ELSEVIER

Gene 274 (2001) 1–13

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Review

Biochemistry and biology of the inducible multifunctional transcription factor TFII-I

Ananda L. Roy[a,b,c,*]

[a]*Department of Pathology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA*
[b]*Programs in Immunology and Genetics, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA*
[c]*Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA*

Received 31 May 2001; received in revised form 6 July 2001; accepted 18 July 2001
Received by A.J. van Wijnen

Abstract

An animal cell has the capability to respond to a variety of external signals through cell surface receptors. The response is usually manifested in terms of altered gene expression in the nucleus. Thus, in modern molecular and cell biology, it has become important to understand how the communication between extracellular signals and nuclear gene transcription is achieved. Originally discovered as a basal factor required for initiator-dependent transcription *in vitro*, recent evidence suggests that TFII-I is also an inducible multifunctional transcription factor that is activated in response to a variety of extracellular signals and translocates to the nucleus to turn on signal-induced genes. Here I review the biochemical and biological properties of TFII-I and related proteins in nuclear gene transcription, signal transduction and genetic disorders.

Keywords: Transcription; Signaling; Growth factors; Kinase; Nuclear translocation

1. Introduction

The signals generated outside a cell are transduced to the nucleus through a series of complicated biochemical steps, ultimately resulting in spatial and/or temporal activation of specific sets of genes (Pawson and Nash, 2000). Thus, there must exist specific protein(s) that serves to direct signal transduction pathways to cell type-specific genes and thereby provide a molecular link between signal transduction and growth, proliferation or developmental programs in a given cell. Transcription factors play a critical role in these processes in general and often serve as links between signal transduction and cell type-specific gene activation. TFII-I is such a ubiquitously expressed multifunctional transcription factor that is activated in response to various extracellular signals and links signal transduction events to transcription.

TFII-I was originally discovered as a basal transcription factor that binds and functions through a core promoter element, initiator (Inr), *in vitro* (Roy et al., 1991). But at the same time it was also realized that TFII-I has additional capability of binding an unrelated upstream element (E-box) that is usually recognized by a family of helix-loop-helix (HLH) proteins viz., USF, and that TFII-I cooperates in binding to both E-box and Inr elements with USF (Roy et al., 1991). These initial observations raised the exciting possibility that TFII-I is a unique transcription factor that can simultaneously function both as a basal factor and as an activator and thus facilitates communication between the basal machinery assembled at the core promoter and the activator complexes assembled at upstream regulatory site(s) (Roy et al., 1991).

In addition to these unique transcription properties, it has been shown that TFII-I is phosphorylated at both serine and tyrosine residues and that tyrosine phosphorylation of TFII-I is required for its transcriptional functions (Novina et al., 1998). Equally interesting is the observation that a variety of extracellular signals mediating through cell surface receptors, including growth factor receptors, lead to enhanced tyrosine phosphorylation and increased transcriptional activity of TFII-I raising the possibility that apart from its transcriptional roles, TFII-I may mediate receptor-mediated signal transduction events (Kim et al., 1998; Novina et al.,

Abbreviations: BEN, binding factor for early enhancer; BR, basic region; Btk, Bruton's tyrosine kinase; DPE, downstream promoter element; ERSE, endoplasmic reticulum stress response element; GRP, glucose regulated protein; HLH, helix-loop-helix; LZ, leucine zipper; PDGF, platelet-derived growth factor; SRF, serum response factor; TAF, TBP-associated factor; TBP, TATA-binding protein; WBS, Williams–Beuren syndrome

* Tel.: +1-617-636-6715; fax: +1-617-636-2990.
E-mail address: ananda.roy@tufts.edu (A.L. Roy).

0378-1119/01/$ - see front matter
PII: S0378-1119(01)00625-4

1998, 1999; Yang and Desiderio, 1997). In B cells a significant fraction of TFII-I is associated constitutively with Bruton's tyrosine kinase (Btk) (Novina et al., 1999; Yang and Desiderio, 1997), mutations which lead to X-linked immune deficiency in humans and mice (Rawlings et al., 1993; Thomas et al., 1993; Tsukada et al., 1993; Vetrie et al., 1993). TFII-I is tyrosine phosphorylated by Btk *in vitro* and upon immunoglobulin receptor cross-linking in B cells (Novina et al., 1999; Yang and Desiderio, 1997). These observations suggest that TFII-I mediates signaling events and links the resulting signal responsive activator complexes to the general transcription machinery.

Recent genetic and biochemical data suggest that TFII-I belongs to a family of protein each having the I-repeat, first identified in the founding member TFII-I (Bayarsaihan and Ruddle, 2000; Franke et al., 1999; O'Mahoney et al., 1998; Osborne et al., 1999; Perez Juardo et al., 1998; Tassabehji et al., 1999; Roy et al., 1997; Yan et al., 2000). Interestingly, both TFII-I and the related protein have been mapped to the breakpoint regions of the 7q11.23 Williams–Beuren syndrome (WBS) deletion (reviewed in Francke, 1999). Furthermore, genetic and biochemical analyses show that each of these proteins has multiple isoforms (Cheriyath and Roy, 2000; Perez Juardo et al., 1998; Tussié-Luna et al., 2001). The tissue and species distribution of these isoforms suggests that they may not have redundant functions. Moreover, recent evidence suggests that the function of these isoforms, even when present simultaneously, may be regulated by their mutual interactions (Cheriyath and Roy, 2000).

How can the basal function of TFII-I be reconciled with its signal dependent inducible transcription functions? Although a complete answer to this problem awaits further study, it appears that one of the isoforms of TFII-I is found constitutively in the nucleus that might selectively function in basal transcription. Thus, different isoforms of TFII-I might serve different transcription functions on different promoters. Additionally, because the isoforms of TFII-I also interact with each other, the homomeric and heteromeric interactions amongst them also might regulate basal versus signal-induced transcription functions. In the following sections, the transcription functions, signaling properties and possible genetic implications of TFII-I and its relative are discussed.

2. Role of TFII-I in transcription

2.1. Inr-dependent function

TFII-I was originally discovered as an Inr-dependent protein 10 years ago (Roy et al., 1991). Accurate transcription initiation in metazoan protein coding genes requires core promoter elements that comprise of a TATA-box, the Inr element and the downstream promoter element (DPE) (Burke and Kadonaga, 1996, 1997; Lee and Young, 2000; Roeder, 1996; Smale and Baltimore, 1989). These elements can be present either individually (distinct promoter) or in combination (composite promoter) (Novina and Roy, 1996). It has been shown that the heterogeneity in core promoter elements allows alternate initiation strategies utilized by eukaryotic genes in response to specific regulatory signals and thus the mechanisms of transcription initiation mediated by the various core promoter elements are also distinct (Novina and Roy, 1996). For example, it is universally accepted that the TATA-box-mediated basal (activator-independent) transcription begins with TATA recognition by the TBP component of TFIID which is sufficient to nucleate the assembly of additional general transcription factors (GTFs) and RNA polymerase II into a functional preinitiation complex (Lee and Young, 2000; Roeder, 1996). However, the corresponding preinitiation complex assembly for Inr-directed basal transcription appears to be more complex (Lee and Young, 2000; Roeder, 1996; Smale, 1997). The Inr-mediated basal transcription appears to require several factors, including TBP-associated factors (TAFs) that are not required for TATA-directed basal transcription (Cheriyath et al., 1998; Martinez et al., 1998; Kaufmann et al., 1998; Roeder, 1996; Smale, 1997). Three different models have been proposed for Inr-dependent transcription in the absence of a TATA-box. The first one proposes direct recognition of the Inr by a TAF component, and/or the DPE when present, followed by stable TFIID binding and subsequent initiation complex formation (reviewed in Roeder, 1996; Lee and Young, 2000). Although any isolated TAF may not exhibit sequence-specific interactions at the Inr element in the absence of a TATA-box, a combination of TAFs may bind sequence specifically to the Inr element regardless of the TATA-box and/or DPE (Chalkley and Verrijzer, 1999). Thus, TAFs may function as core promoter 'selectors' through interactions via any of the core promoter elements (Verrijzer and Tjian, 1996). A second model of Inr recognition implicates independent Inr binding proteins (IBPs). It is postulated that their initial binding to the Inr, followed by secondary interactions with TFIID or components thereof, nucleates assembly of the GTFs at the core promoter. Consistent with the latter model, several factors (e.g. YY1 and TFII-I) have been shown to bind at or adjacent to Inr elements (Roy et al., 1991, 1993a; Usheva and Shenk, 1994). A third model proposes recognition of Inr by RNA polymerase II in the absence of both TAFs and IBPs (Weis and Reinberg, 1997). These observations most likely reflect diversity in Inr elements and corresponding biochemical interactions, especially given that consensus sequences for such elements are loose (Javahery et al., 1994; Roeder, 1996; Roy et al., 1997).

Following the initial *in vitro* studies using partially purified TFII-I and other GTFs on the potent TATA- and Inr-containing composite adenovirus major late (AdML) promoter, more recent biochemical purification, cDNA cloning and functional expression of recombinant TFII-I substantiated that it also functions through the Inr element

in AdML promoter *in vivo* (Roy et al., 1991, 1993a, 1997). cDNA cloning further revealed that it is widely expressed in tissues and the highest expression is observed in brain (Roy et al., 1997). It is also highly expressed in B lymphoid cells (Roy et al., 1997). Although moderate effects of ectopically expressed TFII-I were observed on Inr-dependent transcription most likely reflecting a high endogenous level of TFII-I, co-expression of USF1 markedly enhanced the transcriptional activity of TFII-I through an Inr (Roy et al., 1997). However, the mechanistic basis for the functional synergy between TFII-I and USF1 is not clear at present. It is possible that the synergistic functions of USF1 and TFII-I could be due to synergistic DNA binding mediated through interactions of the DNA-binding/dimerization domains of USF1 and TFII-I with subsequent interactions of activation domains in these proteins with components of the general transcription machinery.

Inr-dependent function of recombinant TFII-I was also shown on a naturally occurring TATA-less Inr-containing murine Vβ 5.2 promoter both *in vitro* and *in vivo* (Manzano-Winkler et al., 1996; Cheriyath et al., 1998). Immunodepletion of nuclear extracts, which were expected to contain all factors necessary for Inr function, with an anti-TFII-I antibody directed against the cDNA-encoded protein completely abrogated transcription of the $TATA^{-}Inr^{+}$ Vβ promoter *in vitro*. Transcription was restored by addition of purified or recombinant TFII-I (Manzano-Winkler et al., 1996). Equally important was the observation that ectopically expressed TFII-I markedly stimulates the expression of the Vβ promoter in *in vivo* (transient transfection) assays (Cheriyath et al., 1998). Moreover, an intact Inr element was required for TFII-I functions since either mutations in the Inr or lack of an Inr element resulted in a lack of transcriptional activation by TFII-I or its artificial derivatives. Furthermore, a fragment of TFII-I (p70) that showed specific Inr binding properties but lacked the Inr-dependent transcriptional activation behaved in a dominant negative fashion when co-expressed with the wild-type TFII-I in these assays (Cheriyath et al., 1998). However, the Inr-specific activation functions of p70 were restored when the activation domain of GAL4 was fused to it (Cheriyath et al., 1998). These studies thus demonstrated that TFII-I has at least two separable domains: an N-terminal DNA-binding domain and a C-terminal activation domain and that the Inr-dependent specificity is dictated by its DNA-binding properties.

2.2. Role of TFII-I in Inr-like or non-consensus Inr elements

The transcriptional function of TFII-I has also been shown on a variety of Inr-like elements (that do not correspond to the sequence of the classical Inr sequence) from both cellular and viral genes. For example, it has been demonstrated that TFII-I binds and functions through the KDR/flk-1 Inr element (Wu and Patterson, 1999). KDR/flk-1 is an endothelial specific growth factor receptor that is regulated during endothelial cell development. Although the sequence of this Inr element deviates from the classical Inr, both *in vitro* and *in vivo* experiments clearly showed that TFII-I functions trough the KDR/flk-1 Inr (Wu and Patterson, 1999). TFII-I was also shown to bind to a functional Inr element of the human Mullerian-inhibiting substance (MIS) and the murine ribonuclease reductase R1 promoters (Johansson et al., 1995; Morikawa et al., 2000). Like the KDR/flk-1 Inr element, the MIS Inr element although fully functional does not correspond to the classical Inr consensus (Morikawa et al., 2000). Similarly, TFII-I functions through a core element termed transcription start site core (TSSC), that bears a high degree of sequence homology to the classic Inr element in the Rous sarcoma virus (RSV) long terminal repeat promoter (Mobley and Sealy, 2000). Like many other viral promoters, RSV is a composite promoter containing both TATA and Inr-like elements and both core elements are required for efficient function. Binding of TFII-I at the TSSC of RSV and corresponding transcription function was convincingly demonstrated (Mobley and Sealy, 2000). Taken together, these data suggest that TFII-I can function through both classical and non-classical Inr elements. It will be interesting to determine whether the function of TFII-I through non-classical Inr elements also requires the same structural domains required for its function through the classical Inr.

2.3. Role of TFII-I through upstream/activator sites

Along with the demonstration that TFII-I binds to the Inr element both individually and synergistically with the HLH activator protein USF, these interactions were also observed at the AdML-derived upstream E-box element, a natural and high affinity binding site for USF (Roy et al., 1991). This was the first demonstration that a transcription factor can bind specifically to both activator sites and core promoter elements apparently through distinct DNA-binding domains. More importantly, simultaneous binding of both of these proteins at both sites raised the notion that they facilitate efficient communication between the upstream activator and basal machinery formed at the core promoter (Roy et al., 1991). Consistent with these observations, it was shown that recombinant TFII-I, together with USF1, functions as an activator through an upstream E-box element in the absence of a functional Inr (Roy et al., 1997). Moreover, these proteins interact physically, both on and off the DNA, further providing a mechanistic basis for the observed functional synergy at both sites (Roy et al., 1997). Interestingly, the functional synergism requires both A and B domains of USF1 that are not required for independent function of USF1 suggesting a separate co-regulatory function of TFII-I through oligomerization with USF1 (Roy et al., 1997; Luo and Sawadogo, 1996). Detailed structure-functional analysis is required to determine which domains of TFII-I are required either for physical association or for its functional synergy with USF1.

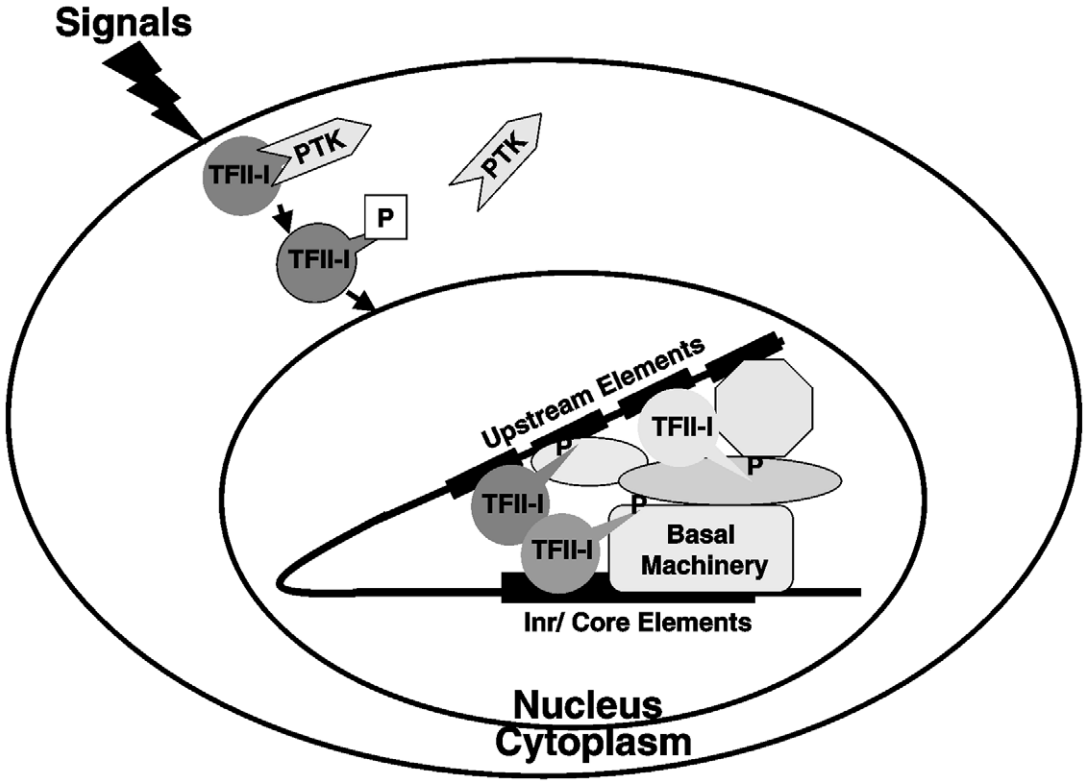

Fig. 1. TFII-I is a multifunctional transcription factor that binds to both a core promoter element (Inr) and various upstream elements thereby facilitating communication between the basal machinery and upstream activators. Both transcription functions of TFII-I are dependent on its phosphorylation status. In resting primary cells, TFII-I remains sequestered in the cytoplasm with a non-receptor protein tyrosine kinase (e.g. Btk or c-Src). TFII-I gets tyrosine phosphorylated and released from the kinase in response to extracellular signals. The tyrosine phosphorylated TFII-I translocates to the nucleus for gene activation.

Apart from its interactions with the upstream E-box and corresponding interactions with USF1, TFII-I also interacts with c-Myc at least in *in vitro* assays (Roy et al., 1993b). Like USF, these interactions are observed both at the Inr element and at the E-box. However, unlike USF1, interactions of TFII-I with c-Myc resulted in an inhibition of TFII-I and Inr-dependent transcription (Roy et al., 1993b). These results indicated for the first time that c-Myc can directly inhibit genes via the Inr element (Roy et al., 1993b). Because many of the cell adhesion molecules have TATA-less but Inr-containing genes that could be repressed by c-Myc, these results may provide a possible rationale for loss of cell adhesion that is presumably associated with transformation by c-Myc. Indeed, Inr-dependent inhibitory function of c-Myc has now been shown on a number of genes and is now widely accepted as one of the physiological functions of c-Myc (reviewed in Dang, 1999). However, it is not yet known in how many instances the repression is mediated through TFII-I under physiological conditions and further studies are required to elucidate the precise molecular mechanisms of Myc-mediated repression via the Inr element.

One of the most surprising and perhaps major breakthroughs came when Grueneberg et al. (1997) independently cloned TFII-I. Although homeodomain protein Phox1 interacts functionally with serum response factor (SRF) to activate *c-fos* gene in transient transfection experiments, a stable complex of these proteins bound to serum response element (SRE) could not be observed *in vitro* (Grueneberg et al., 1997). In an attempt to reconstitute this stable binding that presumptively occurs *in vivo*, these authors undertook a biochemical complementation assay by fractionating HeLa nuclear extract. This approach resulted in the identification of a protein named SPIN (*S*RF-*P*hox1 *In*teracting protein) that when exogenously added to SRF and Phox1 reconstituted a stable SRE-dependent complex (Grueneberg et al., 1997). cDNA cloning and sequence comparison revealed that SPIN is identical to TFII-I (Grueneberg et al., 1997). It was shown that TFII-I interacts with SRF and Phox1 *in vitro* and *in vivo*. Moreover, TFII-I binds to sites overlapping the SRE and the c-*cis*/platelet-derived growth factor (PDGF)-inducible element (SIE) and together with Phox1 and SRF stimulates the transcription of the *c-fos* promoter *in vivo* (Grueneberg et al., 1997). Together with the fact that TFII-I/SPIN binds to the Inr element, these data corroborated the notion that TFII-I is a multifunctional transcription factor that promotes formation of gene-specific activator complexes and facilitates the communication of such complexes with the basal transcription machinery (Grueneberg et al., 1997) (Fig. 1).

More recently, TFII-I was also isolated as a protein that binds to an upstream element called the endoplasmic reticulum stress response element (ERSE) in the promoter regions of glucose regulated protein (GRP) genes (Parker et al., 2001). Depletion of stored endoplasmic calcium results in a stress signal that causes induction of a family of GRPs that encode chaperones (Kaufman, 1999; Parker et al., 2001). These proteins play a crucial role in unfolded protein response (UPR) that is linked to several genetic diseases (reviewed in Kaufman, 1999). TFII-I was discovered by biochemical fractionation and subsequent microsequencing in an attempt to isolate ERSE-binding factor (ERSF) (Parker et al., 2001). Subsequent studies showed that TFII-I binds to ERSE and upregulates transcription of grp78 in response to depletion of endoplasmic reticulum stored calcium by thapsigargin treatment. The level of TFII-I transcript was also shown to increase in response to thapsigargin treatment and provided an explanation for the increase in ERSF binding upon endoplasmic reticulum stress (Parker et al., 2001). Interestingly, TFII-I was found to interact physically and functionally with ATF6, a signal-induced transcriptional activator that is also required for optimal transcriptional induction of grp78. Because ATF6 can also interact with SRF, these observations suggest that TFII-I has the potential of interacting with a variety of signal-induced transcriptional activators in a promoter context-dependent or signal-dependent manner (Parker et al., 2001). Along these lines it is also worth noting that TFII-I has been shown to interact with components of NFκB to facilitate binding of the latter, although the physiological implications of these interactions are not yet known (Montano et al., 1996).

2.4. Structure-function analysis of TFII-I

The binding of TFII-I at multiple sequence elements together with its transcriptional activation through these sites suggested unique DNA-binding potentials. Moreover, as mentioned above, TFII-I also interacts with several transcriptional activators. Thus, it is critical to understand its structure-function relationships.

The DNA-binding domain of TFII-I was first discovered through use of limited protease digestion (Cheriyath et al., 1998). Given that it is a relatively large protein, protease resistance of TFII-I was remarkable. For example, digestion with thrombin even for 1 h showed only one cleavage site at amino acid 677 that divided the protein into two domains: a 70 kDa N-terminal domain that retained DNA binding and a 43 kDa C-terminal domain. This pattern of cleavage was unchanged even when TFII-I was bound to DNA (Cheriyath et al., 1998). Based on these biochemical experiments, the N-terminal half (p70) was cloned, expressed and isolated in a recombinant form (Cheriyath et al., 1998). The p70 mutant of TFII-I showed specific Inr binding *in vitro* but lacked any detectable transcription functions in transient transfection assays using Vβ as a reporter. More important, the p70 mutant behaved as a dominant negative mutant of wild-type TFII-I function when both proteins were co-expressed (Cheriyath et al., 1998). Although p70 lacked an independent activation domain, fusion of the GAL4 activation domain to p70 rescued its transcriptional potentials. Hence, the C-terminal domain of TFII-I must contain or be part of an activation domain. Interestingly, this activation domain is not required for Inr-specific transcription since the GAL4 activation domain imparts Inr-specific function suggesting that the Inr specificity is largely dictated by the DNA-binding domain of TFII-I. On the other hand, when the C-terminal 280 amino acids were fused to the DNA-binding domain of GAL4 (1–147), it failed to impart any detectable transcriptional responses from a promoter that contained five GAL4 binding sites upstream of a TATA-box (Cheriyath et al., 1998). Thus, this C-terminal domain of TFII-I may be necessary but not sufficient for activation function and requires other portions of TFII-I for appropriate transcriptional responses.

The primary amino acid structure of TFII-I is quite unique and consistent with its multifunctional properties. It comprises six direct reiterated I-repeats, R1–R6, each containing a putative HLH motif, but apparently only one basic region (BR) just before R2 (Roy et al., 1997). The BR in traditional HLH proteins has been shown to constitute a sequence-specific DNA-binding domain (Ferre-D'Amare et al., 1994). Each of the 90 amino acid I-repeats, by virtue of having the potential HLH motifs, may represent protein–protein interaction surfaces (Ferre-D'Amare et al., 1994). Thus, there may be potentially six tandem HLH domains in TFII-I, each perhaps mediating a distinct protein–protein interaction (Roy et al., 1997) (Fig. 2).

Recent studies revealed a further glimpse of its structure-function relationship. As anticipated, the BR (amino acids 301–306) is indeed necessary for DNA binding since deletion of this region leads to a loss of DNA binding by TFII-I on both Vβ-derived Inr element and *c-fos*-derived upstream element overlapping the SRE (Cheriyath and Roy, 2001). However, it is not yet determined whether the BR is both necessary and sufficient for DNA binding. Moreover, it is not yet clear whether there are any other DNA-binding domains in TFII-I in addition to the BR. In this regard, it should be noted that while the TFII-I recognition element from the *c-fos* promoter lies upstream of the transcription start site, the sequence matches a consensus Inr element (Grueneberg et al., 1997). Thus, the BR may recognize only this sequence albeit in a context-dependent fashion and there might be other DNA recognition surfaces present in TFII-I.

Surprisingly, deletion of the N-terminal 90 amino acids, that includes a putative leucine zipper (LZ), led to a loss of binding to the Vβ Inr and the *c-fos* upstream sites, despite the fact that this mutant protein (ΔN90) contained an intact BR (Cheriyath and Roy, 2001). Although it is not known why ΔN90 fails to bind DNA, the lack of DNA binding may reflect its ability to efficiently oligomerize due to lack of the

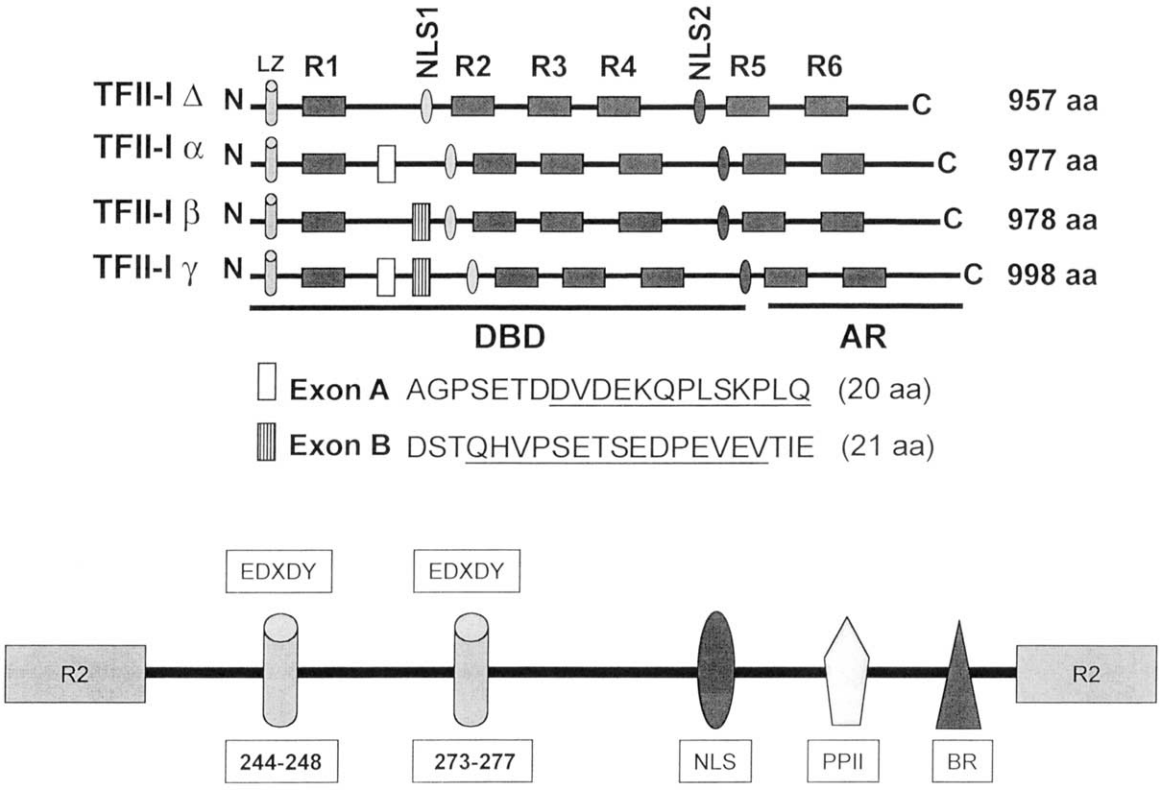

Fig. 2. Schematics of TFII-I isoforms. The closed boxes represent the direct I-repeats, R1–R6, and the open boxes represent the additional amino acids encoded by either exon A or exon B. LZ denotes the putative LZ and NLS1 and 2 denote the nuclear localization signals. Of these only the NLS1 appear to be functional in cell types tested. The two functional domains, the DNA-binding domain (DBD, p70) and the activation region (AR, p43), are also indicated. The sequence of the exons A and B are denoted below while the peptides used to raise isoform-specific antibodies are underlined. The other functional sites in between R1 and R2 are shown in the bottom panel. The two Src phosphorylation sites (EDXDY) with their amino acid positions, the putative PPII helix and the BR/DNA-binding region are indicated in this scheme.

LZ. However, oligomerization per se may not be sufficient for DNA binding since the ΔN90 mutant dimerizes with itself (Cheriyath and Roy, 2001, see also below). It is very likely that the N-terminal region containing part or all of the 90 amino acids may directly contact DNA or indirectly contribute to DNA binding by altering the native structure. Thus, removal of this region may expose a negative inhibitory domain that covers the true DNA-binding domain.

Consistent with the lack of DNA-binding ability of the BR mutant (ΔBR) and ΔN90 mutants, they failed to activate the TFII-I-dependent Vβ and *c-fos* reporters (Cheriyath and Roy, 2001). The nuclear translocation of these mutants was unaffected. Therefore, together with the data showing that the p70 mutant exhibits DNA binding but lacks an activation domain (Cheriyath et al., 1998), proper transcriptional function of TFII-I requires both its DNA-binding capabilities and its transcriptional activation domain. Hence, TFII-I behaves like a classical transcription factor with an N-terminal DNA-binding domain and a separable C-terminal activation domain. In the future, it will be interesting to determine what component(s) within the basal machinery interact with the activation domain of TFII-I.

2.5. Isoforms of TFII-I: homomeric and heteromeric interactions

In addition to the well-characterized form of TFII-I, recent genetic mapping data indicated additional alternatively spliced isoforms of TFII-I in human and in mice (Perez Juardo et al., 1998; Wang et al., 1998). Thus, besides the 957 amino acid form of TFII-I (referred to as Δ), three other alternatively spliced isoforms exist in human and perhaps an additional one exists in mice. These additional isoforms are called: α (977 amino acids), β (978 amino acids) and γ (998 amino acids) (Cheriyath and Roy, 2000). In comparison to the Δ-isoform, the α-isoform contains an additional 20 amino acids (encoded by exon A, Fig. 2), the β-isoform contains an additional 21 amino acids (encoded by exon B, Fig. 2) and the γ-isoform, that arises by the presence of both exons A and B, contains 41 additional amino acids (Fig. 2) (Cheriyath and Roy, 2000; Perez Juardo et al., 1998; Wang et al., 1998). Of the four isoforms, the γ-isoform is most likely expressed predominantly, if not exclusively, in neuronal cells (Perez Juardo et al., 1998). Our preliminary data also indicate the presence of at least two additional isoforms (V. Cheriyath and A.L.R.,

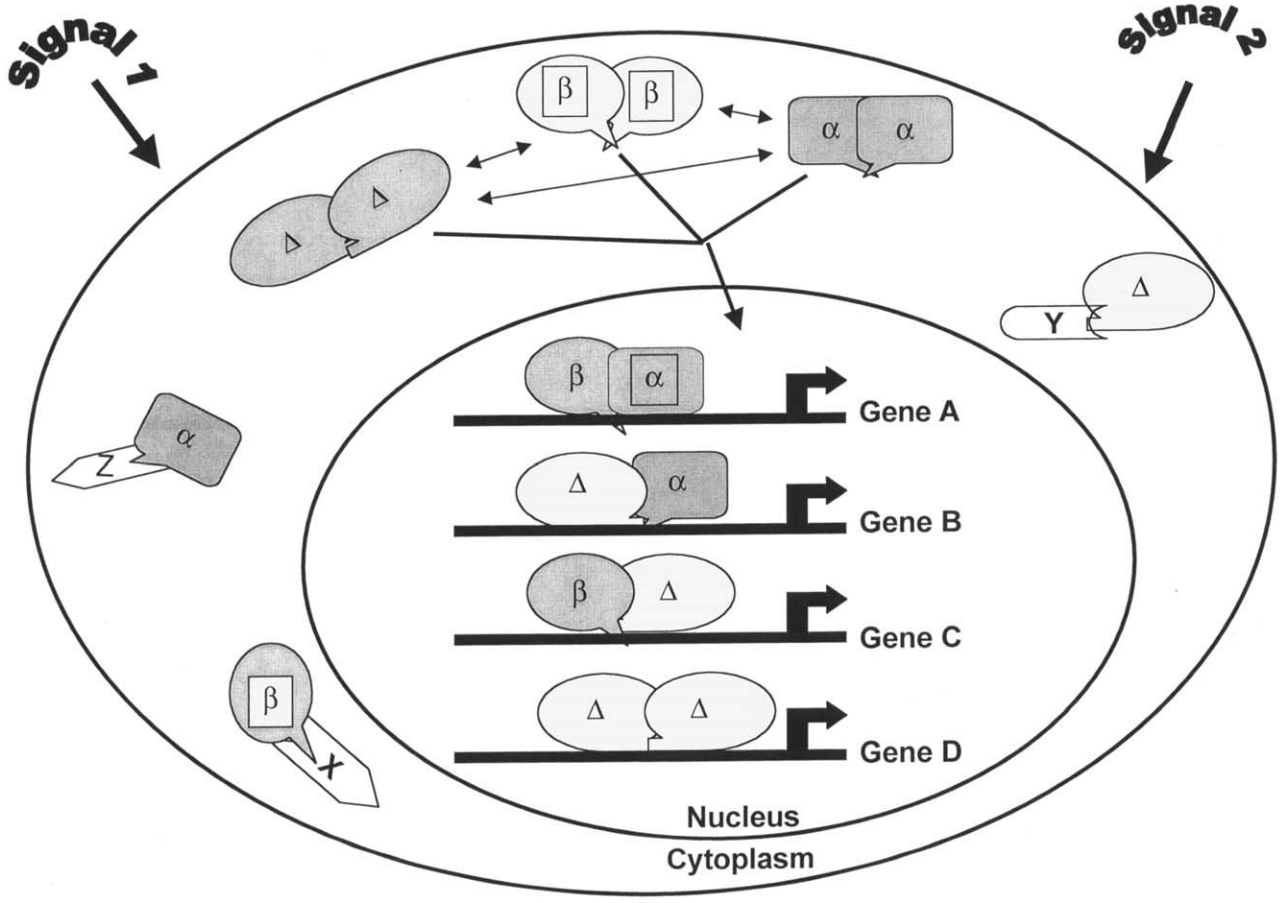

Fig. 3. Model indicating homomeric and heteromeric interactions between TFII-I isoforms. TFII-I isoforms may largely exist as homomers or heteromers (in association with other proteins, e.g. X, Y, and Z) in the cytoplasm. These latter interactions may prevent them from forming either homomeric or heteromeric complexes with themselves. In the nucleus, TFII-I exists predominantly in heteromeric complexes and such nuclear entry might also be regulated via complex or sub-complex formation amongst the isoforms. Depending on the cell type or species, a sub-population of homomers might also be found in the nucleus. This entire process could be signal-dependent such that different signals (e.g. 1 versus 2) may induce formation of distinct sub-complexes. The nuclear sub-complexes may in turn translocate to the nucleus to mediate differential gene regulation (the different genes are denoted as A, B, C and D) perhaps through differential protein–protein and/or protein–DNA interactions. Thus, spatial and/or temporal control of sub-complex formation may allow additional regulatory steps for the TFII-I network.

unpublished data). The β-isoform is expressed much higher in murine cells than in human cells (Cheriyath and Roy, 2000) and the α-isoform appears to be lacking in murine cells (Wang et al., 1998). The expression levels and patterns of these isoforms in various cell types and species suggest non-redundant functions (Cheriyath and Roy, 2000). Each isoform contains all the I-repeats, the BR and the putative LZ and a functional nuclear localization signal (amino acids 297–304 with respect to the Δ-isoform) (Cheriyath and Roy, 2000). More interestingly, these isoforms exhibit both homomeric and heteromeric interactions with themselves that lead to their preferential nuclear localization (Cheriyath and Roy, 2000). Employing isoform-specific antibodies, it was shown that the isoforms form a stable complex *in vivo* in the nucleus. Moreover, a nuclear localization deficient mutant of the Δ-isoform, which on its own remains exclusively in the cytoplasm, was found in the nucleus when co-expressed with any of the wild-type isoforms of TFII-I (Cheriyath and Roy, 2000). Thus, either homomeric or heteromeric interactions among the isoforms lead to preferential nuclear localization. Although each isoform individually bound to DNA and activated transcription both from the Vβ and *c-fos* promoters, co-expression of different combinations of TFII-I isoforms leads to enhanced basal activity of the Vβ promoter and attenuated signal responsive activity of the *c-fos* promoter (Cheriyath and Roy, 2000). These data collectively suggest that TFII-I has the potential of differentially regulating its target genes via homo- or heteromerization of its isoforms (Fig. 3). However, it is not yet known whether these isoforms form dimers or higher order oligomers.

One of the interesting features of these interactions is that ΔN90 readily interacts with the β-isoform but not with the Δ-isoform from which it was originally derived (Cheriyath and Roy, 2001). However, ΔN90 interacts with itself leading to the suggestion that although the LZ may represent a primary interaction domain, secondary interactions are perhaps mediated by the I-repeats. In agreement with this

and the anticipation that the I-repeats represent interaction modules, both I-repeats R1 and R2, either in combination or individually, were shown to mediate homomeric and heteromeric interactions, although the extent of these interactions is weaker than that mediated by the full-length protein (Cheriyath and Roy, 2001). Hence, although the individual repeats can mediate homomeric interactions, they alone are insufficient and require the N-terminal 90 amino acids for optimal interaction. Deletion of either the BR or repeats R6 through R3 from the C-terminal end had no significant effects on homomeric interactions when compared to the wild-type TFII-I (Cheriyath and Roy, 2001).

Based on these observations, it is proposed that the LZ (conserved in the TFII-I family) in Δ-isoform remains in a 'closed conformation'. The 'availability' of the LZ is regulated in a signal-induced fashion to interact with other partners: either another molecule of Δ or another molecule of β or an unrelated molecule. The β-isoform, in contrast, can assume a constitutively 'open conformation'. Thus, the regulated availability of the LZ might control the extent of homomeric or heteromeric interactions and consequently basal versus induced transcription function of TFII-I (Cheriyath and Roy, 2001).

3. Role of TFII-I in signaling

3.1. Induced phosphorylation

TFII-I is phosphorylated 'basally' (in the absence of apparent extracellular signals) at both serine/threonine and tyrosine residues (Novina et al., 1998). Based on the DNA-binding properties of both the bacterially expressed recombinant TFII-I and *in vitro* dephosphorylation of native TFII-I, it is concluded that phosphorylation is dispensable for its specific DNA-binding activity (Novina et al., 1998). However, dephosphorylated native TFII-I *in vitro* and the putative Src tyrosine phosphorylation site mutant (Y-F248) *in vivo* failed to support Vβ-dependent transcription, suggesting that at least tyrosine phosphorylation of Y248 is required for its transcriptional activity (Novina et al., 1998). Although tyrosine phosphorylation of TFII-I is required for its transcription function, the precise reason for such requirement is not yet elucidated. The possibility that tyrosine phosphorylation of TFII-I is required for its nuclear localization is ruled out since the Y-F248 mutant TFII-I readily translocates to the nucleus upon ectopic expression. Because phosphorylation is not also required for its specific DNA binding, the protein–protein interactions of TFII-I with the basal machinery may be dependent upon its phosphorylation status (Novina et al., 1998). It is also important to emphasize that 'basal' phosphorylation of TFII-I may not be so under cell culture conditions since the effects of cell cycle and/or growth factors present in serum are not taken into consideration (Novina et al., 1998). It is interesting to note that the same mutant Y-F248 also failed to transcriptionally activate the GRP promoter in response to thapsigargin treatment in transient transfection assays, suggesting that activation of TFII-I by stress response also involves tyrosine phosphorylation of TFII-I (Parker et al., 2001).

3.2. Role of TFII-I in B cell signaling

Btk is a non-receptor protein tyrosine kinase that is preferentially expressed in hematopoietic cells of B and myeloid lineages (Desiderio, 1997). Btk belongs to a family of Src-like tyrosine kinases called the Tec family that are characterized by the presence of a pleckstrin homology (PH) domain (Desiderio, 1997; Lemmon et al., 1996). Mutations in Btk result in X-linked agammaglobulinemia (XLA) in humans in which there is a near absence of B cells and thus a failure to produce serum immunoglobulin (Ig) (Rawlings et al., 1993; Thomas et al., 1993; Tsukada et al., 1993; Vetrie et al., 1993). A spontaneous mutation in mice (R28C) produces X-linked immunodeficiency (xid) with milder defects. Because Btk appears to be critical for multiple signaling pathways important for B cell differentiation and proliferation, it is of considerable interest to determine its downstream target (Desiderio, 1997). In such an attempt, Desiderio and colleagues immunoprecipitated and subsequently cloned a protein factor, BAP-135, that is constitutively associated with Btk in a human B cell line (Yang and Desiderio, 1997). Interaction of BAP-135 with Btk required the PH domain of Btk and tyrosine phosphorylation of BAP-135 *in vitro* by Btk required the kinase domain of Btk. Most significantly, B cell antigen receptor (BCR) engagement leads to transient tyrosine phosphorylation of BAP-135 (Yang and Desiderio, 1997). Although the sequence of TFII-I was not yet published at that time, BAP-135 turned out to be identical to TFII-I Δ-isoform. Thus, TFII-I was independently cloned from three different sources (Grueneberg et al., 1997; Roy et al., 1997; Yang and Desiderio, 1997).

Following these remarkable studies by Desiderio and colleagues, subsequent experiments from our laboratory demonstrated a functional interaction between TFII-I and Btk (Novina et al., 1999). When co-expressed with wild-type Btk, both its tyrosine phosphorylation and transcriptional activation were markedly enhanced in transient transfection assays (Novina et al., 1999). Either the xid mutant Btk (R28C) or the kinase domain mutant (K430E) compromises its ability to enhance both the tyrosine phosphorylation and the transcriptional activity of TFII-I in transient transfection assays. Constitutive association of TFII-I *in vivo* with wild-type Btk required an intact PH domain since the xid/R28C mutant failed to interact with TFII-I, although the kinase dead/K430E mutant interacts with TFII-I like wild-type Btk. Most significantly, TFII-I was found to be associated with Btk in B cell cytoplasm but membrane IgM cross-linking led to dissociation of TFII-I from Btk. While TFII-I was found both in the nucleus and

cytoplasm of wild-type and xid primary resting B cells, nuclear TFII-I was approximately three-fold higher in xid B cells. Importantly, receptor cross-linking of wild-type (but not xid) B cells results in increased nuclear import of TFII-I. Thus, the PH domain of Btk is primarily responsible for its physical interaction with TFII-I. In addition, an intact kinase domain of Btk is required to enhance transcriptional activity of TFII-I (Novina et al., 1999). These data suggest that Btk tethers TFII-I in the cytoplasm in resting B cells. BCR signaling results in tyrosine phosphorylation of TFII-I by Btk and the subsequent release of TFII-I for nuclear import. The regulated nuclear import and thus transcriptional activity of TFII-I by Btk implies that TFII-I-dependent transcription may be required for proper B cell function such that disruptions in the Btk–TFII-I pathway may lead to defective TFII-I-dependent gene activation and defective immune function (Novina et al., 1999) (Fig. 1).

3.3. TFII-I-mediated signaling in non-lymphoid cells

While Btk has restricted expression in myeloid and B cell lineages (Desiderio, 1997), TFII-I is ubiquitously expressed and tyrosine phosphorylated in a variety of cell types suggesting that TFII-I can be tyrosine phosphorylated by other ubiquitous kinases (Novina et al., 1998). Consistent with these expectations, TFII-I undergoes induced tyrosine phosphorylation in response to pervanadate treatment and PDGF and epidermal growth factor (EGF) stimulation in fibroblasts and epithelial cells (Kim et al., 1998; Novina et al., 1998). In addition, induction of the *c-fos* promoter in response to potent mitogenic signals including serum, PDGF, EGF, lysophosphatadic acid (LPA) and tetradecanoyl phorbol acetate (TPA) was enhanced in the presence of ectopic TFII-I, suggesting that TFII-I is either directly downstream of or synergies with these signaling pathways (Kim et al., 1998). Compatible with observations that TFII-I binds sites overlapping the SRE and SIE on the *c-fos* promoter, TFII-I interacts with both SRF and STAT1 and STAT3 *in vitro* (Kim et al., 1998). Finally, the TFII-I-dependent transcriptional activation of the *c-fos* promoter in response to growth factors was severely impaired in the presence of ectopically expressed dominant negative N17 Ras, suggesting that the TFII-I function requires an intact *Ras* signaling pathway (Kim et al., 1998).

These observations beg a major question: what tyrosine kinase(s) phosphorylate TFII-I in non-lymphoid cells? Based on the fact that TFII-I contains two Src phosphorylation sites and a polyproline type II (PPII) helix that can potentially interact with a Src-homology domain 3 (SH3) (Fig. 2, Roy et al., 1997), it may be argued that Src could physically interact with TFII-I and phosphorylate it. Furthermore, one of the Src tyrosine phosphorylation sites (Y248) is required for transcriptional activity of TFII-I (Novina et al., 1998). Indeed, it has now been shown that Src associates with TFII-I through the SH3 domain of Src and the PPII helix of TFII-I and that TFII-I undergoes Src-mediated induced tyrosine phosphorylation (Cheriyath et al., submitted). Furthermore, the major Src-mediated tyrosine phosphorylation site of TFII-I appears to be Y248, a site required for its induced transcriptional activation of the *c-fos* promoter in response to either EGF or ectopically expressed Src. Lastly, Src controls the nuclear translocation of endogenous TFII-I in fibroblasts (lacking Btk) in a signal-dependent fashion (Cheriyath et al., submitted). These features of TFII-I are completely lost in $Src^{-/-}$ cells or when a dominant negative Src is stably expressed in fibroblasts. Collectively, these data demonstrate that c-Src physically and functionally phosphorylates TFII-I and controls its nuclear localization in fibroblasts and perhaps in other non-lymphoid cells. However, it should be pointed out that recent indirect transient transfection experiments suggest that TFII-I is phosphorylated by JAK2 in fibroblasts (Kim and Cochran, 2001). The site of phosphorylation is suggested to be Y248. Yet, both the tyrosine phosphorylation and nuclear translocation of TFII-I in $JAK2^{-/-}$ cells appear to be identical to the wild-type fibroblasts (Cheriyath et al., submitted). The most likely explanation for this discrepancy could be the systems used. It will be important in the future to unambiguously assign direct JAK2-dependent tyrosine phosphorylation of TFII-I. Nevertheless, it is most likely that multiple tyrosine kinases might target TFII-I in multiple cell types or under distinct physiological conditions.

3.4. Other signaling pathways involving TFII-I

While the tyrosine phosphorylation of TFII-I is important and functionally significant, the role of serine phosphorylation in TFII-I is less clear at present. Because there are numerous potential S/T phosphorylation sites, it is difficult to assess which one of these are utilized inside the cell (Novina et al., 1998; Roy et al., 1997). Despite these obvious difficulties, some of the predictions regarding its serine/threonine phosphorylation sites could be made. Thus, a consensus (PXSP) mitogen activated protein kinase MAPK/extracellular signal regulated kinase (ERK) phosphorylation site between amino acids 631 and 634 was observed based on sequence analysis (Novina et al., 1998; Roy et al., 1997). Consistent with these predictions, it has been shown that serine 633 is phosphorylated by MAPK *in vitro* and a mutation of serine to alanine (S-A 633) diminishes TFII-I-dependent activation of the *c-fos* promoter in transient transfection experiments (Kim et al., 1998; Kim and Cochran, 2000; Novina et al., 1998). Furthermore, Ras and RhoA synergies with TFII-I for activation of the *c-fos* promoter. TFII-I binds to MAPK through a presumptive D-box that is known to mediate MAK interactions, although it is not clear from these transient transfection experiments whether such interactions are constitutive or inducible and whether they occur in the cytoplasm or in the nucleus (Kim and Cochran, 2000). Further experiments are required to

clarify these issues and determine the physiological relevance of serine phosphorylation in TFII-I.

4. Chromosomal location of TFII-I and the related gene

4.1. Haploinsufficieny in WBS

The human *TFII-I* gene (called *GTF2I*) was mapped in a region of chromosome 7 (7q11.23) that is deleted in WBS (Perez Juardo et al., 1998). WBS is a neurodevelopmental disorder with multisystem manifestations, including supravalvar aortic stenosis, hypercalcemia in infancy, mild to moderate mental retardation, cognitive defects and characteristic facial features (Francke, 1999; Keating, 1997). The frequency of this genetic haploinsufficiency is estimated to be 1 in 20,000 live births (Perez Juardo et al., 1998). *GTF2I* is duplicated in this region and although the centromeric and telomeric copies are nearly identical, the centromeric copy seems to be missing the initiating codon and most likely reflects a pseudogene (*GTF2IP1*) with no predicted protein product (Perez Juardo et al., 1998). The telomeric copy has a total of 32 exons and is predicted to have several alternatively spliced variants (Perez Juardo et al., 1998). This prediction together with the sequencing data paved the way for cDNA cloning and subsequent biochemical characterization of these isoforms (Cheriyath and Roy, 2000). Interestingly, only the telomeric copy is deleted in WBS and the centromeric pseudogene is unaffected (Perez Juardo et al., 1998). Despite the fact that the pseudogene is readily detected in tissues or normal cells, the pseudogene is rarely detected in transformed cells (A.L.R., unpublished data). Whether this is directly related to the process of transformation or simply a byproduct is currently unknown. The high degree (99.9%) of sequence conservation between the telomeric *GTF2I* and centromeric *GTF2IP1* suggests that this duplication is evolutionarily recent (Perez Juardo et al., 1998). Because there exists only a single copy of the *TFII-I* gene (*GtfI2*) in mice and other non-primate mammals, it also suggests that the duplication must have occurred after the primates diverged from other species (Wang et al., 1998). Interestingly, some genes that are associated with the WBS deletion in human chromosome 7 also appear to cluster at the same position on corresponding mouse chromosome 5 (Wang et al., 1998). *Gtf2I* is about 88% identical to *GTF2I* at the DNA level and more than 97% identical at the protein level (Wang et al., 1998). In addition to the corresponding β, γ, and Δ human isoforms, there appears to be an additional spliced variant and the isoform corresponding to the human α-isoform missing in mice (Wang et al., 1998). Both human and mouse genes are fairly ubiquitously expressed and *in situ* studies reveal an early embryonic expression (7 dpc) in mice (Wang et al., 1998). Although it is likely that the broad phenotypic spectrum associated with WBS is the consequence of deletion of several genes, the haploinsufficiency of the *TFII-I* gene suggests a potential link between TFII-I function and one or more of the WBS phenotypes (Perez Juardo et al., 1998). Many human multisystem disorders appear to be correlated with hemizygosity for transcriptional activators or co-activators and DNA-binding proteins and thus their effects may be dosage-sensitive in general. It should be noted, however, that the B cell function appears to be within normal limits in WBS individuals heterozygous for deletions of *GTF2I* (Perez Juardo et al., 1998).

4.2. TFII-I-related gene and its protein products

The presence of a TFII-I-related gene was reported by several groups as a gene located within the 1.4–1.6 Mb heterozygous WBS deletion and was called *WBSCR11* (Osborne et al., 1999), *GTF2IRD1* (Franke et al., 1999) or *GTF3* (Tassabehji et al., 1999). This gene is located just centromeric to *GTF2I* in the same transcriptional orientation and appears to be a single copy gene and has been deleted in all WBS patients tested thus far (Francke, 1999). There are altogether 27 exons in the *GTF3* gene and the possibility of several spliced isoforms. Others isolated it as a functional protein involved in transcription of diverse genes. For example, BEN (binding factor for early enhancer) was isolated in an elegant one-hybrid screen as a transcription factor that binds to early enhancer of the *Hoxc*8 gene (Bayarsaihan and Ruddle, 2000). The open reading frame of murine BEN encodes 1072 amino acid protein that contains six HLH domains (I-repeats), an N-terminal hydrophobic LZ-like motif, and a serine-rich repeat. Although BEN was isolated and cloned as a novel DNA-binding transcription factor, whether it behaves as an activator or repressor in functional assays was not determined (Bayarsaihan and Ruddle, 2000). The BEN protein is structurally similar to human TFII-I and is an ortholog of the human WBSCR11/GTF2RD1 (Tussié-Luna et al., 2001). An identical protein is also reported as a retinoblastoma protein-associated nuclear factor (CREAM-1) with potential transcriptional activator functions (Yan et al., 2000). A TFII-I-related protein was first reported as a novel troponin I enhancer binding protein, MusTRD1 (muscle TFII-I repeat domain-containing protein 1), that was highly enriched in muscle tissues (O'Mahoney et al., 1998). However, the predicted MusTRD1 protein was 458 amino acids long while the WBSCR11, GTF2IRD1, GTF3 and CREAM-1 were either 944 or 959 amino acids long and their corresponding protein products were all ubiquitously expressed (O'Mahoney et al., 1998). The latter difference in size arises due to alternative splicing (reviewed in Tussié-Luna et al., 2001). Subsequent sequence re-analysis of the MusTRD1 cDNA and ectopic expression of its protein product suggested a 120 kDa protein with 944 amino acids containing five repeats (Tussié-Luna et al., 2001). Despite these confusing nomenclature and apparent size differences, it is clear that MusTRD1/GTF3 is the human ortholog of the corresponding mouse BEN/GTF2IRD1 and there are

several spliced variants. This protein has been referred to as MusTRD1/BEN (Tussié-Luna et al., 2001) (Fig. 4).

The function of MusTRD1/BEN has not yet been well-characterized biochemically. It has been reported as a muscle-specific activator of the troponin I gene (O'Mahoney et al., 1998). It also seems to function as an activator in yeast one-hybrid assays (Yan et al., 2000). However, clear demonstration of its activator function is lacking. Surprisingly, it has also been shown to be a specific repressor of TFII-I function (Tussié-Luna et al., 2001). Although each protein when expressed individually is predominantly in the nucleus, TFII-I is excluded from the nucleus when MusTRD1/BEN is co-expressed with it in eukaryotic cells. Nuclear exclusion of TFII-I results in the repression of the TFII-I responsive *c-fos* gene. Importantly, such nuclear exclusion and concomitant transcriptional repression are specific for TFII-I since MusTRD1/BEN failed to exert these effects on other transcriptional activators. A key to this novel nuclear exclusion function appears to be the serine stretch in MusTRD1/BEN because deletion of this stretch results in co-occupancy of both proteins in the nucleus (Tussié-Luna et al., 2001). Because these experiments have relied on ectopic expression systems, it is not clear yet whether such mechanisms operate *in vivo* (Tussié-Luna et al., 2001). However, it appears that under physiological conditions, expression of MusTRD1/BEN is much lower than that of TFII-I and as such may provide a rationale as to how and why TFII-I might overcome such potent repression mechanisms (Tussié-Luna et al., 2001). Furthermore, preliminary data suggest that BEN might behave as a repressor of the *Hoxc*8 gene (Bayarsaihan and Ruddle, pers. commun.; Tussié-Luna et al., 2001). It would be important in the future to determine the conditions under which MusTRD1/BEN expression might be elevated and whether deletion of both of these proteins in WBS in any way contributes to one or more of the phenotypes associated with the disorder. Taken together, these diverse data suggest that MusTRD1/BEN might behave both as an activator and as a repressor. However, clear assignment of the physiological and biochemical functions of MusTRD1/BEN is required for future studies.

5. Future perspectives

We have gathered a substantial amount of information about TFII-I and its relative MusTRD1/BEN. However, this field of study is in its nascent stages. Although the biochemical function of TFII-I is better understood and its mechanism of signal-induced activation is beginning to be elucidated, the corresponding function of MusTRD1/BEN is still not clear. It is worthwhile mentioning here that all isoforms of TFII-I might not behave in an identical signal-dependent fashion. Indeed, the preliminary data suggest a constitutive nuclear form of TFII-I that might help to explain its basal transcription functions (V. Cheriyath and A.L.R., unpublished data). But more importantly, much needs to be learned about the physiological functions and target genes of TFII-I and MusTRD1/BEN. When and how are these regulators themselves regulated? Do the same signaling pathways target these proteins? Do the opposing functions of these proteins suggest a reciprocal regulation during development and/or differentiation in a given lineage? Some of these exciting questions are about to be explored and the future availability of 'knock-out' and/or transgenic animals should reveal the biological importance of this novel system of transcriptional regulation. In this latter regard, it is worth noting that in an interesting analysis of c-*myc* transgenic mice, the transgene integration induced a deletion that included part of the *GTF2IRD1* gene (Durkin et al., 2001). The levels of the *GTF2IRD1* transcripts were significantly reduced in mice homozygous for the transgene-induced deletion, leading to the postulation that *GTF2IRD1* (coding for MusTRD1/BEN) may not be essential for viability (Durkin et al., 2001). However, abnormalities associated with neurodevelopmental disorders or other non-lethal defects could be looked at more carefully in these mice (Durkin et al., 2001). Moreover, it would also be inter-

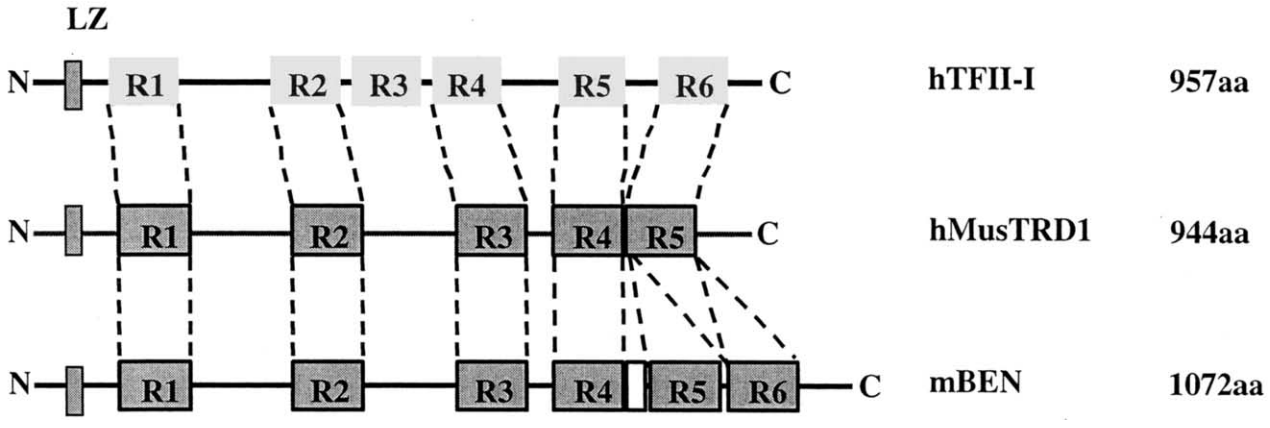

Fig. 4. Schematic organization of the TFII-I family members: human TFII-I (hTFII-I), human MusTRD1 (hMusTRD1) and murine BEN (mBEN). The position of the conserved LZ is shown. mBEN is an ortholog of hMusTRD1.

esting to determine whether such deletions directly contribute to enhanced neoplasia observed in these transgenic mice. Although elegant genetic mapping studies are breaking new grounds, undoubtedly, the animal models will provide future clues to the possible involvement of the TFII-I family of proteins in human genetic disorders like WBS and XLA and perhaps cancer.

Acknowledgements

The work done in my laboratory is supported largely by a NIH grant (AI 45150). I wish to thank members of my laboratory for many exciting discussions and arguments and for critically reading the manuscript. I also thank many of our collaborators for their contributions and input that made this field vigorous and exciting.

References

Bayarsaihan, D., Ruddle, F.H., 2000. Isolation and characterization of BEN, a member of the TFII-I family of DNA-binding proteins containing distinct helix-loop-helix domains. Proc. Natl. Acad. Sci. USA 97, 7342–7347.

Burke, T.W., Kadonaga, J.T., 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. 10, 711–724.

Burke, T.W., Kadonaga, J.T., 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by $TAF_{II}60$ of Drosophila. Genes Dev. 11, 3020–3031.

Chalkley, G.E., Verrijzer, C.P., 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J. 18, 4835–4845.

Cheriyath, V., Roy, A.L., 2000. Alternatively spliced isoforms of TFII-I: complex formation, nuclear translocation and differential gene regulation. J. Biol. Chem. 275, 26300–26308.

Cheriyath, V., Roy, A.L., 2001. Structure-function analysis of TFII-I: roles of the N-terminal end, basic region, and I-repeats. J. Biol. Chem. 276, 8377–8383.

Cheriyath, V., Novina, C.D., Roy, A.L., 1998. TFII-I regulates Vβ promoter activity through an initiator element. Mol. Cell. Biol. 18, 4444–4454.

Dang, C.V., 1999. c-Myc target genes involved in cell growth, apoptosis, and metabolism. Mol. Cell. Biol. 19, 1–11.

Desiderio, S., 1997. Role of Btk in B cell development and signaling. Curr. Opin. Immunol. 9, 534–540.

Durkin, M.E., Keck-Waggoner, C.L., Popescu, N.C., Thorgeirsson, S.S., 2001. Integration of a c-myc transgene results in disruption of the mouse Gtf2ird1 gene, the homologue of the human GTF2IRD1 gene hemizygously deleted in Williams-Beuren syndrome. Genomics 73, 20–27.

Ferre-D'Amare, A.R., Pognonec, P., Roeder, R.G., Burley, S.K., 1994. Structure and function of the b/HLH/Z domain of USF. EMBO J. 13, 180–189.

Francke, U., 1999. Williams-Beuren syndrome: genes and mechanisms. Hum. Mol. Genet. 8, 1947–1954.

Franke, Y., Peoples, R.J., Francke, U., 1999. Identification of GTF2IRD1, a putative transcription factor within the Williams-Beuren syndrome deletion at 7q11.23. Cytogenet. Cell Genet. 86, 296–304.

Grueneberg, D.A., Henry, R.W., Brauer, A., Novina, C.D., Cheriyath, V., Roy, A.L., Gilman, M., 1997. A multifunctional DNA-binding protein that promotes the formation of serum response factor/homeodomain complexes: identity to TFII-I. Genes Dev. 11, 2482–2493.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., Smale, S.T., 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol. Cell. Biol. 14, 116–127.

Johansson, E., Skogman, E., Thelander, L., 1995. The TATA-less promoter of mouse ribonucleotide reductase R1 gene contains a TFII-I binding initiator element essential for cell cycle-regulated transcription. J. Biol. Chem. 270, 30162–30167.

Kaufman, R.J., 1999. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. Genes Dev. 13, 1211–1233.

Kaufmann, J., Ahrens, K., Koop, R., Smale, S.T., Muller, R., 1998. CIF150, a human cofactor for transcription factor IID-dependent initiator function. Mol. Cell. Biol. 18, 233–239.

Keating, M.T., 1997. On the trail of genetic culprits in Williams syndrome. Cardiovasc. Res. 36, 134–137.

Kim, D.-W., Cochran, B.H., 2000. Extracellular signal-regulated kinase binds to TFII-I and regulates its activation of the *c-fos* promoter. Mol. Cell. Biol. 20, 1140–1148.

Kim, D.-W., Cochran, B.H., 2001. JAK2 activates TFII-I and regulates its interaction with extracellular signal-regulated kinase. Mol. Cell. Biol. 21, 3387–3397.

Kim, D.-W., Cheriyath, V., Roy, A.L., Cochran, B.H., 1998. TFII-I enhances activation of the *c-fos* promoter through interactions with upstream elements. Mol. Cell. Biol. 18, 3310–3320.

Lee, T.I., Young, R.A., 2000. Transcription of eukaryotic protein-coding genes. Annu. Rev. Genet. 34, 77–137.

Lemmon, M.A., Ferguson, K.M., Schlessinger, J., 1996. PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface. Cell 85, 621–624.

Luo, X., Sawadogo, M., 1996. Functional domains of the transcription factor USF2: atypical nuclear localization signals and context-dependent transcriptional activation domains. Mol. Cell. Biol. 16, 1367–1375.

Manzano-Winkler, B., Novina, C.D., Roy, A.L., 1996. TFII-I is required for transcription of the naturally TATA-less but initiator-containing Vβ promoter. J. Biol. Chem. 271, 12076–12081.

Martinez, E., Ge, H., Tao, Y., Yuan, C.X., Palhan, V., Roeder, R.G., 1998. Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex. Mol. Cell. Biol. 18, 6571–6583.

Mobley, C.M., Sealy, L., 2000. The Rous sarcoma virus long terminal repeat promoter is regulated by TFII-I. J. Virol. 74, 6511–6519.

Montano, M.A., Kripke, K., Novina, C.D., Achacoso, P., Herzenberg, L.A., Roy, A.L., Nolan, G.P., 1996. NFkB homodimer binding within the HIV-1 initiator region and interactions with TFII-I. Proc. Natl. Acad. Sci. USA 93, 12376–12381.

Morikawa, N., Clarke, T.R., Novina, C.D., Watanabbe, K., Haqq, C., Weiss, M., Roy, A.L., Donahoe, P.K., 2000. Human Mullerian-inhibiting substance promoter contains a functional TFII-I-binding initiator. Biol. Reprod. 63, 1075–1083.

Novina, C.D., Roy, A.L., 1996. Core promoter mediated transcription. Trends Genet. 12, 351–355.

Novina, C.D., Cheriyath, V., Roy, A.L., 1998. Regulation of TFII-I activity by phosphorylation. J. Biol. Chem. 273, 33443–33448.

Novina, C.D., Kumar, S., Bajpai, U., Cheriyath, V., Zhang, K., Pillai, S., Wortis, H.H., Roy, A.L., 1999. Regulation of nuclear localization and transcriptional activity of TFII-I by Bruton's tyrosine kinase. Mol. Cell. Biol. 19, 5014–5024.

O'Mahoney, J.V., Guven, K.L., Lin, J., Joya, J.E., Robinson, C.S., Wade, R.P., Hardeman, 1998. Identification of a novel slow-muscle-fiber enhancer binding protein, MusTRD1. Mol. Cell. Biol. 18, 6641–6652.

Osborne, L.R., Campbell, T., Daradich, A., Scherer, S.W., Tsui, L.-C., 1999. Identification of a putative transcription factor gene (WBSCR11) that is commonly deleted in Williams-Beuren syndrome. Genomics 57, 279–284.

Parker, R., Phan, T., Baumeister, P., Roy, B., Cheriyath, V., Roy, A.L., Lee, A.S., 2001. Identification of TFII-I as the endoplasmic reticulum stress

response element binding factor ERSF: its autoregulation by stress and interaction with ATF6. Mol. Cell. Biol. 21, 3220–3233.

Pawson, T., Nash, P., 2000. Protein-protein interactions define specificity in signal transduction. Genes Dev. 14, 1027–1047.

Perez Juardo, L.A., Wang, Y.-K., Peoples, R., Coloma, A., Cruces, J., Franke, U., 1998. A duplicated gene in the breakpoint regions of the 7q11.23 Williams-Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of Btk. Hum. Mol. Genet. 7, 325–334.

Rawlings, D.J., Saffran, D.C., Tsukada, S., Largaespada, D.A., Grimaldi, J.C., Cohen, L., Mohr, R.N., Bazan, J.F., Howard, M., Copeland, N.G., Jenkins, N.A., Witte, O.N., 1993. Mutation of the unique region of Bruton's tyrosine kinase in immunodeficient XID mice. Science 261, 358–361.

Roeder, R.G., 1996. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. 21, 327–335.

Roy, A.L., Meisterernst, M., Pognonec, P., Roeder, R.G., 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature 354, 245–248.

Roy, A.L., Malik, S., Meisterernst, M., Roeder, R.G., 1993a. An alternative pathway for transcription initiation involving TFII-I. Nature 365, 355–359.

Roy, A.L., Carruthers, C., Gutjahr, T., Roeder, R.G., 1993b. Direct role for MYC in transcription initiation mediated by interactions with TFII-I. Nature 365, 359–361.

Roy, A.L., Du, H., Gregor, P.D., Novina, C.D., Martinez, E., Roeder, R.G., 1997. Cloning of an Inr- and E-box binding protein, TFII-I, that interacts physically and functionally with USF1. EMBO J. 16, 7091–7104.

Smale, S.T., 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta 1351, 73–88.

Smale, S.T., Baltimore, D., 1989. The "initiator" as a transcription control element. Cell 57, 103–113.

Tassabehji, M., Carette, M., Wilmot, C., Donnai, D., Read, A.P., Metcalf, K., 1999. A transcription factor involved in skeletal muscle gene expression is deleted in patients with Williams syndrome. Eur. J. Genet. 7, 737–747.

Thomas, J.D., Sideras, P., Smith, C.I.E., Vorechovsky, I., Chapman, V., Paul, W.E., 1993. Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. Science 261, 355–358.

Tsukada, S., Saffran, D.C., Rawlings, D.J., Parolini, O., Allen, R.C., Klisak, I., Sparkes, R.S., Kubagawa, H., Mohandas, T., Quan, S., Belmont, J.W., Cooper, M.D., Conley, M.E., Witte, O.N., 1993. Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. Cell 72, 279–290.

Tussié-Luna, I.M., Bayarsaihan, D., Ruddle, F.H., Roy, A.L., 2001. Repression of TFII-I-dependent transcription by nuclear exclusion. Proc. Natl. Acad. Sci. USA 98, 7789–7794.

Usheva, A., Shenk, T., 1994. TATA-binding protein-independent initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA. Cell 76, 1115–1121.

Verrijzer, C.P., Tjian, R., 1996. TAFs mediate transcriptional activation and promoter selectivity. Trends Biochem. 21, 338–342.

Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C.I., Bentley, D.R., 1993. The gene involved in X-linked agammaglobulinemia is a member of the *src* family of protein-tyrosine kinases. Nature 361, 226–233.

Wang, Y.-K., Perez Juardo, L.A., Francke, U., 1998. A mouse single-copy gene, Gtf2I, the homologue of GFT2I, that is duplicated in the Williams-Beuren syndrome deletion region. Genomics 48, 163–170.

Weis, L., Reinberg, D., 1997. Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. Mol. Cell. Biol. 17, 2973–2984.

Wu, Y., Patterson, C., 1999. The human KDR/flk-1 gene contains a functional initiator element that is bound and transactivated by TFII-I. J. Biol. Chem. 274, 3207–3214.

Yan, X., Zhao, X., Qian, M., Guo, N., Gong, X., Zhu, X., 2000. Characterization and gene structure of a novel retinoblastoma-protein-associated protein similar to the transcription regulator TFII-I. Biochem. J. 345, 749–757.

Yang, W., Desiderio, S., 1997. BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. Proc. Natl. Acad. Sci. USA 94, 604–609.