Volume 307, number 3, 245–248 FEBS 11341 August 1992
© 1992 Federation of European Biochemical Societies 00145793/92/$5.00

Hypothesis

DNA intervention in transcriptional activation

Purnima Bhargava and Dipankar Chatterji

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

Received 26 May 1992

Accurate initiation of eukaryotic mRNA synthesis takes place as a result of the interplay between general transcription factors and RNA polymerase II. Activation of transcription from the basal level involves a number of promoter-specific *trans*-acting factors which interact with *cis* elements in the promoter DNA. In this paper we have emphasized the importance of even those portions of the promoter stretch which do not have any identifiable binding sites for regulatory proteins. The length and structure of the DNA between cognate binding sites of *trans*-acting factors may interfere with the level of transcriptional activation. Depending upon the length of the intervening DNA we describe three cases of transcriptional activation. In addition, based on this classification we propose a new third domain, the other two being DNA binding and transcriptional activation domains, which is involved in bending the intervening DNA so that activation from a distance can take place successfully.

Transcription; Activation; Looping; Bending; Mediator; Co-activator

1. INTRODUCTION

Eukaryotic protein encoding genes are subjected to a wide range of regulatory mechanisms that lead to the production of mature, biologically active proteins. The enzyme RNA polymerase II responsible for transcription of these genes does not transcribe them faithfully without the help of various essential 'general' or 'promoter-specific' transcription factors [1]. While general transcription factors (TFII A, B, D, E and F) are proposed to promote accurate initiation of basal transcription, specific factors are responsible for the fine tuning of the overall rate of expression [2]. The effects of the specific factors are transmitted through the finally assembled transcription initiation complex and its ensuing activity via basically two types of interactions, i.e. DNA–protein and protein–protein interactions. The involvement of a relatively large number of polypeptides (including 10–12 subunits of RNA polymerase) in the overall transcription process makes eukaryotic transcription mechanism complex. While simplicity of basal transcription in prokaryotes made the rapid growth of research on their regulatory mechanisms possible, the research on both levels of eukaryotic transcription is still in its infancy. There are currently two lines of thinking: one suggesting novel mechanisms unique to eukaryotes and the other one depicting the same principles of regulatory mechanisms as in prokaryotes but at a higher level of evolutionary scale. Two recent articles [3,4] aptly describe the lead which available knowledge on prokaryotic transcription mechanisms can give to the future exploration of mechanisms exploited by eukaryotic cells.

Correspondence address: D. Chatterji, Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India. Fax: (91) (842) 851 195.

2. ROLE OF DNA IN TRANSCRIPTIONAL ACTIVATION

Recent research has established the key importance of general transcription factors [5,6] in assembling an active transcription initiation complex and an approximate order of assembly of these factors with RNA polymerase II to form the initiation complex at the 'core promoter' (TATA box and/or initiator element) is slowly being worked out [7,8]. Binding of various promoter specific factors at specific sites activates transcription, presumably by facilitating the assembly and/or stability of the active initiation complex at the core promoter elements. It is proposed that protein–protein contacts between activators and basal factors are needed for this activation. Thus, DNA is thought to play only a passive role in transcriptional activation, major credit going to the *trans*-acting factors. However, the recent research suggests that not only the cis-elements but also their relative positions play an important role in transcriptional activation. Both the sequence and structure of these DNA stretches have significant effects on protein binding [9,10]. The binding sites for promoter specific factors are found located at variable distances from the site of initiation complex assembly at core promoter elements. For example, the binding sites

for mammalian transcription activator Spl are normally found at 30–70 bp away from the TATA box [11] while binding site of CBP (CCAAT binding protein), is usually located 50–100 bp upstream from the mRNA initiation site [12]. Upstream activation sequences (UASs) of yeast can be found at 20–1000 bp upstream from the TATA box [13] while their mammalian counterparts, the enhancers, are found even thousands of bases upstream or downstream the transcription start site. In order to achieve a fruitful activation through interplay of protein factors binding at these sites, they have to be brought together. Several possible mechanisms have been proposed [14] of which 'looping of intervening DNA' has emerged as the most likely one. This mechanism has been experimentally demonstrated to operate in several cases and has been used to describe interactions of proteins bound at distant DNA sites in several cases, even when no direct experimental demonstration of 'looping' is presented. However, by analogy with examples from prokaryotes and eukaryotes it is likely that 'looping' is not an entirely random process and is probably used by cells as a well-planned mechanism governed by certain rules which decide whether looping of intervening DNA will be used to facilitate a particular protein–protein interaction. These 'rules' are described by properties of individual proteins, the distance between their binding sites and probably some sequence elements in intervening DNA.

3. PROMOTER ORGANIZATION AS A MEANS OF REGULATION OF GENE EXPRESSION

Eukaryotic constitutive gene expression is like prokaryotic $\sigma 70$-dependent transcription. Assembly of initiation complex at the strong basal elements of these promoters gives a high level of basal transcription. In all regulated genes, where activation is required, regulatory sequences are placed at distances which increase as further tuning is required. Thus a greater number of proteins and types of protein–protein interactions are required. A look at some of the known examples of transcriptional activation, cited in Table I, shows that in spite of being from a wide variety of sources, the length of DNA, between the binding sites of trans-acting factors and the transcription machinery at the site of initiation complex formation falls in the same range. A more systematic analysis of such protein–protein interactions of all the known examples [2] in addition to the ones given in Table I, allows one to distinguish partially overlapping three groups. Thus, it seems that in order to be an efficient activator, a protein needs to have certain properties which depend on the structure and organization of the target promoter DNA.

Table I

Involvement of a third protein in transcriptional activation by some activators via protein–protein interaction

Protein 1	Protein 2	Intervening DNA length (bp)	A mediator?	Ref.
CAP (*lac*)	*E. coli* RNA polymerase	51	–	16
CAP (*gal*)	„	31	–	16
CBP	TFIID	20-70	?	32
Spl	TFIID	30-70	Yes	21
USF	TFIID	60-70	Yes	22
NTF1	TBP	65-85	Yes	23
ATF	TFIID	40-80	Yes (Ela)	33,34
GAL 4	TFIID	110-210	Direct (or GAL 11)	28,29
Oct. 1	TFIID	175-185	Yes (VP16)	35,36

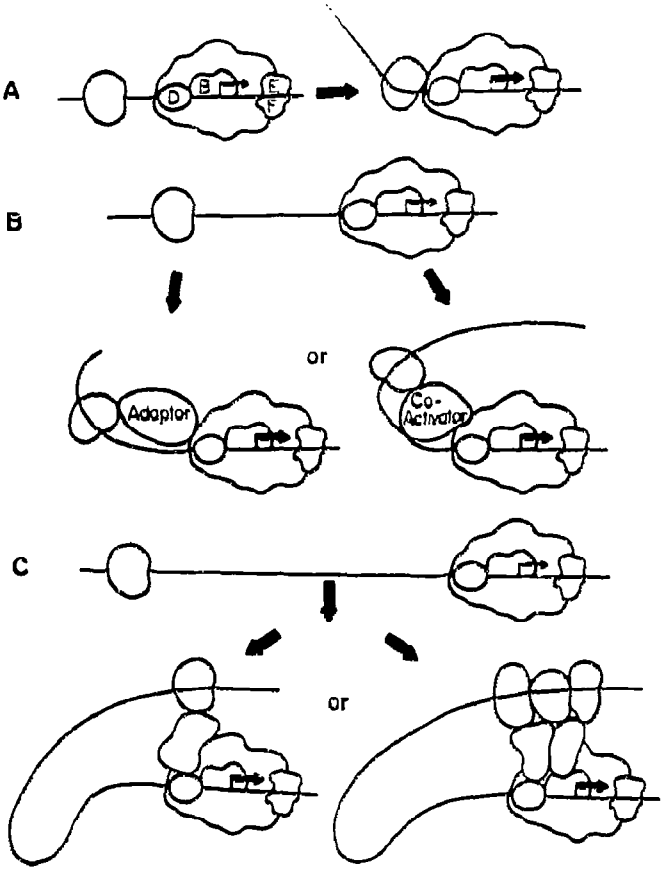

Fig. 1. Bending of intervening DNA and transcriptional activation. Size and shape of the proteins are chosen at random and do not mean to represent any special conformation. Initiation complex of RNA polymerase II with general transcription factors, TFII B, D, E, F is shown sitting near mRNA initiation site. A, B, C differ in length of intervening DNA. Single activator protein is shown at an upstream position and arrow represents the direction of transcription from +1 site.

3.1. *Proteins making direct contact with the initiation complex*

A slight bend of a very short intervening sequence may be enough to bring two very closely spaced DNA bound proteins, close enough to make contact (Fig.

1A). Examples include interaction of the general factors involved in the formation of the initiation complex with the RNA polymerase II [15] and the interaction of prokaryotic RNA polymerase with transcription activators like CAP [16,17]. In fact it is now known that there is a fixed phase relationship along the DNA between the binding of CAP and RNA polymerase [18] and only functional spacings are exactly those found in natural promoters [19].

3.2. *Proteins making contact with the initiation complex from a moderate distance*

Most of the known examples of eukaryotic transcription activators fall in this category. Binding sites of these promoter specific activators is usually located within 70–80 bp away from the TATA box (Table I) and most of these activators are proposed to need a mediator as an 'adaptor' or a 'co-activator', to contact the transcription apparatus [20–23]. In these cases a simple bending of DNA at the binding site of the activator or intervening sequence curvature may not be enough to bring the activator close to the initiation complex. Since DNA of ten to a few hundred base pairs behaves more like stiff rods than a flexible chain [24], bending needs to be induced either through protein binding or by inherent properties of the DNA chain. Thus, a mediator could help establish not only protein–protein contact by bridging the gap (adaptor) but also induce DNA bending (co-activator). The activator itself may also induce DNA bending but DNA bending may not suffice to bring the activator close to the initiation complex. Mediator may induce bending either through complexing with activator or by simultaneously contacting the intervening sequence to induce the bend (Fig. 1B).

3.3. *Proteins making contact with initiation complex from a large distance*

Best examples of looping mechanism are the interactions between enhancer binding proteins and proximal or core promoter binding proteins. Long DNA fragments of several hundred to thousands of base pairs behave like flexible chains, making spontaneous looping feasible [24] but an induced 'bending' per se may be required to initiate the looping. Enhancers usually have multiple binding sites for multiple proteins. For example, the best characterised SV40 enhancer has an overlapping array of a number of binding sites for various protein factors. Each of these proteins can have its own effect on DNA conformation or on proteins with which it interacts. A mosaic of protein–protein interactions can be established which not only leads to activation but also regulates the degree of activation. Because of the long distance involved, multiple proteins probably act in a cooperative manner to induce sufficient bending (Fig. 1C). Thus requirement of additional bending may be used to increase the scope of the further tuning of regulation through further diversification of the proteins involved. In these cases it would seem that a mediator must be involved in activation. The role of the mediator may be the same as discussed under the second case above. However, to bridge the larger steric gap in this case, more than one mediator may be involved in series. This possibility can be exploited to further regulate transcriptional activation, since a distribution of a number of contact domains over an array of proteins which may be needed for final activation is achieved this way. Thus, activation through enhancers, though from a larger distance, probably provides a means of very fine regulation, The intermediatory mediators may be tissue- or physiological state-specific; enhancers themselves have been shown to be responsible for tissue specific expression of certain genes [25].

Thus, we find that requirement for a mediator and bending of intervening DNA sequence vary in opposite manner. Looping can be spontaneous for an activation from far off sites but activation from shorter distances needs induced bending. For moderate distances, one mediator may be enough but longer distance may need more than one mediator. Thus a mediator is a protein required to establish successful contact between an activator and transcription machinery either through co-activation or through action as an adaptor (Fig. 1B). Activator itself may need to have in addition to known DNA binding and activation domains, a domain in its structure which induces a bend in DNA (or the DNA binding domain itself may induce bending) or a domain which becomes functional for bending after complexing with the mediator. This is a property for activators and mediators which has not been given due consideration till now. However, recently some reports have appeared which show some activator proteins have DNA bending properties [26,27].

The well-characterized *trans*-activator of gal family genes of *Saccharomyces cerevisiae*, GAL4 is proposed to activate transcription by directly making contact with TFIID [28]. Gal 11 is proposed to work as co-activator for GAL 4 and it is also demonstrated to interact with zinc finger DNA binding domain of some yeast activator proteins [29]. Thus it is quite possible that GAL 11 works through imposing DNA bending properties over the activators for which it works as co-activator. This hypothesis however needs to be demonstrated experimentally. Incidentally GAL 4 is the only example cited in Table I which does not have a mediator suggested for its action and which binds more than 70 bp away from TATA box. Similarly, with further research more and more number of examples are appearing which show that many regulatory proteins and transcription factors can bend DNA after binding with it [26,27,30,31]. Most recent is the study showing that the general transcription factor TFIID can bend the DNA at its cognate binding site [15].

Therefore we hypothesize:

(1) for activation from a distance below 200 bp induced

bending of DNA seems to be a more likely mechanism with or without involvement of a mediator;

(2) for activations from far off distances (~1000 bp) looping is more easy, although bending may be a pre-requisite. Involvement of one or more mediators is highly probable;

(3) most of activators may have a property to induce a bend over DNA either by themselves or after complexing with a mediator. On the other hand, there may be certain sequences present in intervening DNA which give it a curvature or make it bendable, especially in case of activation from shorter distances. This may thus explain promoter context dependence of activation by certain activators.

Future research may thus focus on these aspects of transcriptional activation in eukaryotes and new mechanisms of activation and repression of transcription may be described when above-mentioned hypotheses are experimentally proved or disproved!

Lastly, we would like to point out that looping is only a simple possibility the DNA can adopt as a consequence of bending in two dimension. However, other conformations may arise if one considers DNA bending in three-dimensional space. It is yet to be seen if any theoretical approach that would account for this type of bending, can give rise to interesting situation in DNA–protein recognition.

REFERENCES

[1] Sawadogo, M. and Sentenac, A. (1990) Annu. Rev. Biochem. 59, 711–754.
[2] Johnson, P.F. and McKnight, S.L. (1989) Annu. Rev. Biochem. 58, 799–839.
[3] Gralla, J.D. (1991) Cell 66, 415–418.
[4] Jachning, J.A. (1991) Science 253, 859.
[5] Greenblatt, J. (1991) Cell 66, 1067–1070.
[6] Hawley, D. (1991) Trends Biochem. Sci. 16, 317–318.
[7] Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Cell 56, 549–561.
[8] Buratowski, S., Sopta, M., Greenblatt, J. and Sharp, P.A. (1991) Proc. Natl. Acad. Sci. USA 88, 7509–7513.
[9] Travers, A.A. (1989) Annu. Rev. Biochem. 58, 427–452.
[10] Hagerman, P.J. (1990) Annu. Rev. Biochem. 59, 755–781.
[11] Dynan, W.S. and Tjian, R. (1985) Nature 316, 774–778.
[12] Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) Trends Biochem. Sci. 11, 20–23.
[13] Struhl, K. (1989) Annu. Rev. Biochem. 58, 427–452.
[14] Ptashne, M. (1986) Nature 322, 697–701.
[15] Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T. and Roeder, G. (1992) Proc. Natl. Acad. Sci. USA 89, 1060–1064.
[16] Zinkel, S.S. and Crothers, D.M. (1991) J. Mol. Biol. 219, 201–215.
[17] Schultz, S.C., Shields, G.C. and Seitz, T.A. (1991) Science 253, 1001.
[18] Straney, D.C., Straney, S.B. and Crothers, D.M. (1989) J. Mol. Biol. 206, 41–57.
[19] Gaston, K., Bell, A., Kolb, A., Buc, H. and Busby, S. (1990) Cell 62, 733–743.
[20] Kelleher III, R.J., Flanagan, P.M. and Kornberg, R.D. (1990) Cell 61, 1209–1215.
[21] Pugh, B.F. and Tjian, R. (1990) Cell 61, 1187–1197.
[22] Meisterernst, M., Roy, A.L., Lieu, H.M. and Roeder, R.G. (1991) Cell 66, 981–993.
[23] Dynlacht, B.D., Hoey, T. and Tjian, R. (1991) Cell 66, 563–576.
[24] Wang, J.C. and Giaever, G.N. (1988) Science 240, 300–304.
[25] Lübbe, A. and Schaffner, W. (1985) Trends Neuro Sci. 8, 100–104.
[26] Kerppola, T.K. and Curran, T. (1991) Cell 66, 317–326.
[27] Su, W., Jackson, S., Tjian, R. and Echols, H. (1991) Genes Dev. 5, 820–826.
[28] Horikoshi, M., Carey, M.F., Kakidans, H. and Roeder, R.G. (1988) Cell 54, 665–669.
[29] Himmelfarb, H.J., Pearlberg, J., Last, D.H. and Ptashne, M. (1990) Cell 63, 1299–1309.
[30] Schreck, R., Zorbas, H., Winnacker, E.-L. and Baeuerle, P.A. (1990) Nucleic Acids Res. 18, 6497–6502.
[31] Verrijzer, C.P., van Oosterhout, J.A.W.M., van Weperen, W.W. and van der Vliet, P.C. (1991) EMBO J. 10, 3007–3014.
[32] Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) Nature 334, 218–224.
[33] Horikoshi, M., Hai, T., LIn, Y.-S., Green, M.R. and Roeder, R.G. (1988) Cell 54, 1033–1042.
[34] Horikoshi, M., Maguire, K., Kralli, A., Maldonado, E., Reinberg, D. and Weinmann, R. (1991) Proc. Natl. Acad. Sci. USA 88, 5124–5128.
[35] Croston, G.E., Kerrigon, L.A., Lira, L.M., Marshak, D.R. and Kadonaga, J.T. (1991) Science 251, 643–649.
[36] Lin, Y.-S. and Green, M.R. (1991) Cell 64, 971–981.