Sigma Factor Relatives in Eukaryotes

JUDITH A. JAEHNING

Sigma factor is a protein required by the bacterial core RNA polymerase for promoter recognition and initiation of transcription (*1*). Bacteria and bacteriophage use multiple sigma factors to modify RNA polymerase and control temporal and spatial expression of subsets of genes (*1, 2*). There is little overall sequence similarity among sigma factors, but shared functions have been localized to conserved subdomains (*1*). These functions include interaction with core RNA polymerase, contribution to enzyme stability, suppression of nonspecific initiation, recognition of promoter elements (−10 and −35 base pairs from the initiation site), and melting of the DNA at the transcriptional start site. Certain structural and functional features of sigma factor have been conserved in the eukaryotic nuclear and mitochondrial transcription apparatus (Fig. 1).

Three different nuclear mRNA transcription factors have sequence similarity with different parts of the major *Escherichia coli* sigma factor, σ70. The protein RPO24 is the fourth largest subunit of yeast RNA polymerase II (*3*). RPO24 has a domain with sequence similarity to a domain unique to σ70 (*1*). Although the function of the domain is unknown, a mutation in this region of σ70 leads to RNA polymerase instability at extreme temperatures (*1*). Deletion of the gene in yeast that encodes RPO24 leads to a similar conditional phenotype (*3*), implying that the function of this domain may be similar in its role as a RNA polymerase subunit.

The human transcription factor RAP30 was isolated as part of a complex of two proteins that bind tightly to RNA polymerase II (*4*). RAP30 binds to core bacterial polymerase and to RNA polymerase II with comparable affinities (*5*). In addition, the rat homolog of the RAP30-containing complex suppresses nonspecific initiation events (*6*). Subdomains important for these two functions have been localized in σ70 (Fig. 1). Regions similar to these subdomains are also found in RAP30 (*4*).

TFIID, the eukaryotic nuclear protein that binds to the TATA element in promoters of mRNA-encoding genes, has similarity to domain 2 of σ70 (*7*), which is important for recognition of the −10 promoter sequence (Fig. 1). Although the similarity is not striking, the amino acids in this part of TFIID are critical for function (*8*).

The fact that eukaryotic transcription factors resemble prokaryotic sigma factors is consistent with the observation that the two largest subunits of eukaryotic nuclear RNA polymerases are homologs of the large subunits of bacterial RNA polymerases (*9*). In contrast, the core polymerase found in yeast mitochondria is similar to RNA polymerases from phage T7 and T3 (*10*). Despite the differences in the structure of these enzymes, the mitochondrial RNA polymerase has a specificity factor, MTF1, with sequence similarity to domains 2 and 3 of the family of sigma factors (*11*) (Fig. 1). MTF1 confers promoter recognition and selective initiation properties on the yeast mitochondrial RNA polymerase (*12*), consistent with the −10 promoter recognition and DNA melting properties of the σ70 sequences it shares.

That both TFIID and MTF1 contain regions similar to the domain of σ70 that recognizes the −10 element is of interest as their consensus DNA sequences are related to each other (TATAAA and ATATAAGTA, respectively) and to the consensus sequence of *E. coli* (TATAAT). Examples of cross-recognition of promoters by these RNA polymerases have been observed, including a case where a single sequence serves as a promoter for all three enzymes (*12*). This apparent conservation of protein structure and target sequence may be important in the horizontal movement of genes between prokaryotes and eukaryotes, or from organelles to the nucleus.

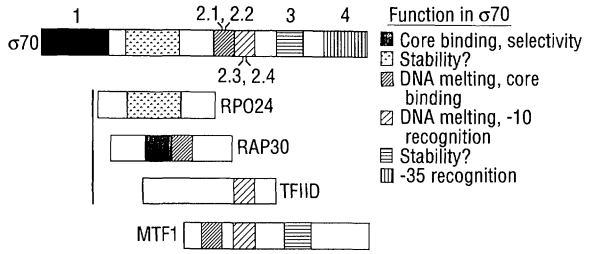

Fig. 1. The boxed domains in *E. coli* σ70, and their possible functions, are adapted from (*1*). The variously shaded boxes indicate regions of amino acid similarity in RPO24 (*3*), RAP30 (*4*), TFIID (*7*), and MTF1 (*11*); the proteins are drawn to scale with σ70 (620 amino acids).

Although mitochondrial RNA polymerase does not require additional promoter elements or protein factors (*12*), RNA polymerase II transcription is affected by many protein factors bound to sites upstream and downstream from the start site of transcription (*9*). This dependence on additional protein factors has been used as an argument for fundamental differences between mechanisms of prokaryotic and eukaryotic gene expression. However, the fact that several eukaryotic transcription factors resemble sigma and the demonstration that some prokaryotic genes are regulated by upstream elements that act like eukaryotic enhancers (*2*) suggest that the mechanisms may be more similar than different.

The distribution of sigma-like functions among a number of polypeptides is not unique to eukaryotes. Many bacterial sigma factors lack one or more of the conserved domains (*1*). The *E. coli* σ54 protein, which is involved in the expression of genes with diverse physiological functions (*2*), lacks amino acid sequences similar to domains 2.1 and 2.3 of σ70 (Fig. 1). These domains are important for melting the DNA at the start site of transcription. Transcription of σ54-dependent promoters requires binding of the phosphorylated form of an activator protein to a distant regulatory site. Interaction between this activator protein and σ54 leads to promoter melting and initiation (*2*).

By distributing the functions of sigma factor among several polypeptides, organisms gain increased potential for regulation. Having three or more eukaryotic proteins involved in template commitment, isomerization, and start site selection predicts that different regulatory factors may synergistically control different steps of the initiation process. Multiple bacterial sigma factors were not identified until diverse promoters were used for transcription reconstitution experiments (*1*). Because relatively few eukaryotic promoters have been studied in highly purified systems, it seems likely that multiple forms of general eukaryotic transcription factors involved in the control of subsets of eukaryotic genes may be identified.

Department of Biology, Indiana University, Bloomington, IN 47405.

REFERENCES

1. J. D. Helman and M. J. Chamberlin, *Annu. Rev. Biochem.* **57**, 839 (1988).
2. S. Kustu *et al.*, *Micro. Rev.* **53**, 367 (1989).
3. N. A. Woychik and R. A. Young, *Mol. Cell. Biol.* **9**, 2854 (1989).
4. M. Sopta, Z. F. Burton, J. Greenblatt, *Nature* **341**, 410 (1989).
5. S. McCracken and J. Greenblatt, *Science* **253**, 900 (1991).
6. J. W. Conaway and R. C. Conaway, *ibid.* **248**, 1550 (1990).
7. M. Horikoshi *et al.*, *Nature* **341**, 299 (1989).
8. P. Reddy and S. Hahn, *Cell* **65**, 349 (1991).
9. M. Sawadago and A. Sentenac, *Annu. Rev. Biochem.* **59**, 711 (1990).
10. B. S. Masters, L. L. Stohl, D. A. Clayton, *Cell* **51**, 89 (1987).
11. S. H. Jang and J. A. Jaehning, unpublished data.
12. C. S. Winkley, M. J. Keller, J. A. Jaehning, *J. Biol. Chem.* **260**, 14214 (1985).