Biochemical Society Transactions (2001) Volume 29, part 4

Basal and regulated transcription in Archaea
S. D. Bell1, C. P. Magill and S. P. Jackson
Wellcome Trust and Cancer Research Campaign Institute of Cancer and Developmental Biology, Tennis Court Road, Cambridge, CB2 1QR, U.K., and Department of Zoology, University of Cambridge, U.K.

Abstract
The basal transcription machinery of Archaea is fundamentally related to the eucaryal RNA polymerase (RNAP) II apparatus. In addition to a 12-subunit RNAP, Archaea possess two general transcription factors, the activities of which are required for accurate and efficient in vitro transcription. These factors, TBP and TFB, are homologues of the eucaryal TATA-box binding protein and TFIIB respectively. Archaea also possess TFE, a homologue of the eucaryal RNAP II general transcription factor TFIIE. Although not absolutely required for transcription in vitro, TFE nonetheless plays a stimulatory role under conditions where promoter recognition by TBP is sub-optimal. The basal transcription apparatus of Archaea is closely related to that of Eucarya but archaeal transcriptional regulators resemble those of bacteria. The mode of action of two such regulators has been characterized to determine how these 'bacterial-like' regulators impinge on the 'eucaryal-like' basal machinery.

The archaeal basal transcription machinery
With the characterization of an archaeal RNA polymerase (RNAP) over 20 years ago, it became apparent that aspects of archaeal transcription more closely resembled those of Eucarya than those of bacteria [1]. Subsequent characterization of archaeal promoters and general transcription factors has underlined further this high degree of similarity [2-4]. However, the archaeal basal transcription machinery is vastly more simple than the homologous RNAP II system in Eucarya [5]. As indicated in Figure 1, the archaeal general transcription factors constitute three proteins. This is in contrast with the more than 25 polypeptides involved in the RNAP II basal apparatus. Specifically, Archaea possess homologues of the eucaryal RNAP II general transcription factors TBP, TFIIB and TFIIE that are named TBP, TFB and TFE respectively in Archaea. This simplicity, coupled with the ease of biochemical analysis of proteins from hyperthermophilic Archaea, has greatly facilitated dissection of the molecular mechanisms underpinning the process of transcription initiation in these organisms. Initial studies revealed that the minimal requirements to reconstitute basal transcription in vitro from archaean promoters are TBP, TFB and RNAP [6,7]. These three proteins can mediate transcription initiation from the majority of promoters tested. Remarkably, and in contrast with Eucarya, this minimal subset of factors and RNAP is able to catalyse transcription initiation, even on topologically unfavourable templates, in vitro [8].

Our understanding of the architecture of the archaeal transcription pre-initiation complex has been greatly aided by work in the late Paul Sigler's lab, in which the three-dimensional crystal structure of TBP and a C-terminal 'core' domain of TFB were co-crystallized on archaeal promoter DNA [9,10]. From these structural studies, combined with functional analyses, we now know that recognition of the promoter is mediated by recognition of two adjacent promoter elements by TBP and TFB [11,12]. These elements, termed the TATA-box and BRE respectively, are well conserved in archaeal promoters, and are also found in many eucaryal RNAP II promoters. The juxtaposition of these two elements imparts a

Figure 1
Cartoon of the archaeal basal transcription factors
TBP contains an imperfect direct repeat and, in some species, a highly acidic short C-terminal tail [20]. TFB has two principal domains, an N-terminal zinc ribbon and a C-terminal 'core domain' containing an imperfect direct repeat of a cyclin fold. These two domains are separated by a highly conserved linker region. TFE possesses an N-terminal leucine-rich region containing a helix-turn-helix and a C-terminal zinc ribbon motif.

Key words: RNA polymerase, transcription factors.
Abbreviations used: RNAP, RNA polymerase; MDRI, metal-dependent repressor 1; ABC, ATP-binding cassette.
1To whom correspondence should be addressed (e-mail sdb@mole.bio.cam.ac.uk) at the present address: The Hutchinson/MRC Research Centre, Hills Rd, Cambridge CB2 2QH, U.K.
Molecular Communications

unique geometry to the pre-initiation complex, and ensures that the RNAP is appropriately recruited to the initiation site, located approx. 22 bp downstream of the TATA-box, on the opposite side of the BRE [12]. The molecular events undergone during the subsequent recruitment of the RNAP itself to the promoter remain less well characterized. Deletion analyses of TFB have revealed that its N-terminal domain is required for recruitment [13], and structural studies of this domain have revealed that it forms a zinc ribbon [14]. Work is currently underway to identify which subunit(s) of the RNAP are contacted by this region of TFB. Once recruited, the RNAP is positioned over the start site, generating a DNase I footprint that extends from -43 to at least +8 bp [13]. Following formation of the pre-initiation complex, the RNAP catalyses isomerization of the promoter DNA, melting a region of about 13 bp extending from approx. +3 to -10 bp relative to the transcription start site [8,13]. Subsequently, the RNAP initiates RNA synthesis. Between 5 and 6 nt into the transcript, the RNAP–DNA–RNA complex undergoes a sharp transition, becoming resistant to the polyanion heparin. At this point, a concomitant alteration in the position of the melted region is observed, as adjudged by permanganate footprinting (S. D. Bell, unpublished work). The process of promoter clearance, defined as the transition from a promoter-bound form to an elongating form of the RNAP, appears to be facilitated by a highly conserved motif in TFB, located between the N-terminal zinc ribbon and C-terminal core domain [13]. Thus TFB appears to play multiple key roles in the transcription-initiation process. First, by binding the TBP–DNA complex and making sequence-specific contacts with the BRE, TFB establishes a unique polarity to the pre-initiation complex. Secondly, the N-terminal zinc-ribbon domain plays a pivotal role in recruiting the RNAP to the promoter. Thirdly, following the establishment of the stable pre-initiation complex it appears that the motif between the zinc ribbon and core domains of TFB plays a role in facilitating the escape of the RNAP from this complex, and allowing the RNAP to escape the promoter (Scheme 1).

Although TBP, TFB and RNAP are necessary and sufficient to mediate transcription initiation from most archaeal promoters in vitro, it has been apparent for some time that all sequenced archaeal genomes contain an open reading frame homologous with the α subunit of TFIIE [15]. However, until recently, it has been unclear what role, if any, the encoded protein might play in transcription. Recent work has demonstrated that, on some promoters or under certain conditions, TFE can play a stimulatory role in transcription initiation [16,17]. Specifically, TFE appears to facilitate pre-initiation complex formation under in vitro conditions, where TBP–TATA-box interactions are sub-optimal [16]. Although this effect is relatively modest in vitro (maximally, a 3-fold stimulation), it seems likely that TFE will play a more substantial role in vivo, where TBP must compete with archaeal chromatin proteins for binding to the TATA-box.

Scheme I

Diagram of the steps in archaeal transcription initiation

The sequential assembly of archaeal general transcription factors on a promoter is shown. The binding of TBP to the TATA-box is facilitated by TFE. TFB then binds the TBP–DNA complex, making sequence-specific contacts with the BRE. The N-terminal domain of TFB then interacts with RNAP to form the pre-initiation complex. Subsequently, the escape of RNAP from the promoter is facilitated by a conserved motif located between the zinc ribbon and core domain of TFB. This interaction is indicated by an arrow (bottom panel). Shown on the right of the Figure are the points at which the transcriptional repressors Ls14 and MDR1 act to block pre-initiation complex assembly on their own promoter.
Regulation of transcription in Archaea

Given that the archaeal basal transcription machinery resembles that of Eucarya, it might be predicted that archaeal transcriptional regulators would also be 'eucaryal-like'. It was a considerable surprise, therefore, to discover that archaeal genomes encode many homologues of bacterial regulators. Indeed, the relative proportion of these regulator homologues in archaeal and bacterial genomes is roughly equivalent, suggesting that these regulators were established before the divergence of the archaeal and bacterial lineages [18]. Accordingly, we have coined the description 'BA regulators' to describe these putative regulatory molecules shared between the bacterial and archaeal lineages [18]. We have now studied how two such BA regulators impinge on the archaeal basal machinery. The first is metal-dependent repressor 1 (MDR1) from Archaeoglobus fulgidus. This regulator is co-transcribed with open reading frames encoding homologues of a tripartite-metal-importing ABC (ATP-binding-cassette) transporter. Transcription of this operon is regulated by the availability of metal ions in the growth medium. Specifically, in the presence of metal ions transcription is reduced and, conversely, under limiting metal ion concentrations, transcription is induced around 20-fold. MDR1 is a homologue of bacterial metal-dependent repressors, typified by the diphtheria toxin repressor, DtxR. Purified MDR1 protein was demonstrated to bind cooperatively to three operator sequences in the promoter of the operon in a metal-dependent manner. This was found to prevent RNAP recruitment to the promoter [19]. Intriguingly, however, the recognition of the TATA-box and BRE by TBP and TFB is unaffected by MDR1. This suggests that, under repressing conditions, the promoter is poised to rapidly recruit the RNAP via the DNA-bound basal factors, allowing an extremely rapid response to changing environmental conditions [19].

A second BA regulator, Lrs14, also uses a steric hindrance mechanism to regulate its own expression. However, the details of the repression mechanism differ considerably from those employed by MDR1. Lrs14 binds to at least two operators in its own promoter. In marked contrast with MDR1, Lrs14-binding sites overlap the core promoter elements of its own promoter. Indeed, base pairs within the TATA-box itself are of key importance for Lrs14 binding [18] (Scheme 1). This therefore suggests that the regulation of Lrs14 levels might not be as rapid as for MDR1. This potential for differential kinetics for regulation of these two systems might reflect differing requirements for the cell to respond to changing conditions. In the case of MDR1, the level of metal ions in the cell will be of key importance to cellular survival, and thus a rapid modulation of expression of the metal-importing ABC transporter is required. In contrast, cells might be able to tolerate varying levels of the Lrs14 regulator. It will be of considerable interest to identify downstream targets of Lrs14 and determine their mode of regulation.

Future prospects

Although considerable detail about the archaeal transcription machinery is now known, many questions remain. In particular, the roles of the many small subunits conserved between the archaeal and eucaryal RNAPs remain poorly understood. To address this central question in transcription, an important goal for future experiments will be the in vitro reconstitution of the archaeal RNAP and sub-assemblies thereof from recombinant subunits. Analysis of the properties of sub-assemblies lacking either entire subunits or domains of these will provide insights into the molecular function of these subunits. Many additional questions remain in the field of transcriptional regulation. For example, while a number of repressors of archaeal transcription have been identified, it will be of great interest to discover how activators of archaeal transcription function. Will these function by simple recruitment mechanisms, as do many bacterial activators, and, if so, which components of the basal machinery will they contact? Finally, to date, the majority of experiments in vitro have been performed on naked DNA templates. It will be extremely informative to examine the effect of archaeal chromatin proteins on transcription. Thus studies of transcription in Archaea with their properties of biochemical tractability and relative simplicity, combined with their unique position in the evolutionary tree of life, will allow a high-resolution dissection of the fundamental molecular processes of transcription initiation, and provide insight into this process in all three domains of life.

This work has been supported by the Wellcome Trust and Cancer Research Campaign. We would also like to thank John Reeve (Ohio State University) for communicating data prior to publication.
Molecular Communications

References


Received 23 February 2001

HMG1 and 2: architectural DNA-binding proteins
J. O. Thomas
Cambridge Centre for Molecular Recognition and Department of Biochemistry, 80 Tennis Court Road, Cambridge CB2 1GA, U.K.

Abstract

HMG1 and 2 (high mobility group proteins 1 and 2; renamed HMGB1 and 2) contain two DNA-binding HMG-box domains (A and B) and a long acidic C-terminal domain. They bind DNA without sequence specificity, but have a high affinity for bent or distorted DNA, and bend linear DNA. The individual A and B boxes (which, although broadly similar, show both structural and functional differences) exhibit many of the structure-specific properties of the whole protein. The acidic tail modulates the affinity of the tandem HMG boxes in HMG1 and 2 for a variety of DNA targets, including four-way junctions, but not distorted DNA minicircles, to which the proteins bind with very high affinity. HMG1 and 2 appear to play important architectural roles in the assembly of nucleoprotein complexes in a variety of biological processes, for example, V(D)J recombination, the initiation of transcription, and DNA repair.

Key words: DNA distortion, DNA minicircles, high mobility group (HMG) proteins, HMG box.

Introduction

HMG1 and 2 (high mobility group proteins 1 and 2; recently redesignated HMGB1 and 2 [1]) are ubiquitous proteins in vertebrates. They are relatively abundant (about 1 molecule per 10-15 nucleosomes on average) and, like histones, bind to DNA without sequence specificity, being initially regarded as probable chromatin structural components [2]. However, more recent studies have revealed other important roles for HMG1 and 2, which result from their distinctive 'structure-specific', rather than sequence-specific, DNA-binding properties. They bend linear DNA and bind preferentially to bent or distorted DNA in vitro, and appear to have a role in vivo in the assembly of nucleoprotein complexes [2-4].

HMG1 and 2 have two distinguishing features: two HMG boxes (A and B), homologous folded domains of about 80 amino acid residues which mediate DNA binding, and a long acidic tail containing 30 (HMG1) or 20-24 (HMG2a and 2b) aspartic or glutamic acid residues, linked to the boxes by an overall basic region of about 20 residues (Figure 1). Abundant HMG-box proteins in organisms other than vertebrates (e.g. HMG-D and HMG-Z in Drosophila melanogaster and Nhp6ap and Nhp6bp in Saccharomyces cerevisiae),...