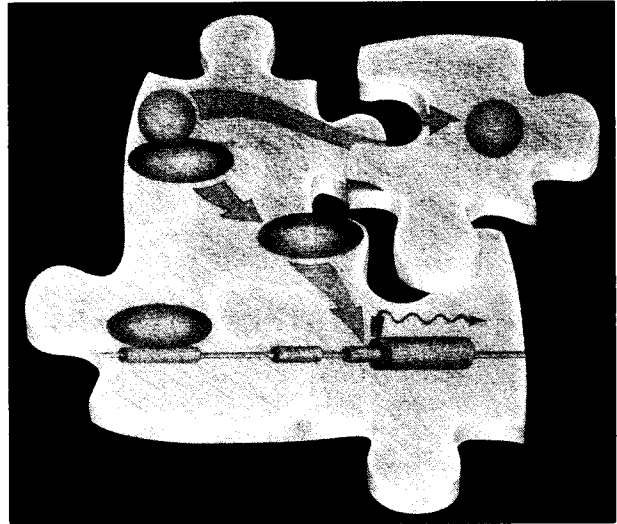


Transcriptional regulation in lymphocytes

Understanding the mechanisms that regulate the development of complex mammalian organisms is one of the most fascinating problems in biology. The coordinated transcriptional regulation of specific sets of genes defines the framework upon which cell determination and differentiation are established. Thus, an understanding of tissue-specific gene transcription will significantly advance our knowledge of mammalian development. Lymphoid differentiation, the process by which committed progenitor cells give rise to both the T- and B-cell lineages, has long been used as a paradigm for studies of both tissue-specific and developmentally regulated gene expression. The following two papers review different aspects of transcriptional control during T- and B-cell development, focusing on two of the best-understood systems, namely the α TCR gene and the immunoglobulin genes. Both conclude that lymphocyte-specific gene expression is controlled by complex regulatory circuits involving interactions among multiple families of ubiquitous and lineage-specific transcription factors.



Transcriptional regulation during T-cell development: the α TCR gene as a molecular model

Jeffrey M. Leiden

The regulation of gene expression during lymphocyte differentiation is a complex process involving interactions between multiple positive and negative transcriptional regulatory elements. In this article, transcriptional regulation of the archetypal T-cell-specific gene, α TCR, is discussed. Major recent developments, including the identification of novel families of transcription factors that regulate multiple T-cell genes during thymocyte ontogeny and T-cell activation, are described.

The differentiation of mature B and T cells from committed lymphoid progenitors is an excellent model system for understanding transcriptional regulation during mammalian development^{1,2}. Studies of murine and human lymphopoiesis have allowed the identification of sets of tissue-specific genes that are transcriptionally regulated at distinct developmental time points^{3,4}. In addition, tumor cell lines, representing specific stages of B- and T-cell development, have been characterized and grown *in vitro*. Some of these cell lines can be stimulated to differentiate *in vitro* by the appropriate extracellular signals⁴.

The α TCR gene has been studied as a paradigm of lineage-specific and developmentally regulated gene expression in T cells. α TCR gene rearrangement and

expression are restricted to $\alpha\beta$ TCR T cells⁵⁻⁷. Moreover, the expression of this gene is developmentally regulated in thymocytes: α TCR gene expression is first detected on fetal day 17 in mouse thymocytes, several days after δ , β and γ TCR gene expression^{8,9}. Finally, the expression of the α TCR gene is specifically upregulated during the transition from immature CD3^{dull} cortical thymocytes to the more mature CD3^{bright} cells found in the thymic medulla and the peripheral lymphoid system. This review summarizes the current state of knowledge concerning the molecular mechanisms that regulate the transcription of the α TCR gene, with particular emphasis on several recently cloned lymphoid-specific transcription factors that modulate α TCR gene expression. In addition, the structure of the α TCR enhancer is compared with those

REVIEW

of other recently identified T-cell-specific transcriptional regulatory elements. This comparison suggests that the same transcription factors that control α TCR gene expression regulate the expression of multiple T-cell genes during both development and activation.

Identification and characterization of the α TCR transcriptional enhancer

As first demonstrated by Davis and co-workers¹⁰, the α and δ TCR genes are tightly linked in humans and mice; the δ TCR gene lies embedded within the α TCR locus between V_α and J_α , such that rearrangements of the α TCR gene result in the deletion of the entire δ TCR locus (Fig. 1). Despite this tight linkage, the rearrangement and expression of these two genes is lineage specific, that is the α TCR gene is expressed almost exclusively in $\alpha\beta$ TCR cells, whereas δ TCR gene rearrangement and expression is limited to cells of the $\gamma\delta$ lineage¹¹. Thus, the α TCR gene represents an excellent model system for studies of lineage-specific gene expression in T cells. Initial studies of several TCR V_α promoters demonstrated that these promoters lack significant transcriptional activity in T cells¹². This suggested the existence of intragenic transcriptional regulatory sequences that control the expression of the α TCR gene. Potent T-cell-specific transcriptional enhancers were subsequently identified 3' to both the mouse and human α TCR genes (Fig. 1 and Refs 12,13). These enhancers, which display greater than 90% sequence identity, are both located 3–4.5 kilobases (kb) 3' of the single C_α gene segment. Because the 5'-most J_α gene segment is located as much as 75 kb 5' to C_α ¹⁴, these enhancers must be able to act over long distances. Further studies of the human α TCR enhancer localized its activity to a 275 bp fragment that was preferentially active in $\alpha\beta$ TCR T cells, and which was required for high-level transcription from a TCR V_α promoter. This enhancer also stimulated high-level T-cell-specific transcription from heterologous promoters such as the minimal SV40 or *c-fos* promoters¹².

DNaseI footprint analyses of the human α TCR enhancer allowed the identification of four functionally significant nuclear protein-binding sites, called T α 1–T α 4 (Fig. 2 and Ref. 12). Given the pattern of DNaseI footprinting, it was suggested that T α 2 might be composed of two distinct nuclear protein-binding sites. A similar analysis of the mouse enhancer identified nuclear protein-binding sites corresponding to T α 1 (NF α 2) and T α 2 (NF α 3 and NF α 4)¹³. The T α 3 and T α 4 sites were not detected in the analysis of the mouse enhancer because the fragment used for footprinting did not extend 3' of T α 2. Extensive mutational and deletional analyses of the human α TCR enhancer demonstrated that, at least in the Jurkat T-cell line, the minimal α TCR enhancer is contained within a 116 bp fragment containing only the T α 1 and T α 2 nuclear protein-binding sites¹². A fragment containing the T α 3 and T α 4 elements alone displayed low-level enhancer activity (6–10-fold compared with 50–100-fold for the T α 1- and T α 2-containing fragment). Both T α 1 and T α 2 were required for the activity of the minimal α TCR enhancer because mutation and/or deletion of either site abolished transcriptional regulatory activity¹⁵. In addition, the spacing between T α 1 and T α 2 was critical for enhancer function;

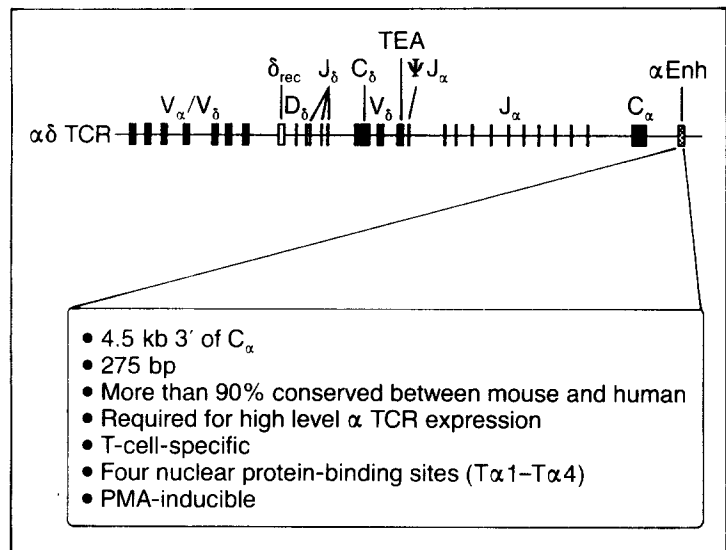


Fig. 1. A schematic representation of the human T-cell receptor $\alpha\delta$ locus. The general organization of the locus and the location of δ_{rec} , TEA and ΨJ_α are from Hockett et al.¹⁴ Variable (V), diversity (D), joining (J) and constant (C) gene segments are shown as boxes. The T-cell receptor α enhancer (α Enh) is shown as a cross-hatched box. The precise number and organization of V_α , V_δ and J_α gene segments are unknown. kb: kilobase pairs; bp: base pairs; PMA: phorbol myristate acetate.

reducing the normal 20 bp spacing to 15 bp eliminated enhancer activity, as did increasing the spacing to more than 85 bp. In contrast, the phasing of the T α 1 and T α 2 elements on the DNA helix did not appear to be critical; increasing the spacing from 20 to 25 bp had a relatively small effect on enhancer activity. Finally, the T α 3 and T α 4 elements, while themselves not possessing high-level enhancer activity, could compensate for mutations in T α 1 or T α 2 that abolished enhancer activity¹⁵. Thus, T α 3 and T α 4 represent redundant elements, at least as assayed in Jurkat T cells. It should be emphasized, however, that these results did not rule out the possibility that T α 3 and/or T α 4 play an important role in modulating α TCR enhancer function at some earlier stage in T-cell development, or in response to specific extracellular signals. Taken together, these functional analyses revealed that the α TCR enhancer is composed of multiple, partially redundant, nuclear protein-binding sites. Interactions between proteins binding to these different sites appear to be important for enhancer function.

Transcription factors that regulate the activity of the α TCR enhancer

An examination of the sequences of the nuclear protein-binding sites of the mouse and human α TCR enhancers revealed several previously described transcriptional regulatory elements (Fig. 2). For example, T α 1 contains an evolutionarily-conserved cAMP response element (TGACGTCA)¹². Similarly, T α 3 contains overlapping AP-2 (CCCCAGGC)-like^{16,17} and κ E2 (E box) (AGGCCACGTGCCGA)-like¹⁸ motifs. In contrast, T α 2 and T α 4 do not contain previously identified nuclear protein-binding sites. Electrophoretic mobility shift assays (EMSA), using synthetic oligonucleotides corresponding to T α 1–T α 4, demonstrated that T α 1 binds a set of four to six ubiquitously expressed cAMP response element binding (CREB) proteins^{15,19}. Binding of each of these proteins was inhibited by mutations

REVIEW

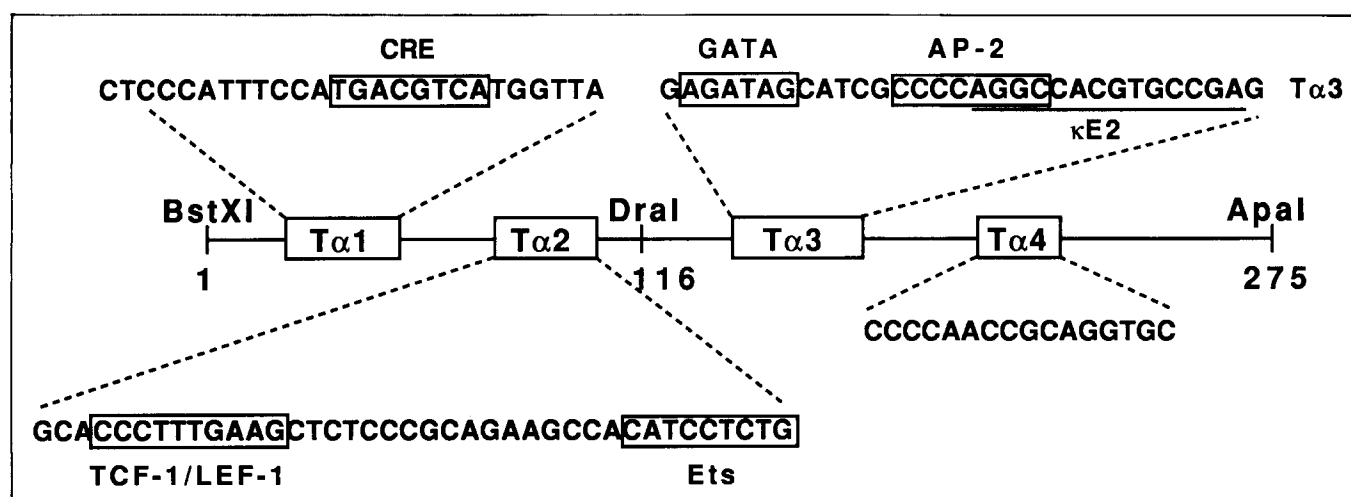


Fig. 2. A schematic illustration of the 275 bp BstXI/ApaI human α TCR enhancer¹². The location of the Dral restriction endonuclease site, and the locations and sequences of the four nuclear protein-binding sites within the enhancer ($T\alpha 1$ – $T\alpha 4$) are shown¹². Sequences corresponding to the cAMP response element (CRE), and the GATA, TCF-1/LEF-1, Ets and AP-2-binding sites are boxed. The potential κ E2-binding site in $T\alpha 3$ is underlined.

within the core of the CRE and no $T\alpha 1$ -binding proteins were detected that recognized non-CRE sequences (Ref. 19 and I-C. Ho and J.M. Leiden, unpublished). Similar EMSAs demonstrated that $T\alpha 2$ binds at least four nuclear protein complexes. One of these is T cell specific while a second is present in all T-cell, and some B-cell, nuclear extracts. $T\alpha 3$ binds three nuclear protein complexes, one of which is T cell specific. Finally, $T\alpha 4$ binds three ubiquitously expressed nuclear protein complexes. Interestingly, none of these experiments identified $\alpha\beta$ TCR lineage-specific nuclear protein complexes.

More recent studies have focused on the cloning of transcription factors that bind to the α TCR enhancer. These factors are described individually below and their properties are summarized in Table 1.

CREB proteins bind to $T\alpha 1$

Southwestern screening of a λ gt11 cDNA expression library from human Jurkat T cells with a radiolabelled

$T\alpha 1$ oligonucleotide probe allowed the identification of three clones that encode $T\alpha 1$ -binding proteins (B.A. Karpinski *et al.*, submitted). DNA sequence analysis of these clones demonstrated that they correspond to three previously identified cDNAs of the CREB or ATF families of transcription factors. The first clone corresponded to CREB, a 43 kDa transcription factor that has been shown to bind to CRE sites in several viral and cellular enhancers, and to confer cAMP-dependent transcriptional activation on these enhancers^{20,21}. The second clone corresponded to a cDNA called CRE-BP1 (Ref. 22), mXBP²³ or ATF-2 (Ref. 24); this protein also binds with high affinity to CRE sites in several adenovirus promoters, as well as to the X box of mouse MHC class II promoters²⁵. Recent studies have demonstrated that binding of this protein to CRE sites in several adenovirus promoters confers responsiveness to the viral E1A protein²⁶. The final clone identified was a full-length homolog of a partial-length cDNA, previously called ATF-4

Table 1. Transcription factors that regulate the α TCR enhancer

Transcription factor	Binding site	Molecular mass	DNA-binding domain	Recognition sequence	Lineage specificity	Refs
CREB	$T\alpha 1$	43 kDa	Basic domain/leucine zipper	TCACGTCA	Ubiquitous	20,21
CRE-BP1/ATF-2/mXBP	$T\alpha 1$	55 kDa	Basic domain/leucine zipper	TCACGTCA	Ubiquitous	22–24
ATF-4	$T\alpha 1$	38 kDa	Basic domain/leucine zipper	TCACGTCA	Ubiquitous	24
TCF-1	$T\alpha 2$	30 kDa	HMG box	C ₂ C ₁ T GAA T T TTT	T cells	31,32
TCF-1 α /LEF-1	$T\alpha 2$	53–57 kDa	HMG box	C ₂ C ₁ T GAA T T TTT	T cells/ Pre-B cells	29,30
Ets-1	$T\alpha 2$	54–60 kDa	Basic domain/ α helix	GA _{GGA} AG ^T _{TC} G	T cells/B cells	33
GATA-3	$T\alpha 3$	47 kDa	2 Cx zinc fingers	A _T GATA ^A _G	T cells/kidney/ brain	50–52

REVIEW

(Ref. 24). The function of ATF-4 remains unknown; however, unpublished studies from our laboratory have demonstrated that overexpression of ATF-4 in monkey CV-1 cells specifically downregulates cAMP-dependent transcription from the proenkephalin promoter/enhancer. Thus, at least in certain cell types, ATF-4 may represent a repressor of CRE-dependent transcription. The identities of the additional one to three CREB proteins that bind to T α 1, and the function of each of the CREB proteins in regulating T-cell transcription are unknown. As described in more detail below, a large number of T-cell enhancers and promoters contain CRE sites that are important for their transcriptional activity. Moreover, all T cells contain multiple CRE-binding proteins, although no T-cell-specific CREB protein has so far been identified.

It is worth mentioning that the name CREB may be a misnomer. Although members of the CREB/ATF family of transcription factors are important in regulating transcription in response to alterations in intracellular cAMP levels, they also play important roles in modulating basal levels of transcription, and have recently been shown to upregulate transcription in response to alterations in intracellular calcium levels in neural cells²⁷. Thus, their precise role in regulating transcription in T cells awaits studies using *trans*-dominant inhibitors and embryonic stem (ES) cell gene knock-out experiments, both *in vitro* and *in vivo*.

TCF/LEF proteins bind to T α 2

Several recent reports have described the purification and cloning of two related transcription factors that bind to a pyrimidine-rich sequence (PyCTTTG/TT/AT/A) from the 5' end of T α 2 (as well as to a related motif in the CD3 ϵ enhancer)²⁸⁻³². These two factors, called TCF-1 (Ref. 31) and TCF-1 α ²⁹ (or LEF-1 (Ref. 30)), are approximately 30 kDa and 53-57 kDa proteins containing closely related 80-90 amino acid DNA-binding domains that display significant homology to a region present in the high mobility group (HMG) family of DNA-binding proteins. Whereas LEF-1 appears to be the mouse homolog of TCF-1 α , TCF-1 is apparently encoded by a distinct gene. Studies of the expression patterns of these genes have demonstrated that LEF-1 is expressed during all stages of T-cell development, in both the $\alpha\beta$ TCR and $\gamma\delta$ TCR lineages, and in mouse pre-B cells³⁰. TCF-1 was reported to be expressed exclusively in T cells³¹ but a comprehensive examination of TCF-1 expression at different stages of B-cell development has not yet been reported, and it is possible that it is also expressed transiently in early B cells. Overexpression of both TCF-1 and TCF-1 α /LEF-1 in B cells, HeLa cells or COS cells leads to low level (3-20-fold) *trans*-activation of reporter constructs containing TCF-1-binding sites from the α TCR or CD3 ϵ enhancers²⁹⁻³¹. However, *trans*-activation by LEF-1 and TCF-1 α appears to require the presence of additional nuclear protein-binding sites because overexpression of these factors cannot *trans*-activate a reporter construct containing only multimerized TCF/LEF-1-binding sites upstream of a minimal promoter^{28,30}.

Multiple forms of both the TCF-1 and TCF-1 α /LEF-1 proteins appear to be generated by alternative splicing. These include variants of TCF-1, containing or lacking a

proline-rich domain amino-terminal to the HMG box, as well as multiple forms of the protein with distinct carboxy-terminal regions, including one which is rich in serine and threonine residues^{31,32}. Similarly, two splice variants of TCF-1 α have been described, which contain or lack a serine-threonine-rich domain, immediately amino-terminal to the HMG box²⁹. The DNA-binding activities and *trans*-activation potentials of these alternatively spliced variants have not been studied in detail but it seems likely that the activities of these transcription factors may be regulated at both post-transcriptional and post-translational levels.

Ets proteins bind to T α 2

Using a λ gt11 expression screening approach, Ho *et al.*³³ demonstrated that the protein product of the *Ets*-1 proto-oncogene binds specifically to a purine-rich sequence (GAGGATGTG) located at the 3' end of T α 2. Along with several simultaneous reports³⁴⁻³⁸, this was the first demonstration that the *Ets* proto-oncogenes encode sequence-specific DNA-binding proteins. Binding of Ets-1 to the α TCR enhancer was especially intriguing because Ets-1 is known to be expressed preferentially in thymocytes and T cells (as well as B cells)³⁹, and because its expression is developmentally regulated during thymic ontogeny, with a time course that is almost identical to that of the α TCR gene: both are first expressed on days 17 and 18 in the fetal mouse thymus⁴⁰. Thus, Ets-1 is a candidate trigger factor for α TCR gene expression. The importance of the Ets-1-binding site for α TCR enhancer function was demonstrated by experiments showing that mutations in the core Ets-1-binding site in T α 2 abolished the activity of the minimal α TCR enhancer in Jurkat cells³³. Thus far, however, overexpression of *Ets*-1 in B cells and in nonlymphoid cells has not resulted in the *trans*-activation of the α TCR enhancer (I-C. Ho and J.M. Leiden, unpublished). This suggests that other T-cell-specific factors may be required for *trans*-activation of the α TCR enhancer by Ets-1 (for example, TCF/LEF factors or GATA-3) or that other *Ets* family members may regulate the activity of the α TCR enhancer in T cells. (Evidence from several sources suggests that *Ets* proteins usually require additional transcription factors to communicate with the basal transcription machinery³³.) Alternatively, the transcriptional activity of Ets-1 may require post-translational modification, and Ets-1 is known to exist in both phosphorylated and nonphosphorylated forms in T cells⁴¹.

Each of the *Ets* family members contains a conserved basic domain and an adjacent region with predicted α -helical structure⁴². Using deletion and mutation analyses, we have demonstrated that both of these conserved regions are necessary and sufficient for DNA binding (C-Y. Wang *et al.*, submitted). Only two *Ets* family members, Ets-1 and Ets-2, were previously known to be expressed in T cells³⁹. However, by using low stringency hybridization with a cDNA fragment encompassing the DNA-binding domain of Ets-1, we have recently isolated two additional and novel *Ets* family members from a Jurkat T-cell cDNA library. One of these clones, called Elf-1, contains a DNA-binding domain that is almost identical to that of the *Drosophila* transcription factor, E74 (Ref. 36) (hence the name E74-like factor-1) (C.B.

REVIEW

Thompson *et al.*, submitted). Interestingly, this factor binds to several T-cell transcriptional regulatory regions, including the IL-2 enhancer and the CD3R site of the HIV-2 enhancer⁴², but not to the Ets-binding sites within the α TCR or β TCR enhancers. The other novel factor, which we have called Ets-4, contains a DNA-binding domain that is more similar to those of Ets-1 and Ets-2, and binds to both the α and β TCR enhancers⁴³. These results demonstrate that multiple Ets family members are expressed in T cells, and that these family members display distinct DNA-binding specificities that allow them to interact independently with, and regulate, multiple T-cell genes. Mutational analyses of several different Ets-binding sites have demonstrated that nucleotides at the 3' end of the Ets-binding sites are responsible for the differential binding activities of some of the different Ets family members. As is the case for the CREB and TCF/LEF proteins, a complete understanding of the functions of each of the Ets family members awaits gene knock-out and *trans*-dominant inhibitor experiments, in both cell lines and mice.

GATA proteins that bind to T α 3

GATA-1 (previously called Eryf-1 and GF-1) is a zinc finger protein that is expressed exclusively in erythroid cells, megakaryocytes, mast cells and their common progenitors⁴⁴⁻⁴⁷. This factor is a transcriptional activator that binds to a sequence, A/TGATAA/G, which is present in the transcriptional regulatory regions of a number of erythroid-specific genes including the α , β and δ globin genes⁴⁸. Recent experiments, in which GATA-1 gene expression has been abolished by homologous recombination in embryonic stem cells, have demonstrated that GATA-1 is required for the development of the erythroid lineage in mice⁴⁹. Thus, GATA-1 belongs to the small family of lineage-specific determination genes.

An examination of the T α 3 motif of the α TCR enhancer revealed the presence of a consensus binding site (AGATAG) for GATA-1. Since GATA-1 is not expressed in T cells, this finding raised the possibility that there is a distinct T-cell GATA protein that binds to this site. To address this question, a Jurkat T-cell cDNA library was screened by low stringency hybridization with a cDNA probe from the zinc finger region of GATA-1. These experiments resulted in the isolation of a novel cDNA called hGATA-3 (Ref. 50). The chicken (cGATA-3) and mouse (mGATA-3) homologs of hGATA-3 were isolated using a similar approach^{51,52}. hGATA-3 is a 48 kDa polypeptide that contains two zinc fingers that are more than 90% identical to those of GATA-1 (Ref. 50). In contrast, the remainder of the protein is less than 12% identical to GATA-1. Studies of the expression pattern of hGATA-3 showed that, within hematopoietic cells, its expression is restricted to the $\alpha\beta$ and $\gamma\delta$ T-cell lineages⁵⁰. Interestingly, a slightly larger GATA-3 mRNA is expressed in kidney and in some mesangial cell lines. The relationship between the kidney and T-cell GATA-3 mRNAs has not been established. Low level GATA-3 expression was also detected in whole brain. Overexpression of hGATA-3 (as well as cGATA-3 or mGATA-3) in nonlymphoid cells resulted in *trans*-activation of reporter constructs containing multimerized GATA sites from T α 3 (Ref. 50) or from a related

sequence from the δ TCR enhancer (δ E4)⁵² upstream of a minimal promoter. Thus, unlike the TCF-1 α /LEF-1 and the Ets proteins, hGATA-3 can *trans*-activate transcription from its cognate binding site in the absence of additional nuclear protein-binding sites.

An examination of mGATA-3 expression during thymocyte development has demonstrated expression of this gene in the earliest thymocytes tested (fetal day 14 in the mouse) (B.K. Oakley and J.M. Leiden, unpublished). Thus, like TCF-1 and TCF-1 α /LEF-1, GATA-3 is a lineage-restricted transcription factor that is expressed very early in T-cell development. hGATA-3 was also shown to be expressed at low levels in CEM cells, an immature T-cell line that does not express α TCR mRNA. However, after stimulation with PMA both hGATA-3 and α TCR mRNA levels increased dramatically⁵⁰. (Similar findings have been reported for TCF-1 α in this system²⁹.) These findings, when considered with the previously described results concerning the role of GATA-1 in erythroid differentiation⁴⁹, suggested that GATA-3 is a good candidate for a T-cell-determining gene. This possibility is currently being tested by using homologous recombination to eliminate GATA-3 expression in embryonic stem cells and mice.

Shared nuclear protein-binding sites in multiple T-cell enhancers and promoters

The studies of the α TCR enhancer described above were predicated on the hypothesis that transcription factors that regulate α TCR gene expression might also regulate the expression of additional T-cell genes. This should be particularly true for factors that play important roles in regulating lymphoid and T-cell development. To address this question, the sequences of the known nuclear protein-binding sites from the α TCR enhancer have been compared with those of several other T-cell-specific transcriptional regulatory regions, including the human β ¹⁹ and δ ⁵³ TCR enhancers, the CD2 enhancer⁵⁴, and the CD3 ϵ ⁵⁵ and δ ⁵⁶ enhancers. As shown in Fig. 3 and Table 2, each of these enhancers contains different combinations of binding sites for CREB, TCF-1/LEF-1, Ets, GATA and E box proteins. In some cases, these sites have been shown to bind the relevant proteins. Thus, the α TCR gene is an excellent paradigm of T-cell-specific transcriptional regulation, and it appears likely that each of these families of transcription factors plays an important role in regulating multiple T-cell genes during development. In addition, as shown in Table 2, potential Ets-binding sites can also be found in the transcriptional regulatory regions of multiple lymphokine genes, suggesting that Ets family members may also play an important role in regulating gene expression during the process of T-cell activation; for example, at least one functional Ets-binding site is required for IL-2 enhancer function.

Many of the genes shown in Fig. 3 are expressed at different times in development, in some cases in different cell lineages, and, in others, only after cell activation. How, then, are these genes differentially regulated in response to distinct developmental and activational signals, if they contain common sets of nuclear protein-binding sites? While the answer to this question is complex and not yet fully understood, there are several

REVIEW

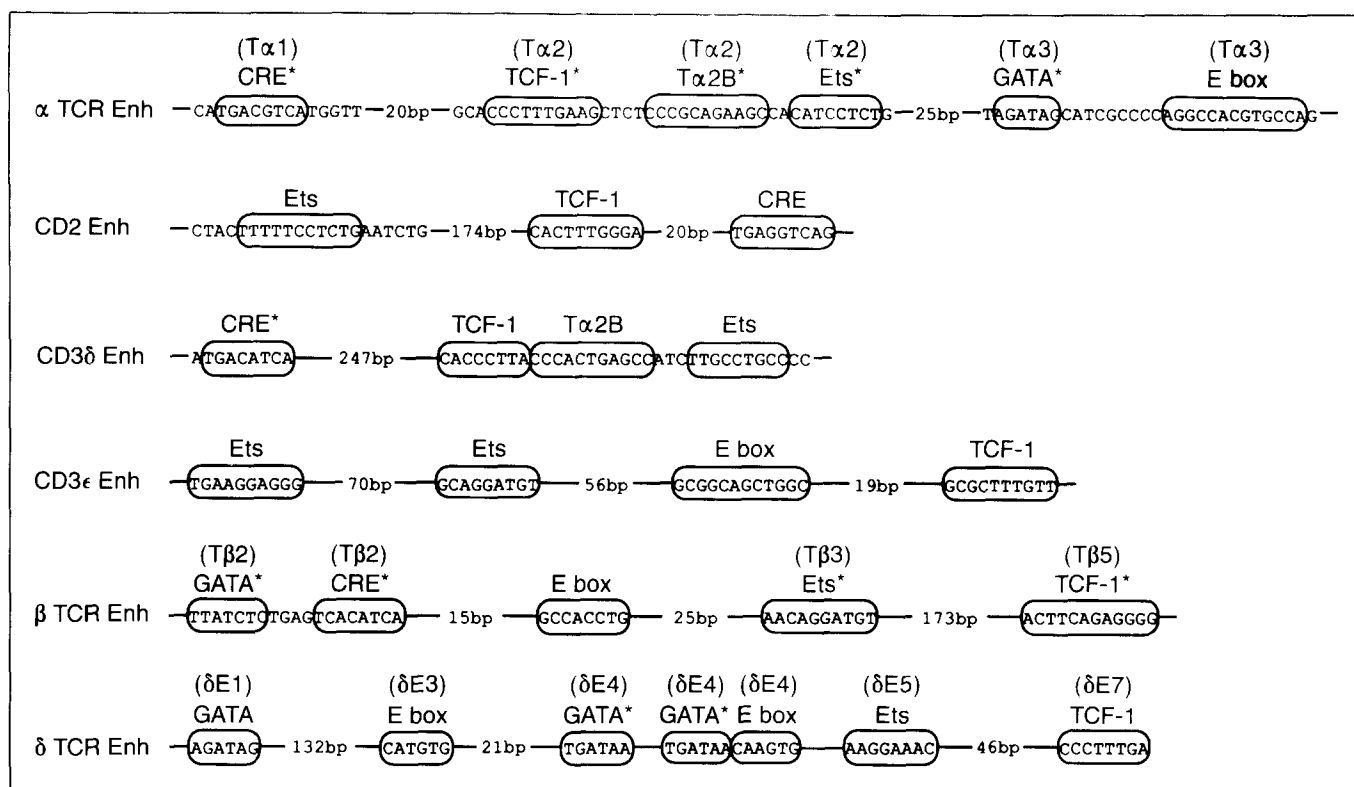


Fig. 3. A comparison of potential nuclear protein-binding sites in six T-cell transcriptional enhancers. Potential nuclear protein-binding sites are boxed. Previously described names of these binding sites within each enhancer are shown in parentheses. Binding sites that have been demonstrated to interact with the relevant protein are denoted with an asterisk.

models that allow a significant redundancy in the *cis*-acting regulatory sequences in different genes, while retaining the potential for differential expression.

First, different combinations of common, or overlapping, sets of nuclear protein-binding sites may allow flexible patterns of expression with a minimum number of different transcription factors. Such combinatorial diversity is clearly an important mechanism underlying the generation of antigen receptor diversity. This type of model is supported by the finding that interactions between transcription factors are required for the activity of many of the T-cell-specific enhancers.

Second, different members of a single family of transcription factors may display different DNA-binding specificities (which may be further modified by post-transcriptional or post-translational modifications) or may bind preferentially to a given enhancer in combination with an adjacent nuclear protein. Such differences in DNA-binding specificities have already been demonstrated for the Ets family members and may explain how this single family of transcription factors regulates the expression of different genes in resting and activated T cells.

Third, it is possible that negative regulatory elements or silencers play an important role in restricting the lineage specificity of certain T-cell genes. This type of model is supported by the inability to detect lineage-specific transcriptional activator proteins; all of the previously described α TCR enhancer binding proteins are expressed in both αβ TCR and γδ TCR T cells. Such transcriptional silencers may be distinct negative regulatory elements within these genes. This type of transcriptional silencer has been reported to downregulate the expression of the α TCR gene in non-αβ T cells⁵⁷ and the

γ TCR⁵⁸ gene in non-γδ T cells⁵⁹. Alternatively, the same multipartite transcriptional regulatory element may function as both a silencer and an enhancer, depending upon the particular set of nuclear proteins expressed by a given cell type. Such bifunctional transcriptional regulatory elements have been described in yeast⁵⁹. Moreover, a recent study has suggested that the Tα2 motif from the human α TCR enhancer can act as both a positive and negative transcriptional regulatory element, depending upon the precise DNA context in which it is located⁶⁰.

Future directions

During the last few years, a great deal has been learned about the molecular mechanisms that regulate α TCR gene expression during thymocyte development. These data are summarized schematically in Fig. 4. The important *cis*-acting transcriptional activator sequences have been identified and characterized, and a set of novel transcription factors, some of which are T cell specific, have been cloned. Several of these factors, including the CREB, TCF/LEF, Ets and GATA proteins, also appear to play important roles in regulating the expression of multiple genes during T-cell development and activation. Because they are T cell specific and expressed very early during thymocyte ontogeny, the TCF/LEF and GATA-3 proteins are particularly good candidates for T-cell-determining genes.

Although these studies have addressed a number of important questions concerning T-cell development, they have raised at least as many questions as they have answered. First, several of the transcription factors that bind to the α TCR enhancer have not yet been identified. In particular, there is evidence for a third factor that binds to Tα2 (Tα2B, Fig. 4) (I.C. Ho and J.M. Leiden,

REVIEW

Table 2. Shared nuclear protein-binding sites in multiple transcriptional enhancers

Transcriptional enhancer	Binding site		Transcriptional enhancer	Binding site	
Ets-binding sites			TCF-1-binding sites		
Tested			Tested		
α TCR Enh	CAGAGGATGTG	(Tα2)	α TCR	GGCACCCCTTTGAA	(Tα2)
MSV LTR	GAGCGGAAGCG		HIV-1*	AGCAGTCTTTGTA	
Polyoma Enh	AGCAGGAAGTG	(PEA3)	lck	GGCCTCCTGTGAA	
MHC class II Pr	AAGAGGAAGTT		lck	AAAAGCCTGTTTG	
HTLV-1 LTR	GGGAGGAAATG		lck	GAAACTCTCTGAA	
IL-2 Enh	AGGAGGAAAAA	(NF-AT-1)	lck	GGATGTCTCATGT	
	AAGAGGAAAAA	(IL-2B)	CD3γ	TGCCTTCTCTCAA	
Ig K 3' Enh	TTCAGGAACTG		CD3γ	AGATGCCTTTTGT	
HIV-2 LTR	GACAGGAACAG	(CD3R)	CD3δ	GAAACACTTTTCAA	
E74	AACCGGAAGTA		CD3ε	AGAGCGCTTTTGT	
	ATCAGGAATTA		Ig H Enh	GGTCCTCTTTTAA	
	AGTAGGAACAA		Ig K 3' Enh	AAGACCCTTTGAG	
TCR β Enh	AACAGGATGTG	(Tβ3)	β TCR	CCTCCCCTCTGAA	(Tβ5)
Consensus	AGG _A AGGA _A AATG GACC _T TGAA C		Consensus	C _{CT} C _T TAA T _T T _{GTT}	
Untested			Untested		
GM-CSF Pr	CAGAGGAAATG		ADA	AGTCTCCTTTGTT	
	CACAGGAACAT		α TCR	GTCCCCCTTTGAA	
IL-4 Pr	TCCAGGAGAAA		δ TCR	AAAGCCCTTTGAA	
IL-3 Pr	CAGGAGAAA		CD3ε	AGCACACTCTGCT	
CD2 Enh	CAGAGGAAAAA		Thy-1	GAGGTTCTGTTAC	
CD3δ Enh	TTGAGGATGAG		CD2	CTAACACTTTGGG	
CD3ε Enh	TGAAGGAGGGA		CD3δ	AGCACCCCTTACCC	
δ TCR Enh	GAAAGGATTAG				
	GCAAGGAAACC	(δE5)			
	ACAAGGAGATA	(δE1)			
HIV-1 LTR	CAAAGGAGAGA				
HIV-2 LTR	CAGAGGAAGAG				
	AAGAGGAATAC				
	GGCAGGAAGTA				
	CAGAGGAATTT				
GATA-binding sites			CREs		
Tested			Tested		
α TCR Enh	TAGATAGC	(Tα3)	α TCR Enh	TGACGTCA	(Tα1)
β TCR Enh	GAGATAAA	(Tβ2)	β TCR Enh	TCACATCA	(Tβ2)
Ig κ 3' Enh	AAGATAGC		V _β TCR Pr	TGATGTCA	
δ TCR Enh	TTGATAAC	(δE4)	CD3δ Enh	TGACATCA	
	TTGATAAC	(δE4)	CD2 Enh	TGAGGTCA	
Consensus	A _T GATA _A G _A		Consensus	TGACGTCA	
			E boxes		
			α TCR Enh	AGGCCACGTGCCGA	(Tα3)
			β TCR Enh	CAGCCACCTGCCCT	
			Ig κ 3' Enh	GGCACATCTGTTGC	
			CD3ε Enh	GCGGCAGCTGGCGG	
			Consensus	CANNTG	

REVIEW

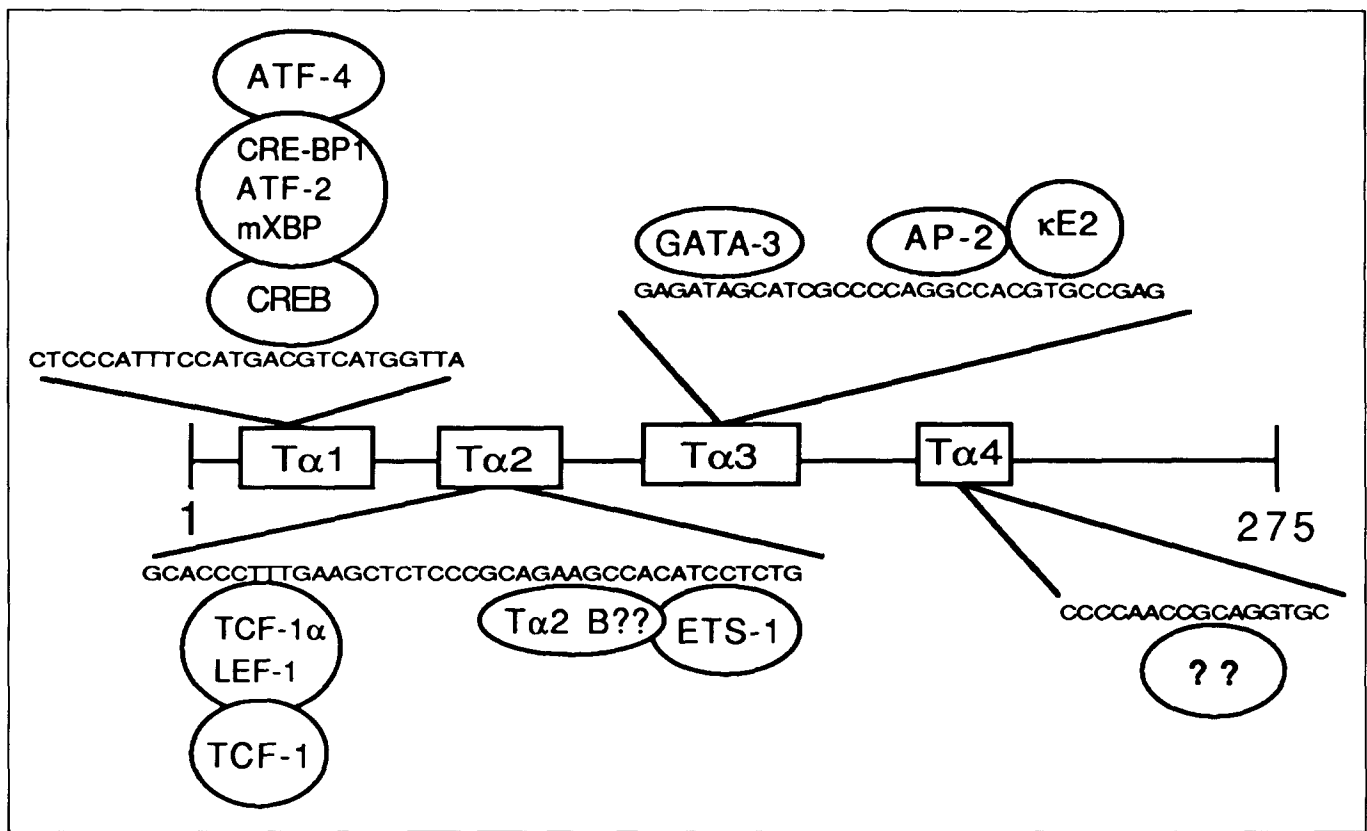


Fig. 4. Schematic illustration of the human T-cell receptor α enhancer and its cognate nuclear binding proteins. These proteins are discussed in detail in the text. The identity of the T α 2 B protein is unknown as are the identities of the T α 4-binding proteins.

unpublished) and almost nothing is known about the identity of the T α 4-binding proteins. Moreover, it is not yet known which E-box-binding proteins are expressed in T cells. Given the importance of this family of transcription factors (the bHLH proteins such as E12/E47 and MyoD) in other developmental systems⁶¹, this remains an important issue. Because each of the factors that has been shown to bind to the α TCR enhancer belongs to a family of nuclear proteins, it remains unclear which of the family members is involved in regulating which gene at a given point in T-cell development. In addition, the available evidence suggests that many of these factors may themselves be regulated at the post-transcriptional and post-translational levels, and currently, relatively little about these processes is understood. The role of the enhancers and of silencers in determining lineage-specific gene expression is also not clear. Finally, almost nothing is known about what regulates the regulators, that is the mechanisms that regulate the expression of Ets-1, GATA-3, TCF-1 or TCF-1 α /LEF-1 during thymocyte differentiation.

Over the next few years, gene knock-out experiments, using homologous recombination along with the expression of *trans*-dominant inhibitors *in vitro* and *in vivo*, should help to shed light on the function of the individual transcription factors. Biochemical studies should help to elucidate the post-transcriptional and post-translational regulation of these factors. Finally, studies of the transcriptional regulation of these transcription factors themselves may move us one step closer to identifying the true master switch gene(s) for the T-cell lineage. Given the current level of interest in these problems and the reagents now available, further insights into

the molecular mechanisms underlying T-cell determination and differentiation can be expected in the near future.

Jeffrey M. Leiden is at the Howard Hughes Medical Institute and the Dept of Internal Medicine, University of Michigan Medical Center, Ann Arbor, MI 48109, USA.

References

- 1 Metcalf, D. (1989) *Nature* 339, 27–30
- 2 Sawyers, C.L., Denny, C.T. and Witte, O.N. (1991) *Cell* 64, 337–350
- 3 Strominger, J.L. (1989) *Science* 244, 943–950
- 4 Haynes, B.F., Denning, S.M., Singer, K.H. and Kurtzberg, J. (1989) *Immunol. Today* 10, 87–91
- 5 Saito, H., Kranz, D.M., Takagaki, Y. *et al.* (1984) *Nature* 312, 36–40
- 6 Chien, Y.-H., Becker, D.M., Lindsten, T. *et al.* (1984) *Nature* 312, 31–35
- 7 Sim, G.K., Yagüe, J., Nelson, J. *et al.* (1984) *Nature* 312, 771–775
- 8 Raulat, D.H., Garman, R.D., Saito, H. and Tonegawa, S. (1985) *Nature* 314, 103–107
- 9 Snodgrass, H.R., Kisielow, P., Kiefer, M., Steinmetz, M. and von Boehmer, H. (1985) *Nature* 313, 592–595
- 10 Chien, Y., Iwashima, M., Kaplan, K.B., Elliott, J.F. and Davis, M.M. (1987) *Nature* 328, 677–682
- 11 Winoto, A. and Baltimore, D. (1989) *Nature* 338, 430–432
- 12 Ho, I.-C., Yang, L.-H., Morle, G. and Leiden, J.M. (1989) *Proc. Natl Acad. Sci. USA* 86, 6714–6718
- 13 Winoto, A. and Baltimore, D. (1989) *EMBO J.* 8, 729–733
- 14 Hockett, R.D., de Villartay, J.-P., Pollock, K. *et al.* (1988) *Proc. Natl Acad. Sci. USA* 85, 9694–9698

REVIEW

- 15 Ho, I-C. and Leiden, J.M. (1990) *Mol. Cell. Biol.* 10, 4720-4727
- 16 Mitchell, P.J., Wang, C. and Tjian, R. (1987) *Cell* 50, 847-861
- 17 Imagawa, M., Chiu, R. and Karin, M. (1987) *Cell* 51, 251-260
- 18 Sen, R. and Baltimore, D. (1986) *Cell* 46, 705-716
- 19 Gottschalk, L.R. and Leiden, J.M. (1990) *Mol. Cell. Biol.* 10, 5486-5495
- 20 Gonzalez, G.A., Yamamoto, K.K., Fischer, W.H. *et al.* (1989) *Nature* 337, 749-752
- 21 Hoeffler, J.P., Meyer, T.E., Yun, Y., Jameson, J.L. and Habener, J.F. (1988) *Science* 242, 1430-1433
- 22 Maekawa, T., Sakura, H., Kanei-Ishii, C. *et al.* (1989) *EMBO J.* 8, 2023-2028
- 23 Ivashkiv, L.B., Liou, H-C., Kara, C.J. *et al.* (1990) *Mol. Cell. Biol.* 10, 1609-1621
- 24 Hai, T., Liu, F., Coukos, W.J. and Green, M.R. (1989) *Genes Dev.* 3, 2083-2090
- 25 Liou, H-C., Boothby, M.R. and Glimcher, L.H. (1988) *Science* 242, 69-71
- 26 Liu, F. and Green, M.R. (1990) *Cell* 61, 1217-1224
- 27 Sheng, M., Thompson, M.A. and Greenberg, M.E. (1991) *Science* 252, 1427-1430
- 28 Waterman, M.L. and Jones, K.A. (1990) *New Biologist* 2, 621-636
- 29 Waterman, M.L., Fischer, W.H. and Jones, K.A. (1991) *Genes Dev.* 5, 1-14
- 30 Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991) *Genes Dev.* 5, 880-894
- 31 van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H. (1991) *EMBO J.* 10, 123-132
- 32 Oosterwegel, M., van de Wetering, M., Dooijes, D. *et al.* (1991) *J. Exp. Med.* 173, 1133-1142
- 33 Ho, I-C., Bhat, N.K., Gottschalk, L.R. *et al.* (1990) *Science* 250, 814-818
- 34 Gunther, C.V., Nye, J.A., Bryner, R.S. and Graves, B.J. (1990) *Genes Dev.* 4, 667-679
- 35 Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C. and Maki, R.A. (1990) *Cell* 61, 113-124
- 36 Urness, L.D. and Thummel, C.S. (1990) *Cell* 63, 47-61
- 37 Bosselut, R., Duvall, J.F., Gégonne, A. *et al.* (1990) *EMBO J.* 9, 3137-3144
- 38 Waslylyk, B., Waslylyk, C., Flores, P. *et al.* (1990) *Nature* 346, 191-193
- 39 Bhat, N.K., Fisher, R.J., Fujiwara, S., Ascione, R. and Papas, T.S. (1987) *Proc. Natl Acad. Sci. USA* 84, 3161-3165
- 40 Bhat, N.K., Komschlies, K.L., Fujiwara, S. *et al.* (1989) *J. Immunol.* 142, 672-678
- 41 Pognonec, P., Boulukos, K.E. and Ghysdael, J. (1989) *Oncogene* 4, 691-697
- 42 Karim, F.D., Urness, L.D., Thummel, C.S. *et al.* (1990) *Genes Dev.* 4, 1451-1453
- 43 Davis, M.M. and Bjorkman, P.J. (1988) *Nature* 334, 395-401
- 44 Evans, T. and Felsenfeld, G. (1989) *Cell* 58, 877-885
- 45 Tsai, S-F., Martin, D.I.K., Zon, L.I. *et al.* (1989) *Nature* 339, 446-451
- 46 Martin, D.I.K., Zon, L.I., Mutter, G. and Orkin, S.H. (1990) *Nature* 344, 444-447
- 47 Romeo, P-H., Prandini, M-H., Joulin, V. *et al.* (1990) *Nature* 344, 447-449
- 48 Evans, T., Reitman, M. and Felsenfeld, G. (1988) *Proc. Natl Acad. Sci. USA* 85, 5976-5980
- 49 Pevny, L., Simon, M.C., Robertson, E. *et al.* (1991) *Nature* 349, 257-260
- 50 Ho, I-C., Voorhees, P., Marin, N. *et al.* (1991) *EMBO J.* 10, 1187-1192
- 51 Yamamoto, M., Ko, L.J., Leonard, M.W. *et al.* (1990) *Genes Dev.* 4, 1650-1662
- 52 Ko, L.J., Yamamoto, M., Leonard, M.W. *et al.* (1991) *Mol. Cell. Biol.* 11, 2778-2784
- 53 Redondo, J.M., Hata, S., Brocklehurst, C. and Krangel, M.S. (1990) *Science* 247, 1225-1229
- 54 Lake, R.A., Wotton, D. and Owen, M.J. (1990) *EMBO J.* 9, 3129-3136
- 55 Clevers, H., Lonberg, N., Dunlap, S., Lacy, E. and Terhorst, C. (1989) *EMBO J.* 8, 2527-2535
- 56 Georgopoulos, K., van den Elsen, P., Bier, E., Maxam, A. and Terhorst, C. (1988) *EMBO J.* 7, 2401-2407
- 57 Winoto, A. and Baltimore, D. (1989) *Cell* 59, 649-655
- 58 Ishida, I., Verbeek, S., Bonneville, M. *et al.* (1990) *Proc. Natl Acad. Sci. USA* 87, 3067-3071
- 59 Brand, A.H., Micklem, G. and Nasmyth, K. (1987) *Cell* 51, 709-719
- 60 Ho, I-C. and Leiden, J.M. (1990) *J. Exp. Med.* 172, 1443-1449
- 61 Murre, C., McCaw, P.S. and Baltimore, D. (1989) *Cell* 56, 777-783



The Budapest Bulletin: Congress News Number 1

Copies of the Final Announcement, which contains registration forms and abstract documents can be obtained from:

Intercongress Ltd, H-1068 Budapest, Dózsa György ut 84/a,
Hungary.

Tel +36-1-1222-203

Fax +36-1-1424-118