Biochimica et Biophysica Acta 1829 (2013) 9–19

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Structural basis of transcription elongation ☆

Fuensanta W. Martinez-Rucobo, Patrick Cramer *

Gene Center and Department of Biochemistry, Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany

ARTICLE INFO

Article history:
Received 12 June 2012
Received in revised form 6 September 2012
Accepted 7 September 2012
Available online 13 September 2012

Keywords:
Gene transcription
RNA polymerase
Elongation complex structure

ABSTRACT

For transcription elongation, all cellular RNA polymerases form a stable elongation complex (EC) with the DNA template and the RNA transcript. Since the millennium, a wealth of structural information and complementary functional studies provided a detailed three-dimensional picture of the EC and many of its functional states. Here we summarize these studies that elucidated EC structure and maintenance, nucleotide selection and addition, translocation, elongation inhibition, pausing and proofreading, backtracking, arrest and reactivation, processivity, DNA lesion-induced stalling, lesion bypass, and transcriptional mutagenesis. In the future, additional structural and functional studies of elongation factors that control the EC and their possible allosteric modes of action should result in a more complete understanding of the dynamic molecular mechanisms underlying transcription elongation. This article is part of a Special Issue entitled: RNA polymerase II Transcript Elongation.

1. Introduction

Structural studies have revealed a conserved architecture, active center, and fundamental transcription mechanism for cellular RNA polymerases [1–8]. Whereas bacteria and archaea contain a single multisubunit RNA polymerase (Pol), eukaryotic transcription relies on three essential RNA polymerases, Pol I, Pol II and Pol III, which synthesize mainly ribosomal, messenger, and transfer RNA, respectively. In plants, two additional RNA polymerases, Pol IV and Pol V, generate non-coding RNA transcripts and mediate gene-silencing processes [9,10]. Although Pol IV and Pol V have not been characterized structurally, both enzymes are highly related to Pol II since they contain subunits that are identical or homologous to the Pol II subunits [11]. All cellular RNA polymerases contain a highly conserved catalytic center implemented onto a conserved double psi β-barrel in the largest subunit that is shared with RNA-dependent RNA polymerases, indicative of a common ancestor [12–15]. Eukaryotic polymerases comprise a conserved ten-subunit core and up to seven additional subunits on the periphery. For transcription initiation, the eukaryotic enzymes form topologically conserved initiation complexes on promoter DNA [16]. Initiation involves promoter binding, DNA opening, and initial RNA synthesis, and is completed when the polymerase dissociates from initiation factors, escapes from the promoter, and forms a stable elongation complex (EC) [17–21].

Most structural studies of RNA polymerases have been carried out with Pol II. Structural information on Pol II was first provided by electron microscopy (EM) [22–25], and later by low-resolution X-ray analysis [26]. Determination of the crystal structure of the ten-subunit core of *Saccharomyces cerevisiae* Pol II in free form at 2.8 Å resolution [3] and in complex with a tailed DNA template and RNA at 3.3 Å resolution [27] provided detailed information on polymerase structure and its interactions with nucleic acids. Backbone models of the complete 12-subunit Pol II were subsequently obtained [28,29] and revealed that the Rpb4/7 subcomplex protrudes from the core enzyme. A refined structure of the complete, 12-subunit Pol II was later obtained at 3.8 Å resolution by combining the core enzyme structure with a 2.3 Å structure of free Rpb4/7 [30]. The structures of core and complete Pol II provided powerful instruments for structure solution of Pol II complexes with nucleic acids and additional factors through molecular replacement [1]. Similarly, structures of bacterial RNA polymerases [7,8] enabled modeling [31–33] and structure determination [34,35] of the bacterial EC.

Here we summarize structural studies of complexes of cellular RNA polymerases with nucleic acids, and present our current understanding of the mechanisms of transcription elongation that resulted from these studies. We concentrate on cellular transcription, and thus structural studies of ECs of bacteriophage RNA polymerases [36–40] are not discussed here, although they provided many fundamental insights into the principles of transcription elongation.

2. Elongation complex structure and maintenance

The first structural information on a Pol II EC was provided by an EM reconstruction that revealed downstream DNA entering the active

☆ This article is part of a Special Issue entitled: RNA polymerase II Transcript Elongation.

* Corresponding author.
E-mail address: cramer@LMB.uni-muenchen.de (P. Cramer).

1874-9399/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagrm.2012.09.002

center cleft of core Pol II [41]. Subsequent X-ray crystallography of core Pol II transcribing a tailed-template DNA confirmed the location of downstream DNA in the cleft, and visualized a 9-base pair DNA–RNA hybrid duplex emanating at right angles to the incoming DNA from the active site at the floor of the cleft [27]. These observations were consistent with results for the bacterial EC obtained by atomic force microscopy [42], and with structural models for the EC [31,32] that were based on the first crystal structure of a bacterial RNA polymerase [8] and on crosslinking [43–45]. Crosslinking also located downstream DNA in the cleft of an archaeal RNA polymerase, albeit in an initiation complex [46]. Subsequent higher-resolution EC structures of core Pol II [47,48] and complete Pol II [49] with synthetic nucleic acid scaffolds provided a comprehensive view of the EC, including nucleotides upstream of the hybrid and interactions of nucleic acids with polymerase domains involved in EC maintenance (Fig. 1).

In the EC, incoming (downstream) DNA is unwound before the active center, allowing the template single strand to reach the active site, and is rewound to form the exiting (upstream) DNA duplex (Fig. 1). Within the unwound region (the transcription bubble), the growing RNA is attached to the catalytic site with its 3′-end and forms an 8–9 base pair hybrid duplex with the DNA template strand [19,48–50]. The upstream DNA duplex and the non-template strand in the bubble are mobile, but were located with the use of fluorescence resonance energy transfer experiments (Fig. 1A) [51,52]. The arrangement of nucleic acids in the Pol II EC is very similar in crystal structures of a bacterial EC that additionally revealed exiting upstream RNA in a previously proposed exit channel [34,35]. Many mutations in the genes encoding RNAP subunits were rationalized with the EC structure [53]. The DNA–RNA hybrid duplex is the distinguishing feature of the EC, and binds within the active center cleft to adopt a conformation that is close to an A-form duplex. The hybrid-binding site cannot bind a B-form DNA–DNA duplex, but an A-form RNA–RNA duplex is accommodated [14]. This explains how cellular RNA polymerases can have some RNA template-dependent RNA synthesis activity, and indicates that all cellular RNA polymerases might have evolved from an ancient replicase that duplicated RNA genomes.

During elongation, the polymerase maintains the transcription bubble. This requires unwinding of the downstream DNA duplex, and separation of the RNA product from the DNA template at the end of the hybrid duplex. DNA unwinding is apparently achieved by positively charged residues in the polymerase switch 2 region that 'pull' the

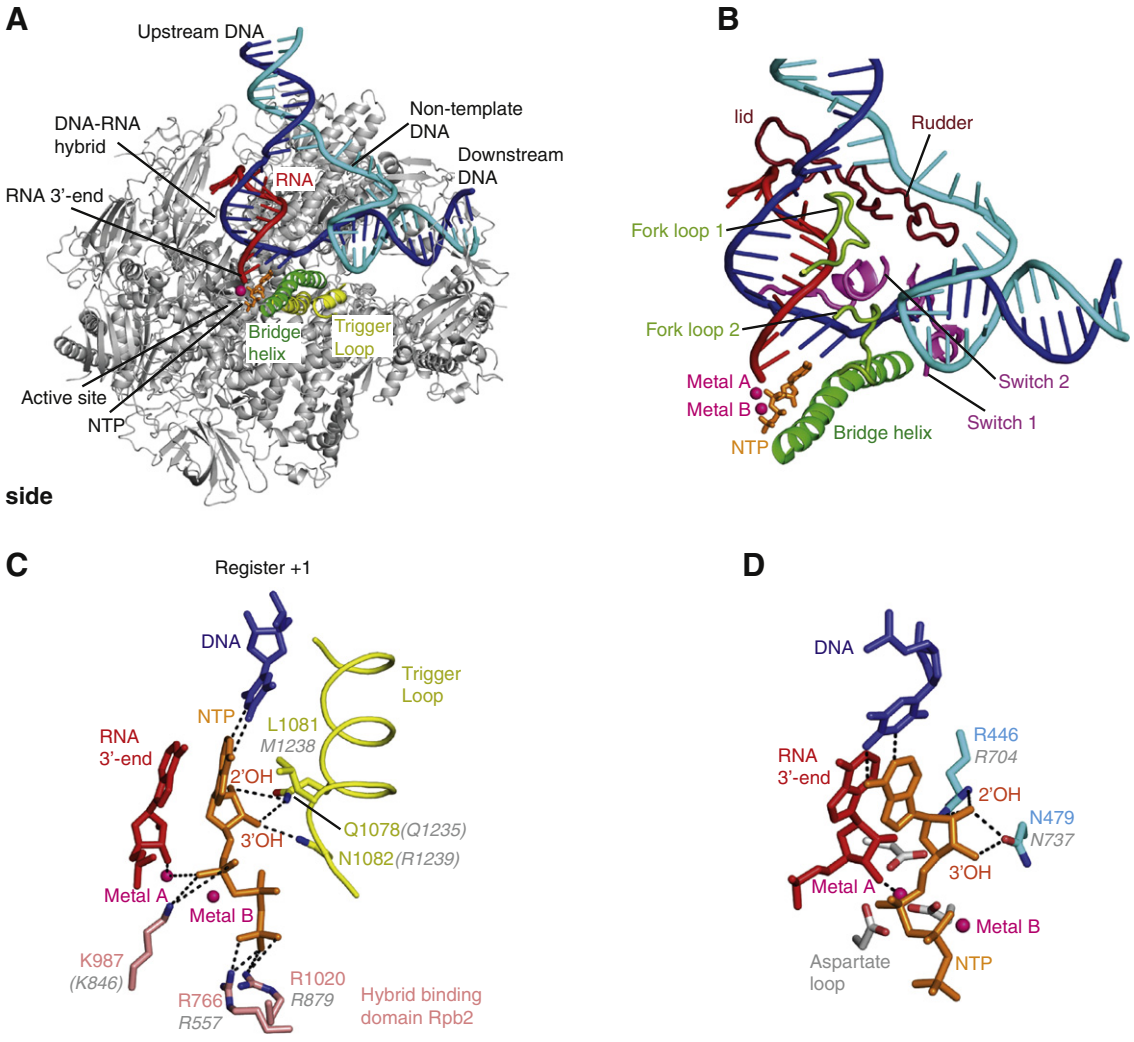

Fig. 1. Elongation complex structure and maintenance. (A) Overview of the Pol II elongation complex with bound NTP (PDB 4A3F [60]). The exiting RNA, upstream DNA duplex and the non-template strand in the bubble were modeled according to [51,52]. The bridge helix (green), trigger loop (yellow), and metal A (pink) are highlighted. DNA template, DNA non-template, RNA, and NTP are shown in blue, cyan, red, and orange, respectively. (B) Nucleic acids and Pol II structural elements involved in transcription bubble opening and maintenance. Refer to the text for details. Functional elements are highlighted in color and labeled. (C, D) Contacts of the NTP substrate with Pol II residues, DNA template, RNA and metal A. (D) is related to (C) by an approximate 60° clockwise rotation along the y-axis. Pol II residues are colored and labeled. Corresponding bacterial RNA polymerase residues are labeled in gray [35]. Bacterial residues that do not contact the nucleotide are indicated in brackets. Metal B was modeled in (B, C, and D) by superposition of a bacterial EC (PDB 2O5J) [35].