Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A Structure/Function Analysis of Escherichia coli RNA Polymerase

Carol A. Gross, Cathleen L. Chan and Michael A. Lonetto

Phil. Trans. R. Soc. Lond. B 1996 **351**, 475-482
doi: 10.1098/rstb.1996.0045

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

This journal is © 1996 The Royal Society

Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A structure/function analysis of *Escherichia coli* RNA polymerase

CAROL A. GROSS, CATHLEEN L. CHAN AND MICHAEL A. LONETTO

Departments of Microbiology and Stomatology, UCSF Box 0512, University of California-San Francisco, San Francisco, California 94143 U.S.A.

SUMMARY

Control of RNA polymerase is a common means of regulating gene expression. A detailed picture of both the structure and how the structural details of RNA polymerase encode function is a key to understanding the molecular strategies used to regulate RNA polymerase. We review here data which ascribes functons to some regions of the primary sequence of the subunits (α,ββ′σ) which make up *E. coli* RNA polymerase. We review both genetic and biochemical data which place regions of the primary sequence that are distant from one another in close proximity in the tertiary structure. Finally we discuss the implications of these findings on the quaternary structure of RNA polymerase.

1. INTRODUCTION

After the determination of the structure of DNA in 1954 (Watson & Crick 1953), the entire question of the control of the flow of genetic information from DNA to proteins was vigorously attacked on several fronts. The years around 1960 showed an enormous outpouring of information concerning this connection from several different scientific disciplines. Jacob, Monod and their colleagues published reports indicating that the activity of structural genes was controlled by the activity of regulatory genes (Jacob & Monod 1961). Experiments from phage-infected cells established that mRNA was likely to be the active molecule in transferring information from gene to protein (Brenner *et al*. 1961). Finally, reports from several laboratories working on systems as diverse as rat liver, peas and bacteria indicated that possession of a DNA-dependent transcriptase was likely to be a universal characteristic of living organisms (Weiss & Gladstone 1959; Huang *et al*. 1960; Hurwitz *et al*. 1960; Stevens 1960; Weiss & Nakamoto 1961). This latter finding opened the door to a study of how this enzyme worked and the mechanisms regulating its activity.

Early studies of the bacterial transcriptase indicated that it was a multisubunit enzyme that came in two forms: holoenzyme ($\alpha_2\beta\beta'\sigma$) specialized for initiation at promoter sequences and core RNA polymerase ($\alpha_2\beta\beta'$), specialized for elongation and termination (Burgess *et al*. 1969). Moreover, despite the complexity of the three eukaryotic RNA polymerases (pol I, pol II, pol III) which each contain upwards of 10 subunits (Greenleaf 1992; Sentenac *et al*. 1992), it is clear that there is a close relation between prokaryotic and eukaryotic transcription. First, eukaryotic RNA polymerases contain subunits homologous to all of the prokaryotic core RNA polymerase subunits. Second, recent reports from several laboratories suggest that eukaryotic cells contain RNA polymerase holoenzyme. Holoenzyme (consisting of RNA polymerase and more than 20 associated proteins), together with the TATA binding protein is capable of initiating basal transcription and responding to activators at eukaryotic promoters (Thompson *et al*. 1993; Kim *et al*. 1994; Koleske & Young 1994; Ossipow *et al*. 1995). Thus although pol II initation is a highly complicated process, requiring stepwise assembly of a large number of basal transcription factors in conjunction with RNA polymerase, the structural motifs within the enzyme that carry out catalysis of RNA synthesis are probably the same for both eukaryotes and prokaryotes (Buratowski & Sharp 1992; Maldonado & Reinberg 1995). Further, examples of protein–protein communication within the initiating transcription complex that modulate the activity of RNA polymerase suggest that similar mechanisms operate in establishing the pol II pre-initiation complex. Thus understanding transcription activation at the molecular level in *Escherichia coli* will provide insight into gene regulation in more complex systems.

An understanding of the structure–function relations of RNA polymerase is critical to determining how regulators modulate the intrinsic functions of the enzyme. Because RNA polymerase is both large and essential, genetic and physical approaches to this question have been limited. Biochemical analysis of both initiation and elongation of transcription has led to the view that the transcribing complex contains sites that bind the DNA template and allow for formation and maintenance of a melted region of DNA (the transcription bubble), bind the RNA (a product binding site) and bind the incoming NTP (a substrate site). Thus there are several functional 'sites' in RNA polymerase defined by biochemical properties. A

© 1996 The Royal Society

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Philosophical Transactions: Biological Sciences*. STOR® www.jstor.org

Downloaded from rstb.royalsocietypublishing.org

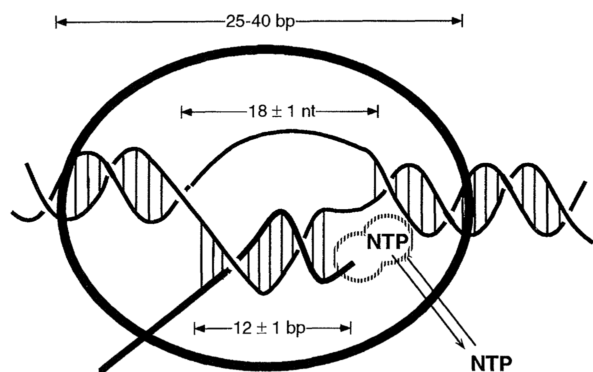


Figure 1. Structure of the elongation complex.

representation of the elongation complex is shown in figure 1. Each segment of nucleic acid shown in this figure could potentially interact with a unique binding site(s) in RNA polymerase. The challenge is to identify which regions of the protein constitute each of these different sites and determine the spatial proximity of the different sites to each other.

Several complementary approaches are currently being used to identify and understand the structure and function of these sites. First, the overall shape of the enzyme is being determined by low resolution electron crystallography. Both holoenzyme and core RNA polymerase have been examined by this technique. The notable difference between the two is the relation of a thumblike projection to an adjacent channel with the appropriate dimensions to bind DNA. In holoenzyme, the channel is open, whereas in core RNA polymerase the thumb closes over the channel. The 'open conformation' may represent initating polymerase, whereas the 'closed conformation' may represent an elongating form of the enzyme (Polyakov *et al.* 1995). The two-dimensional crystal structure of yeast pol II resembles that of *E. coli* core RNA polymerase (Darst *et al.* 1991). Second, crosslinking studies have been utilized to determine proximity of various portions of polymerase to the transcript. Third, functions and structures of subdomains of the enzyme are being determined. Finally, both selective and directive genetic approaches, in conjunction with the development of sophisticated assays for analysis of mutant phenotypes, have proved informative. The current state of our knowledge concerning the structural and functional organization of prokaryotic RNA polymerase is summarized below.

2. THE α SUBUNIT OF RNA POLYMERASE

The α subunit is implicated in three distinct processes: it is required for assembly of RNA polymerase, it provides an auxillary DNA binding site at some very strong promoters and is involved in activation of transcription by activator proteins (Igarashi & Ishihama 1991; Ross *et al.* 1993; Busby & Ebright 1994; Chen *et al.* 1994; Kimura & Ishihama 1995).

The first function of α to be elucidated was its role in assembly of RNA polymerase. Early studies indicated that α was the key to this process. Initial formation of an α_2 dimer is followed by formation first of an $\alpha_2\beta$ subassembly and finally of the completed $\alpha_2\beta\beta'$ core enzyme (Zillig *et al.* 1976). More recent studies have localized this assembly function to the N-terminal domain (NTD) of α, as proteins truncated about 2/3 of the way through the subunit are competent for assembly (Igarashi & Ishihama 1991). Very recently, mutational analysis has suggested that discrete portions of α are involved in each assembly step (Kimura & Ishihama 1995). Each of these regions is conserved between eukaryotic and prokaryotic polymerases, suggesting that the assembly process exhibited by prokaryotic core RNA polymerase may be a common feature of multisubunit RNA polymerases.

The role of α in mediating activation is also a story with a long history. For many years, people working on a variety of activator-controlled systems noticed that mutations preventing activation mapped in α. However, this relation was not systematically investigated until it was found that the holoenzyme containing only the α-NTD was able to perform basal but not activated transcription at some promoters (Igarashi & Ishihama 1991). Further detailed genetic investigation indicated that the α-C-terminal domain (α-CTD) was required for activation by a number of different activators. In at least one case, these genetic interactions were verified by cross-linking analysis. A crosslinker positioned at the activating domain of the CRP activator crosslinks with very high efficiency to the α-CTD, indicating their close proximity (Chen *et al.* 1994).

The DNA binding function of α is also located in the CTD. A minimal prokaryotic promoter contains two conserved hexamer sequences at −10 and −35 b.p. (base pairs) from the start-site of transcription. α-binds to DNA sites called UP elements, which are located upstream from these two conserved sites (Ross *et al.* 1993). Current models of activation suggest that binding of α to activator may either substitute for interaction with UP-element or position α so that it can interact with non-consensus UP elements. These issues are discussed much more fully in the article by S. Busby, this volume. Note that the α-CTD is not conserved between prokaryotes and eukaryotes.

3. THE TWO LARGEST SUBUNITS OF RNA POLYMERASE

The two largest subunits of RNA polymerase comprise the catalytic core of the enzyme, are involved in the DNA binding, as well as the elongation and termination activities of the enzyme. However, the relation of these activities to the structure of the enzyme has yet to be established. Instead, our understanding is at a more primitive level. As will become clear from the discussion below, at the present time we have identified certain landmarks in the primary sequence, and are beginning to understand the overall topology of the enzyme, but have yet to define the relation between these features and enzyme function.

Downloaded from rstb.royalsocietypublishing.org

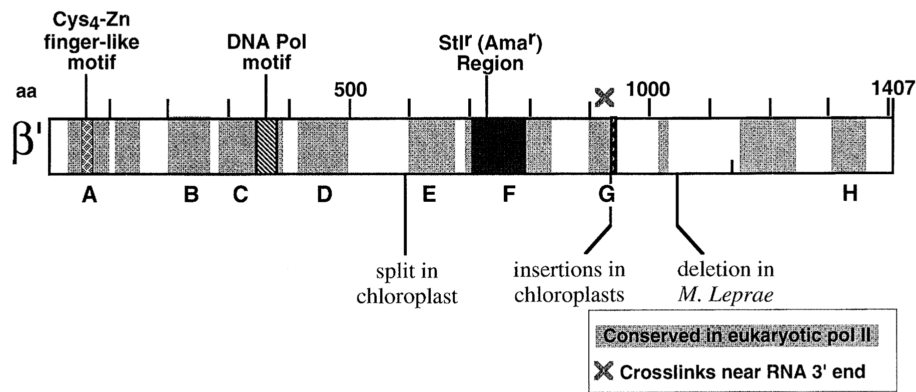


Figure 2. Conserved features of β′, the largest subunit of *E. coli* RNA polymerase.

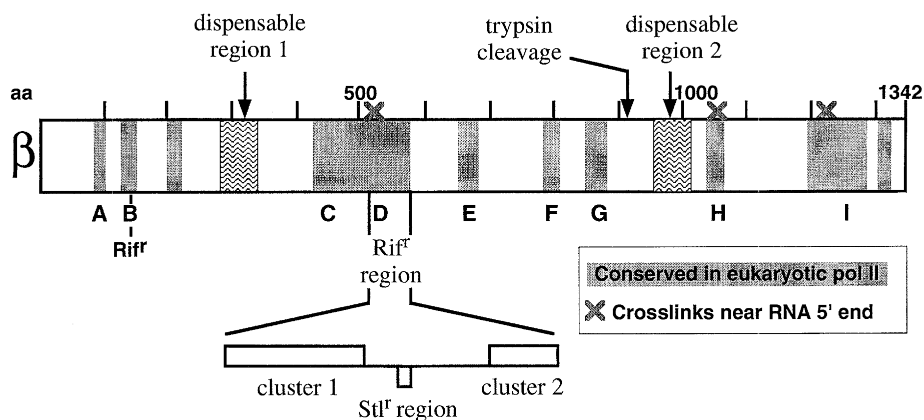


Figure 3. Conserved features of β, the second largest subunit of *E. coli* RNA polymerase.

(*a*) ***Landmarks in the two largest subunits***

The two largest subunits of *E. coli* RNA polymerase, termed β′ and β, are homologous respectively with the largest and second largest subunits of all eukaryotic multisubunit RNA polymerases (Weilbaecher *et al.* 1994). Both β′ and β share eight or nine colinear segments of sequence homology with their eukaryotic counterparts. The average amino acid identity in these segments is 35% and the average similarity is 70%. These relations are presented for β′ in figure 2 and β in figure 3. Intererstingly, many mutations specifically selected for their effects on termination lie in the universally conserved domains (Landick *et al.* 1990; Weilbaecher *et al.* 1994). At present, the interpretation of the distribution of such mutations is unclear. Note that the C-terminal domain of the largest subunit of eukaryotic pol II, which is probably an interaction site required for assembly of a pre-initiation complex capable of responding to activators, is missing in prokaryotic polymerase.

Four notable landmarks have been identified in β′ (figure two). A Cys_4-Zn finger-like motif is located near the N-terminus of the subunit in conserved region A (Weilbaecher *et al.* 1994). Such motifs are often involved in binding nucleic acid. Some evidence exists to support this idea. An *E. coli* bacteriophage, called HK022, has a unique antitermination system allowing readthrough of some of its terminators. Antitermination can be recreated *in vitro* with purified RNA polymerase (Clerget *et al.* 1995). Mutations abolishing antitermination map solely in the Cys_4-Zn finger-like motif, and have not been found to exhibit a general alteration in host transcription. Current thinking is that an RNA structure, unique to the mRNA of this bacteriophage interacts with the Cys_4-Zn finger-like motif and by an unknown mechanism, prevents termination by RNA polymerase. In addition, two mutations in this region of eukaryotic Pol I exhibit a ts phenotype, indicating the importance of this region for polymerase function (Wittekind *et al.* 1988). A weak homology to the DNA binding region of the Klenow fragment of DNA pol I has been noted in conserved region C (Allison *et al.* 1985). Indeed, two mutations altering residues within the region of homology destabilize the binding of RNA polymerase to promoter DNA during initiation, suggesting that this region of RNA polymerase may well interact with DNA (L. M. Heisler, G. Feng, D. J. Jin, C. A. Gross & R. Landick, unpublished results). Streptolydigin is a drug that binds to RNA polymerase and inhibits elongation. Until recently, mutations conferring streptolydigin resistance (stl^r) had been located exclusively in β (Heisler *et al.* 1993; Severinov *et al.* 1993). However, mutations conferring resistance to streptolydigin have now been found in conserved region F of β′ as well

Downloaded from rstb.royalsocietypublishing.org

(Severinov *et al.* 1995). Interestingly, these mutations are located close to mutations conferring amanitin resistance in eukaryotic pol II (Bartolomei & Corden 1987). Finally, the 3′ end of the nascent transcript crosslinks to conserved region G, suggesting that this region of RNA polymerase may form one half of the active site (Borukhov *et al.* 1991).

The structure–function analysis of β is considerably more advanced than that of β′, primarily because of the early finding that rifampicin inhibited RNA polymerase. All mutations conferring resistance to the antibiotic rifampicin are located in this subunit. Because these mutants had interesting phenotypes, obtaining and characterizing Rif^r mutants allowed directed mutagenesis before the recombinant DNA technology era. Mutations conferring Rif^r were found to be located in four distinct regions of β, including conserved Regions B and D (see figure 3) (Jin & Gross 1988; Severinov *et al.* 1994*b*). In the presence of rifampicin, RNA polymerase forms an open complex but adds only a few nucleotides. However, elongating RNA polymerase is resistant to rifampicin. These characteristics suggested that the rifampicin binding site was located close to the site of nucleotide addition. Indeed, recent physical experiments have verified this assumption. Rifampicin derivatives containing nucleotides spaced at different lengths from the drug indicate that bound rifampicin is within 15 Å of the initiating nucleotide and within 2 Å of the −2 and −3 position of template DNA (Mustaev *et al.* 1994). The positioning of rifampicin, close to and upstream of the nucleotide binding site suggests that it may block the channel leading nascent RNA from the active site. These experiments also suggest that the rifampicin binding site on RNA polymerase may interact with DNA. Consistent with this idea, one Rif^{r}mutation located in conserved Region D destabilizes RNA polymerase binding to the promoter (L. M. Heisler, G. Feng, D. J. Jin, C. A. Gross & R. Landick, unpublished results). Early experiments indicated that Stl^r mutations were also located in β,-although one report suggested that such mutations in *B. megaterium* were in β′. Stl^r mutations have now been mapped to four contiguous amino acids between two clusters of Rif^r mutations (figure 3) (Heisler *et al.* 1993). A variety of experiments suggest that this region of β is looped out and non-essential for enzyme function (Severinov *et al.* 1993). We return below to the implications of the finding that the binding site for streptolydigin spans both subunits.

A second major impetus for detailed study of β was the finding that three distinct regions in β are in close proximity to the 5′ end of the nascent transcript. Crosslinking indicates that lys1065 in conserved region H, his1237 in conserved region I (Mustaev *et al.* 1991) and a residue in between amino acids 516 and 540 (Cluster 1 of the Rif region; see figure 3) are close to the 5′ end of the nascent RNA transcript (Severinov *et al.* 1995*b*). Crosslinking indicates that lys1065 in conserved region H, his1237 in conserved region I (Mustaev *et al.* 1991) and a residue in conserved region D are close to the 5′ end of the nascent RNA transcript (A. Goldfarb, personal communication). Amino acids in the second largest subunit of eukaryotic RNA polymerases that are homologous to lys1065 also crosslink to the 5′ face of the nascent transcript, indicating that the relation between polymerase and transcript described in prokaryotes has been preserved throughout evolution. Although lys1065 is within 2 Å of the initiating nucleotide, it does not seem to be part of the catalytic site per se. Substituting alanine for lys1065 has only a minor effect on catalysis. Likewise, mutants with a variety of single and multiple substitutions between aa 1063 to 1073 still have significant $\geqslant 15\%$ transcriptional activity in multiple round transcription assays (Sagitov *et al.* 1993). However, such mutants do exhibit altered transcriptional properties including weakened binding of nucleotides at both the initiation and elongation sites of the enzyme, and aberrant patterns of abortive initiation, promoter clearance and pausing. Such defects are consistent with the idea that this region of polymerase is in close proximity to the active site although not directly part of this site.

(*b*) *Topology of the two largest subunits*

Currently we are faced with the challenge of going from information concerning the primary sequence to an understanding of the organization of RNA polymerase. A variety of experiments are leading to the viewpoint that the two largest subunits consist of independent modules which assemble to form RNA polymerase. Moreover, each of these modules is a conserved region and these may interact providing a framework for assembling the enzyme.

Homology alignments among prokaryotic and eukaryotic RNA polymerases, as well as other studies, suggested that each of the two largest subunits might be composed of at least three independently folding domains. This has now been experimentally verified for the β subunit of *E. coli* RNA polymerase. Fragments of β consisting of the N-terminal, central and C-terminal regions, engineered so that the splits occurred at so-called ‘dispensable’ regions of the subunit, assemble to RNA polymerase holoenzyme and carry out initiation and limited elongation (Severinov *et al.* 1994*a*). Similar experiments are underway to test the subdomain structure of β′. Thus our current view of RNA polymerase must include the idea that multiple independent domains can assemble into a functional RNA polymerase.

Other studies suggest that regions distant from each other in the linear map of the polypeptide are located close together in the tertiary structure. This is most clearly established for the regions of RNA polymerase that form the rifampicin binding site. Mutations at three different locations can mutate to give strong Rif^r, suggesting that these sites are close together in RNA polymerase (Jin & Gross 1988; Severinov *et al.* 1994*b*). One of these sites consists of a single position in region B where a mutational change at aa 142 confers Rif^r. The other two sites where mutational changes confer Rif^r occur in conserved region D, in the central fragment of β. The idea that regions D and B interact is further supported by the observation that an allele-

Downloaded from rstb.royalsocietypublishing.org

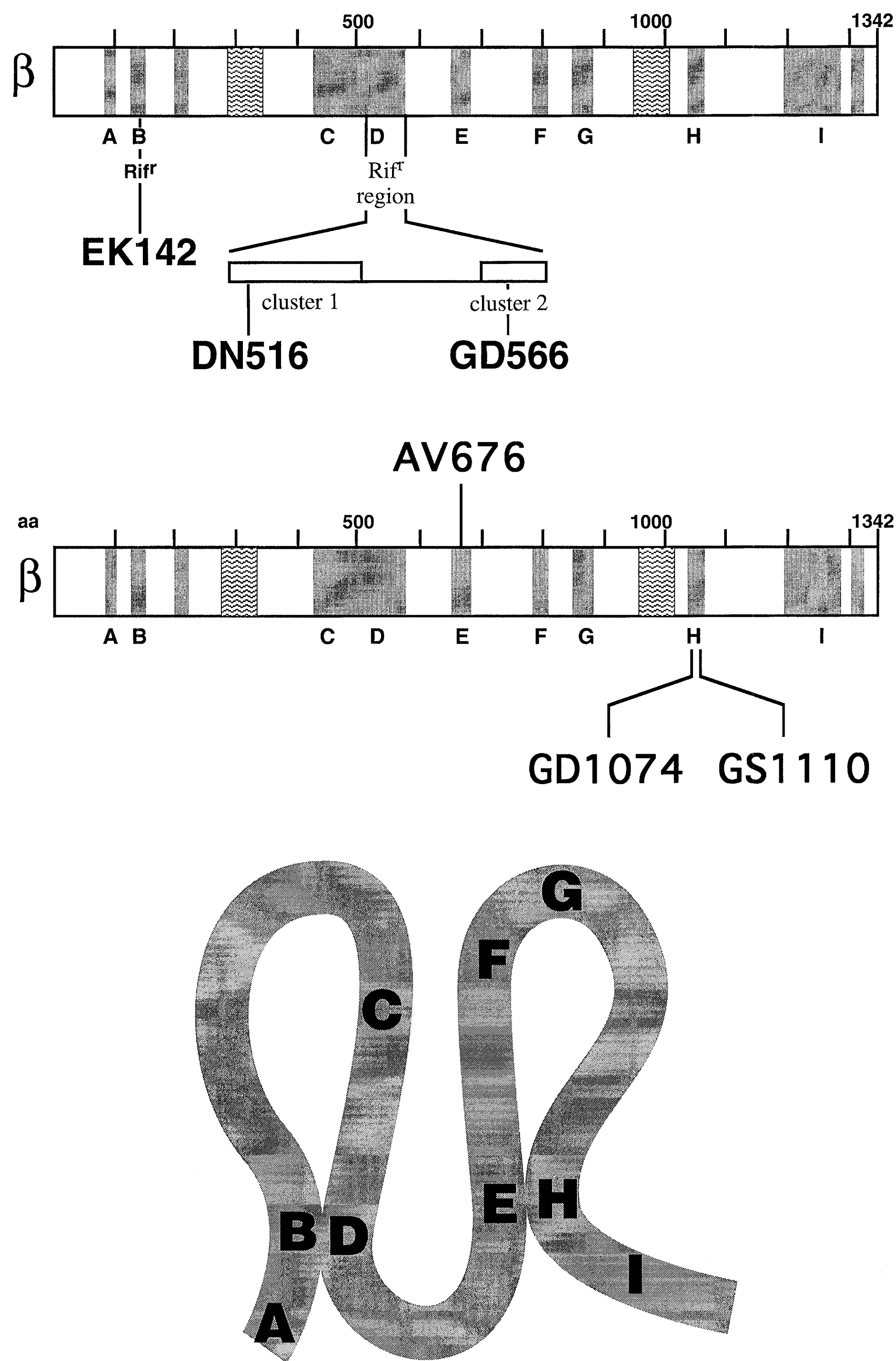


Figure 4. (*a*) Allele specific revertants of GD566. (*b*) Allele specific reverants of AV676. (*c*) Conserved regions of RNA polymerase may interact.

specific second-site revertant of the lethal mutation GD566 maps to aa 146, just 4 amino acids upstream of the site in the N-terminus at aa 142 that confers Rif^r (figure 4*a*) (Tavormina *et al.* 1996). Similar studies suggest that some portions of conserved regions E and H may also be in close proximity with each other (figure 4*b*) (Tavormina *et al.* 1996). The existing evidence for interactions within the β subunit is summarized schematically in figure 4*c*.

Evidence is now accumulating to suggest that regions of β and β′ are in close proximity as well. This is perhaps most clearly demonstrated for streptolydigin, where the occurrence of resistant mutants in both β and β′ suggests that the binding site is partially located in each subunit (Severinov *et al.* 1995). However, crosslinking data indicating that conserved portions of both β and β′ are located close to the 3′ face of the nascent transcript also suggests the cooperation of both

Downloaded from rstb.royalsocietypublishing.org

subunits in construction of the catalytic site (Borukhov *et al.* 1991). Taken together, these studies suggest that a variety of conserved regions in both subunits may interact to form the core of the enzyme. There is no doubt that both physical and genetic studies will explore these relations in greater detail in the near future.

4. THE σ SUBUNIT OF RNA POLYMERASE

The σ subunit of RNA polymerase converts the core enzyme capable of elongation and termination to a holoenzyme capable of specific initiation (Burgess *et al.* 1969; Gross *et al.* 1992). Thus understanding its function is crucial for understanding the transcription process. Prokaryotic organisms contain multiple σ factors. The primary σ, serves a housekeeping function, transcribing most of the genome. In contrast, the 5–10 alternative σ's generally serve a more specialized function, coordinating expression of small groups of genes in response to a variety of environmental stresses (Gross *et al.* 1992). Comparing the numerous sequenced members of this family indicates that most σ's share four conserved regions, which themselves can be further subdivided (figure 5) (Helmann & Chamberlin 1988; Lonetto *et al.* 1992). As opposed to β and β′, σ appears to be a modular protein, making both genetic and biochemical analysis of function easier. Although σ is unique to prokaryotes, it plays a role analogous in some respects to that of TATA binding protein in eukaryotes, which first binds to the TATA box and, in concert with several other proteins, directs RNA polymerase to promoter regions of the DNA.

From its discovery in 1969, it was clear that σ was a specificity factor permitting promoter recognition. Moreover, the fact that alternative σ's directed polymerase to promoters that differed at both the −10 and −35 conserved hexamers suggested that σ had two different DNA recognition domains (Losick & Pero 1981). Genetic analysis conducted on two different σ's were in accord with this idea. Mutations in a helix-turn-helix (HTH) motif located in region 4.2 of σ altered recognition of the −35 consensus hexamer whereas mutations in region 2.4 altered recognition of the −10 consensus hexamer. Taken together, these results allowed σ to be aligned linearly with respect to conserved sequence elements in the prokaryotic promoter (figure 5) (Helmann & Chamberlin 1988; Gardella *et al.* 1989; Siegele *et al* 1989; Zuber *et al.* 1989; Waldburger *et al.* 1990).

One problem with this view is that many σ's do not bind to DNA in the absence of core RNA polymerase, calling into question the simplest version of σ as a DNA binding protein. This discrepancy was solved by the finding that the N-terminal domain of several σ's inhibits their DNA binding (Dombroski *et al.* 1992, 1993). In the absence of this inhibitory domain, σ fragments exhibit the specific binding expected from the genetic experiments described above. Presumably, upon binding to core RNA polymerase, σ undergoes a conformational change, which results in exposing its DNA binding domains.

Primary σ's exhibit conservation downstream of the HTH in Region 4.2. This conserved region extends essentially to the end of the protein and is rich in basic amino acids, suggesting that it is surface exposed. The alignment of σ and the promoter presented in figure 5 suggests that this region is in proximity to sequences immediately upstream of the −35 region of the promoter. Several activators have binding sites in this vicinity, suggesting the possibility that the CTD of σ, like the CTD of α, might interact with some activator proteins. Indeed, the very first mutation described in σ, which changed an arginine residue in this region, is best explained by suggesting that it affects the interaction between σ and the AraC positive activator of transcription (Silverstone *et al.* 1972; Hu & Gross 1985). Further evidence suggests that this same residue (arg 596) also interacts with the cI activator encoded by bacteriophage λ (Li *et al.* 1994). To examine this point systematically, we have made a series of alanine substitution mutations in the C-terminus of σ. This work is currently in progress, but preliminary results indicate that several different activation proteins may interact with a number of positions in the σ-CTD (M. Lonetto & C. Gross, unpublished results).

At this point, very little is known about how σ binds to core RNA polymerase, although deletion analysis strongly implicates Region 2.1 in this process (Lesley & Burgess 1989). Binding to core not only exposes the DNA binding domain of σ, but also alters core RNA polymerase so that it binds less tightly to non-specific DNA. Much more needs to be learned about these interactions and their consequences. Interestingly, the RAP30 protein, which is one component of the eukaryotic transcription factor TFIIF, appears to decrease nonspecific DNA binding by eukaryotic PolII (Conaway & Conaway 1990). RAP30 shares several parallels with σ. The region of RAP30 that binds to PolII has been identified and it shows weak similarity to region 2.1 of σ (Sopta *et al.* 1989). Moreover, RAP30 has a structural plan similar to that of σ in that it has a C-terminal cryptic DNA binding domain. This domain is exposed when the N-terminal inhibitory domain is removed and is essential for RAP30 activity (Tan *et al.* 1994).

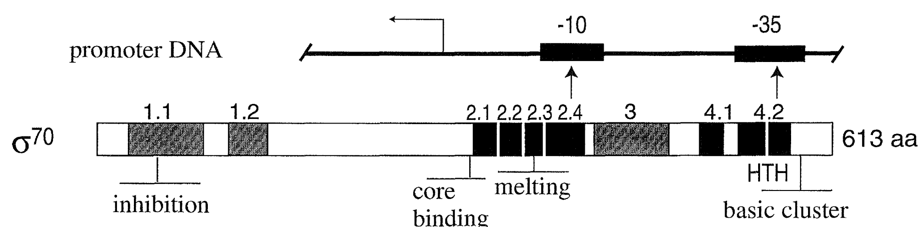


Figure 5. Conserved features and promoter contacts of sigma 70.

Downloaded from rstb.royalsocietypublishing.org

BIOLOGICAL SCIENCES THE ROYAL SOCIETY PHILOSOPHICAL TRANSACTIONS OF

5. SUMMARY AND PROSPECTS

This brief review of RNA polymerase structure-function relations indicates that a great deal of activity is currently directed at understanding the unique features of multisubunit transcriptases. As a result, we know something about the localization of a variety of activities of the enzyme. The challenge for the future will be to understand how the different parts of this protein machine fit together to carry out the processes of initiation, elongation and termination that are crucial to the flow of genetic information from genes to proteins.

REFERENCES

Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. 1985 Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**, 599–610.

Bartolomei, M. S. & Corden, J. L. 1987 Localization of an alpha-amanitin resistance mutation in the gene encoding the largest subunit of mouse RNA polymerase II. *Molec. Cell Biol.* **7**, 586–594.

Borukhov, S., Lee, J. & Goldfarb, A. 1991 Mapping of a contact for the RNA 3′ terminus in the largest subunit of RNA polymerase. *J. biol. Chem.* **266**, 23932–5.

Brenner, S., Jacob, F. & Meselson, M. 1961 An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature, Lond.* **190**, 576–580.

Buratowski, S. & Sharp, P. A. 1992 Initiation of transcription by RNA polymerase II. In *Transcriptional regulation* (ed. S. L. McKnight & K. R. Yamamoto), pp. 227–246. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Burgess, R. R., Travers, A. A., Dunn, J. J. & Bautz, E. K. F. 1969 Factor stimulating transcription by RNA polymerase. *Nature, Lond.* **221**, 43–46.

Busby, S. & Ebright, R. H. 1994 Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**, 743–746.

Chen, Y., Ebright, Y. W. & Ebright, R. H. 1994 Identification of the target of a transcription activator protein by protein-protein photocrosslinking. *Science, Wash.* **265**, 90–2.

Clerget, M., Jin, D. J. & Weisberg, R. A. 1995 A zinc-binding region in the β′ subunit of RNA polymerase is involveed in antitermination of early transcription of phage HK022. *J. Molec. Biol.* **248**, 768–780.

Conaway, J. W. & Conaway, R. C. 1990 An RNA polymerase II transcription factor shares functional properties with *Escherichia coli* sigma 70. *Science, Wash.* **248**, 1550–3.

Darst, S. A., Edwards, A. M., Kubalek, E. W. & Kornberg, R. D. 1991 Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell* **66**, 121–128.

Dombroski, A. J., Walter, W. A. & Gross, C. A. 1993 Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev.* **7**, 2446–55.

Dombroski, A. J., Walter, W. A., Record, M. T. Jr, Siegele, D. A. & Gross, C. A. 1992 Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**, 501–12.

Gardella, T., Moyle, H. & Susskind, M. M. 1989 A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J. Molec. Biol* **206**, 579–90.

Greenleaf, A. L. 1992 RNA polymerase II subunits, transcription factors, and kinases: Investigations in *Drosophila melanogaster* and *Sacharomyces cerevisiae*. In *Transcriptional regulation* (ed. S. L. McKnight & K. R. Yamamoto), pp. 55–80. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Gross, C. A., Lonetto, M. & Losick, R. 1992 Bacterial sigma factors. In *Transcriptional regulation* (ed. S. L. McKnight & K. R. Yamamoto), vol. 1, pp. 129–178. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Heisler, L. M., Suzuki, H., Landick, R. & Gross, C. A. 1993 Four contiguous amino acids define the target for streptolydigin resistance in the beta subunit of *Escherichia coli* RNA polymerase. *J. biol. Chem.* **268**, 25369–75.

Helmann, J. D. & Chamberlin, M. J. 1988 Structure and function of bacterial sigma factors. *A. Rev. Biochem.* **57**, 839–72.

Hu, J. C. & Gross, C. A. 1985 Mutations in the sigma subunit of *E. coli* RNA polymerase which affect positive control of transcription. *Molec. Gen. Genet.* **199**, 7–13.

Huang, R. C., Maheshwari, N. & Bonner, J. 1960 Enzymatic synthesis of RNA. *Biochem. Biophys. Res. Comm.* **3**, 689.

Hurwitz, J., Bresler, A. & Diringer, R. 1960 The enzymatic incorporation of ribonucleotides into plyribonucleotides and the effect of DNA. *Biochem. Biophys. Res. Comm.* **3**, 15.

Igarashi, K. & Ishihama, A. 1991 Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65**, 1015–22.

Jacob, F. & Monod, J. 1961 Genetic regulatory mechanisms in the synthesis of proteins. *J. Molec. Biol.* **3**, 318–356.

Jin, D. J. & Gross, C. A. 1988 Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J. Molec. Biol.* **202**, 45–58.

Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. 1994 A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599–608.

Kimura, M. & Ishihama, A. 1995 Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: insertion analysis of the amino-terminal assembly domain. *J. Molec. Biol.* **248**, 756–67.

Koleske, A. J. & Young, R. A. 1994 An RNA polymerase II holoenzyme responsive to activators [see comments]. *Nature, Lond.* **368**, 466–9.

Landick, R., Stewart, J. & Lee, D. N. 1990 Amino acid changes in conserved regions of the β-subunit of *Escherichia coli* RNA polymerase alter transcription pausing and termination. *Genes Dev.* **4**, 1623–36.

Lesley, S. A. & Burgess, R. R. 1989 Characterization of the *Escherichia coli* transcription factor sigma 70: localization of a region involved in the interaction with core RNA polymerase. *Biochemistry* **28**, 7728–34.

Li, M., Moyle, H. & Susskind, M. M. 1994 Target of the transcriptional activation function of phage lambda cI protein. *Science, Wash.* **263**, 75–77.

Lonetto, M., Gribskov, M. & Gross, C. A. 1992 The sigma 70 family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**, 3843–9.

Losick, R. & Pero, J. 1981 Cascades of sigma factors. *Cell* **25**, 582–584.

Maldonado, E. & Reinberg, D. 1995 News on initiation and elongation of transcription by RNA polymerase II. *Curr. Opin. Cell. Biol.* **7**, 352–61.

Mustaev, A., Kashlev, M., Lee, J. Y., Polyakov, A., Lebedev, A., Zalenskaya, K., Grachev, M., Goldfarb, A. & Nikiforov, V. 1991 Mapping of the priming substrate contacts in the active center of Escherichia coli RNA polymerase. *J. biol. Chem.* **266**, 23927–31.

Mustaev, A., Zaychikov, E., Severinov, K., Kashlev, M.,

Downloaded from rstb.royalsocietypublishing.org

Polyakov, A., Nikiforov, V. & Goldfarb, A. 1994 Topology of the RNA polymerase active center probed by chimeric rifampicin-nucleotide compounds. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12036–40.

Ossipow, V., Tassan, J. -P., Nigg, E. A. & Schibler, U. 1995 A mammalian RNA polymerase II holenzyme containing all components required for promoter-specific transcription initiation. *Cell* **83**, 137–146.

Polyakov, A., Severinova, E. & Darst, S. 1995 Three-dimensional structure of *E. coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* **83**, 365–373.

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. & Gourse, R. L. 1993 A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science, Wash.* **262**, 1407–13.

Sagitov, V., Nikiforov, V. & Goldfarb, A. 1993 Dominant lethal mutations near the 5′ substrate binding site affect RNA polymerase propagation. *J. biol. Chem.* **268**, 2195–202.

Sentenac, A., Riva, M., Thuriaux, P., Buhler, J. -M., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., Mann, C., Chiannilkulchai, N., Stettler, S. & Mariotte, S. 1992 Yeast RNA polymerase subunits and genes. In *Transcriptional regulation* (ed. S. L. McKnight & K. R. Yamamoto), pp. 27–54. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Severinov, K., Soushko, M., Goldfarb, A. & Nikiforov, V. 1993 Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of *Escherichia coli* RNA polymerase. *J. biol. Chem.* **268**, 14820–5.

Severinov, K., Kashlev, M., Severinova, E., Bass, I., McWilliams, K., Kutter, E., Nikiforov, V., Snyder, L. & Goldfarb, A. 1994*a* A non-essential domain of *Escherichia coli* RNA polymerase required for the action of the termination factor Alc. *J. biol. Chem.* **269**, 14254–9.

Severinov, K., Soushko, M., Goldfarb, A. & Nikiforov, V. 1994*b* RifR mutations in the beginning of the *Escherichia coli rpoB* gene. *Molec. Gen. Genet.* **244**, 120–6.

Severinov, K., Markov, D., Severinova, E., Nikiforov, V., Landick, R., Darst, S. A. & Goldfarb, A. 1995*a* Streptolydigin-resistant mutants in an evolutionarily conservered region of the β′ subunit of *Escherichia coli* RNA polymerase. *J. biol. Chem.* **270**, 23926–23929.

Severinov, K., Mustaev, A., Severinova, E., Kozlov, M., Darst, S. A. & Goldfarb, A. 1995*b* The β-subunit Rif cluster I is only angstroms away from the active centre of *Escherichia coli* RNA polymerase. *J. biol. Chem.* **270**, 29428–29432.

Siegele, D. A., Hu, J. C., Walter, W. A. & Gross, C. A. 1989 Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. *J. Molec. Biol.* **206**, 591–603.

Silverstone, A. E., Goman, M. & Scaife, J. G. 1972 ALT: a new factor involved in the synthesis of RNA by *Escherichia coli*. *Molec. Gen. Genet.* **118**, 223–234.

Sopta, M., Burton, Z. F. & Greenblatt, J. 1989 Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature, Lond.* **341**, 410–414.

Stevens, A. 1960 Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B. *Biochem. Biophys. Res. Comm.* **3**, 92.

Tan, S., Garrett, K. P., Conaway, R. C. & Conaway, J. W. 1994 Cryptic DNA-binding domain in the C terminus of RNA polymerase II general transcription factor RAP30. *Proc. natn. Acad. Sci. U.S.A.* **91**, 9808–12.

Tavormina, P. L., Reznikoff, W. S. & Gross, C. A. 1996 Identifying interacting regions in the β subunit of *Escherichia coli* RNA polymerase. *J. Molec. Biol.* (In the press.)

Thompson, C. M., Koleske, A. J., Chao, D. M. & Young, R. A. 1993 A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361–75.

Waldburger, C., Gardella, T., Wong, R. & Susskind, M. M. 1990 Changes in conserved region 2 of Escherichia coli sigma 70 affecting promoter recognition. *J. Molec. Biol.* **215**, 267–76.

Watson, J. D. & Crick, F. H. C. 1953 A structure for DNA. *Nature, Lond.* **171**, 736–738.

Weilbaecher, R., Hebron, C., Feng, G. & Landick, R. 1994 Termination-altering amino acid substitutions in the β′ subunit of *Escherichia coli* RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev.* **8**, 2913–27.

Weiss, S. B. & Gladstone, L. 1959 A mammalian system for the incorporation of cytidine and triphosphate into ribonucleic acid. *J. Am. Chem. Soc.* **81**, 4118.

Weiss, S. B. & Nakamoto, T. 1961 A net syntheis of ribonucleic acid with a microbial enzyme requiring deoxyribonucleic acid and four ribonucleoside triphosphates. *J. biol. Chem.* **236**, PC 19.

Wittekind, M., Dodd, J., Vu, L., Kolb, J. M., Buhler, J. M., Sentenac, A. & Nomura, M. 1988 Isolation and characterization of temperature-sensitive mutations in RPA190, the gene encoding the largest subunit of RNA polymerase I from *Saccharomyces cerevisiae*. *Molec. Cell Biol.* **8**, 3997–4008.

Zillig, W., Palm, P. & Heil, A. 1976 Function and reassembly of subunits of DNA-dependent RNA polymerase. In *Polymerase* (ed. R. Losick & M. Chamberlin), pp. 101–125. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Zuber, P., Healy, J., Carter, H. L. D., Cutting, S., Moran, C. P. Jr & Losick, R. 1989 Mutation changing the specificity of an RNA polymerase sigma factor. *J. Molec. Biol.* **206**, 605–14.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

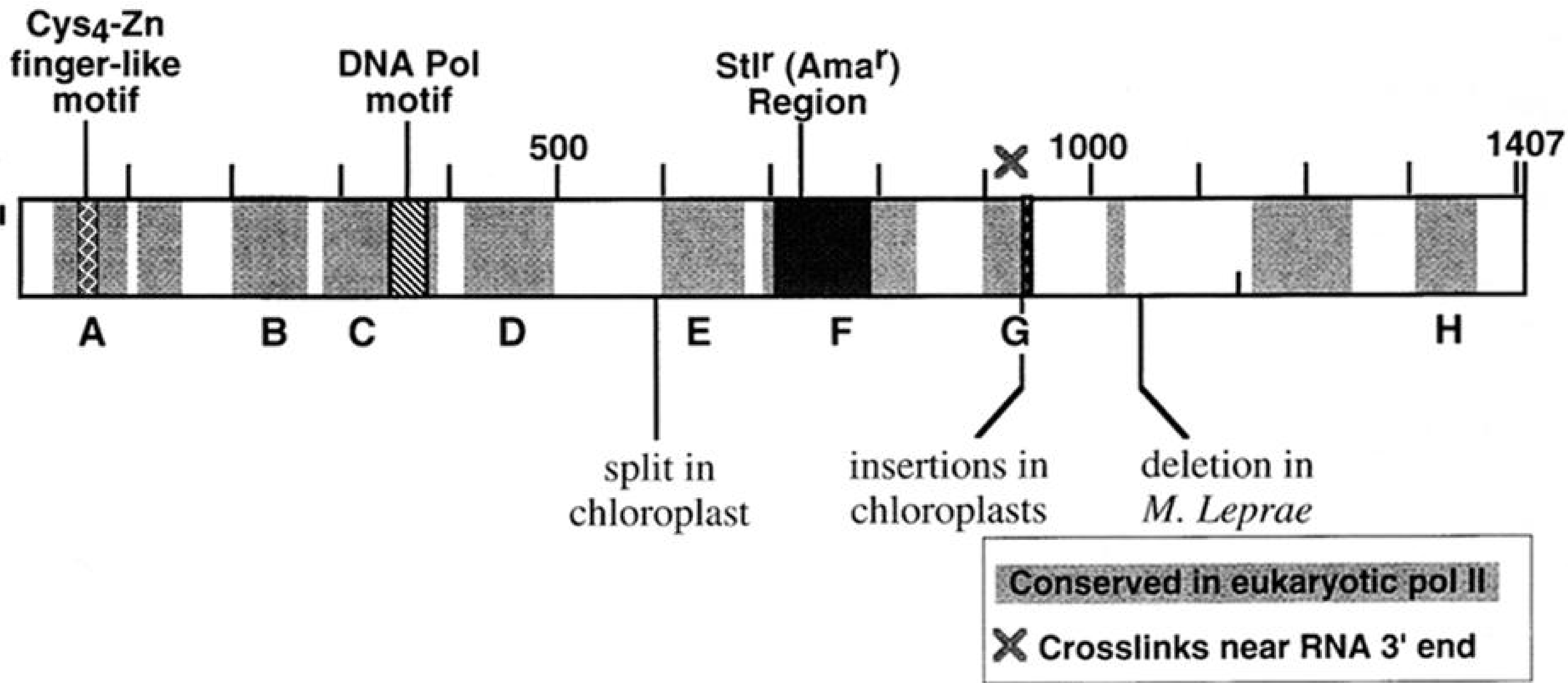
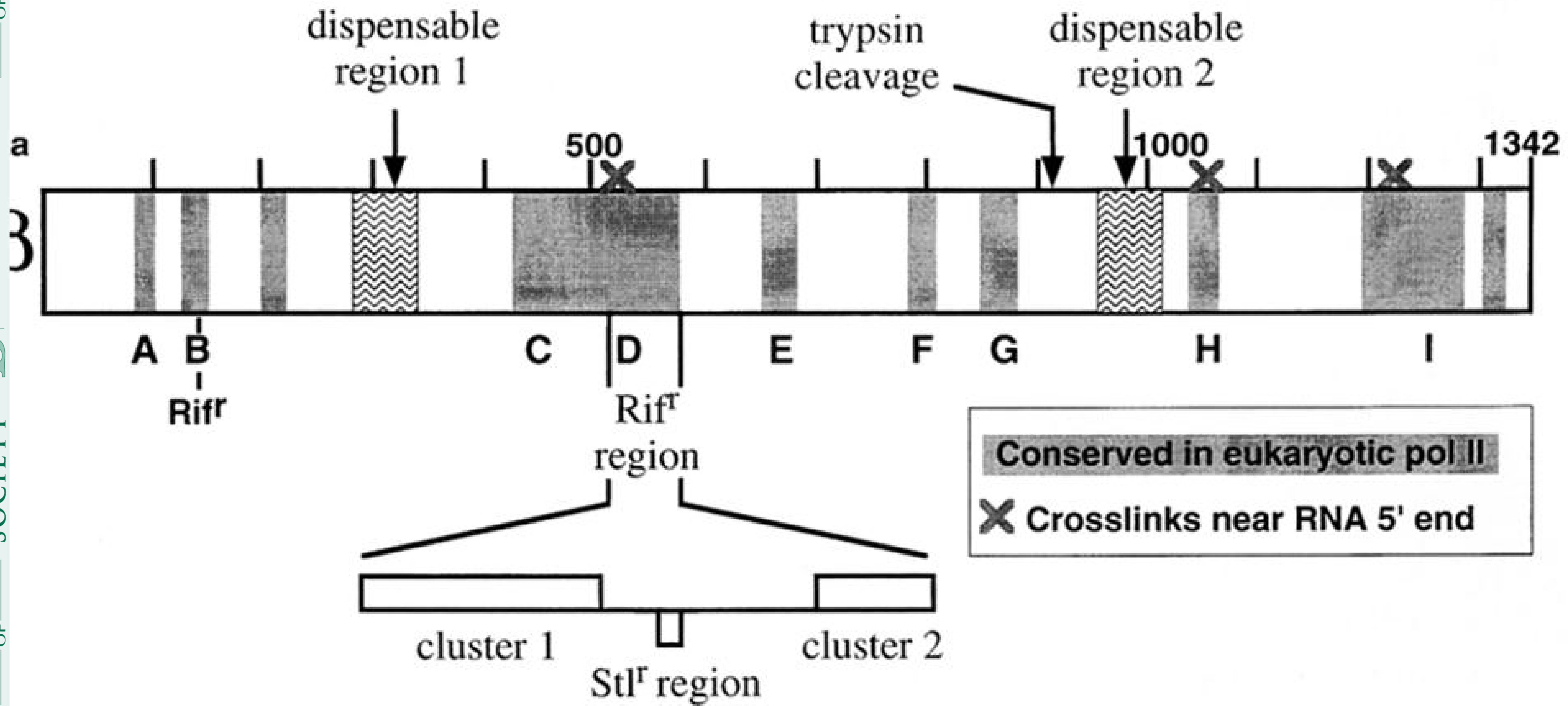


Figure 2. Conserved features of β′, the largest subunit of *E. coli* RNA polymerase.

PHILOSOPHICAL TRANSACTIONS OF
THE ROYAL SOCIETY
BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF

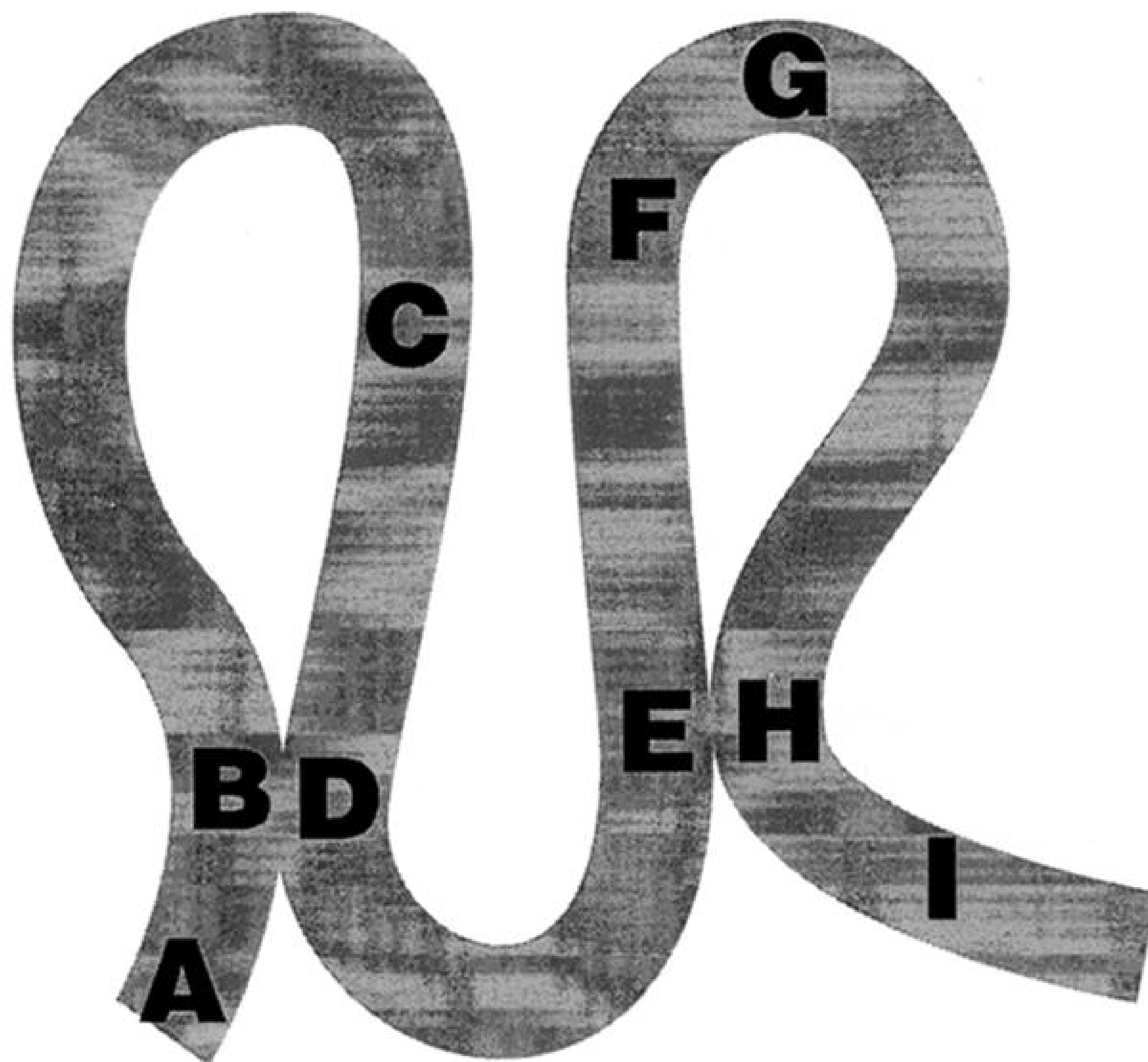


igure 3. Conserved features of β, the second largest subunit of *E. coli* RNA polymerase.

500 1000 1342
β
A B C D E F G H I
Rif^r
Rif^r region
EK142
cluster 1 cluster 2
DN516 GD566

AV676
aa 500 1000 1342
β
A B C D E F G H I
GD1074 GS1110

Downloaded from rstb.royalsocietypublishing.org



PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

gure 4. (*a*) Allele specific revertants of GD566. (*b*) Allele specific reverants of AV676. (*c*) Conserved regions of
NA polymerase may interact.