Int. J. Exp. Path. (1993) **74**, 417–422

Current Status Review

Transcription factors: an overview

DAVID S. LATCHMAN

Division of Molecular Pathology, University College London Medical School, The Windeyer Building, Cleveland Street, London W1P 6DB, UK

Function of transcription factors

The process of transcription is the first stage of gene expression resulting in the production of a primary RNA transcript from the DNA of a particular gene. It therefore represents a critical first step in gene expression which is followed by a number of post-transcriptional processes such as RNA splicing and translation. These lead ultimately to the production of a functional protein. Moreover, as well as this central role in gene expression in general, transcription also plays a part in specificity, since it is the primary target for the process of gene regulation, which results in different proteins being produced in different tissues. In fact, whilst some cases of regulation after transcription do exist, in most cases selection occurs at this level by deciding which genes will be transcribed into the primary RNA transcript (Darnell 1982; Latchman 1990a). It is only following the gene transcription process that the other stages of gene expression such as RNA splicing occur automatically and result in the production of the corresponding protein.

Both basal transcription and its regulation are dependent upon specific protein factors known as transcription factors. These bind to particular DNA sequences in gene regulatory regions and control their transcription. These transcription factors are commonly classified into families on the basis of the precise protein structure which they use to mediate binding to DNA or to cause factor dimerization which is often essential for DNA binding. A list of DNA binding and dimerization motifs is given in Table 1 (for more detailed reviews see Latchman 1990b; 1991). Rather than consider the structure of these motifs in detail, this review will focus primarily on the manner in which DNA binding of these factors affects gene expression and on how their activity is regulated to produce tissue specific patterns of gene transcription.

Activity of transcription factors

Although binding to DNA is obviously a necessary prerequisite for a factor to affect transcription, it is not in itself sufficient. Thus, following DNA binding the factor must interact with other factors or the RNA polymerase itself in order to modulate transcription. Although such an interaction very often results in the activation of transcription, a number of cases have now been described in which factor binding results in transcriptional repression. Activation and repression of transcription will therefore be discussed in turn.

Activation

The ability of a transcription factor to activate transcription has been shown to be dependent upon specific regions of the protein which are distinct from the region mediating DNA binding and are known as activation domains (Ptashne 1988). Several types of activation domain have been described which are rich in either acidic amino acids, proline residues or glutamine residues (for review see Mitchell & Tjian 1989).

It is likely that the different activation domains act by interacting with other protein factors in order to facilitate transcription. Although this may occur by direct interaction with the RNA polymerase itself (Sigler 1988), at least in the case of the acidic activation domain it seems more likely that its effect is mediated via contact with other transcription factors which form a basal transcriptional complex that then interacts with the RNA polymerase. One potential contact factor is called TATA binding protein or TBP. This factor is a component of the TFIID transcription factor, which binds to the DNA sequence known as the TATA box present in many gene promoters. Thus, for example, binding of yeast GAL4, or of the mammalian transcription factor known as ATF (both of which contain an acidic activation domain) to their specific binding sites in regulated promoters has been shown to change the conformation of TFIID that is already bound so that, instead of contacting the TATA box alone, it contacts both the TATA box and the site at which transcription will start (Horikoshi *et al.* 1988). In turn, this altered binding facilitates the binding of other factors such as TFIIC, TFIIE and the RNA polymerase itself into the stable transcriptional complex that is

Table 1. Transcription factor domains

Domain	Role	Factors containing domain	Comments
Homeobox	DNA binding	Numerous *Drosophila* homeotic genes, related genes in other organisms	DNA binding mediated via helix–turn–helix motif
Cysteine–histidine zinc finger	DNA binding	TFIIIA, Kruppel, Sp1, etc.	Multiple copies of finger motif
Cysteine–cysteine zinc finger	DNA binding	Steroid-thyroid hormone receptor family	Single pairs of fingers, related motifs in adenovirus E1A and yeast GAL4, etc.
Basic element	DNA binding	C/EBP, c-*fos*, c-*jun*, GCN4	Often found in association with leucine zipper
Leucine zipper	Protein dimerization	C/EBP, c-*fos*, c-*jun*, GCN4, c-*myc*	Mediates dimerization which is essential for DNA binding by adjacent domain
Helix–loop–helix	Protein dimerization	c-*myc*, *Drosophila* daughterless MyoD, E12, E47	Mediates dimerization which is essential for DNA binding by adjacent domain
Amphipathic acidic alpha-helix	Gene activation	Yeast GCN4, GAL4, steroid-thyroid receptors, etc.	Probably interacts directly with TFIID
Glutamine-rich region	Gene activation	SP1	Related regions in Oct-1, Oct-2, AP2, etc.
Proline-rich region	Gene activation	CTF/NF1	Related regions in AP2, c-*jun*, Oct-2

necessary for transcription to occur. The altered conformation of TFIID is directly dependent upon the activation domain of GAL4 or ATF, since the binding of truncated factors lacking the activation domain does not produce the change in TFIID binding.

Hence specific regions of transcription factors can activate transcription following DNA binding, by altering the conformation of another factor that is already bound, thus facilitating the assembly of a stable transcriptional complex (Figure 1).

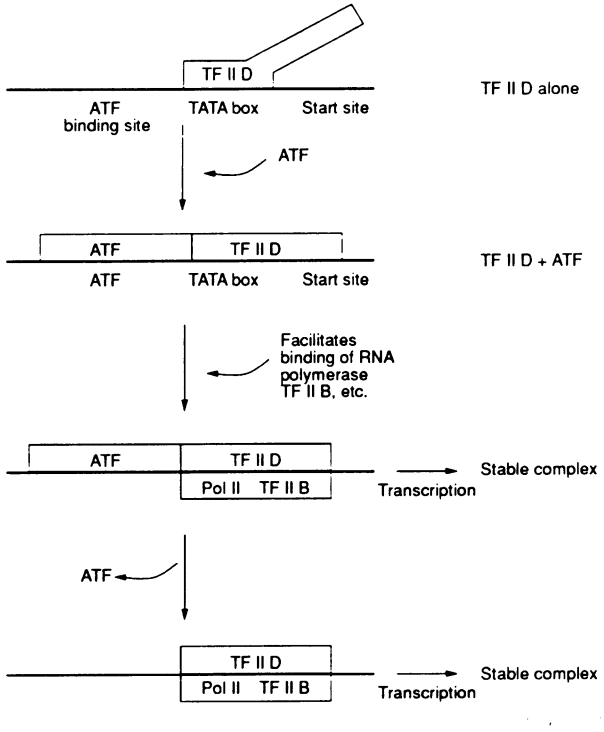

Figure 1. Schematic diagram of the interaction of the transcriptional activator ATF with the TATA box binding protein TFIID to create a stable transcription complex and activate transcription.

Repression

Although the majority of transcription factors act in a positive manner, a number of cases have now been described in which a transcription factor exerts an

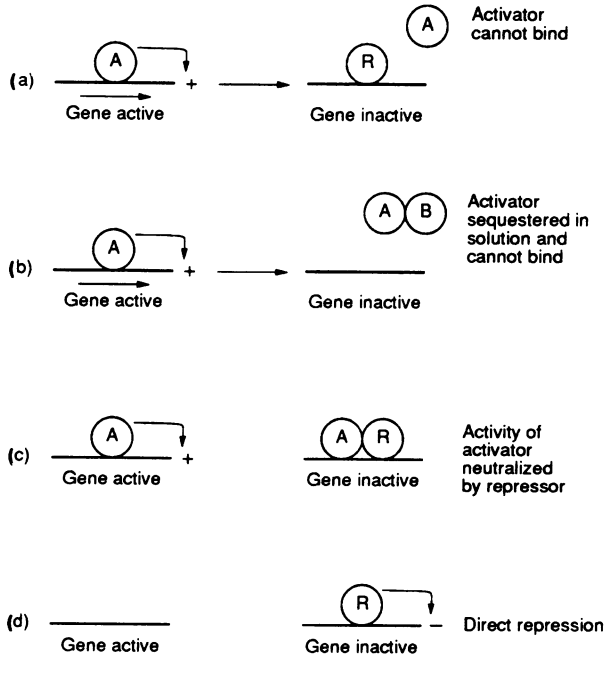

Figure 2. Repression of transcription by a, competition for binding, b, sequestration in solution, c, quenching of activity or d, direct repression.

inhibiting effect on transcription and several possible mechanisms have been described by which this may be achieved (Levine & Manley 1989; Goodbourn 1990) (Figure 2).

The simplest means of achieving repression is seen in the β-interferon promoter, where the binding of two positively acting factors is necessary for gene activation. Another factor acts negatively by binding to this region of DNA and simply preventing the binding of the positively acting factors (Figure 2a). In response to viral infection, this negative factor is inactivated, allowing the positively acting factors to bind and transcription then occurs (Goodbourn *et al.* 1986).

As well as directly interfering with DNA binding, a negatively acting factor can also indirectly prevent DNA binding by interacting with a positively acting factor in solution and forming a complex which cannot bind to DNA (Figure 2b). This is seen in the case of the I-POU member of the POU family which cannot itself bind to DNA but which interacts with the positively acting POU factor CF1a and prevents it binding to DNA (Treacey *et al.* 1991).

Another way that an inhibitory transcription factor can act is by interfering with the activation of transcription mediated by a bound factor in a phenomenon known as quenching (Figure 2c). This is seen in the case of the yeast mating type α_2 protein which binds to a site adjacent to that bound by the MCM1 transcriptional activator protein. This masks the activation domain and prevents it from activating the α-specific genes (Keleher *et al.* 1988).

In all these cases the negative factor exerts its inhibiting effect by neutralizing the action of a positively acting factor, preventing either its DNA binding or its activation of transcription. It is likely, however, that some factors may have an inherently negative effect and may directly inhibit transcription. Possibly this operates via a discrete domain, analogous to the activation domain(s) but with the opposite effect on the formation or stability of the basal transcriptional complex (Figure 2d).

Regulation of transcription

Clearly the interplay of positively and negatively acting transcription factors has a critical role in controlling the process of gene expression. It is evident, however, that if transcription is to be regulated in different cell types and tissues, some means must exist of modulating the activity of specific factors in order that they may produce the correct pattern of tissue specific gene expression. This modulation is achieved either by regulating the synthesis of the particular transcription factor so that it is present only in a specific cell type, by regulating its activity so that it is present in an active form only in a specific cell type or by ensuring that activation occurs only in response to a specific signal (Figure 3a–c).

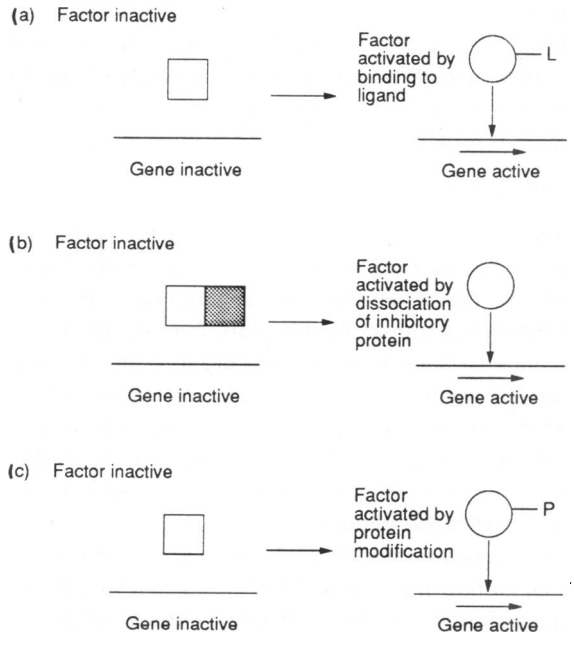

Figure 3. Activation of transcription factors by a, ligand binding, b, dissociation of an inhibitory protein or c, protein modification.

Regulation of synthesis

A simple example of the role of regulated synthesis in controlling tissue specific gene transcription is provided by the Oct-2 transcription factor. This factor is involved in the stimulation of immunoglobulin gene expression in B cells. It is found commonly in B cells but is absent from cell types such as HeLa cells which do not express immunoglobulin genes (Scheidereit *et al.* 1987). The expression of the gene encoding Oct-2 in HeLa cells results in the transcriptional activation of immunoglobulin genes introduced into these cells (Muller *et al.* 1988). Hence the synthesis of Oct-2 only in specific cell types allows it to activate the expression of particular genes in such cells specifically.

Another example is seen in the case of genes such as that encoding creatine kinase. These genes are expressed only in muscle cells and are regulated by the MyoD transcription factor which is also present only in muscle cells. This is an even more dramatic example than immunoglobulin, however, because the artificial expression of MyoD in non-muscle cells such as fibro-

blasts is sufficient to convert them into muscle cells, indicating that MyoD activates transcription of all the genes whose protein products are necessary to produce a differentiated muscle cell (Edmondson & Olson 1993).

Regulation of activity

Although regulation of transcription factor synthesis is widely used as a method of gene regulation, it suffers from the deficiency that the transcription of the genes encoding the transcription factors themselves must be regulated. Hence it only sets the problem one step back. It is not surprising therefore that many transcription factors are regulated in a different way. The factor is synthesized in all tissues, and is converted into an active form only in a specific cell type or in response to a specific stimulus (Figure 3a–c).

A simple example of such modulation occurs in the yeast transcription factor ACE1. This factor activates transcription of the metallothionein gene in response to copper. In this case the protein undergoes a conformational change in the presence of copper which allows it to bind to regulatory sites in the metallothionein gene and activate transcription (Furst *et al.* 1988; Figure 3a).

A similar dependence on an activating ligand is also observed in mammalian cells in members of the steroid/thyroid hormone receptor family. These receptors must bind the appropriate steroid hormone in order to activate transcription of target genes. A region at the C-terminus of each receptor which binds the appropriate hormone has been identified. It was thought therefore that binding of the hormone to the receptor activates its ability to bind to DNA and thereby switch on transcription. It is now clear that although *in vivo* the receptors bind to DNA only in the presence of hormone, *in vitro* they can do so even in its absence. This suggests that *in vivo* such DNA binding is prevented by some other anchorage protein, with the hormone acting to release the receptor from this factor, thus allowing it to bind DNA. In support of this hypothesis several steroid receptors have been shown to be associated with the 90-kDa heat-inducible protein hsp90 prior to hormone treatment. The receptor dissociates following the addition of hormone (Sanchez *et al.* 1987).

This activation may be mediated not simply by a ligand-induced change in conformation but by disruption of an inhibitory protein–protein interaction (Figure 3b). A similar disruption of an interaction with an inhibitory protein is also responsible for the activation of the NF kappa B protein in response to phorbol ester treatment of T cells or HeLa cells (Baeurele & Baltimore 1988).

In addition to protein–protein interaction, activation of transcription factors can also be achieved by protein modification, providing a direct means of activating a particular factor in response to a specific signal (Figure 3c). One example of this is provided by the CREB transcription factor. This factor mediates the activation of several cellular genes following cyclic AMP treatment. Cyclic AMP is known to stimulate the protein kinase A enzyme. In turn this enzyme phosphorylates CREB, which stimulates the activity of a transcriptional activation domain adjacent to the site of phosphorylation. Thus stimulation of gene expression by cyclic AMP is mediated via its stimulation of protein kinase A and the consequent phosphorylation of CREB (Yamamoto *et al.* 1990). Similar activation by phosphorylation is also involved in the activation of the NF kappa B protein by phorbol esters. In this latter case however it is the inhibitory protein associated with NK kappa B (see above) which is phosphorylated causing it to dissociate from the NF kappa B protein and thus allowing NF kappa B to move to the nucleus, bind to DNA and activate transcription (Ghosh & Baltimore 1990).

Therefore, a variety of mechanisms involving both increased synthesis and protein activation by ligand binding, protein modification or disruption of protein–protein interaction (Figure 3) coexist allowing specific factors to become active in response to a particular signal or in a particular cell type, and thereby regulate gene expression in individual cell types.

Transcription factors and disease

Failure of transcription factor function

In view of their critical importance in gene regulation, it is not surprising that failure of transcription factor function can result in disease. This has been described to occur in one of two ways.

Firstly, an abnormality in a transcription factor itself can result in a lack of specific gene expression leading to disease. Thus mutations in the POU family transcription factor Pit-1 result in a failure of pituitary gland development and a lack of growth hormone gene expression. This results in dwarfism (Radovick *et al.* 1992). Similarly, one type of congenital severe combined immunodeficiency is caused by a failure of HLA class II gene transcription, resulting in the absence of these proteins from the surface of immunocompetent cells. This failure of transcription is dependent on the lack of a specific factor necessary for the transcription of these genes (Reith *et al.* 1988).

The alternative means by which failure of transcription factor function can lead to disease is seen in haemophilia B. In this disease the factor necessary for transcription of

the factor IX gene is present, but it fails to bind to the gene promoter, owing to a mutation in the DNA sequence to which it would normally bind. This results in failure of gene transcription (Crossley & Brownlee 1990). Similarly, a mutation in the promoter of the plasminogen activator inhibitor 1 gene (PAI-1) results in a reduction in the ability to bind an inhibitory transcription factor. This leads to an increase in PAI-1 levels with consequent enhanced risk of heart disease (Dawson *et al.* 1993).

Malregulation of transcription factor activity

As well as the examples discussed above, in which disease results from a failure of transcription factor function, disease can also occur if a transcription factor is synthesized or becomes active at the wrong time or in the wrong place, with consequent ectopic activation of gene expression. This form of malregulated gene expression is central to the development of certain cancers in which the mutation or over-expression of specific cellular genes, known as proto-oncogenes, results in their conversion into 'cancer-causing' oncogenes (for review see Bishop 1987). Thus some proto-oncogenes encode growth factors or their receptors and others such as *erbA*, *fos*, *myb*, and *myc*, encode cellular transcription factors that are involved in regulating the expression of specific genes. After the conversion of these proto-oncogenes into oncogenes, which can occur either by mutation or by over-expression, corresponding alterations occur in the expression of the genes which they regulate and this results in cancer.

One example of this effect is provided by *fos* and *jun* transcription factors (Curran & Franza 1988). These are normally synthesized transiently in response to growth promoting signals and act to activate the genes encoding specific proteins required for cellular growth. If for any reason these proteins are made continuously, they act to promote the continuous growth in the absence of growth factors which is characteristic of the cancer cell.

Hence the *fos* and *jun* genes can be important as proto-oncogenes whose products have a critical role in the growth of normal cells, but which can be converted into oncogenes capable of transforming cells. Moreover, in contrast to the other diseases discussed above, in this case malregulation of gene expression and disease is caused not by failure of transcription factor function but rather by failure to regulate correctly the activity of specific factors so that they activate gene expression inappropriately. This indicates that, as with other cellular processes, gene expression is subject to complex regulatory mechanisms, the failure of which can be as disastrous as the failure of the basic process itself.

Conclusion

This review is intended to provide a brief overview of the function, activity and regulation of transcription factors and their malregulation in disease. Much remains to be understood about how these factors act to promote or inhibit gene transcription. However, this is a very active and productive research field at present and, as indicated in the accompanying article, progress is being made all the time, in an attempt to gain a more complete picture of the role that these factors play both in normal cellular function and in disease processes.

References

Baeurele P.A. & Baltimore D. (1988) I kappa B: a specific inhibitor of the NF kappa B transcription factor. *Science* **242,** 540–546.

Bishop J.M. (1987) The molecular genetics of cancer. *Science* **235,** 305–311.

Crossley M. & Brownlee G.G. (1990) Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B. *Nature* **345,** 444–446.

Curran R. & Franza B.R. (1988) Fos and Jun: the AP-1 connection. *Cell* **55,** 315–397.

Darnell, J.E. (1982) Variety in the level of gene control in eukaryotic cells. *Nature* **297,** 365–371.

Dawson S.J., Wiman B., Hansten A., Green F., Humphries S. & Henney A.M. (1993) The two allele sequences of a common polymorphism in the promoter of the plasminogen activator inhibitor-1 (PAI-1) gene respond differently to IL-1 in HEP G2 cells. *J. Biol. Chem.* (in press).

Edmondson D.G. & Olson E.N. (1993) Helix-loop-helix proteins as regulators of muscle-specific transcription. *J. Biol. Chem.* **268,** 755–758.

Furst P., Hu S., Hackett R. & Hamer D. (1988) Copper activates methallothionein gene transcription by altering the conformation of a specific DNA binding protein. *Cell* **55,** 705–717.

Ghosh, J. & Baltimore D. (1990) Activation of NF kappa B by phosphorylation of its inhibitor I kappa B. *Nature* **344,** 677–682.

Goodbourn S. (1990) Negative regulation of transcriptional initiation in eukaryotes. *Biochim. Biophys. Acta* **1032,** 53–77.

Goodbourn S., Burstein J. & Maniatis T. (1986) The human beta-interferon enhancer is under negative control. *Cell* **45,** 601–610.

Horikoshi M., Hai T., Lin Y-S., Green M.R. & Roeder R.G. (1988) Transcription factor ATF interacts with the TATA box factor to facilitate establishment of a preinitiation complex. *Cell* **54,** 1033–1042.

Keleher C.A., Gouthe C. & Johnson A.D. (1988) The yeast cell-type specific repressor alpha 2 acts co-operatively with a non cell-type specific protein. *Cell* **53,** 927–936.

Latchman D.S. (1990a) *Gene Regulation: A Eukaryotic Perspective.* Unwin Hyman, London.

Latchman D.S. (1990b) Eukaryotic transcription factors. *Biochem. J.* **270,** 281–289.

Latchman D.S. (1991) *Eukaryotic Transcription Factors.* Academic Press, London, San Diego.

Levine M. & Manley J.L. (1989) Transcriptional repression of eukaryotic promoters. *Cell* **59,** 405–408.

MITCHELL P.J. & TJIAN R. (1989) Transcriptional regulation in mammalian cells by sequence specific DNA binding proteins. *Science* **245,** 371–378.

MULLER M.M., RUPPERT S., SCHAFFNER W. & MATTHIAS P. (1988) A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature* **336,** 544–551.

PTASHNE M. (1988) How eukaryotic transcriptional activators work. *Nature* **335,** 683–689.

RADOVICK S., NATIONS M., DU Y., BERG L.A., WEINTRAUB B.D. & WONDISFORD F.E. (1992) A mutation in the POU-homeodomain of Pit-1 responsible for combined pituitary hormone deficiency. *Science* **257,** 1115–1118.

REITH W., SATOLA S., SANCHEY C.H., AMALDI I., LISOWSKA-GROSPIERRE B., GRISCELLI C., HADAM M.R. & MACH B. (1988) Congenital immunodeficiency with a regulatory defect in MRC Class II gene expression lacks a specific HLA-DR promoter binding protein RF-X. *Cell* **53,** 897–906.

SANCHEZ E.R., MESHINCHI S., TIENRUNGROJ W., SCHLESINGER M.J., TOFT D.O. & PRATT W.B. (1987) Relationship of the 90-kDa murine heat shock protein to the untransformed and transformed states of the L cell glucocorticoid receptors. *J. Biol. Chem.* **262,** 6986–6991.

SCHEIDEREIT C., HEGUY A. & ROEDER R.G. (1987) Identification and purification of a human lymphoid-specific octamer-binding protein (OTF-2) that activates transcription of an immunoglobulin promoter *in vitro*. *Cell* **51,** 783–793.

SIGLER P.B. (1988) Acid blobs and negative noodles. *Nature* **333,** 210–212.

TREACEY M.N., HE X. & ROSENFELD M.G. (1991) I-POU: a POU domain protein that inhibits neuron-specific gene activation. *Nature* **350,** 577–584.

YAMAMOTO K.R., GONZALEZ G.A., MENZEL P., RIVIER J. & MONTMINY M.R. (1990) Characterization of a bipartite activation domain in transcription factor CREB. *Cell* **60,** 611–617.