Title
Protein–nucleic acid interactions: unlocking mysteries old and new

Journal
Current Opinion in Structural Biology, 15(1)

ISSN
0959-440X

Authors
Aggarwal, Aneel K
Doudna, Jennifer A

Publication Date
2005-02-01

DOI
10.1016/j.sbi.2005.01.017

Published in final edited form as:

Curr Opin Struct Biol. 2005 February ; 15(1): 65–67. doi:10.1016/j.sbi.2005.01.017.

Protein-nucleic acid interactions: unlocking mysteries old and new

Aneel K Aggarwal,

Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, NY 10029, USA

Jennifer A Doudna

Department of Molecular and Cell Biology, Department of Chemistry, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA

"Is there no end to the versatility of RNA?" asks Cech in a recent review. The versatility of RNA is encapsulated in several articles in this section of *Current Opinion of Structural Biology* — running the gamut of cellular processes from mRNA capping and translational regulation to the intricacies of RNA interference. DNA and the proteins that work on it have their own bag of tricks, typified by reviews on the initiation of replication, the synthesis of damaged DNA, SF3 helicases and non-specific nucleases. Together, these reviews capture the excitement and pace of structural biology: embodying themes of structural and chemical specificity, and the coupling between recognition and catalysis.

Cunningham and Berger, in the first review, highlight recent insights into the initiation and control of DNA replication. In prokaryotes, chromosome replication typically begins at a single origin site through binding of the DnaA initiator protein to several repetitive, non-palindromic sequences called DnaA boxes, leading to unwinding of a flanking AT-rich region to enable assembly of the replisome complex. Comparison of archaeal DNA replication origins suggests the presence of multiple origins within a single chromosome; this is more similar to eukaryotic replication and perhaps reflects the closer structural relationship of archaeal replication proteins to those of eukaryotes, rather than to their bacterial counterparts. Structures of the bacterial RepA dimer and the origin repressor protein SeqA bound to its DNA target site reveal two different mechanisms by which cells control replication initiation — sequestering the initiator protein or blocking access to origin sites, respectively. The roles of ATP binding and hydrolysis in altering initiator structure and function are beginning to be elucidated, suggesting interesting parallels in replication machinery across phylogeny.

Hickman and Dyda review the recent wave of SF3 viral helicase structures — all determined within the past two years. Although much is known about the biological activities of SF3 helicases, which are encoded by small DNA and RNA viruses, an assessment of their DNA

aggarwal@inka.mssm.edu .

The Aggarwal group studies interactions with nucleic acids mediated by enzymes, and transcriptional and translational regulators. The Doudna laboratory studies the molecular structure and function of RNA, focusing on ribozymes and large RNAs involved in protein synthesis initiation and protein trafficking.

unwinding mechanism in viral replication has been hindered by the lack of structural data. This has changed with the first molecular views of the helicase regions of the large T antigen (LTag) of simian virus 40 (SV40), the Rep protein of adenoassociated virus 2 (AAV2) and the E1 protein of human papillomavirus serotype 18 (HPV18). Hickman and Dyda describe the basic architecture of these SF3 helicases: a bi-modular arrangement, wherein an α/β domain (as occurs in AAA+ proteins) is preceded by a small helical domain. The authors note the subtle variations on this structural framework in accord with the accessory roles of AAV Rep and HPV18 E1 in viral packaging and interaction with the 'helicase loader' E2 protein, respectively. How do these SF3 helicases distort and/or unwind DNA? This is the million-dollar question; the most compelling insights emerge from the structure of SV40 LTag, which is the only one of the three proteins to crystallize as a hexamer. SV40 LTag hexamer is shaped like a double-tier 'wedding cake', with a central channel through which DNA can pass. On cue, structures of SV40 LTag hexamers in the presence of ATP and ADP were reported recently, revealing the basis of the interdomain motion for melting DNA. SF3 helicases can at long last join their more pervasive SF1 and SF2 helicase cousins for a family portrait.

Hogg, Wallace and Doublié take us on a bumpy ride on the road to DNA synthesis. The past year was particularly bounteous in revealing how replicative DNA polymerases are stopped in their tracks by some lesions and yet can zoom past oxidatively damaged DNA. The most prevalent form of oxidative DNA damage is 8-oxoguanine (8oxoG), which is particularly noxious because of its ability of pair with either C or A — leading to mutagenic G•C to T•A transversions. The authors review recently determined crystal structures of several high-fidelity polymerases (belonging to the A, B and X families) in the act of replicating 8oxoG. Amongst these, structures of *Bacillus stearothermophilus* DNA polymerase I are particularly poignant, revealing not only the basis of the 8oxoG•C and 8oxoG•A pairings but also a reversal of genetic logic, such that mutagenic 8oxoG•A mimics a cognate base pair and the correct 8oxoG•C pair registers as a mismatch. Together, these and other structures (with T7 DNA polymerase, for example) provide an elegant basis for why mismatched 8oxoG•A is more readily extended than the correct 8oxoG•C. The authors also describe new structural insights into the blockage of DNA synthesis by UV-induced *cis-syn* cyclobutane pyrimidine dimers (CPDs) and hydrolytically generated abasic sites — derived from the recent structures of T7 and RB69 DNA polymerases in complex with such 'bumps'. If next year is anything like the past year — put on your seatbelt.

In his review, Brown addresses a two-decade-old mystery, namely the molecular mechanism by which the eukaryotic transcription factor TFIIA recognizes RNA. TFIIIA is the epitome of the classical Cys_2His_2 (C2H2) zinc fingers that have tantalized over the years: fingers that wrap around the DNA major groove, fingers that act as straight spacer elements and fingers that mediate protein–protein interactions. The modularity of TFIIIA-like zinc fingers has also made them the motif *par excellence* for the engineering of new specificities. Most puzzling of all, TFIIIA interacts not only with DNA but also with RNA, or as Brown puts it "What kind of structural arrangement could allow zinc fingers to bind to both?". This long-standing puzzle is resolved by the recent crystal structure of the 'central' portion of a TFIIIA–5S rRNA complex. Brown spotlights this remarkable structure and the mechanism by which the zinc fingers shun the deep and narrow major groove of the 5S rRNA double

helix and, instead, latch onto the sugar–phosphate backbone and/or onto bases exposed on the surface of loops. Surprisingly, the juxtaposition of two of the three fingers is the same as when TFIIIA binds DNA — a case of "one structure fits all". Brown also highlights the recent NMR structure of the translation regulator TIS11d bound to a single-stranded AU-rich RNA element. TIS11d is composed of tandem CCCH zinc fingers, which, as the structure reveals, are folded differently to TFIIIA-like fingers. A common feature of the TIS11d–AU and TFIIIA–5S RNA complexes, one that may apply to other zinc finger–RNA complexes, is the hydrophobic stacking of aromatic residues with accessible RNA bases. There are few motifs as endearing to structural biologists as zinc fingers as they continue to surprise and to notch-up new functions, their dexterity well tuned to the complexities of both DNA and RNA structure.

The life and death of an mRNA are largely determined by the controlled addition and removal of its 5’ cap structure, a distinctive 5’ N7-methyl guanosine linked through an inverted 5’-5’ triphosphate connector to the first nucleotide of the transcript. Gu and Lima discuss recent structural studies that illuminate the elaborate suite of activities required to attach the 5’ cap to the ends of mRNAs and to couple this activity to that of RNA polymerase II, the enzyme responsible for mRNA biosynthesis. Complementing previous studies of the first two enzymatic steps of nascent pre-mRNA capping, the RNA methyltransferase responsible for adding a methyl group to the N7 of guanine in the third and final step of the capping process has now been investigated in exciting structural detail. Structural insights into co-transcriptional mRNA capping come from a yeast guanylyltransferase, the second enzyme in the capping pathway, complexed with a phosphopeptide derived from the RNA polymerase II C-terminal domain. At the other end of the mRNA life-span, decapping enzymes hydrolyze the cap as part of two separate pathways. The Dcp2 hydrolase, in complex with the stimulatory factor Dcp1, removes the cap during 5’-3’ mRNA decay. A recent structure of the Dcp1 protein reveals probable interaction sites for Dcp2, but leaves open the question of how stimulation of Dcp2 activity is achieved. During 3’-5’ mRNA decay, the enzyme DcpS catalyzes cap hydrolysis and may also eliminate 7-methyl guanosine intermediates that might otherwise inhibit mRNA translation, export and processing. Recent structures of DcpS in different ligand-bound states show coordination between the two active sites of the homodimer. With these capping and decapping players in place, the stage is now set for future investigations of the cooperation between mRNA 5’ end processing and the transcriptional machinery, an exciting prospect.

The discovery of RNA interference (RNAi) and the explosion of interest in its role in gene regulation have focused increasing attention on the mechanisms of enzymes involved in the RNAi pathway. Lingel and Sattler review structural studies of PAZ, Piwi and RNase III domains derived from proteins central to RNAi in different organisms. Together, these structures highlight the ingenious ways that proteins specifically recognize the hallmark 5’ monophosphate and dinucleotide overhangs of the short interfering RNAs (siRNAs) and micro RNAs (miRNAs) produced by enzymes related to RNase III during induction of the RNAi pathway. Future work will focus on determining the molecular basis of different biochemical activities associated with RNAi-related enzymes bearing these domains, including the role of RNA structure in initial recognition and targeting during RNAi.

Nucleases are everywhere — from fabled restriction enzymes to the newly discovered RNases of the RNA silencing pathway. However, one group of nucleases has received less attention than others, the so-called sugar non-specific nucleases that play key roles in host cell defense. Hsia, Li and Yuan survey recent crystal structures of two such nucleases bound to DNA: Vvn from *Vibrio vulnificus* and ColE7 from *Escherichia coli*. Although the two nucleases are unrelated in sequence, they share a topologically similar active site and possess a kindred basic surface that adheres to the DNA minor groove. As would be expected for a non-specific DNA-binding protein, many of the contacts to the DNA are restricted to the sugar–phosphate backbone. Like restriction enzymes, Vvn and ColE7 require metals for catalysis. The authors evaluate the potential one-metal mechanism of catalysis of these non-specific nucleases, wherein a histidine may function as a general base. The authors draw structural and mechanistic connections between these bacterial nucleases and the eukaryotic apoptotic endonucleases, such as CAD/DFF40 and endoG, which degrade chromosomal DNA during apoptosis. From host defense to programmed cell death — nucleases are a fact of life!

To ensure that newly synthesized proteins get to their correct cellular location, the signal recognition particle (SRP) binds ribosomes as they are translating proteins containing an N-terminal signal sequence, which identifies secretory or integral membrane polypeptides. Halic and Beckmann review several important structures of the different targeting states of the SRP that have come from X-ray crystallographic and cryo-EM studies, leading to the first models of interactions between the SRP, its receptor protein and the ribosome. Comparison of a 4 Å crystal structure of SRP54, the conserved SRP protein that binds both the SRP RNA and the signal peptide, and an ~12 Å cryo-EM reconstruction of the full SRP bound to the ribosome suggests that a conformational change in SRP54 may accompany signal sequence binding. Crystal structures of the GTPase domain heterodimer that is part of the SRP interaction with its receptor show how GTP binds at the dimer interface and how GTPase activation may occur. Future studies will focus on understanding how signal peptide binding is communicated to the GTPase domains, coupling GTP hydrolysis to the peptide release that occurs upon docking of the SRP–ribosome complex with a transmembrane pore. Also, the role of the evolutionarily conserved and functionally critical SRP RNA remains to be fully elucidated.

Together, these reviews demonstrate the power of structural biology in unlocking mysteries, old and new. The role of RNAi in gene silencing has only begun to emerge over the past few years and yet here we are with a set of structures at the heart of this novel pathway. The mechanism of DNA unwinding by SV40 LTag and the basis of RNA recognition by zinc fingers are amongst the oldest mysteries. Thus, it is particularly satisfying to see a structural framework for these age-old conundrums. Equally impressive, we now have molecular views of the synthesis of oxidatively damaged DNA, as well as the events underlying replication initiation and control, and the life and death of mRNA. The structures of bacterial non-specific nucleases and the signal recognition particle further reflect the recent progress in the study of protein–nucleic acid complexes. One message to emerge from these reviews is the need — more than ever — to connect structures with biochemistry and genetics. The structures of SV40 LTag and the TFIIIA–RNA complex are all the more splendid because of the detailed biochemistry on these two systems. Thus, as we applaud the era of ’somes’, the

challenge of solving such structures must be met by similar advances in biochemistry and genetics.