important contribution to the origin of the disease.

In the third chapter, the applicability of several computational approaches for understanding protein–DNA interactions is discussed from different perspectives—development of databases for the protein–DNA interactions and binding specificity of the complexes, structural analysis of protein–DNA complexes, discriminating DNA-binding proteins from amino acid sequence, prediction of DNA-binding sites and protein–DNA binding specificity using sequence and/or structural information, and understanding the recognition mechanism of protein–DNA complexes. Furthermore, the available online resources for understanding protein–DNA interactions are collectively listed, which can be of a great benefit to researchers working in this field.

Last two chapters summarize the current advances of the chromatin immunoprecipitation (ChIP)-based approaches to decipher gene regulatory and epigenetic network in the cells. The limitations of the method are discussed and the future ChIP-based developments are explored with special emphasis on ChIP-Seq which is presented in details in the fifth chapter.

Authors demonstrate here how sequence analysis of ChIP-Seq derives novel biological knowledge on multiple levels, from individual transcription factor binding sites to genome segments operating as regulatory modules. An overview of existing software in the field is also provided.

It is my sincere hope that this overview of some widely used methods for studying protein–nucleic acid interactions and the involvement of some dysfunctional interactions in a number of diseases would inspire future translational studies focusing on targeting protein–nucleic acid regulatory mechanisms as a strategy for therapies in various diseases/disorders.

Rossen Donev
Institute of Life Sciences, College of Medicine
Swansea University, United Kingdom

CONTENTS

Preface *vii*

1. Proliferating Cell Nuclear Antigen Structure and Interactions: Too Many Partners for One Dancer? **1**

Alfredo De Biasio and Francisco J. Blanco

1. DNA Sliding Clamps 2
2. PCNA Structure, Stability, and Dynamics 5
3. Interactions of PCNA with Other Proteins 9
4. Crystal Structure of FEN1 Bound to PCNA and Observation of the Interaction in Solution by NMR 16
5. Interactions of PCNA with PIP-Box Peptides 18
6. Binding of MCL-1, Gadd45α, and CDK2 Proteins to PCNA is Not Direct or Extremely Weak in Solution 21
7. Revisiting the PIP-Box Pattern Definition 24
8. More Caution Should Be Exerted When Characterizing Direct Protein Interactions with PCNA 25
9. The Problem of Binding Competition Among PCNA Partners 27
10. PCNA and Disease and Its Potential as a Pharmacological Target 29

Acknowledgments 30

References 30

2. Anomalous Protein–DNA Interactions Behind Neurological Disorders **37**

Sergio Camero, María J. Benítez, and Juan S. Jiménez

1. Introduction 38
2. Proteins Involved in DNA Packaging 38
3. Proteins Interacting with Nucleic Acids in Transcription 39
4. Protein–DNA Interactions in Cell Division 41
5. Polyglutamine Diseases 42
6. Aggregation 46
7. Nuclear Location 48
8. DNA Interaction 49

Acknowledgment 52

References 52

3. Computational Approaches for Predicting the Binding Sites and Understanding the Recognition Mechanism of Protein–DNA Complexes **65**

M. Michael Gromiha and R. Nagarajan

1. Introduction 66
2. Databases for Protein–DNA Complexes/Interactions 68
3. Structural Analysis of Protein–DNA Complexes 73
4. Discrimination of DNA-Binding Proteins 75
5. Prediction of DNA-Binding Sites and Protein–DNA Binding Specificity 76
6. Recognition Mechanism of Protein–DNA Complexes 81
7. Conclusions 93

Acknowledgments 94
References 94

4. Detecting DNA–Protein Interactions in Living Cells—ChIP Approach **101**

Rossitza Christova

1. Introduction 102
2. Chromatin Immunoprecipitation 103
3. ChIP-Based High-Throughput Technologies 111
4. Integration of ChIP-seq Profiles with Other DNA and Protein-Related Approaches 120
5. Genome-Wide View of Transcriptional Regulation 124
6. Summary and Future Directions 127

Acknowledgments 128
References 129

5. DNA Sequence Motif: A Jack of All Trades for ChIP-Seq Data **135**

Ivan V. Kulakovskiy and Vsevolod J. Makeev

1. Introduction 136
2. ChIP-Seq Peaks: What Do They Actually Show? 145
3. *De Novo* Motif Discovery in ChIP-Seq Data 151
4. Further Biological Questions to Be Addressed by Sequence Analysis of ChIP-Seq Data 161
5. ChIP-Seq Software Survey 162
6. Final Remarks 165

Acknowledgments 165
References 166

Author Index *173*
Subject Index *195*

CHAPTER ONE

Proliferating Cell Nuclear Antigen Structure and Interactions: Too Many Partners for One Dancer?

Alfredo De Biasio*, Francisco J. Blanco*,†,1

*Structural Biology Unit, CIC bioGUNE, Derio, Spain
†IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[1]Corresponding author: e-mail address: fblanco@cicbiogune.es

Contents

1. DNA Sliding Clamps 2
2. PCNA Structure, Stability, and Dynamics 5
3. Interactions of PCNA with Other Proteins 9
4. Crystal Structure of FEN1 Bound to PCNA and Observation of the Interaction in Solution by NMR 16
5. Interactions of PCNA with PIP-Box Peptides 18
6. Binding of MCL-1, Gadd45α, and CDK2 Proteins to PCNA Is Not Direct or Extremely Weak in Solution 21
   - 6.1 MCL-1/PCNA interaction 21
   - 6.2 Gadd45α/PCNA interaction 22
   - 6.3 CDK2/PCNA interaction 23
7. Revisiting the PIP-Box Pattern Definition 24
8. More Caution Should Be Exerted When Characterizing Direct Protein Interactions with PCNA 25
9. The Problem of Binding Competition Among PCNA Partners 27
10. PCNA and Disease and Its Potential as a Pharmacological Target 29

Acknowledgments 30
References 30

Abstract

PCNA is the DNA sliding clamp found in eukaryotes and archaebacteria. Sliding clamps were first described as processivity factors in DNA replication. They consist of multimeric, toroidal-shaped structures with pseudo-sixfold symmetry that encircle the DNA duplex and tether the replicative polymerases to the genomic template. Later, it was found that PCNA serves as a docking platform where other proteins dock to carry out different DNA metabolic processes. The structure of the bacterial clamp bound to a short primed DNA shows a tilted duplex in the central channel, which is lined by α-helices with net positive charges. Many of the proteins reported to interact with PCNA do so via the PCNA Interaction Protein sequence (PIP-box). The structures of several

Advances in Protein Chemistry and Structural Biology, Volume 91
ISSN 1876-1623
http://dx.doi.org/10.1016/B978-0-12-411637-5.00001-9

proteins and peptides bound to PCNA show a common binding mode, but it is still unknown how the many different partners compete for binding and exert their enzymatic and regulatory functions. Furthermore, the literature contains many reports on proteins that directly bind to PCNA as detected by different methods, but only few of the putative complexes have been examined in detail by quantitative analytical techniques or high-resolution structural methods. Some of the reported interactions are not observed in solution using the pure proteins, indicating that the direct interaction is nonexistent or very weak and is likely mediated by other factors. We review here the current knowledge on PCNA interactions from a structural point of view, with a focus on human proteins and highlighting the questions that remain to be answered.

1. DNA SLIDING CLAMPS

DNA sliding clamps are multimeric proteins with toroidal-shaped structures that encircle the DNA duplex and slide bidirectionally along it (Kuriyan & O'Donnell, 1993). They are central components of the replisome, the chromosomal–DNA replication machinery (Pomerantz & O'Donnell, 2007), and act as replication processivity factors by tethering the polymerases to the genomic template (Moldovan, Pfander, & Jentsch, 2007). Sliding clamps are loaded onto DNA by specialized protein complexes known as clamp loaders (Kelch, Makino, O'Donnell, & Kuriyan, 2012). The three domains of life, and even viruses, have functionally and structurally related sliding clamps (Indiani & O'Donnell, 2006). Protein clamps may be homodimeric, homotrimeric, or heterotrimeric assemblies, with the protomers binding each other in a head-to-tail fashion forming a ring. Despite limited sequence homology between the protomers of different organisms, all sliding clamps consist of six structural domains that form a characteristic fold: an outer layer of 6 β-sheets and an inner layer of 12 α-helices lining the central channel (Jeruzalmi, O'Donnell, & Kuriyan, 2002). The bacterial clamp is formed by the homodimeric association of two protomers, each one with three topologically similar domains. In contrast, the archaeal (Matsumiya, Ishino, & Morikawa, 2001) and eukaryotic clamps assemble into homo- or heterotrimeric rings in which each protomer contains two similar domains and a long interdomain-connecting loop (IDCL). As examples of clamp structures, the *Escherichia coli* DNA polymerase III β subunit (Kong, Onrust, O'Donnell, & Kuriyan, 1992) and the Homo sapiens PCNA (hPCNA) ring (Kontopidis et al., 2005) are illustrated in Fig. 1.1.

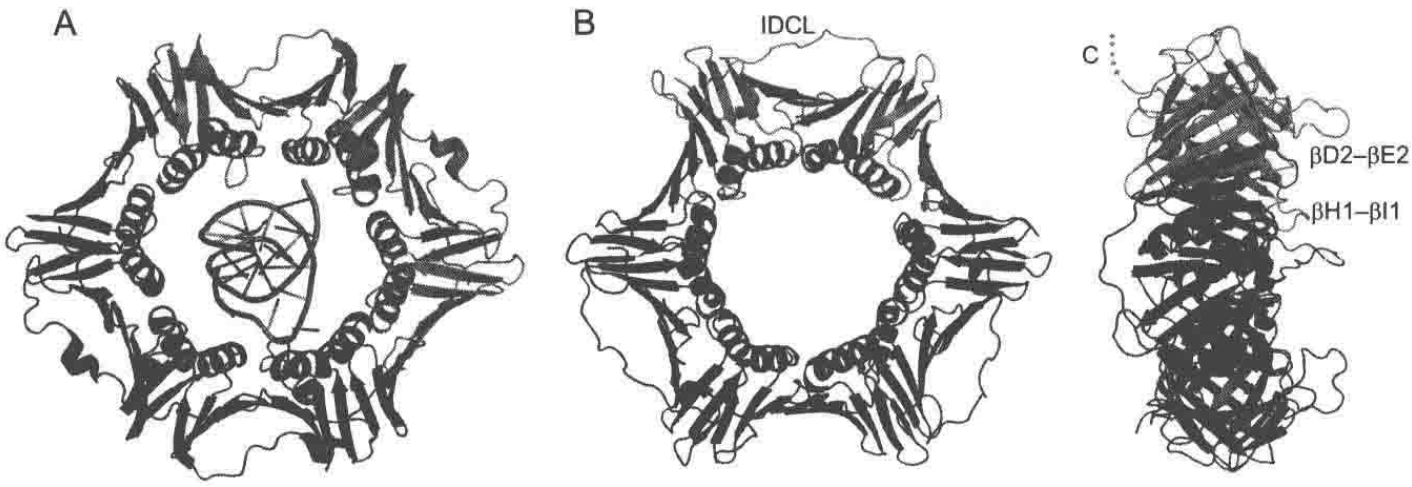


Figure 1.1 Ribbon representations of the crystallographic structures of bacterial and eukaryotic DNA sliding clamps. (A) The *E. coli* DNA polymerase III β subunit prokaryotic β-clamp with a primed DNA in the central channel (3BEP). The front side is shown and the two protein protomers are colored in green and red. (B) Two views of the human PCNA trimeric ring (1VYM) are shown with the three protomers colored in green, red, and blue. The view on the left shows the front side and the view on the right is the previous one rotated 90° along a vertical axis. The IDCL (residues 117–134), the βD2–βE2 loop (184–195), the βH1–βI1 (residues 105–109), and the C-terminus are labeled on the protomer colored in green using the original nomenclature. This indicates with a Greek letter the type of secondary structure that is connected by the loop, with a capital Latin letter the order of those elements along the PCNA sequence and with a number the corresponding pseudodomain of the PCNA protomer. The five C-terminal residues (residues 256–261) not seen in the crystal structure are indicated by circles in the green protomer. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this chapter.)

The clamp ring has two distinct faces: the front side, containing the hydrophobic pocket where polymerases and other proteins bind, and the back side. The front side is also known as the C-side after the localization of the C-termini of the protomers. Although the charge distribution in the outer surface of the ring differs among the clamps, the inner surface is always rich in positively charged amino acids generating a positive electrostatic potential and thus favoring the passing of DNA. The ring has an internal diameter of approximately 35 Å, which is much larger than the approximately 24 Å lateral extent of a double-stranded DNA (dsDNA) helix in a canonical B-form (Wing et al., 1980). Molecular dynamics simulations of a DNA duplex within the channel of human PCNA suggested that the helical axis would be tilted with respect to the plane of the ring (Ivanov, Chapados, McCammon, & Tainer, 2006). The only available high-resolution structure of a sliding clamp bound to DNA is the crystal structure of the *E. coli* DNA polymerase III β subunit bound to a 10-bp long dsDNA with one four-base 5′-overhang of single-stranded DNA (ssDNA) on one end (Georgescu et al., 2008). This DNA molecule represents a DNA strand

primed for being copied by a polymerase. In the crystal, the dsDNA portion was found within the ring and tilted, with the dsDNA helical axis and the clamp C2 rotation axis making an angle of 22°. This tilting may be caused by interactions of the DNA with amino acids on the front side of the ring but may be also influenced by the interaction of the ssDNA with another symmetry related ring in the crystal lattice. In fact, the primed DNA was found with the ssDNA overhang protruding through the back side, which is the reverse orientation relative to the physiological orientation, likely an artifact of crystallization. The authors show evidence that in solution the interaction is intramolecular, and that the binding site of ssDNA on PCNA is the same where the polymerase binds. These results indicate that in solution the orientation is the physiological one and that the ssDNA binding to the protein binding pocket may serve as a "placeholder" keeping the clamp at the primed site after loading instead of sliding off. Because the polymerase contains two regions that bind the clamp (Lopez de Saro, Georgescu, Goodman, & O'Donnell, 2003), displacing the ssDNA from its binding site allows the complex to slide and diffuse along the DNA.

A crystal of the same primed DNA bound to a single-chain chimera of Saccharomyces cerevisiae PCNA (scPCNA) did not provide a high-quality structure of the DNA moiety because of partial disorder or low occupancy (McNally, Bowman, Goedken, O'Donnell, & Kuriyan, 2010). The ssDNA overhang was not observed, but the model derived from the visible electron density data indicates that the dsDNA is tilted within the central channel optimizing the contacts between the negatively charged DNA and the positively charged inner side of the ring.

It has been progressively realized that sliding clamps are involved in DNA metabolism and in other DNA template-based processes besides chromosomal replication. These processes include sister chromatid cohesion, DNA repair, chromatin assembly and remodeling, and cell cycle control and survival. Several reviews cover the mode of action of sliding clamps in these processes and other aspects of sliding clamp function (Bloom, 2009; Green, 2006; Lopez de Saro, 2009; Moldovan et al., 2007), although they have not been studied up to the same level of detail as DNA replication has.

Because of their central activity around DNA-related processes, sliding clamps interact with many other proteins besides polymerases. In particular, a large number of proteins have been reported to interact with the eukaryotic PCNAs, the most studied sliding clamps (Maga & Hubscher, 2003), posing the question on how do they access PCNA in a coordinated manner and how does PCNA organize their cellular functions in space and time (Warbrick, 2000). Although structural and biochemical studies provide

some answers to these questions, the PCNA interactors are so numerous that it looks an impossible task to experimentally unravel all the implications of such a diverse set of binders. In this review, we focus on the structural properties of human PCNA, its interactions with other proteins for which there is some experimental support, and how robust is the evidence for the interaction with some of its partners in light of recent studies questioning the previously reported binding events. In the text hereafter, PCNA designates the human PCNA, unless specified otherwise.

2. PCNA STRUCTURE, STABILITY, AND DYNAMICS

Human PCNA is predominantly trimeric in solution at concentrations above 50 nM (Yao et al., 1996) and is stable over a 25–45 °C temperature range (Schurtenberger et al., 1998). This is why PCNA rings need to be opened, at the expense of ATP hydrolysis, by the clamp loader replication factor C to assemble the ring onto DNA (Tsurimoto & Stillman, 1991).

The crystal structure of the PCNA ring shows the protomers arranged in a head-to-tail fashion forming a ring with the front side containing the three IDCLs linking the two domains of each protomer, and the back side with prominent loops that protrude into the solvent (Fig. 1.1B). There are two available PCNA structures (PDB entries 1VYM and 1W60) corresponding to different crystallographic spatial groups (monoclinic C2 and trigonal P3, respectively), with a total of five crystallographically independent PCNA protomer structures (Kontopidis et al., 2005). Although the two structures are essentially the same, the C^{α} atom RMSD between the five protomer structures ranges from 0.7 to 2 Å, and the overall fit of the trimer rings shows a bucking of the trimer in the C2 crystal compared with the planar ring in the P3 one. These results suggest a conformational plasticity of the PCNA molecule. The three-dimensional structure of PCNA in solution has not been determined, but its NMR spectrum has been assigned (Sanchez, Torres, Prieto, Blanco, & Campos-Olivas, 2007), and the structural information derived from the measured backbone chemical shifts and long-range NOEs is consistent with the crystal structure, indicating that the overall structure in solution is the same (De Biasio et al., 2011).

It has been reported that PCNA can form double trimers inside the cells, as detected by cross-linking experiments in intact cells and cell extracts (Naryzhny, Zhao, & Lee, 2005) and by size-exclusion chromatography of recombinant pure PCNA (Naryzhny, Desouza, Siu, & Lee, 2006). Mutational analysis indicated that the double trimer is formed by a back-to-back

association of two trimers. Although this kind of association still allows the front sides to bind the polymerases and/or other proteins, it will reduce the tilting of the DNA duplex and the optimization of its contacts with the positive-charged residues in the inner surface of the ring, in case that the model inferred from the bacterial clamp applies to PCNA (Ivanov et al., 2006). The double trimer will also be incompatible with the proposed mechanism of PCNA loading, which imposes a defined orientation of PCNA on the primed DNA (Kazmirski, Zhao, Bowman, O'Donnell, & Kuriyan, 2005; Miyata et al., 2005).

In our hands, recombinantly purified PCNA behaves as one single trimer, as observed by analytical ultracentrifugation and confirmed by size-exclusion chromatography coupled to multiangle laser light scattering, which is very sensitive to large aggregates because of the nonlinear dependence of scattering with the particle size. These measurements showed that the trimeric form is the predominant one even at concentrations of 3 g/L (on a monomer basis), with less than 1% of species with the molecular weight of a dimer of trimers (De Biasio et al., 2011). This is consistent with the crystal structures and NMR analyses, done at much higher concentrations, which do not show any indication of ring association. Although PCNA may exist transiently as a loosely bound double trimer inside the cells, its functional relevance is unclear.

The thermal denaturation of PCNA followed by far-UV circular dichroism does not show a cooperative transition, possibly due to simultaneous secondary, tertiary, and quaternary structure rearrangements, while the chemical denaturation by urea or GuHCl show two sigmoidal transitions (De Biasio et al., 2011). The urea-induced chemical denaturation shows a biphasic behavior consistent with a partially unfolded intermediate state which is significantly populated in the transition zone (Fig. 1.2A). The first transition is dependent on protein concentration while the second is not, suggesting that PCNA unfolding is coupled to the dissociation of the ring and that the intermediate is dissociated. Assuming that the formation of the initial partially unfolded-dissociated species and its subsequent unfolding are two independent processes, the curve could be split into two parts and analyzed separately. With this assumption, the equilibrium in the range of 0–2.5 M urea could be fitted to a two-state model of trimer dissociation into monomeric intermediates. The second transition (between 2.5 and 8.5 M urea) could be fitted to a two-state unimolecular reaction of intermediate unfolding. In comparison, the scPCNA is more resistant to urea denaturation, showing a single transition starting at urea concentrations where the first transition of human PCNA is complete.

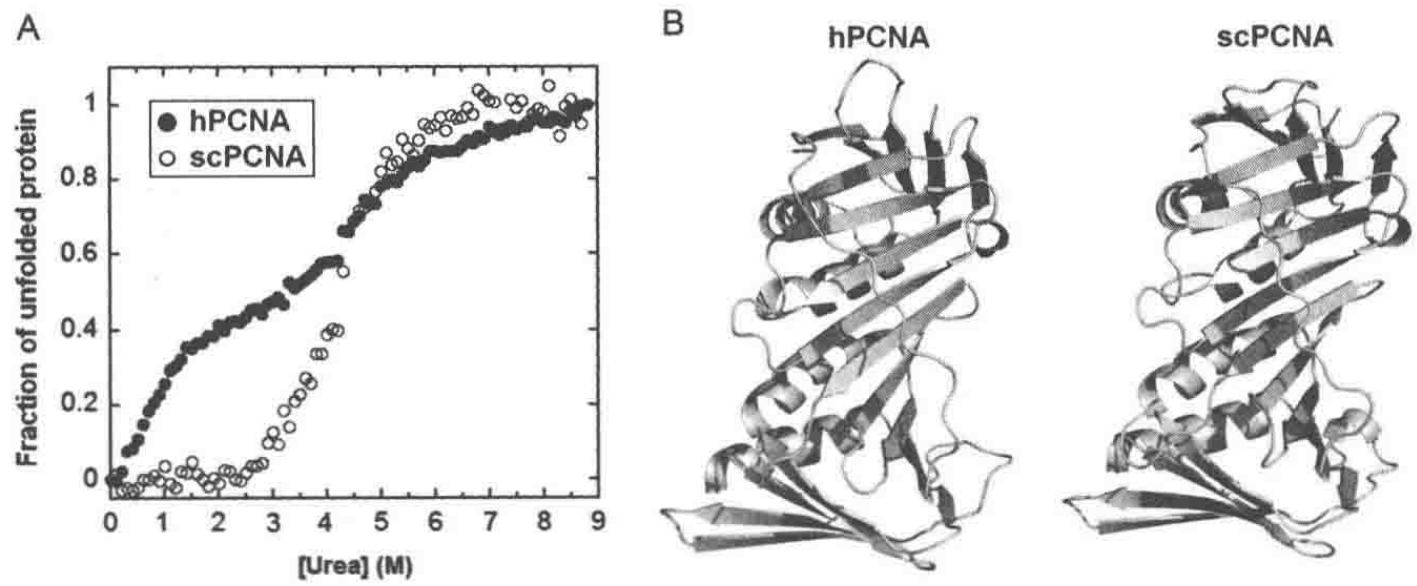


Figure 1.2 Stability and dynamics of human and yeast PCNA rings. (A) Urea-induced denaturation of hPCNA (solid circles) and scPCNA (open circles) at 1 μM and 35 °C. The fraction of unfolded protein was computed from the normalized change in circular dichroism signal at 222 nm. (B) Faster backbone dynamics in hPCNA relative to scPCNA based on H/D amide exchange NMR data. Ribbon diagrams represent the hPCNA (left) or scPCNA (right) protomer. Residues of scPCNA whose NMR signal persisted 1 h after D_2O exchange and for which an exchange constant could be measured and residues of hPCNA whose signal was still detected immediately after D_2O exchange are highlighted in bright red over the pale violet color of the corresponding protomer. Residues colored in wheat represent either prolines, or residues whose assignment in the NMR experiment was ambiguous or (in the case of scPCNA) residues for which an exchange constant could not be measured. These figures were adapted from previously published results (De Biasio et al., 2011). (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this chapter.)

The human and yeast PCNA homologs share 35% of sequence identity, and their crystal structures are highly superimposable over the elements of secondary structure, with an RMSD of 1.0 Å for the C^{α} atoms (Krishna et al., 1994). Differences exist regarding the overall content of different types of amino acids: the yeast protein contains larger and smaller percentages of charged and uncharged polar residues, respectively, as compared to the human homolog. The same trend is observed between the thermophilic *Pyrococcus furiosus* homolog and the yeast homolog, in agreement with the general trend observed for thermophilic proteins (Zhou, Wang, Pan, & Li, 2008).

The internal dynamics of hPCNA has been investigated by the observation of local and global unfolding events under native conditions through NMR measurements of the backbone amide proton exchange with the solvent deuterons (De Biasio et al., 2011). These measurements provide a mean for observing slow main-chain motion events, as the exchange in folded polypeptide chains may occur in timescales from milliseconds to years.

While most of the backbone amide resonances of hPCNA are visible in the NMR spectrum at 35 °C in protonated aqueous buffer, only 15 out of 253 resonances could be seen after 1 h in deuterated buffer, and all of them exchanged beyond detection after a few hours. In contrast, for scPCNA at least 90 amide protons were detected after 1 h, and some of them remained visible after several days (Fig. 1.2B).

The internal dynamics of the hPCNA backbone on faster timescales has been characterized by ^{15}N NMR relaxation measurements (De Biasio et al., 2012). As illustrated in Figure 1.3, there are several regions in the polypeptide chain of hPCNA with increased flexibility with respect to the α-helices and β-strands, which are generally very rigid. These regions are mainly located in the exposed loops connecting secondary structure elements, and particularly in the IDCL and the βD2–βE2 loop, which are the longest loops in the protein. The last eight residues in the C-terminal tail present the highest flexibility, consistent with most of them not appearing in the electron density map of the crystallographic structures. The relative high flexibility observed in the loops is in general agreement with the high crystallographic B-factors for residues in the corresponding regions of the X-ray structures of free hPCNA (Kontopidis et al., 2005). However, there is a large variability in the B-factors reported for the two available crystal

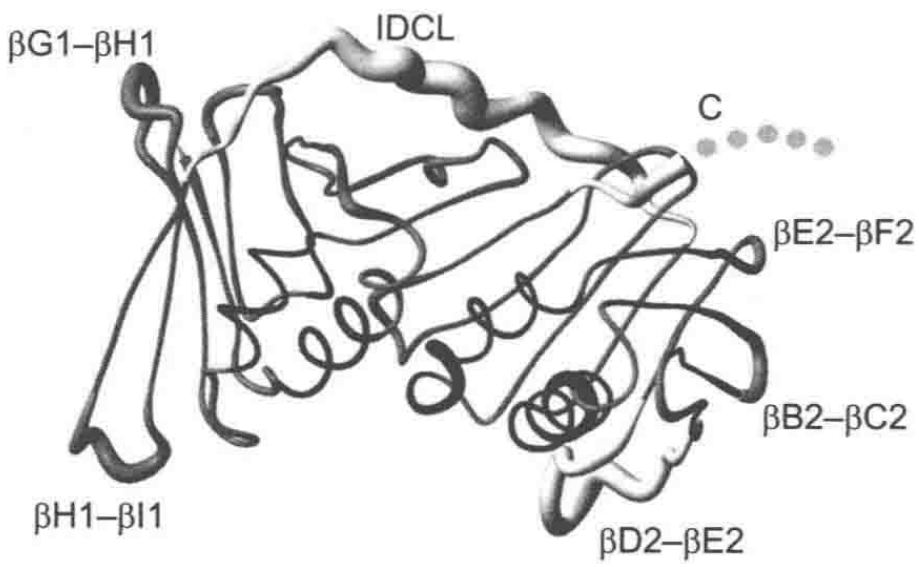


Figure 1.3 Fast backbone dynamics of human PCNA derived from ^{15}N NMR relaxation measurements. Representation of the backbone structure of one protomer of the hPCNA ring as a coil whose thickness illustrates the relative dynamics of the corresponding backbone NH bond. Helices and strands are colored in red and blue, respectively, and the three most flexible regions are colored in yellow. The loops with high relative disorder are labeled. *Adapted from previously published results (De Biasio et al., 2012).* (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this chapter.)

structures. In the monoclinic form, both the IDCL and the βD2–βE2 loop display very large B-factors as compared with the average, while in the trigonal form only the βD2–βE2 loop shows relative high values. This discrepancy is very likely due to the larger number of crystal contacts involving the IDCL residues in the latter crystal structure than in the former one.

The differences in the two crystal structures of hPCNA, the fast backbone amide proton exchange with the solvent and the susceptibility to chemical denaturation point to a conformational plasticity of the hPCNA structure that probably plays a role in its function as docking platform for a large number of proteins. The more complex interactome regulating DNA replication, repair, and cell cycle control may require a higher tolerance by hPCNA to accommodate the binding of a larger variety of ligands than the less dynamic scPCNA.

3. INTERACTIONS OF PCNA WITH OTHER PROTEINS

In addition to the replicative function, PCNA directs other important cellular processes through the interaction with a host of DNA-processing proteins and cell cycle regulators (Warbrick, 2000). Many of the proteins that interact with PCNA contain a conserved sequence known as PIP-box (PCNA-Interacting Protein-box). The pattern of the PIP-box sequence is *QXXhXXaa*, where *h* is an aliphatic hydrophobic residue, *a* is an aromatic hydrophobic one (F, W, or Y), and *X* is any of the 20 proteinogenic amino acids (Warbrick, Heatherington, Lane, & Glover, 1998). However, the current scientific literature reports many other proteins that interact with PCNA and that do not have a PIP-box in their sequences.

We have examined the PCNA–protein interactions listed in different databases that search in the scientific literature for protein–protein interactions with some sort of experimental support. Among them, the Human Protein Reference Database (HPRD) was shown to be the most complete and less redundant (Mathivanan et al., 2006). Therefore, Table 1.1 was compiled using the HPRD. A literature search was done for each protein annotated in the HPRD as a PCNA interactor in order to confirm the direct interaction and to know the sort of experimental evidence supporting it.

A total of 52 human proteins can be counted by this criterion to directly interact with PCNA. However, it is well known that not all the experimental techniques to detect an interaction are equally reliable, with some of them having a higher probability than others of yielding false-positives. In our