Recent structural studies of RNA polymerases II and III

P. Cramer

Gene Center Munich, Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany

Abstract

Here, I review three new structural studies from our laboratory. First, the crystal structure of RNA polymerase (Pol) II in complex with an RNA inhibitor revealed that this RNA blocks transcription initiation by preventing DNA loading into the active-centre cleft. Secondly, the structure of the SRI (Set2 Rpb1-interacting) domain of the histone methyltransferase Set2 revealed a novel fold for specific interaction with the doubly phosphorylated CTD (C-terminal repeat domain) of Pol II. Finally, we obtained the first structural information on Pol III, in the form of an 11-subunit model obtained by combining a homology model of the nine-subunit core enzyme with a new X-ray structure of the subcomplex C17/25.

RNA synthesis in the eukaryote nucleus is carried out by the multisubunit RNA polymerases (Pol) I, II and III. Whereas Pol I and Pol II synthesize ribosomal and mainly mRNA respectively, Pol III transcribes small RNAs, including transfer RNAs, 5 S ribosomal RNA and U6 small nuclear RNA. Structural studies have so far concentrated on Pol II (reviewed in [1–5]). X-ray structures are known of the ten-subunit Pol II core and its complexes [6–11], and of the complete 12-subunit Pol II and its complexes [12–16].

A recently discovered mechanism for gene regulation involves non-coding RNAs that bind and inhibit cellular RNA polymerases. Mouse B2 RNA binds Pol II and inhibits transcription of protein-coding genes during heat shock [17,18]. In Escherichia coli, 6 S RNA binds RNA polymerase and inhibits transcription of housekeeping genes upon entry of stationary cell growth [19–21]. Understanding how these RNAs inhibit transcription requires their three-dimensional structures in complex with the target polymerases. To this end, we used Pol II from the yeast Saccharomyces cerevisiae, which is well suited for structural studies, and a previously described RNA aptamer referred to as ‘FC’ RNA [22] that binds and inhibits yeast Pol II in a manner similar to the two natural RNAs described above. The crystal structure of yeast Pol II in complex with the central region of FC RNA (Figure 1), together with biochemical data and published results, suggests the inhibitory mechanisms of polymerase-binding RNAs [23]. The structure shows that the central region of FC RNA binds in the polymerase active-centre cleft above the bridge helix [7]. The RNA forms two stem–loop structures with the two double strands twisted against each other. The 5′-stem and the 3′-stem contain four and six base pairs respectively of near A-form duplex. One RNA strand extends over both stems with continuous base stacking. Comparison of the structure with Pol II elongation complex structures [11,16] reveals substantial overlap of the RNA inhibitor with nucleic acids in an elongation complex, and explains why FC RNA inhibits initiation, but not elongation [22]. Binding of FC RNA to free Pol II blocks DNA entry into the polymerase cleft during initiation; hence, the template cannot reach the active centre, and transcription cannot start. Topological considerations suggest that FC RNA prevents DNA entry into the cleft during open complex formation. In a preformed elongation complex, however, nucleic acids exclude FC RNA from the cleft, since they are tightly bound to an overlapping, but not identical site. Like FC RNA, B2 RNA inhibits initiation, but not elongation [18], although these two RNAs do not share sequence homology. Mouse B2 RNA binds to yeast Pol II as tightly as FC RNA, indicating that the B2 RNA-binding site on Pol II is conserved among eukaryotes. Competition experiments revealed that FC RNA efficiently displaces B2 RNA from Pol II. Likewise, B2 RNA displaces FC RNA prebound to Pol II. Thus FC RNA and B2 RNA may bind to overlapping sites in the Pol II cleft. Taken together, the present study presents the first RNA polymerase structure in complex with RNA alone, and suggests a simply topological mechanism for transcription inhibition by RNAs.

Gene transcription by Pol II is physically and functionally coupled with other nuclear events [24–30]. Transcription-coupled events generally depend on the CTD (C-terminal repeat domain) of the largest Pol II subunit, which binds many nuclear factors during transcription elongation. The CTD forms a mobile extension from the structural core of Pol II [7], and consists of heptapeptide repeats of the consensus sequence Y1-S2-P3-T4-S5-P6-S7, which can be phosphorylated at residues Ser2 and Ser5. Recently, it emerged that transcription is coupled with the alteration of chromatin structure. The histone methyltransferases Set1 and Set2, which catalyse methylation of histone H3 lysine residues

Key words: histone methyltransferase, RNA polymerase, Set2 Rpb1-interacting (SRI) domain, X-ray crystallography, yeast.

Abbreviations used: CTD, C-terminal repeat domain; CID, CTD-interacting domain; HRDC, helicase and RNase D C-terminal; Pol, RNA polymerase; SRI, Set2 Rpb1-interacting.

email cramer@lmb.uni-muenchen.de

©2006 Biochemical Society
Lys⁴ and Lys³⁶ respectively, are associated with Pol II during elongation (reviewed in [31,32]). Set2 directly interacts with the phosphorylated CTD of Pol II, and is observed throughout the coding region of genes [33–36]. Set2 interacts with the Pol II CTD via a novel domain, the SRI (Set2 Rpb1-interacting) domain [37,38]. In vitro, the yeast SRI domain binds specifically and with high affinity to the CTD doubly phosphorylated at Ser² and Ser⁵ [37]. In vivo, deletion of the SRI domain abolishes histone H3 Lys³⁶ methylation, and impairs transcription elongation [37], suggesting that the SRI domain is responsible for coupling transcription with histone methylation by Set2. We determined the solution structure of the yeast SRI domain by multidimensional NMR [39]. The structure revealed three α-helices arranged in a left-handed bundle (Figure 2). Comparison with the five known structures of CTD-binding domains reveals that the SRI domain defines a novel CTD-binding fold. Other CTD-binding domains include FF domains, CIDs (CTD-interacting domains), WW (protein–protein interaction domain containing two conserved tryptophan residues) domains, BRCT (BRCA1 C-terminal) domains, and a domain in the Cgt1 subunit of the 5′-capping enzyme (reviewed in [30]).

Of these, FF and CID domains also form helical bundles [40,41], but, in contrast with the SRI domain, the superhelical arrangement in these two domains is right-handed. NMR titration experiments with Ser²/Ser⁵-phosphorylated CTD peptides characterized the CTD-binding determinants of the SRI domain. A phosphopeptide consisting of a single CTD repeat did not bind significantly. However, titration with a peptide that comprised two CTD repeats and three flanking N-terminal residues resulted in many strong chemical-shift perturbations, indicating specific binding. Residues that show strong chemical-shift perturbations of their backbone NH groups cluster in two regions on the SRI domain structure, which are conserved among fungal Set2 homologues. The observation of two putative CTD-binding regions and the finding that two CTD repeats are required for SRI domain binding indicate that the phospho-CTD extends over a long distance along helices α1 and α2, and the connecting linker. The peptide titration experiments also revealed that the two-repeat CTD peptide binds to the SRI domain via its tyrosine residues. In the three known CTD–protein complex structures, the Tyr¹ side chain is also involved in hydrophobic contacts [41–43], suggesting that Tyr¹ binding is a general feature of CTD recognition. Unexpectedly, the SRI domain resembles a region in bacterial σ factors. Taken together, the present study revealed that the SRI domain adopts a novel CTD-binding fold, demonstrated that at least two repeats of the doubly phosphorylated CTD are needed for binding, and suggested that CTD tyrosine residues are generally involved in CTD interactions.

Pol III is the largest nuclear RNA polymerase (for reviews, see [44–46]). It has a total molecular mass of approx. 0.7 MDa and comprises 17 subunits. Nine Pol III subunits form a structural core. The two largest subunits, C160 and C128, show substantial homology to the Pol II subunits Rpb1 and Rpb2 respectively. Five Pol III subunits are identical in Pol I and Pol II (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12). The Pol III subunits AC40 and AC19 are identical in Pol I and homologous with the Pol II subunits Rpb3 and Rpb11 respectively.
On the periphery of the core enzyme, Pol III contains eight additional subunits, which form three distinct subcomplexes. The subcomplex C82/34/31 [47,48] and the subcomplex C53/37/11 [49,50] are Pol III-specific, although subunit C11 shows limited homology to the Pol II subunit Rpb9 and to the Pol II elongation factor TFIIS (transcription factor IIS) [15,51]. Finally, the Pol III subcomplex C17/25 has been suggested to be the counterpart of subcomplexes Rpb4/7 in Pol II [50,52,53], Rpa14/43 in Pol I [54–56], and RpoF/E in archaeal RNA polymerase [57]. Based on the Pol II structure, we constructed a homology model for the Pol III nine-subunit core (Figure 3) [58]. Modelling involved inspection of the three-dimensional environment of each residue, and manual structure-guided realignment of weakly conserved stretches. The Pol II subcomplex Rpb4/7 could not be used in the Pol III modelling because of weak or apparently lacking sequence conservation with its Pol III counterpart C17/25. We therefore determined the C17/25 structure by X-ray crystallography. The structure of C25 resembles that of its counterparts Rpb7 and RpoE. The structure of C17 reveals a compact N-terminal ‘tip-associated’ domain, which packs mainly against the C25 tip domain, and not between the tip and OB domains as in Rpb4/7 and RpoF/E (Figure 3). The C17 tip-associated domain connects via a flexible linker to a C-terminal HRDC (helicase and RNase D C-terminal) domain, a fold that occurs in RecQ helicases and ribonucleases [55,59]. The C17 HRDC fold resembles the corresponding domains in Rpb4 and RpoF, although the sequence conservation is very weak or absent, but its position is very different from that in Rpb4 and RpoF. Compared with Rpb4 or RpoF, the HRDC domain of C17 is translated by approx. 35 Å (1 Å = 0.1 nm) and rotated by approx. 150°, and there is evidence that this domain is mobile. In contrast with Rpb4, C17 is essential for viability of *S. cerevisiae*. Complementation studies in yeast showed that the essential in vivo function requires both structural domains of C17. Similar to Rpb4/7 and Rpa14/43 [55,60,61], C17/25 binds to single-stranded RNA. Comparative EMSA analysis showed that C17/25 bound much more strongly to a tRNA sample, with an apparent affinity in the low micromolar range. Rpb4/7 bound more weakly to tRNA, although it also shows a preference for tRNA relative to a 22 nt single-stranded RNA. The HRDC domain is apparently not involved in nucleic acid binding. Docking of the C17/25 X-ray structure on to the homology model of the Pol III core resulted in a model for the 11-subunit central part of Pol III, which lacks only the subcomplexes C82/34/31 and C53/37/11 (Figure 3). The 11-subunit Pol III model reveals that at least 83.4% of the Pol II fold is conserved in Pol III, although the overall sequence identity is only 39.4%. Comparison with the Pol II elongation complex structures [11,16,62] shows that extended surfaces conserved between Pol III and Pol II are only found in the polymerase cleft, around the incoming DNA, in the active site, around the binding sites for the nucleoside triphosphate substrate and the DNA–RNA hybrid, and at the RNA exit tunnel, reflecting the conservation of the basic mechanisms of RNA elongation in the two polymerases. Thus the similarity of the core fold and active centre of Pol III and Pol II reflects a common basic transcription mechanism, whereas structural differences in a region that includes
C17/25 and directs initiation complex assembly partially accounts for promoter specificity. C17/25 contributes to initiation complex assembly, since C17 binds to the Pol III initiation factor TFIIB (transcription factor IIB) [63] and to the Pol III subcomplex C82/34/31 ([44,45] and references cited therein). Taken together, the present study provides the first structural information on Pol III and is expected to aid further structural and functional analysis of the Pol III transcription machinery.