Methods 86 (2015) 10–18

Contents lists available at ScienceDirect

Methods

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

Review Article

Fluorescently labeled recombinant RNAP system to probe archaeal transcription initiation

Sarah Schulz [a], Kevin Kramm [a], Finn Werner [b], Dina Grohmann [a,*]

[a] Physikalische und Theoretische Chemie – NanoBioSciences, Technische Universität Braunschweig, Hans-Sommer-Straße 10, 38106 Braunschweig, Germany
[b] RNAP Laboratory, University College London, Institute of Structural and Molecular Biology, Division of Biosciences, Gower St, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 14 February 2015
Received in revised form 14 April 2015
Accepted 16 April 2015
Available online 22 April 2015

Keywords:
Fluorescence
Archaea
RNA polymerase
Single-molecule
Transcription

ABSTRACT

The transcriptional apparatus is one of the most complex cellular machineries and in order to fully appreciate the behavior of these protein–nucleic acid assemblies one has to understand the molecular details of the system. In addition to classical biochemical and structural studies, fluorescence-based techniques turned out as an important – and sometimes the critical – tool to obtain information about the molecular mechanisms of transcription. Fluorescence is not only a multi-modal parameter that can report on molecular interactions, environment and oligomerization status. Measured on the single-molecule level it also informs about the heterogeneity of the system and gives access to distances and distance changes in the molecular relevant nanometer regime. A pre-requisite for fluorescence-based measurements is the site-specific incorporation of one or multiple fluorescent dyes. In this respect, the archaeal transcription system is ideally suited as it is available in a fully recombinant form and thus allows for site-specific modification via sophisticated labeling schemes. The application of fluorescence based approaches to the archaeal transcription apparatus changed our understanding of the molecular mechanisms and dynamics that drive archaeal transcription and unraveled the architecture of transcriptional complexes not amenable to structural interrogation.

1. Introduction

The structures of multisubunit RNA polymerases (RNAP) from bacteria, archaea and eukaryotes show that RNAPs share a conserved structural organization [1–3]. The RNA polymerase core is composed of the assembly platform (archaeal subunits Rpo3, 10, 11, 12) and the two largest subunits (archaeal subunits Rpo1 and Rpo2 which are sometimes split and referred to as Rpo1′/Rpo1″ and Rpo2′/Rpo2″ [4]) that harbor the catalytic center at their interface (Fig. 1A) [5]. Despite the mutual crab claw-like architecture of multisubunit RNAPs, there are significant differences between the bacterial and archaeal–eukaryotic lineage. Archaeal and eukaryotic RNAPs show a number of additional subunits at the outer surface providing a platform for the interaction with transcription factors [1]. The archaeal transcription system furthermore shares a highly conserved set of basal transcription initiation factors which exhibit no structural homology to the initiation factors (sigma factors) utilized by the bacterial RNAP [1].

Transcriptional regulation starts with the sequence-specific recognition of unique promoter DNA elements by transcription factors. Archaeal and eukaryotic RNAPs strictly depend on two basal transcription factors TBP (TATA-binding protein, Fig. 1B) and TFB/TF(II)B (transcription factor B, Fig. 1C) to initiate transcription [6–11]. TBP binds to the TATA-element of the promoter and simultaneously bends the promoter DNA [12–15]. TFB is able to recognize the TBP–DNA complex, binds to it and recruits the RNAP yielding the pre-initiation complex (PIC) [16]. TFB is composed of the zinc ribbon (ZR) and the C-terminal core domain, which are connected by a flexible linker [17,18]. The core domain contacts the B-recognition element located upstream of the TATA box in the promoter DNA and spans across the DNA-binding channel. The flexible linker of TFB is involved in the selection of the start site and penetrates deep into the active site of the RNAP thereby extensively interacting with the RNAP [19]. TBP and TFB are necessary and sufficient to recruit the RNAP and to start transcription. In eukaryotes, the RNAPII pre-initiation complex contains a dedicated ATP-dependent helicase (part of the general transcription factor TFIIH) responsible for DNA melting. In contrast [20–22], DNA opening is a spontaneous process in archaea [7,10,23]. However, a third transcription factor (transcription factor E, TFE, Fig. 1D) stabilizes the PIC and stimulates transcript synthesis [23,24]. TFE

* Corresponding author at: Department of Biochemistry, Genetics and Microbiology, Institute of Microbiology, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany. Fax: +49 941 943 2403.
E-mail address: d.grohmann@tu-braunschweig.de (D. Grohmann).

http://dx.doi.org/10.1016/j.ymeth.2015.04.017
1046-2023/

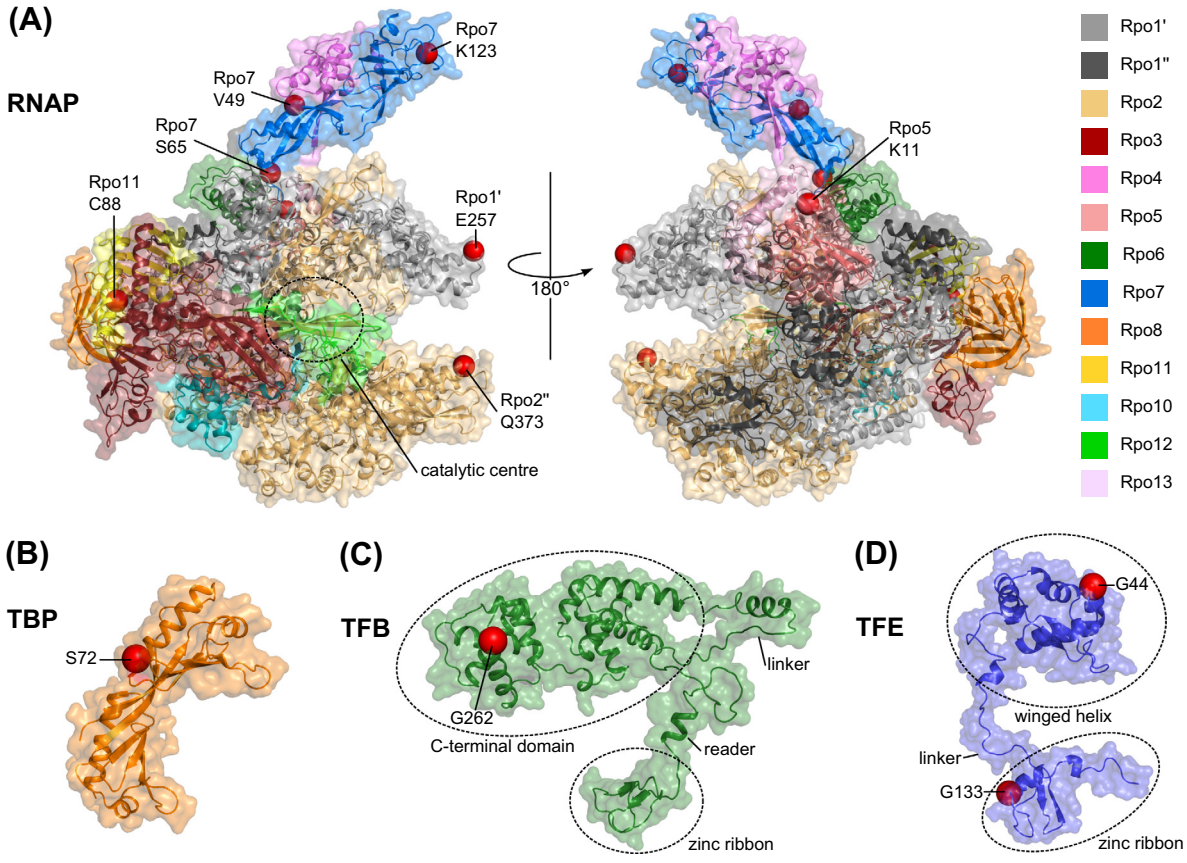

Fig. 1. A fluorescently labeled archaeal transcription system. (A) Composition of the archaeal RNA polymerase from *Sulfolobus shibatae* (PDB: 2WAQ). The subunits are individually color-coded and the seven labeling sites used in a variety of studies are marked with a red sphere. Subunits Rpo4, 5, 7 were genetically engineered to carry a single cysteine at the highlighted positions allowing the site-specific coupling of fluorescent probes via the thiol group. Site-specific coupling of fluorophores to the large subunits Rpo1′ and Rpo2″ was achieved via a bioorthogonal coupling reaction between the unnatural amino acid *p*-azidophenylalanine (AzF) introduced at this site and a phosphine-derivative of a fluorescent dye. (B) Crystal structure of *Methanocaldococcus jannaschii* TBP (PDB: 2Z8U) with a labeling site on position S72. The serine at this position was replaced by a single cysteine. (C) TFB consists of the C-terminal domain, linker, reader helix and the zinc ribbon domain (model of archaeal TFB presented in Nagy et al. [34] based on the TBP-TATA-TFB structure from *Pyrococcus woesei*, PDB: 1D3U, and the eukaryotic TFIIB structure, PDB: 4BBR). The C-terminal domain contacts the DNA whereas the zinc ribbon domain is threaded into the RNAP. The flexible linker protrudes into the active site of the RNAP. Coupling of a fluorophore was achieved via AzF introduced at position G262 (C-terminal domain). (D) Crystal structure of TFE (homology model of *M. jannaschii* TFE [74]) composed of the winged helix domain and the zinc ribbon domain that are connected by a flexible linker. TFE was fluorescently labeled via AzF introduced at position G44 (winged helix domain) and G133 (zinc ribbon domain).

is a two-domain protein comprised of the ZR and winged helix (WH) domain. After the RNAP escapes from the promoter it enters the productive phase of transcription (elongation). Here, the transcription elongation factor Spt4/5 binds to the RNAP thereby locking the DNA in the DNA binding channel [25,26]. This leads to a stimulation of the RNAP processivity [27]. In archaea, the RNAP will terminate RNA synthesis and dissociate from the DNA template once a poly-U stretch is transcribed [28].

The development of recombinant transcription systems has been crucial to analyze the structure–function relationships of the transcriptional apparatus [7,9,29–32]. Based on the recombinant archaeal transcription system from *Methanocaldococcus jannaschii* we developed the fluorescently labeled version of the transcription system. This allowed us to apply a wide range of fluorescence-based methods (see Section 3) to unravel the structure, function and dynamics of archaeal transcription and by inference provided insights into mechanisms of eukaryotic transcription.

2. Preparing fluorescently labeled archaeal RNA polymerases

A highly tractable recombinant form of the transcriptional machinery provides the basis for a specific and efficient fluorescently labeled transcription system. However, only the bacterial and archaeal transcription machinery could be reconstituted in an active form *in vitro* so far [7,9,29–32]. An advantage of the recombinant system is that the RNAP can be assembled from its individual subunits. This opens up the unique possibility to introduce site-specific modifications (e.g., fluorescent dyes, spin labels or biotins) in an isolated subunit first before the complete RNAP is assembled (Fig. 2). Among the numerous advantages of this approach are the free choice of the coupling reaction, the chance to introduce different types of labels, to introduce mutations and the option to assemble labeled RNAPs that differ in their subunit composition. It also means that an RNAP molecule can be site-specifically equipped with multiple modifications as each subunit can contribute its individual label. This way, fluorescent probes could be introduced in the bacterial and archaeal RNAP at numerous sites that are mechanistically informative and minimally perturbative to the biological activity [24,33,34]. Additionally, we also developed protocols for the site-specific labeling of transcription factors TBP, TFB and TFE (Fig. 1B-D) which follow the same principle steps of protein modification outlined for the RNAP subunits (Fig. 2).

Typically, fluorescent dyes are incorporated into purified proteins exploiting the reactivity of single cysteines (Fig. 2A/D). Here, thiol groups of cysteines in a protein spontaneously and specifically undergo a reaction with the maleimide moiety and form a thioether linkage, a reaction that is not reversible [35]. Ideally, labeling sites