Center for Genomics and Bioinformatics
Karolinska Institutet, Stockholm, Sweden

A STUDY ON THE TFIID SUBUNIT TAF4

Adrian Brunkhorst

Stockholm 2005

Cover art: Liv Widell

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden

ISBN 91-7140-206-3

Celebrate! This part is over

-An Emotional Fish

ABSTRACT

Transcription is a fundamental process in the regulation of gene expression. It has long been generally believed that certain so called basal factors of the transcriptional machinery are ubiquitously expressed. During recent years however, this view has changed and several examples of cell- and promoter-specific roles for basal transcription factors have been demonstrated. In this thesis, we provide evidence for cell- and context-specific roles for a component of the basal transcription factor complex TFIID - TAF4 - in transcription and neuronal differentiation.

In the first study, we identified the mouse TAF4 gene and investigated its expression during embryonic development and in the adult. The results demonstrated that TAF4 gene expression was enriched in the central nervous system and the expression pattern showed clear spatial and temporal variations. In addition, reporter assays using transient transfections of TAF4 in cell lines indicated that TAF4 could act as a positive regulator of the activity of several neuronal promoters. In the second paper, the presence and function of several alternative isoforms of TAF4 was characterized. The expression levels of the alternative isoforms varied in a cell- and tissue-specific manner. Overexpression in retinoic acid treated cell lines showed that these truncated forms of TAF4 generally had a dominant negative effect on retinoic acid receptor-mediated transcription. The results suggest a novel, additional level of complexity in TAF4 and nuclear receptor-mediated transcription.

In the third study, we investigated a putative role for TAF4 in differentiation of cortical embryonic neural stem cells (eNSC). We found that the expression of TAF4 protein was high in fibroblast growth factor-2 (FGF2)-expanded primary cortical eNSC. Whereas astrocytic differentiation of eNSC had little effect on TAF4 protein levels or subcellular distribution, platelet-derived growth factor (PDGF) mediated neuronal differentiation was associated with a marked decrease in TAF4 protein levels. However, overexpression of TAF4 did not show any significant effect on the PDGF-mediated differentiation of eNSC. We speculated that the lack of effect could be due to insufficient levels of required co-factors. A protein-protein interaction screen using a domain previously showed to be important for nuclear receptor transcriptional regulation (Paper II) identified the signaling protein RanBPM as a putative co-factor of TAF4. We found that high levels of TAF4 and RanBPM protein co-localized in the nuclei of neural progenitors *in vivo* and *in vitro*. Interestingly, co-transfections of TAF4 and RanBPM led to a significant increase in the number of primary neurite processes but no increase in total neurite length, whereas RanBPM and a TAF4 construct lacking the RanBPM interacting domain had no significant effect. These results demonstrate how two factors considered to be relatively general in function can influence temporally specific events such as neuritogenesis during neuronal differentiation.

To further increase the understanding of the mechanisms underlying TAF4 function, we performed a gene profiling study using TAF4-deficient cells to identify putative downstream targets of TAF4. This study revealed changes in expression levels of a number of genes associated with early and late neuronal differentiation that may underlie cell- and context-specific effects of TAF4 during brain development.

LIST OF PUBLICATIONS

I. Metsis M., Brunkhorst A. and Neuman T. (2001) Cell-type-specific expression of the TFIID component $TAF_{II}135$ in the nervous system. *Exp. Cell Res.* 269:214-221.

II. Brunkhorst A., Neuman T., Hall A., Arenas E., Bartfai T., Hermanson O. and Metsis M. (2004) Novel isoforms of the TFIID subunit TAF4 modulate nuclear receptor-mediated transcriptional activity. *Biochem. Biophys. Res. Commun.* 325:574-579.

III. Brunkhorst A., Karlén M., Shi J., Mikolajczyk M., Nelson M.A., Metsis M. and Hermanson O. (2005) A specific role for the TFIID subunit TAF4 and RanBPM in neural stem cell differentiation. *Submitted.*

IV. Brunkhorst A., Mengus G., Hermanson O. and Davidson I. (2005) Identification of genes affected by TAF4 deficiency using cDNA microarrays. *Manuscript.*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CAD	Constitutive activation domain
cAMP	cyclic adenosine 5'-monophosphate
CAT	chloramphenicol acetyl transferase
CBP	CREB binding protein
CCTD	Conserved C-terminal domain
CDK8	Cyclin-dependent Kinase 8
CNS	Central nervous system
CNTF	Ciliary neurotrophic factor
CREB	cAMP-responsive element-binding protein
dbcAMP	dibutyryl cyclic adenosine 5'-monophosphate
DNA	Deoxyribonucleic Acid
DPE	Downstream promoter element
DRPLA	Dentatorubral and pallidoluysian atrophy
FGF2	Fibroblast growth factor 2
GFP	Green fluorescent protein
GTF	General transcription factors
HD	Huntington Disease
HFD	Histone folding domain
Inr	Initiator
KID	Kinase-inducible domain
MAP2	Microtubule associated protein 2
MJD/SCA3	Machado-Joseph Disease/Spinocerebellar ataxia 3
mRNA	messenger-Ribonucleic acid
NF-κB	Nuclear Factor κB
NII	Neuronal intranuclear inclusions
NSC	Neural stem cell
PCAF	p300/CBP-associated factor
PDGF	Platelet-derived growth factor
PIC	Pre-initiation complex
RA	Retinoic acid
RAR	Retinoic Acid Receptor

RBP1	Retinol-binding protein 1
RNA	Ribonucleic acid
RNAPolII	RNA Polymerase II
RPA	RNase protection assay
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
RXR	Retinoid X Receptor
SAGA	SPT-ADA-GCN5 acetylase
TAF	TBP-associated factor
TBP	TATA-binding Protein
TFII	Transcription Factor II
TFTC	TBP-free TAF-containing complex
TGF-β	Transforming growth factor-β
TNFα	Tumor necrosis factor α
TR	Thyroid hormone receptor
TRRAP	Transformation/transcription domain-associated protein
TRF	TBP-related factors
VDR	Vitamin D receptor

A BRIEF INTRODUCTION TO TRANSCRIPTIONAL REGULATION

One of the central mechanisms of biological life is the occurrence of gene activation and regulation. The complex phenomenon of how one gene becomes active in a regulated manner is a major focus of modern molecular science. Central dogma of biology is that DNA serves as a template for reading genetic information through a mechanism called transcription, which produces smaller RNA molecules. The RNA molecules later get "translated" into proteins which are the major functional entities of the cell (Crick 1970).

The dictionary definition of transcription as, "Synthesis of RNA by RNA polymerases using a DNA template" gives a rather dull and tedious explanation of a major mechanism of life, and one of the earliest and biggest milestones of evolution. Research into transcriptional mechanisms and regulation of transcriptional activity provides not only a higher understanding of a basic cellular mechanism, it is also a major key to survey and manipulate cells. The information obtained from research of transcriptional regulation is today used in clinical applications such as cure diseased- or replace diseased tissue or cells. Turning genes on and off at a correct time and position determines cell fate, metabolic state and division.

For all living cells, regulation of gene expression is a fundamental mechanism in development, homeostasis and adaptation to the environment. In particular, regulation of gene expression by extracellular signals is one of the most important mechanisms responsible for the remarkable degree of plasticity exhibited by neurons, as alterations in gene expression underlie many forms of long-term changes in neuronal functioning.

PROKARYOTIC REGULATORY NETWORKS

In prokaryotes the regulation of transcription is highly dependent on σ-factors, proteins that recognize the region of DNA used for transcriptional initiation (Helmann and Chamberlin 1988). These factors constitute of a family of proteins that are subunits of the prokaryotic transcription complexes. The location of DNA interaction and the repertoire of σ-factors present determine the specificity of RNA polymerase binding, thereby allowing different subsets of genes to be expressed. Activators and repressors then modulate the activity of RNA polymerase in response to environmental and metabolic cues (Ninfa, Reitzer et al. 1987). This combination of control mechanisms allows the cell to integrate responses to different stimuli into the specific cell state programs and represents the evolutionary beginning of complex transcriptional regulatory networks.

EUKARYOTIC COMPLEXITY

In eukaryotes, transcriptional control is far more complex, due to the high complexity of both gene regulatory pathways and DNA organisation. The need for more mechanisms and complexity of transcriptional regulation increases with the complexity of the organisms and during evolution. Niches of transcriptional regulation

have evolved on top of each other as the organisms increased through evolution. For appropriate activation of one eukaryotic gene, coordination of mechanisms ranging from DNA availability to the attraction, assembly and phosphorylation of the RNA polymerase are necessary. Transcription of a single eukaryotic gene involves chromatin remodelling by a battery of chromatin remodelling factors in order to establish an environment that is favourable for transcription. The remodelling of chromatin permits sequence specific activators and promoter binding elements to gain access to their DNA binding site. This is followed by the assembly of a host of general transcription factors (TFIIs) that collectively help attract RNA Polymerase II and initiate transcription (Buratowski 1994; Lemon and Tjian 2000).

The multiprotein enzyme complex called RNA polymerase II carries out the transcription of protein encoding genes into mRNA. This process is often divided into three steps: initiation of RNA synthesis, RNA chain elongation, and chain termination. Although biologically significant regulation may occur at any of these steps, it is at the step of transcription initiation that the cell exerts the most significant control over the processes that gate the flow of information out of the genome.

Regulation of transcription initiation is accomplished by (1) positioning RNA polymerase II at the correct start site of the gene transcription and (2) controlling the efficiency of the initiation of RNA synthesis and thus the rate of transcription. Each of these events involves assembly and coordination of numerous proteins and protein complexes as well as chemical modification and enzymatic activity (Gralla 1996).

THE PROMOTER AND ITS INTERACTIONS

The cascade of events that precede the activation of transcription leads to the assembly of the basal transcription machinery at the DNA stretch contiguous of the start site of a gene called the promoter. The promoter consists of DNA sequences that serve as specific binding sites for transcriptional regulatory proteins (Blackwood and Kadonaga 1998). The “core promoter” is defined as the minimum stretch of DNA sufficient to induce transcription (Struhl 1987; Weis and Reinberg 1997; Burke, Willy et al. 1998). The core promoters of Pol II transcribed genes contains three major elements: The TATA box, the Inititation (Inr) and the downstream promoter element (DPE) in higher eukaryotes (Smale, Schmidt et al. 1990). The TATA box is located 25-30 base pairs upstream of the start site whereas the Inr sequence is found surrounding the transcriptional start site and finally, the DPE sequence is positioned about 30 base pairs downstream of the start site.

To date promoters are viewed upon as much more than simple DNA scaffolds for the basal transcription machinery. Rather, core promoter elements are dynamic and vital participants in the regulation of transcription activity. Therefore, analysis of the core promoter contributes fundamental insights into the mechanisms of transcription in eukaryotes. By regulating the assembly at the promoter and its efficiency, expressional regulation of genes is achieved.

THE PREINITIATION COMPLEX

The set of basal or general transcription factors (GTFs) required for specific promoter binding by Pol II *in vitro* includes TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Conaway and Conaway 1997; Hampsey 1998). Despite the name, all GTFs are not required at all genes *in vivo*. As is the case for TFIIA that is not required for assembly of the preinitiation complex (PIC) and transcription initiation at all genes (Aoyagi and Wassarman 2000). GTFs have been biochemically defined as factors required for correct initiation of Pol II transcription *in vivo* on a promoter with a classical TATA box and a strong initiator sequence (Roeder 1996). The PIC can be assembled on eukaryotic class II promoters in a stepwise fashion *in vitro* (Buratowski 1994) The order of addition indicated is that TFIID binds the TATA box followed by TFIIA and TFIIB before the subcomplex of Pol II and the rest of the TFs. However, it seems unlikely that this ordered assembly of TFs at promoters occurs *in vivo* for reasons discussed later.

In silico studies of the GTF have revealed evolutionary patterns in PIC gene structure. TFIIB, -E, F, -H and Pol II are all encoded by single copy genes while TFIID consists of TATA binding protein (TBP) and up to 14 evolutionary conserved TBP associated factors (TAFs) (Albright and Tjian 2000; Tora 2002)

THE ASSEMBLY OF THE TRANSCRIPTIONAL MACHINERY

Numerous interactions of proteins are needed for formation of the PIC. Closer inspection of each of these proteins involved gives a more detailed map of how the complex is composed and discloses the diversity of proteins implicated in the PIC. But the over all picture of the assembly still remains elusive today. To explain the formation of the PIC two different models have been proposed.

The stepwise model of PIC assembly suggests an ordered assembly of the complex based on results from *in vitro* experiments (Buratowski 1994), where stepwise addition of purified basal transcription factors on naked DNA templates was required for accurate transcription initiation and promoter binding. This model is consistent with the observed biochemically defined steps and could satisfy a biological requirement for dynamic regulation. However, as we know the composition of PIC today, it contains more than 80 polypeptides, which would make the assembly prior to initiation an extremely complex process if a precise order of protein recruitment is required. It would appear to be inefficient for regulators to individually recruit and organize such a complex assembly within the time scales necessary at each promoter in the cell. Further support for this notion is the limited concentration of many of the proteins necessary for the process. This makes the process of the stepwise assembly model insufficient for transcriptional initiation under physiological terms as we know it today.

An alternative model for PIC assembly is the pre-assembly or holoenzyme model that envisions a targeted recruitment of pre-assembled Pol II. This model was first proposed when certain preparations of Pol II were observed to co-purify with subsets of the basal transcription machinery along with some co-regulators. The absence of TFIID in these Pol II preparations raised theories of recruitment of TFIID as a holoenzyme (Koleske and Young 1994; Chao, Gadbois et al. 1996; Maldonado, Shiekhattar et al.

1996; Parvin and Young 1998). Consequently, this model minimally requires two targeted steps for PIC formation as the complexes are recruited as holoenzymes. One advantage of this model is the ability for rapid induction of gene expression from a limited number of required initiation steps that obviates the limited cellular concentration of individual transcription factors. However, from an evolutionary perspective it is hard to explain the need of a vast diversity of transcription factors and co-regulators in animal cells. It would be more favourable to employ multiple regulators that act at different stages of transcription, thereby creating a machinery with several possible points of regulation to reach the complexity needed in gene regulation of eukaryotic cells. Further support for this notion considers the fact that Pol II composition differs during initiation and elongation (Sims, Belotserkovskaya et al. 2004). If Pol II recruitment is strictly holoenzyme targeted, the transcription complex has to shed components that then have to be recycled into new holoenzymes or subject to degradation, *de novo* synthesis, and re-assembly. Such a scenario would require the recruitment of new holoenzymes for each initiation event. This process appears to be disadvantageous for elongation of an earlier formed PIC and is inconsistent with the significantly high number of Pol II molecules when compared to other components of the PIC (Kimura, Tao et al. 1999)

Neither model is proven and it is evident that further evidence is needed for a better understanding of the roles of transcriptional complex formation assembly in transcriptional regulation. A relatively recent theory implicates a compartmentalized structure of DNA organisation in the nucleus (Lemon and Tjian 2000; Isogai and Tjian 2003) The compartmentalized theory builds upon different sets of data that implicate DNA organisation where silent genes are arranged near the periphery of the nucleus and active genes in the center (Andrulis, Neiman et al. 1998; Verschure, van Der Kraan et al. 1999). This and several other studies form the backbone of this theory where different cellular events are spatially separated inside the nucleus to overcome the problems of cellular concentration of individual transcription factors.

TRANSCRIPTION FACTOR IID (TFIID)

TFIID and TAFs

Early initiation events in transcription require a directed and specific assembly of the PIC and Pol II at the promoter. This specificity is directed by TFIID, which is the only GTF with intrinsic DNA sequence specificity. The sequence specificity of TFIID is mediated by the affinity of several TFIID subunits for distinct promoter sequences. TFIID consists of the TATA-box-binding protein (TBP) and 13 evolutionarily conserved TBP-associated factors (TAFs) (Hampsey 1998). The TFIID complex is suggested as the promoter recognition domain of the PIC. Because of its early requirement *in vitro*, together with its capacity to direct PIC assembly on both TATA-containing and TATA-less promoters, TFIID was proposed to be a central component of the PIC.

Even though the role played by TAFs in transcription regulation is not yet fully understood, several functions of TFIID in transcription regulation have been proposed including promoter recognition, coactivator function and several enzymatic activities.

The first identified core promoter-protein interaction to be identified was the binding of TFIID core protein TBP that binds to the TATA box (Horikoshi, Wang et al. 1989). The interaction was thought as the first DNA binding of GTFs in transcription initiation. Efficiency of the interaction was later explained by the crystal structure of TBP that reveals a horseshoe shaped protein fitting the minor groove of the TATA box (DeDecker, O'Brien et al. 1996).

In vitro transcription experiments have suggested that the function of TAFs within TFIID complex function are coactivators, engaging in direct and selective interactions with transactivators and/or core promoter sequences (Bell and Tora 1999).

TAF nomenclature

The nomenclature of TAF proteins was firstly based on their electrophoretic mobility in polyacrylamide gels and distinguished by a number that indicated their apparent or predicted molecular weight in kDa. To further distinguish TAFs of different RNA polymerase systems from each other, each TAF was also given a suffix in subscript with the roman number of the particular polymerase the TAF was associated with. A prefix was added when referring to a protein from a certain specie. This was usually the very first letter of that species. However, as more TAFs were identified, later cross species comparisons revealed a strong sequence conservation between the species (Sanders and Weil 2000; Gangloff, Pointud et al. 2001; Gangloff, Sanders et al. 2001; Walker, Rothman et al. 2001) and in the prevailing nomenclature it was then hard to follow one genes' phylogeny, since the same protein had different names in different species. Moreover, the recent availability of genome databases has led to identification of several TAF-paralogous and related genes, which has resulted in even more uncertainty concerning whether a TAF gene encodes a genuine TAF_{II} or a paralog one.

Comparison of the *drosophila*, human and yeast TFIID indicates that there are 13 essential "core" TBP associated factors that are strongly conserved between these eukaryotic species. With these 13 TAFs as a basis, a new nomenclature has been created to overcome the problems mentioned before (Tora 2002) which consists on numbering the TAFs of the human TFIID in size order, starting with the largest protein, the former $TAF_{II}250$. This new nomenclature is shown in table 1 and complies the guidelines of the *Saccaromyces* Genome Database (SGD) and the human HUGO Gene Nomenclature Committees.

In this nomenclature bona fide paralogous TAFs are given the suffix b. The earlier polymerase suffix has been abolished and the earlier prefix to pin point the particular specie should now be two lower case letters of the species' name in latin. Finally, if an identified TAF-like protein is not (yet) proven to be part of any TBP containing complex, the letter L is used as a suffix for "TAF-like". Unfortunately several TAFL proteins have been characterized as bona fide TAFs after creation of this nomenclature. However, the L suffix is used throughout this thesis for convenience.

As for the C. elegans TAFs the original nomenclature suggested by the Blackwell laboratory (Walker, Rothman et al. 2001) is still used but modifications are made for simpler tracking among other species.

Table 1. **New TAF nomenclature including the corresponding known orthologs and paralogs .** For factors that show significant homology to a certain TAF, but have not yet been published or for those for which no biochemical characterization has been performed, accession numbers indicating their existence has been included in brackets.

New name	H. sapiens (hs)	D. melanogaster (dm)	S. cerevisiae (sc)	S pombe (sp)	C.elegans (ce)	
					previous name	new name
TAF1	$TAF_{II}250$	$TAF_{II}230$	Taf134/130	$TAF_{II}111$	*taf-1*	*taf-1*
TAF2	$TAF_{II}150$	$TAF_{II}150$	Taf150/TSM1	(T38673)	*taf-2*	*taf-2*
TAF3	$TAF_{II}140$	$TAF_{II}155$/BIP2	Taf47		(C11G6.1)	*taf-3*
TAF4	$TAF_{II}130/135$	$TAF_{II}110$	Taf48/MPT1	(T50183)	*taf-5*	*taf-4*
TAF4b	$TAF_{II}105$					
TAF5	$TAF_{II}100$	$TAF_{II}80$	Taf90	$TAF_{II}72$	*taf-4*	*taf-5*
TAF5b				TAFII73		
TAF6	$TAF_{II}80$	$TAF_{II}60$	Taf60	(CAA20756)	*taf-3.1*	*taf-6.1*
TAF6L	PAF65α	(AAF52013)			*taf-3.2*	*taf-6.2*
TAF7	$TAF_{II}55$	(AAF54162)	Taf67	$TAF_{II}62$/PTR6	*taf-8.1*	*taf-7.1*
TAF7L	TAF2Q				*taf-8.2*	*taf-7.2*
TAF8	(BAB71460)	Prodos	Taf65	(T40895)	(ZK1320.12)	*taf-8*
TAF9	$TAF_{II}32/31$	TAFII40	Taf17	(S62536)	*taf-10*	*taf-9*
TAF9L	$TAF_{II}31L$					
TAF10	$TAF_{II}30$	$TAF_{II}24$	Taf25	(T39928)	*taf-11*	*taf-10*
TAF10b		$TAF_{II}16$				
TAF11	$TAF_{II}28$	$TAF_{II}30\beta$	Taf40	(CAA93543)	*taf-7.1*	*taf-11.1*
TAF11L					*taf-11.2*	*taf-11.2*
TAF12	$TAF_{II}20/15$	$TAF_{II}30\alpha$	Taf61/68	(T37702)	*taf-9*	*taf-12*
TAF13	$TAF_{II}18$	(AAF53875)	Taf9/FUN81	(CAA19300)	*taf-6*	*taf-13*
TAF14			Taf30			
TAF15	$TAF_{II}68$					

Modified from Tora 2002

TAFs within transcriptional complexes

As previously mentioned, TFIID consists of a TATA-box-binding protein (TBP) and 13 evolutionarily conserved TAFs. Recent studies indicate that sequences surrounding the TATA box may also contribute to specific TFIID-DNA interaction. In line with these results, TAF1 has been identified as a core promoter selectivity factor. A promoter mapping study confers promoter recognition dependency of TAF1 on the region neighbouring the TA rich region (Shen and Green 1997). However the TAF1 interaction might depend at the same time on several factors. TAF1 dependency has been overcome by changes in promoter structure (Wang, Zou et al. 1997) and introduction of specific co-activators (Wassarman, Aoyagi et al. 2000). This suggests

that the TAF1 interaction in the assembly of transcription initiation complexes is dynamic and can be modulated by specific transcription factors and promoter structure thereby showing less dependency on core promoter specificity of TAF1.

TAF2 sequence specific binding has been located to the Inr sequence (Verrijzer, Yokomori et al. 1994). Individually, neither TAF1 nor TAF2 show a clear specificity in DNA sequence affinity. However, a TAF1-TAF2 complex selects sequences that match the Inr consensus. Also, in a trimeric complex with TBP, these two TAFs select Inr sequences at the appropriate distance from the TATA-box (Chalkley and Verrijzer 1999)

Apart from the TATA box and Inr, TAF binding has also been reported on the third type of the core promoter element, the DPE. The interaction was shown by photocrosslinking in *drosophila* where DNA interaction with TAF6 and TAF9 has been confirmed (Burke and Kadonaga 1997).

In vitro studies of TFIID subunits led to the so-called "coactivator hypothesis" were TAFs are postulated to be obligatory targets of activators. In this hypothesis different activator-coactivator complexes combinations selectively regulate transcription (Burley and Roeder 1996). However, it is important to note that isolated TAFs are not relevant entities for transcription. Instead, according to this hypothesis, they should be viewed upon as subunits of a complex with dynamic composition. Thereby the composition of the complex itself is a target for transcriptional regulation by endorsing the pathways required for the activator.

Coordination of numerous protein interactions must succeed to express a given gene in a spatially and temporally regulated manner. As an integral part of the assembly, co-regulators have an active role in transcriptional regulation. Theories of co-regulator involvement have reached new levels of complexity as recent research has discovered a diversity of alternative complex formations. Variation in arrangement of alternative complexes might involve different protein compositions as well as alternative forms of proteins. Table 2 summarizes the different interactions, properties and the alternative complexes formed by the established TAF subunits of the TFIID.

Table 2

TAF subunits of TFIID: properties, interactions and alternative complexes

TAF	Interactions	Other Complex	Other properties
TAF1	TBP		Protein Kinase
	TFIIF		HAT activity
	Cyclin D1		Double bromodomain binds H4 tails
	TAFs		Paralogue protein TAF1L
TAF2	DNA (Inr)	TFTC	
	Activators		
TAF3			
TAF4	TAF12	TFTC	Histone H2A like
TAF4b	TFIIA		Histone fold partner TAF10
	TAF1		Coactivator of Q-rich regions
	TAFIIE/F		Coactivator for viral activators
	SP1		Coactivator for hormone receptors
	CREB		SAGA related protein ADA1
TAF5	TAF4	TFTC	PCAF related protein is TAF5L
		SAGA	Contains WD-40 repeats
TAF6	TAF9	TFTC	Histone H4-like
	TAF12	SAGA	Histone fold partner TAF9
	TAF1	PCAF	PCAF related protein TAF6L
TAF7	Activators	TFTC	Paralogue protein TAF7L
TAF8			Detected in yeast only
TAF9	Activators	TFTC	Histone H3-like
	DNA (DPE)	SAGA	Coactivator for acidic activators
	Spt	PCAF	Paralogue protein TAF9L
		STAGA	
TAF10		TFTC	Coactivator for estrogen receptors
TAF10b		SAGA	
		PCAF	

TAF subunits of TFIID (continued)

TAF	Interactions	Other Complex	Other properties
TAF11	VDR		Histone like
			Histone fold partner TAF13
			Related protein is Spt3
			Paralogue protein TAF11L
TAF12	TBP	TFTC	Histone H2B-like
	TAF4	SAGA	Histone fold partner is TAF
TAF6		PCAF	
TAF11			
TAF12			
TAF13			Histone H4-like
			Histone fold partner is TAF11
TAF14	TFIIF		Found in yeast only
	SWI/SNF		
	NuA		
TAF15			Found in mammals only

Modified from Pugh 2000

FUNCTIONAL DIVERSITY AND DYNAMIC COMPOSITION OF TRANSCRIPTIONAL COMPLEXES

A common theme that has emerged from studies of transcriptional co-regulators is the dynamic composition of these large complexes. Some polypeptides are shared among complexes with altered properties and biochemical functions. Thus the dynamic composition allows similar or even contradicting functions and biochemical activities. Several isoforms and counterparts of SWI/SNF, TFIID and the mediator complex have been discovered. This dynamic subunit composition allows increased plasticity of structure and function of the co-regulator and their mechanisms. For example, it appears that the human and yeast mediator complexes exist as two distinct, but highly related entities (Liu, Ranish et al. 2001; Taatjes, Naar et al. 2002) called CRSP/MED and ARC-L/MED respectively. The smaller CRSP/MED complex, that contains one specific subunit (MED70) interacts with the biggest subunit of Pol II (RBP1) to promote transcription. While ARC-L/MED has 4 specific subunits, MED240,

MED230, CDK8 and Cyclin D, have been implicated as a negative regulator of transcription by inhibiting the interaction of RPB1 (Zaman, Ansari et al. 2001). As for the SWI/SNF complex BAF250 and BRM are SWI/SNF-A specific polypeptides, whilst BAF180 is SWI/SNF-B specific. Like the SWI/SNF constellation the TFTC complex has many polypeptides in common with TFIID, even though TFTC contains specific polypeptides such as TRRAP, GCN5L and ADA3 (Muller and Tora 2004).

Indications of dynamic composition have also been seen in the formation of TFIID. One alternative TFIID complex has been identified which lack certain TAF9 and contain an alternative form of the TFIID subunit TAF6. The identification of an alternative form of TAF6, TAF6δ in a novel TFIID composition called TFIIDβ (Bell, Scheer et al. 2001) may be the first indication of a more complex TFIID dynamic constellation than first thought. TAF6δ has an in frame deletion of 10 amino acids in the α helix 2 of its histone fold domain. The deletion disturbs the interaction site with the TAF6 histone fold domain partner TAF9, thereby preventing TAF9 incorporation into the TFIIDβ. The TFIIDβ complex has also shown highly specific properties. Feeding the discussion of specific control and cell program regulation by certain forms or certain complexes, the dynamic composition of TFIIβ has been shown to be part of the apoptotic program by targeting promoters- and enhance the expression of proapoptotic genes (Bell, Scheer et al. 2001). Earlier proapoptotic activity of both TAF6 and TAF9 have been shown by direct interaction and coactivator activity of p53 (Lu, Lin et al. 1994; Thut, Chen et al. 1995). Although TAF6δ is found at several genes whose expression is maintained or increased, altered promoter recognition may contribute to the gene specific effects of TAF6δ since both TAF6 and TAF9 are implicated to interact with the DPE (Burke and Kadonaga 1997).

Finally, even though TBP appears as the predominant TATA box binding protein, the discovery of TBP-related factors (TRFs) that are closely related to TBP (Berk 2000) could challenge the common view on TBP and TFIID as GTFs. In humans, there is at least one TRF, TRF2 (also known as TLP, TRF, TRP), whereas in *Drosophila*, there are at least two TRFs (TRF1 and TRF2). TRF1 is not only able to bind TATA box motifs as its counterpart TBP (Crowley, Hoey et al. 1993; Hansen, Takada et al. 1997) but is also able to bind to a motif termed the TC box, which is not bound efficiently by TBP (Holmes and Tjian 2000). This fact, along with other data suggests that TRF1 might mediate transcription at another complex that is distinct from TFIID (Holmes and Tjian 2000). On the contrary, TRF2 does not generally appear to bind TATA boxes motifs (Maldonado 1999; Moore, Ozer et al. 1999; Ohbayashi, Kishimoto et al. 1999; Rabenstein, Zhou et al. 1999; Teichmann, Wang et al. 1999), but seems to be required for expression of a specific set of genes (Dantonel, Quintin et al. 2000; Kaltenbach, Horner et al. 2000). It remains to be determined whether TRF2 participates in basal transcription at specific TRF2-binding sites.

TAF CONTAINING COMPLEXES AS SPECIFIC COREGULATORS

Today, little is known about the specific regulatory effects of functional diversity and dynamic compositions of co-regulators. Recent discovery and expressional mapping of several cell specific subunits of TFIID raise theories of specific TFIID complex composition for specific functions.

TAF4b

The first cell specific TFIID subunit detected was TAF4b, which was described to be transiently expressed during B-cell development (Dikstein, Zhou et al. 1996). Rather than being specifically expressed on an mRNA level, the TAF4b protein was shown to be incorporated in significant amounts in the TFIID complexes of B-cells only. The gene specific effects of TAF4b have been proven analysing reporter genes at several promoters after overexpressing TAF4b, were the TAF showed synergistic effects on promoters containing the octamer element found in many B-cell specific genes (Wolstein, Silkov et al. 2000). Furthermore, TAF4b has shown to be important for antiapoptotic gene activation by the NF-κB in response to the tumour necrosis factor TNFα (Yamit-Hezi, Nir et al. 2000). The direct interaction of TAF4b within NF-κB family proteins is selective and strictly dependent on specific dimers of target proteins. Among them RelA have shown importance for proper B-cell maturation (Grossmann, O'Reilly et al. 2000) in an event downstream of invariant chain (Ii) (Shachar and Flavell 1996) TAF4b has been implicated as the cofactor linking Ii induced B-cell maturation to the NF-κB family protein RelA (Matza, Wolstein et al. 2001). This notion has also been proven *in vivo* in transgenic animals (Silkov, Wolstein et al. 2002)

Further discoveries of specific effects of TFIID subunits include the requirement of TAF4b in ovarian development. The finding that female TAF4b -/- mice are viable but infertile, while male knock out mice are unaffected (Freiman, Albright et al. 2001) suggested a defect in female breeding system. Wild type studies of the ovary revealed an expression pattern of TAF4b restricted to the granulosa cells surrounding the developing oocyte. In this study, the knock out animals showed smaller ovaries and a lack of maturing follicles indicating that TAF4b is critical for development and/or proliferation of the granulosa cells of the ovarian follicle required for proper ovarian function and breeding. Microarray analysis of specific genes affected by TAF4b impairment in the ovaries revealed highly specific effects. The twofold down regulating limit was only reached by approximately 1% of the genes assayed. Among those several genes known to be important for female fertility was downregulated 3- 14-fold in the TAF4b deficient ovaries (Freiman, Albright et al. 2001). The most notably effect was a downregulation of multiple components in the inhibin-activin-follistatin pathway. Inhibins and activins are members of the transforming growth factor-β (TGF-β) that regulate synthesis and release of several hormones of the female menstruation cycle. Also the enzyme aromatase p450, responsible for the conversion of androgens to estrogens in granulosa cells, was dramatically decreased. Another affected gene was cyclin D2, indicating a disturbance in cell proliferation. A noteworthy fact is that no significant phenotype of immunological matter was detected. This notion has also been approved by antigen stimuli of the TAF4b-/- mice (Freiman, Albright et al. 2002). Therefore, TAF4b may show redundancy in the immune system with other factors.

TAFs in development

In *drosophila*, TAFs have been shown to by required for activated transcription of certain genes involved in embryo development. TAF4 and TAF6 are evidently reported as limiting factors for Dorsal mediated activation. Altered expression pattern of mesodermal determinating genes downstream of Dorsal are resulting from mutations in TAF4 and TAF6. These impaired TAFs narrows the expression pattern of the Dorsal activated genes Twist and Snail in early embryogenesis (Zhou, Zwicker et al. 1998). *In vitro* studies reveal that more TFIID subunits is required for activation of Dorsal itself as for the downstream target twist. The mediated activation is totally dependent on TBP and TAF1 indicating a signal pathway through TFIID. Dorsal interacts with both TAF4 and TAF6 (Zhou, Zwicker et al. 1998) and can activate transcription from either TBP-TAF1-TAF4 or TBP-TAF1-TAF6 trimeric complexes while twist is totally dependent on TAF4 for *in vitro* activation by trimeric complexes (Pham, Muller et al. 1999)

Specific TAF function in mammalian development is not so well characterized to date. However, characterization in developmental protein localization of TAF10 divulges an uneven protein distribution and the impairment of TAF10 results in disrupted TFIID and embryonic death at E4.5-E5.5 (Mohan, Scheer et al. 2003). In this study it was also concluded that TAF10 was not required for survival of terminally differentiated trophoblast.

Spermatogenesis: The special TAF case

Spermatogenesis has turned out to be the best characterized model of tissue specific expression of TAF paralogous genes. To date, all functionally characterized, tissue-selective paralogs of TFIID components appear to be engaged in gametogenesis, thereby introducing new theories of special transcriptional mechanisms for gametogenesis. The most evident example is the human TAF1 paralog TAF1L. Since this gene lacks introns its evolutionary origin is believed to be a retroposition of a TAF1 mRNA into the genome (Wang and Page 2002). Given that sex chromosomes are silenced during meiosis, and TAF1 is located on the X-chromosome in humans, TAF1L is hypothesized to replace its paralogous protein during human spermatogenesis (Wang and Page 2002). TAF1L shows 94% nuclear identity to TAF1 and further indications of shared characteristics include direct binding to TBP.

Recent work in characterizing expression of TAF and TAF paralogous genes in spermatogenesis in mice is indicating strictly regulated possibilities for formation of the TAF containing complexes in different stages of maturation of the spermatocytic cells in mammals. Protein localization shows that TAF3, TAF4 and TAF10 are strongly expressed in early spermatogenesis but are downregulated as meisosis procedures starts producing haploid cells (Pointud, Mengus et al. 2003). The germ-cell-specific TAF7 paralog TAF7L (Loriot, Boon et al. 2003) is induced later in maturation of sperm cells (Pointud, Mengus et al. 2003). Interestingly, TAF7L protein in germ-cell-maturation is detected in the cytoplasm and its nuclear import later in maturation coincides with a decrease in TAF7 and a strong increase of TBP (Pointud, Mengus et al. 2003). A

confirmed TAF7L-TBP interaction suggests a replacement of TAF7 in the TBP and TAF containing complexes. The purpose of TAF proteins in the cytoplasm remains unclear but the localization implicates theories of dual TAF functions.

In *drosophila* a battery of TAF paralogous proteins important for male germ-cell maturation have been identified. The paralogous TAFs dmTAF4L, dmTAF5L, dmTAF8L and dmTAF12L all show testis specific expression in the fly and impairment of each of the TAF-like proteins shows similar phenotypes of defects in expression of spermatid differentiation genes in primary spermatocytes. This general phenotype of meiotic arrest indicates the mediation of a common spermatocytic gene expression program (White-Cooper, Leroy et al. 2000; Hiller, Lin et al. 2001; Hiller, Chen et al. 2004).

The dmTAF4L lacks the glutamine rich domains characteristic of the *drosophila* and human homolog TAF4. Furthermore, dmTAF7L is like TAF1L encoded by a single exon. The predicted protein has 259 amino acids and contains a H2A-like histone fold domain in the C-terminal (Hiller, Chen et al. 2004) like TAF4. Even though a striking similarity between the histone fold domains of the TAF4L and TAF12L with its bona fide counterparts, the histone fold interaction does not cross interact between the testis specific TAFs and their paralog partner.

TAF4

Human TAF4 is a 1083 amino acid protein of 135 kDa (Mengus, May et al. 1997). The protein has been described to interact through different binding domains with several proteins including: SP1 (Gill, Pascal et al. 1994), CREB (Ferreri, Gill et al. 1994), TAF12 (Gangloff, Sanders et al. 2001) and E1A (Mazzarelli, Atkins et al. 1995). The TAF4 protein has also shown an essential role for ligand- mediated activation of certain nuclear receptors (Mengus, May et al. 1997). However, since no direct interaction has been proven between the nuclear receptor and TAF4, the nature of this interaction has to be further investigated.

The histone fold domain (HFD)

The interaction between TAF4 and TAF12 has been shown to occur through and Histone Fold Domain (HFD) (Gangloff, Werten et al. 2000). HFDs play an important role in formation and structural organization of the TFIID complex (Gangloff, Romier et al. 2001). Sequence comparison has located HFDs in TAF6, TAF9, TAF4 and TAF12. Classification of the different HFDs based on the core histone amino acid sequence has revealed a possible interaction site and has suggested an interacting partner. In the case of TAF4, that contains an H2A-like HFD, it has been proven to interact with the H2B-like HFD of TAF12 (Gangloff, Werten et al. 2000). Similar studies have located an H3-like HFD in the TAF9 protein that interacts with the H4-like HFD in TAF6 and forms a H3-H4-like heterotetramer (Hoffmann, Chiang et al. 1996; Xie, Kokubo et al. 1996). It has been suggested that the TAF6-TAF9 heterotetramer may associate with the with the TAF4-TAF12 heterodimer to form an ocatameric substructure within the TFIID (Hoffmann, Chiang et al. 1996; Gangloff,

Werten et al. 2000). *In vitro* studies in yeast have given support for an assembly of (yTAF6-yTAF9)$_2$-2(yTAF4-TAF12) consistent with that of a histone-like octamer (Selleck, Howley et al. 2001).

The H2A HFD of TAF4 consists of three α helixes of which α1 and α2 are located in the C-terminal part of the protein and are separated by a short loop. The α3 helix is located within the short and highly conserved C-terminal domain (CCTD) (Thuault, Gangloff et al. 2002) and is separated from the other two helices by a linker region that is conserved and essential for TAF4 function *in vivo*.

Gene expression studies have demonstrated that HFD-containing TAFs have a more wide-range effects on transcription, whereas the effects of other TAFs appear to be more restricted (Holstege, Jennings et al. 1998; Lee, Causton et al. 2000; Kirschner, vom Baur et al. 2002).

The Glutamine-rich regions: SP1 and CREB interaction

The central domain of TAF4 contains four glutamine rich regions designated Q1-4, that are essential for interaction with Sp1 and CREB (Tanese, Saluja et al. 1996; Saluja, Vassallo et al. 1998). Those hydrophobic domains are important for the regulated transcriptional activation of these activators and deletions of these glutamine rich regions interfere with the binding of both, CREB and Sp1. However, even though the Q domains and the adjacent sequence seem to be important for both interactions, CREB seems to be more dependent on Q1 and Q2 and the adjacent sequence of these domains (Saluja, Vassallo et al. 1998).

CREB possesses two activation domains, one glutamine rich constitutive activation domain (CAD) and one phosphorylation dependent kinase-inducible domain (KID). Both of these domains have shown to be required for signal-dependent transcriptional activation *in vitro* (Nakajima, Uchida et al. 1997). The CAD associates CREB to TFIID through an interaction with TAF4, while KID associates with, or recruits Pol II via the coactivator CREB binding protein (CBP). This suggests that multiple interactions between an activator and the transcriptional machinery are required for full activity of CREB.

E1A interaction

The extreme C-terminal 105 amino acids of TAF4 can bind the human adenovirus E1A protein through its conserved region 3 (CR3), which has been shown to mediate transactivation. Furthermore, an interesting fact is that the C-terminal region of TAF4 can block transcriptional stimulation from an E1A-inducible promoter *in vivo* with high specificity (Mazzarelli, Mengus et al. 1997)

TAF4 IN POLYQ RELATED DISEASES: DISTURBANCE OF CREB MEDIATED TRANSCRIPTION

The first indication of the relevance of TAF4 in the nervous system was the discovery that TAF4 implication in the pathological processes of different forms of poly-glutamine (polyQ) related diseases (Shimohata, Onodera et al. 2000). These disorders are genetically distinct and characterized by their specific lesion distribution in the nervous system.

The strictly inherited pattern of these neuronal diseases is caused by genetic expansions of CAG trinucleotide repeats in a specific gene. Since the CAG repeat is located in frame with the gene and inside its coding region, the trinucleotide repeat causes multiple glutamine inserts in the translated protein. It is a common belief that the poly-glutamine stretch changes the protein binding ability towards a more adhesive feature. As part of the pathological processes, the sticky polyQ containing protein forms aggregates in the form of neuronal intranuclear inclusions (NII) in the diseased cell. The formation of NIIs is the common hallmark to all polyQ related diseases except spino cerebral ataxia type 6 (SCA6). Thereby the NII is believed to be the cause of the cells pathological state that at the end leads to apoptosis.

The NII are found in the regions of the CNS known to be affected by each disorder, they are ubiquitinated proteins aggregated in a single heterogenous round structure of mixed granular and filamentous feature. The inclusions usually vary in size from 0.7 to 3.7μm and are normally distributed in a random, but occasionally parallel pattern in the nucleus.

The discovery that TAF4 binds directly to poly-glutamine stretches and that the TAF4 protein is part of the nuclear inclusion bodies (Shimohata, Onodera et al. 2000) raised theories of a transcriptional cause for apoptosis. In the same study, aggregates from diseased patients with the polyQ related diseases DRPLA and MJD/SCA3 were found to contain TAF4, TBP, SP1 and CREB gene products. As TAF4 has been shown to directly interact with Sp1 (Gill, Pascal et al. 1994) and CREB (Nakajima, Uchida et al. 1997) and is normally found in complexes with TBP, the presence of TAF4 could be suggested to be a docking protein for the aggregation event. An interesting fact supporting a role for TAF4 in the pathological process is the phenotype rescue observed when TAF4 is transiently overexpressed in cell culture where transiently expressed polyQ proteins induce NII formation (Shimohata, Onodera et al. 2000).

The cause of apoptosis in the diseased cells remains elusive however, one theory is a disturbance in CREB mediated transcriptional regulation. Supporting this theory is the discovery of a significant suppression of CREB-dependent transcriptional activation by expanded polyQ stretches in cell culture (Shimohata, Onodera et al. 2000)

Neuronal relevance of CREB has earlier been reported as a role of CREB in maintaining long term memory. Several studies involving overexpression of dominant negative CREB also suggest a role for CREB as a survival factor in various cellular models (Jean, Harbison et al. 1998; Riccio, Ahn et al. 1999; Somers, DeLoia et al. 1999). Furthermore, CREB phosphorylation has been shown to be essential for NGF induced survival of peripheral neurons (Riccio, Ahn et al. 1999; Somers, DeLoia et al. 1999). However, while no clear phenotype of the CNS neurons has been reported of CREB impairment alone, double knock out of CREB and another CRE binding

transcription factor CREM, shows extensive apoptosis of postmitotic neurons during development (Mantamadiotis, Lemberger et al. 2002). When CREB and CREM disruption is induced in the postnatal forebrain the resulting phenotype shows progressive neurodegenereation in the hippocampus and striatum (Mantamadiotis, Lemberger et al. 2002). The morphological striatal phenotype, as well as the behaviour, of these mice shows reminiscent remarks of Huntington's Disease (HD), another polyQ related disease. These results are consistent with the postulated role of CREB in this type of diseases.

The mechanism of CAG-repeat induced apoptosis still has to be clarified. However, the results presented by Shimohata et al demonstrate that TAF4-polyQ interaction inhibits CREB dependent transcriptional activation. That suggests suppression of the inhibition of CREB dependent transcriptional activation caused by expanded polyQ stretches may offer opportunities for the development of new therapeutic strategies for treatment of CAG-repeat related neuronal diseases.

A BRIEF INTRODUCTION TO NEURAL DIFFERENTIATION

Neural induction occurs after gastrulation when the ectoderm receives signals derived in particular from the mesoderm and the primitive node. These signals include Fibroblast Growth Factors (FGF) and inhibitors of Bone Morphogenetic Proteins (BMP), such as noggin, chordin and follistatin (Edlund and Jessell 1999). After neural induction, neuronal specification proceeds guided by morphogens and externally derived signaling substances like Sonic Hedgehog, retinoids and Wnt molecules. Also various cytokines and neurotrophins are involved, such as platelet derived growth factor (PDGF), neurotrophin 4/5, brain derived neurotrophic factor (BDNF), nerve growth factor etc (Panchision and McKay 2002). At this early stage, the progenitors only differentiate into neurons and not glia, and are therefore referred to as neuronogenic. In these early stages, glial genes are believed to be methylated and not responsive to glial-inducing factors. During mid-gestation, glial genes become de-methylated via unknown mechanisms (Takizawa, Nakashima et al. 2001) and are now repressed by the transcriptional repressor N-CoR (Hermanson, Jepsen et al. 2002). Since N-CoR repression can be relieved by enzymatic modifications, such as phosphorylation, these neural progenitors are responsive to glial-inducing factors such as ciliary neurotrophic factor (CNTF), leukemia-inhibitory factor (LIF), neuregulins and Notch signaling (Ge, Martinowich et al. 2002; Hermanson, Jepsen et al. 2002). Common for the early neuronogenic and the later multipotent progenitors is that they express the intermediary filament nestin (Lendahl, Zimmerman et al. 1990) but not markers for differentiated neurons or glia. In contrast, during late gestation multipotent progenitors start expressing glial fibrillary acidic protein (GFAP) and are referred to as radial glia or, more correctly, radial cells (Gates, Thomas et al. 1995). The responsiveness and molecular characteristics of radial cells are still very much unknown.

Interestingly, many of the extracellular signaling factors and morphogens exert similar effects on neural progenitors *in vitro*. For example, basic FGF (FGF2), that is instrumental in neural induction, can be used to keep neural progenitors in a mitotic and multipotent state. These neural progenitors that can be differentiated into neurons, astrocytes, oligodendrocytes and in some cases also additional cell fates, are many times referred to as embryonic neural stem cells (eNSC). In line with this reasoning, neuronal differentiation of eNSC can be initiated by factors such as BDNF, retinoic acid and PDGF (Panchision and McKay 2002) whereas astrocyte differentiation can be initiated by CNTF and LIF (Hermanson, Jepsen et al. 2002) and reportedly by soluble ligands of the Notch receptor (Morrison, Perez et al. 2000)

Whereas there a large number of studies on extracellular factors and neural differentiation, much less is known when it comes to the transcriptional mechanisms regulating neural cell fate specification. The most important insight have come from studies on early neuronal specification in the spinal cord where it has been demonstrated that different concentrations of morphogens such as Sonic Hedgehog activate different subsets of transcription factors characteristic for the terminal cell fate (Briscoe and Ericson 2001) Transcription factors regulating this cell fate specification include homedomain factors but also neurogenic basic helix-loop-helix (bHLH) factors

are instrumental for neuronal differentiation. On the other hand, repression of neuronal differentiation in early, neuronogenic progenitors is exerted by Sox proteins and the zinc finger factor REST/NRSF (Palm, Belluardo et al. 1998; Bylund, Andersson et al. 2003). Less is known regarding transcriptional co-regulators. Deletion of the acetyl transferases CBP/p300 is sufficient to cause neural induction in C. Elegans and Xenopus, possibly because CBP/p300 are factors required for proper BMP and Notch signaling (Shi and Mello 1998) In neural progenitors the case is the opposite and CBP/p300 are required for neuronal as well as glial differentiation. In line with these reports, brains of mice harbouring targeted gene deletion of p300 have a dramatic increase in the number of undifferentiated neuroblasts (Yao, Oh et al. 1998). This is in sharp contrast to reports from studies on deacetylase- and other repressor-deficient mice, where depletions of the progenitor pools have been reported (Hermanson, Jepsen et al. 2002).

The identification of both neurogenic factors and repressors of neurogenesis has been utterly dependent on the use of neural markers. The integratory mechanisms of extracellular signaling and transcription factors regulating for example axonal outgrowth are much less clear, especially in later embryonic development. *In vitro* studies have implicated roles for CREB and retinoic acid receptor (RAR) in neurite outgrowth, but most if not all of these studies have been performed on neural or neural-like cell lines, such as PC12 cells. There are furthermore no studies on potential roles for transcriptional co-regulators in neurite or axonal outgrowth or initiation, in spite of the clinical relevance for increased understanding of such mechanisms.

AIMS OF THIS STUDY

The over all aim with the thesis is to study the relevance of TAF4 in neuronal development and the role of TAF4 in transcriptional regulation.

Specific aims of this thesis are:

- To investigate TAF4 expression pattern during mouse embryogenesis and in the adult mouse.
  (Paper I)

- To study the effects of TAF4 on neuronal gene expression.
  (Paper I)

- To characterize TAF4 splice variants and study their effects on transcription.
  (Paper II)

- To identify and study the relevance of TAF4 interaction partners.
  (Paper III)

- To investigate the role of TAF4 and its interaction partners in neural stem cell differentiation.
  (Paper III)

- To identify potential downstream targets of TAF4.
  (Paper IV)

RESULTS AND DISCUSSION

Paper I

TAF4 has previously been identified as a member of the TFIID transcriptional complex (Tanese, Saluja et al. 1996). In this paper, after sequencing the C-terminus of the mouse TAF4 gene, we have investigated TAF4 expression during development. We have also studied the participation of TAF4 in differentiation events in PCC7 cells, where we have detected a marked induction of TAF4 in cell culture after the differentiating treatment with RA and dbcAMP.

The complete human cDNA sequence of the TAF4 gene was accomplished in 1997 (Mengus, May et al. 1997). We designed degenerate primers based upon the human TAF4 sequence for identification of the mouse homolog. After RT-PCR the resulting fragment was cloned and used to screen a mouse cDNA library. The obtained cDNA clone contained 2196 base pairs of the C-terminal part of mouse TAF4. In comparison with the human TAF4 protein, our clone lacks 420 amino acids but cover regions implicated for interaction with the N-terminal activation domain of CREB and the transcription factor Sp1 (Tanese, Saluja et al. 1996; Saluja, Vassallo et al. 1998). Other domains implicated in protein interaction covered by our clone include the histone fold domain for TAF12 interaction (Gangloff, Werten et al. 2000), the domain shown to interact with the adenoviral protein E1A (Mazzarelli, Mengus et al. 1997) and the domain shown to be relevant for AF-2 signal transmission of certain nuclear receptors (Mengus, May et al. 1997).

A predicted full length sequence of mouse TAF4 containing 3396 base pairs similar to the clone described by us is at writing point available in public databases (NCBI ref XM_130764.4). When compared to the human TAF4 sequence the predicted full length sequence lacks 1140 base pairs in the N-terminal part. The fact that no protein size difference between human and mouse proteins are detected strongly argues against this sequence to be the truly full length sequence of mouse TAF4. A probable reason for the shorter predicted mouse sequence can possibly be a mistake in position of the proposed start-codon.

The high homology shown when compared to the human TAF4 is well in line with the theories of evolutionary conservation of TAF proteins (Gangloff, Werten et al. 2000; Sanders and Weil 2000; Walker, Rothman et al. 2001).

In Paper I, we demonstrate that mRNA levels of TAF4 are increased in the developing central nervous system by northern blot analysis of whole brain RNA extracts. While the embryonic expression of TAF4 mRNA is relatively low at E13 and E15, the expression then increases at E17. Our results also show that postnatal mRNA levels are relatively low at early stages (P4) but that there is a later increase in the adult. It has been suggested that a cause for the decrease at P4 could be explained by the increasing numbers of glial cells postnatally that could end up being a larger part of the RNA pool, this explanation agrees with our results using *in situ* hybridization that revealed low expression of TAF4 in glial cells.

At E13, the developing nervous system showed a relatively intensive and uniform labelling in the developing brain and spinal cord. Postnataly at P1 the expression

pattern becomes more distinct where the highest levels of TAF4 mRNA are detected in the cerebellum, cortex, hippocampal formation, striatum, and ventral nuclei of the thalamus. Outside the nervous system high TAF4 levels are also detected in the thymus.

The adult mouse brain showed uneven labelling in which TAF4 mRNA expression was detected in many regions. However, stronger labelling was observed in the cortex, hippocampal formation, and cerebellum with significant labeling was also detected in the thalamus.

The non-ubiquitous distribution of TAF4 mRNA and protein evoked theories of a specific role of TAF during CNS development. To test the hypothesis of whether TAF4 has a direct role in activation of developmental neural genes, we compared the activity of several promoters in differentiated cells versus undifferentiated PCC7 cells overexpressing TAF4 so that we could evaluate the transcriptional significance of increased TAF4 expression during neuronal differentiation. The activity of NF-L, GAP-43, BDNF-IV, ME1, and COUP-TFII promoters was analyzed during induction of neuronal differentiation by retinoic acid (RA) and bdcAMP, and resulted in a significant stimulation (5 to 50 times) of all promoters. An interesting result is that a similar stimulation of promoter activity was obtained by overexpression of TAF4. Since exogenous expression of TAF4 more or less mimicked the promoter activation seen in stimulated cells, these data would support our hypothesis that elevated levels of TAF4 in differentiating PCC7 cells could cause stimulation of promoter activity at similar levels seen in RA- and dbcAMP-treated cells.

Recent data demonstrates that certain TAFs are engaged in tissue specific transcriptional regulation (Perletti, Dantonel et al. 1999; Freiman, Albright et al. 2001; Hiller, Lin et al. 2001; Mohan, Scheer et al. 2003; Hiller, Chen et al. 2004) or have tissue specific expression (Freiman, Albright et al. 2001; Mohan, Scheer et al. 2003). Our results indicate TAF4 is also tissue specific. Earlier studies of specific TAF functions have shown direct and selective interactions between TAFs and transcriptional activator proteins (Goodrich, Hoey et al. 1993; Hoey, Weinzierl et al. 1993; Chen, Attardi et al. 1994; Gill, Pascal et al. 1994). In addition some nuclear receptors have also been shown to be TAF dependent (Petty, Krimkevich et al. 1996; Mengus, May et al. 1997). Even though the function of TAF4 in development still remains elusive, in this study we have shown a tissue specific expression of TAF4 in development, where transcriptional activators and many of the nuclear receptors are relevant for proper development and response to extracellular signals.

Paper II

Analysis of endogenous splice variants of TAF proteins has earlier shown these to have a major impact in activation of cell programs (Bell, Scheer et al. 2001) in paper II we have found and characterized five novel alternative isoforms of mouse TAF4 and analyzed the effects on transcription of these truncated forms.

In an RNAse protection assay (RPA) we investigated mouse TAF4 gene expression in various tissues and at various developmental stages. The detection of several shorter transcripts in the developing brain indicated presence of alternatively spliced TAF4 mRNA. Immunoblotting experiments using undifferentiated and

differentiated neuroectodermal cells further revealed the presence of several mTAF4-immunoreactive proteins of various sizes, suggesting the endogenous expression of mTAF4 protein isoforms. We next characterized the structure of the five mTAF4 truncated forms. Comparison of mTAF4 cDNAs and genomic structure revealed that two of the clones resulted from single exon skipping, one cDNA excluded multiple exons and the remaining two were the outcome of alternative donor site usage.

After identification and analysis of transcriptional regulation we observed that all deleted forms had a negative effect on RA induced transcription to some extent. This result that agrees with the earlier described experimental truncations of TAF4 that have been shown to exert negative effects on TAF4-mediated increase of transcriptional activity (Mengus, May et al. 1997). We therefore demonstrated the existence of TAF4 truncated forms *in vivo* and confirmed and expanded previously described transcriptional effects.

An uneven distribution of the truncations was detected in the RPA. Since earlier TAF truncations identified have shown specific effects in apoptotic pathways (Bell, Scheer et al. 2001), this uneven distribution may raise speculations about tissue specific pathways being affected by different TAF4 truncated proteins.

TAF4 has previously been linked to CREB activity (Felinski and Quinn 1999; Bell, Scheer et al. 2001). In accordance with this, transfection of the mTAF4 isoforms modulated cAMP induced CREB-mediated luciferase activity both positively and negatively. Furthermore, co-transfections with full length mTAF4 revealed that some isoforms also exerted a dominant negative function in cAMP-induced CREB activity, a result that could be explained by dimerization of different isoforms that could therefore lead to differences in the activity of mTAF4.

All these observations strengthen the hypothesis that the mTAF4 isoforms have context-specific roles in transcriptional control that may play an important role in regulating cell type-specific differences in endogenous levels of TAF4.

Paper III

In this paper, we demonstrate a downregulation of the levels of TAF4 protein under neuronal differentiation conditions and postulate that TAF4 interacts with the intracellular protein RanBPM through its C-terminal domain. Furthermore, we show that this interaction is involved in the control of neuritogenesis by affecting initiation of primary neurite branching rather than the neurite elongation processes.

To analyze the role of TAF4 in neuronal differentiation we first detected the levels of expression in embryonic neural stem cells (NSC) under several differentiating conditions. High amounts of TAF4 protein were detected in the nuclei of neural stem cells in culture. However, while protein levels remained unaffected by CNTF-induced astrocytical differentiation, a marked decrease was detected under neuronal differentiation conditions induced by PDGF treatment. These results seem to disagree with the results in paper I that indicated an upregulation of TAF4 protein levels upon neuronal differentiation. We believe that the use of different cell types (PCC7 cells in paper I and NSC in paper III) and different means of neuronal induction (RA and bdcAMP in paper I and PDGF in paper III) may be the cause of these contradictory results.

As overexpression of TAF4 had no effect on NSC state, we speculated in the possibility of a protein-protein interaction that could participate in the regulation of the differentiation processes. To identify interaction partners of TAF4 we performed a yeast two-hybrid screening using the part of the protein that has been shown to be important for TAF4 modulated nuclear receptor activity (Mengus, May et al. 1997). Using this system we were able to identify the intracellular signaling protein RanBPM as a putative co-factor of TAF4.

Even though the function of RanBPM is not well understood, it has been demonstrated that RanBPM can act as a transcriptional co-activator for nuclear receptors (Rao, Cheng et al. 2002). This result is in line with TAF4 activity and further suggests a possible functional interaction between these two proteins.

Double immunohistochemical experiments on mouse embryos at age E11.5 showed colocalization of TAF4 and RanBPM in the nuclei of neural progenitors in the proliferative ventricular zone. Due to the different expression pattern of RanBPM and TAF4, the colocalization was, however, restricted and most pronounced in mitotic cells negative for the neuronal marker TuJ1, which indicated that these two proteins could influence the differentiation of neural progenitors.

Furthermore, and in line with the results obtained previously for TAF4 expression, RanBPM protein was shown to be expressed in the NSC nuclei and its levels decreased under PDGF treatment. Chromatin precipitation experiments also revealed that endogenous TAF4 and RanBPM were both present in chromatin preparations of NSC suggesting that both proteins are chromatin associated at the same time in undifferentiated NSC.

Co-transfection of TAF4 and RanBPM resulted in a major increase in the number of NSC that displayed a higher number of primary neurite processes without change in total neurite length. The fact that a TAF4 deletion construct lacking the RanBPM-interacting domain had no effect demonstrated the relevance of this interaction in the regulation of the neurite branching processes.

Our findings provide an example of functional specificity achieved through controlled spatial and temporal expression of factors considered to be involved in general mechanisms of signaling and transcription. Since the ß-galactosidase reaction after TAF4 bait and RanBPM co-expression in yeast was slow we anticipated the interaction to be weak. Even though both proteins were detected in chromatin precipitate, negative results from co-immunoprecipitation experiments strengthened the observation of a weak interaction. However, since many protein-protein interactions are regulated and/or highly dynamic (Nooren and Thornton 2003) these results are hard to interpret.

The descriptive observation in combination with the overexpression experiments suggest that the TAF4-RanBPM interaction influences neuritogenesis and neurite process number in early neuronal differentiation, but is subsequently down-regulated to allow proper neurite elongation or axonal outgrowth. These results would then suggest that the temporal expression levels of two factors considered to be relatively general in function can influence very specific events of neuronal differentiation.

Several possible models could underlie the observed results, one of which could take into account the participation of the DNA-binding transcription factor CREB. CREB has been shown to regulate neurite branching and elongation downstream of nerve growth factor (NGF) in PC12 cells (Nooren and Thornton 2003), and its activity

has been shown to be influenced by phosphorylation in the kinase-inducible domain (KID) event that appears to be crucial for recruitment of the co-activators and acetyl transferases CBP/p300. However, TAF4 has been shown to bind specifically to the phosphorylation-independent activator part of CREB, the constitutive activation domain (CAD) (Nakajima, Uchida et al. 1997; Felinski, Kim et al. 2001; Felinski and Quinn 2001). RanBPM has further been shown to interact with and inhibit the activity of members of the minibrain kinase/DYRK family (Wang, Marion Schneider et al. 2002; Zou, Lim et al. 2003) which have been shown to phosphorylate and activate CREB during differentiation of hippocampal progenitors (Zou, Lim et al. 2003). Integrated with our data, these findings suggest that an overexpression of TAF4 and RanBPM would skew CREB activity so that the phosphorylation-independent activity would become dominant. Therefore, the effect of TAF4 and RanBPM could be the resulting effect of a combination of enhancing TAF4 and suppressing CBP/p300-mediated CREB activity on specific promoters resulting in a suppression of neuronal maturation and aberrant numbers of primary processes.

Paper IV

The earlier suggested tissue specific expression of TAF4 (Paper I) and reports of TAF specificity in regulatory mechanisms (Mengus, May et al. 1997; Freiman, Albright et al. 2001; Shen, Bhaumik et al. 2003) propose gene specific effects of TAF4. To identify potential target genes affected by TAF4 regulation, in this last manuscript we performed microarray hybridizations of fibroblastic TAF4 -/- cell lines to identify potential TAF4 target genes.

After comparing the knock out cell lines with the corresponding heterozygotes, our results indicated about 70 genes of more than 15000 analysed as being downregulated by two-fold or more. Considering that TAF4 is implicated as being a member of the general transcription machinery (Tanese, Saluja et al. 1996), the number of genes affected was surprisingly low. The fact that higher eukaryotes have evolved a complex transcriptional regulatory network where redundancy of gene regulation could be an explanation for the moderate number of genes affected. In line with our results, TAF4b knock out mice shows only a twofold downregulation of 1% of the genes analysed by microarray hybridisation (Freiman, Albright et al. 2001). Furthermore, the strong phenotype of defected ovaries and follicle development in the TAF4b impaired mouse and several downregulated genes in specific pathways (Freiman, Albright et al. 2001) indicated a tissue specific role for TAF4b.

In the case of TAF4 we found five genes previously described to participate in neural development or cell cycle regulation: GATA-3, elfin/Clim1, aurora related kinase 2 (ARK2), caskin 2, and Synaptogyrin 1 (SYNGR1). According to this specific gene regulation together with our expressional analysis of TAF4 in embryogenesis (Paper I) and during neuronal induction (Papers I and III), we would like to propose TAF4 as a member of several of complexes that target the expression of specific genes in certain temporal stages of cell cycle progression and neuronal differentiation.

In order to characterise precise roles of individual TAFs in gene regulation, several genome-wide approaches have been performed in yeast (Lee, Causton et al. 2000; Shen, Bhaumik et al. 2003; Huisinga and Pugh 2004). The percentage of yeast genome

affected by a single TAF varies between 3% for TAF2, and 61% for TAF9 (Shen, Bhaumik et al. 2003). The TAF4 dependency in yeast was estimated to 11%, a result which may cause speculations of a more general role for TAF4 in lower eukaryotes. However, it is worth noting that stress-induced gene expression has been shown to be preferentially regulated by the SAGA complex rather than by the TFIID complex (Huisinga and Pugh 2004). This could indicate more distinct roles and less redundancy of the TFIID complex in lower eukaryotes.

CONCLUDING REMARKS

Several proteins of the TFIID complex are shown to have specific roles to certain extent. In this thesis we provide evidence for cell- and context specific roles of the TFIID subunit TAF4:

- TAF4 show an progressively more specific expression pattern in the developing CNS with a uniform expression of mRNA the brain and spinal cord at E13. At P1, the expression pattern becomes more distinct and high levels of TAF4 mRNA are detected in the cerebellum, cortex, hippocampal formation, striatum, and ventral nuclei of the thalamus. In the adult high levels are detected in the cortex, hippocampal formation, cerebellum and thalamus.

- RA and cAMP induced differentiation in PCC7 cells result in increased levels of TAF4 mRNA while other TAF levels remain unaffected.

- TAF4 overexpression in PCC7 cells increases the activity from neuronal promoters to similar levels as RA-cAMP induced differentiation.

- Five splice variants of TAF4 have been found and characterized in the mouse.

- Most of the spliced variants show dominant negative effects on RA induced transcription by TAF4.

- TAF4 protein levels are decreased in NSC after PDGF induced neuronal differentiation.

- Using a yeast two-hybrid system, an interaction between TAF4 and RanBPM has been detected. These two proteins co-localize in the ventral marginal zone of the developing CNS and in NSC in vitro.

- In NSC RanBPM levels mimic the TAF4 expression after PDGF induced differentiation.

- Co-expression of TAF4 and RanBPM increases neuritogenesis in NSC. This effect is not seen in transient overexpressions of each protein alone. When the RanBPM interacting domain is deleted no neurogenic effects are observed in the cotransfected NSC.

- Microarray hybridization of TAF4 knock out cells demonstrated a low number of TAF dependent genes. Neuronal and developmental genes downregulated by TAF4 impairment are: GATA-3, elfin/Clim1, aurora related kinase 2, caskin 2, and Synaptogyrin 1.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to everyone that has helped and supported me throughout those years. Even though my intention is to include everyone, I apologize to the ones that deserve to be there but their names are not listed.

I would like to thank my supervisor **Madis Metsis** for introducing me to the field of transcriptional regulation and training of the scientific methodology therein, but also for the life experiences achieved during these years.

I would like to express my greatest gratitude to my co-supervisor **Ola Hermanson** and his lab at CMB for extremely helpful hands and mind during the last year of my thesis: **Joshua Duckworth** for his mind blowing English outspoken and straightforward mind-set. It might get you into trouble some day! **Ana Teixeira** for being almost too honest and kind but still outspoken and not pretending. **Erik Södersten** for scientific and non scientific discussions about life. **Karolina Wallenborg** for a great attitude and kindly sharing biological material.

My former prefects **Karl Tryggvason** and **Claes Wahlstedt**.

My current prefect **Christer Höög** for great problem solving skills and his group for giving support during the last year. **Geert Hamer** and **Lena Ström** for a good but though time in the same office with nice chats. **Daniel Lightfoot** for understanding discussions of cultural differencies. **Emily Hodges**, **Ivana Novak**, **Camilla Sjögren**, **Katarina Gell**, **Anna Kouznetsova**, **Jian Guo Liu**, **Hanna Betts Lindroos, Ensaf Mahdy, Jenny Redeliuz**, **Hong Wang** and **Wellin Wu**.

I would like to thank **Patrik Ernfors** and **Ernest Arenas** for the time at MBB and for good friendship, and their current and former lab members for collaboration and good time through those years: **Pontus Holm** for his positive aura and faithfully standing with AIK in good and bad times at "norra stå", **Anna Hjelmqvist** for always smiling, **Karin Agerman** for showing that some attitude in science takes you further than not questioning. **Åsa Edman** for always laughing and bringing positive feelings into the work place whether it is needed or not. **Jens Hjerling-Leffler** for an honest attitude and appearance of being just him and no one else. **Sten Linnarsson** for the philosophy in science that the usual way might not be the only fun way to go**.**
Anita Hall for being helpful and honest. **Anna Stenquist** for a straight forward mind-set and open attitude, **Richard Warfinge** for interesting discussions**, Andreas Montelius**, **Miwaku Kobayashi, Nina Rawal, Gonçalo Castelo-Branco**, **Kyle Souza, Peter Åkerud**,. **Joseph Wagner, Xavier Rodrigues**, **Beth-Anne Sieber**.

Lena Klevenwall-Fridwall, Annika Ahlsén, Lotta Skoog and **Lena Amaloo** for helping out with billions of those small things that make the work flow in an easier way.

I also would like to thank **Carlos Ibanez** and his group for help and support especially concerning advice in protein methodology. **Andries Blokzijl** for discussions and positive encouragements. **Olov Andersson** for supporting AIK and a cool attitude towards the world. **Henrik Jörnvall** for discussions and interesting experiences in all imaginable and unimaginable places during all hours considerable. **Rizaldy Scott**, **Peter Lönneberg**, **Thomas Ringstedt**, **Miles Trupp** and **Susanne Eketjäll.**

For the time in Strasbourg I would like to show my greatest gratitude towards **Irwin Davidson** for letting me work at IGBMC, giving me and his excellent group: **Gabrielle Mengus** for keeping track of everything, **Jean-Christophe Pointud** for great discussions and attitude, **Phillipe Kieffer-Kwon**, **Igor Martianov, Sylvie Thuault**, **Isabelle Michel** and **Virginie Humbert**.

A Special thanks for my friends in France that stood by me through rough times and gave support**: Aurora Pujol**, **Fabrice** and **Diogo Castro**.

I would also like to thank **Tamas Barfai** and Barfai's lab lab members at The Scripps Reaserch Insttitute for inviting me there and having a wonderful time in San Diego. **Bruno Conti** for scientific and football (YES FOOTBALL!) discussions**. Marga Behrens** for trust and I hope Felix is still alive! **Sebastian Wirz** for partying and cultural discussions**, Amy Guzik, Jessica Cramp** for support in bringing surfing attitude into the lab, **Janell Pemberton & Jennifer Neumann** for help in the administrative world of American science, orientation in the lab, in the institute and in the city and **Janell Laca** for being a good and listening friend.

Other friends at Karolinska Institutet:
Kristian Tryggvason, Malin Sandberg, Peronella Kettunen, Niklas Lindgren, Malin Höistad, Saga Johansson, Haleh Razani, Kairi Tammoja, Katarina Luhr, Johan Ericsson and lab crew: **Mattias Karlén, Elisabet Andersson, Ulrika Marklund, Joanna Klos, Madelene Persson, José Dias, Anna Vallstedt** for showing the way in the end. **Jonas Frisén** and lab crew: **Marie Carlén, Anna Falk, Niklas Holmström, Karin Hamrin, Johan Holmberg, Dinos Meletis, Kirsty Spalding, Joel Zupicich** and finally **Daniel Marcellino, Alexandra Trifunovski, Mathew Abrams, Andrea Carmine, Emil Hansson** and **Magnus Sandberg.**

A big thanks to my sister **Charlotta Brunkhorst** for a bigger support then she has understood.

My beloved San Diegan friends: **Steven Hom** and **Rachel & Michael Kurisu**: Someday we will be back in the line up just like in the old times.

Sebastian Troeng and **Mattias Hallsten**: The brotherhood never dies.

A special thanks goes to **Meritxell Canals** for a great support over the phone.

La Familia in San Diego: **Linda Nyberg** for a great time as room mates**. Gustaf Angelborg** and **Toste Tanua** for great spirit in the wonderful "morgonsurf" tradition, **Frances Liberberg** for adding more Mexican attitude and flavour to California, **Ylva & Lucas Elias** for the trust of letting me be the best man, **Tomas Lövgren** for his positive and helping attitude, **Maria Langdown & Darren Thomson** for their true English approach of life, **Helena Blomqvist, Anne Liljenstrand & Eric Mjoen**, **Johanna Nilsson, David Slifka**, **Tobias Nyberg**, **Brendan Brinkman**, **Julia Clemens**, **Rebecca Cho, Jonathan Maynard** and finally the family **Wilhelmsson** including **Mats**, **Helene**, **Celia**, **Leo** and **David** for great support when needed and a place to stay in Stockholm when I needed it the most.

N3B & co for inspiration lasting a lifetime, a mind blowing three years and good friendships thereafter. Am I making the PhD number in one figure?

A big thanks to the K-na friends that are still around: **Fredrik Bergljung** and **Kristina Burman** for sharing good and bad times, plus a great support of accommodation in Stockholm. I hope I can give it back to you some day**. Jörgen Sjögren** for the spark of life that you give to others.

I would also like to thank:

AIK for giving courage and excitement in life wherever I have been and **Jesper Florén** at www.exilgnagare.com for transmitting those feelings globally and an excellent job on the web.

Club Killers for their feeling and the being the best live entertainment in Stockholm

91X for their wave reports and making traffic to work and "morgonsurf" more enjoyable.

Macintosh for understanding that operative system should be UNIX based and software should be international. Without that my computer analysis in Strasbourg would have been a lot more complicated.

REFERENCES

Albright, S. R. and R. Tjian (2000). "TAFs revisited: more data reveal new twists and confirm old ideas." Gene **242**(1-2): 1-13.

Andrulis, E. D., A. M. Neiman, et al. (1998). "Perinuclear localization of chromatin facilitates transcriptional silencing." Nature **394**(6693): 592-5.

Aoyagi, N. and D. A. Wassarman (2000). "Genes encoding Drosophila melanogaster RNA polymerase II general transcription factors: diversity in TFIIA and TFIID components contributes to gene-specific transcriptional regulation." J Cell Biol **150**(2): F45-50.

Bell, B., E. Scheer, et al. (2001). "Identification of hTAF(II)80 delta links apoptotic signaling pathways to transcription factor TFIID function." Mol Cell **8**(3): 591-600.

Bell, B. and L. Tora (1999). "Regulation of gene expression by multiple forms of TFIID and other novel TAFII-containing complexes." Exp Cell Res **246**(1): 11-9.

Berk, A. J. (2000). "TBP-like factors come into focus." Cell **103**(1): 5-8.

Blackwood, E. M. and J. T. Kadonaga (1998). "Going the distance: a current view of enhancer action." Science **281**(5373): 61-3.

Briscoe, J. and J. Ericson (2001). "Specification of neuronal fates in the ventral neural tube." Curr Opin Neurobiol **11**(1): 43-9.

Buratowski, S. (1994). "The basics of basal transcription by RNA polymerase II." Cell **77**(1): 1-3.

Burke, T. W. and J. T. Kadonaga (1997). "The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila." Genes Dev **11**(22): 3020-31.

Burke, T. W., P. J. Willy, et al. (1998). "The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box." Cold Spring Harb Symp Quant Biol **63**: 75-82.

Burley, S. K. and R. G. Roeder (1996). "Biochemistry and structural biology of transcription factor IID (TFIID)." Annu Rev Biochem **65**: 769-99.

Bylund, M., E. Andersson, et al. (2003). "Vertebrate neurogenesis is counteracted by Sox1-3 activity." Nat Neurosci **6**(11): 1162-8.

Chalkley, G. E. and C. P. Verrijzer (1999). "DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator." Embo J **18**(17): 4835-45.

Chao, D. M., E. L. Gadbois, et al. (1996). "A mammalian SRB protein associated with an RNA polymerase II holoenzyme." Nature **380**(6569): 82-5.

Chen, J. L., L. D. Attardi, et al. (1994). "Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators." Cell **79**(1): 93-105.

Conaway, R. C. and J. W. Conaway (1997). "General transcription factors for RNA polymerase II." Prog Nucleic Acid Res Mol Biol **56**: 327-46.

Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-3.

Crowley, T. E., T. Hoey, et al. (1993). "A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila." Nature **361**(6412): 557-61.

Dantonel, J. C., S. Quintin, et al. (2000). "TBP-like factor is required for embryonic RNA polymerase II transcription in C. elegans." Mol Cell **6**(3): 715-22.

DeDecker, B. S., R. O'Brien, et al. (1996). "The crystal structure of a hyperthermophilic archaeal TATA-box binding protein." J Mol Biol **264**(5): 1072-84.

Dikstein, R., S. Zhou, et al. (1996). "Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130." Cell **87**(1): 137-46.

Edlund, T. and T. M. Jessell (1999). "Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system." Cell **96**(2): 211-24.

Felinski, E. A., J. Kim, et al. (2001). "Recruitment of an RNA polymerase II complex is mediated by the constitutive activation domain in CREB, independently of

CREB phosphorylation." Mol Cell Biol **21**(4): 1001-10.
Felinski, E. A. and P. G. Quinn (1999). "The CREB constitutive activation domain interacts with TATA-binding protein-associated factor 110 (TAF110) through specific hydrophobic residues in one of the three subdomains required for both activation and TAF110 binding." J Biol Chem **274**(17): 11672-8.
Felinski, E. A. and P. G. Quinn (2001). "The coactivator dTAF(II)110/hTAF(II)135 is sufficient to recruit a polymerase complex and activate basal transcription mediated by CREB." Proc Natl Acad Sci U S A **98**(23): 13078-83.
Ferreri, K., G. Gill, et al. (1994). "The cAMP-regulated transcription factor CREB interacts with a component of the TFIID complex." Proc Natl Acad Sci U S A **91**(4): 1210-3.
Freiman, R. N., S. R. Albright, et al. (2002). "Redundant role of tissue-selective TAF(II)105 in B lymphocytes." Mol Cell Biol **22**(18): 6564-72.
Freiman, R. N., S. R. Albright, et al. (2001). "Requirement of tissue-selective TBP-associated factor TAFII105 in ovarian development." Science **293**(5537): 2084-7.
Gangloff, Y. G., J. C. Pointud, et al. (2001). "The TFIID components human TAF(II)140 and Drosophila BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger." Mol Cell Biol **21**(15): 5109-21.
Gangloff, Y. G., C. Romier, et al. (2001). "The histone fold is a key structural motif of transcription factor TFIID." Trends Biochem Sci **26**(4): 250-7.
Gangloff, Y. G., S. L. Sanders, et al. (2001). "Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7." Mol Cell Biol **21**(5): 1841-53.
Gangloff, Y. G., S. Werten, et al. (2000). "The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs." Mol Cell Biol **20**(1): 340-51.
Gates, M. A., L. B. Thomas, et al. (1995). "Cell and molecular analysis of the developing and adult mouse subventricular zone of the cerebral hemispheres." J Comp Neurol **361**(2): 249-66.
Ge, W., K. Martinowich, et al. (2002). "Notch signaling promotes astrogliogenesis via direct CSL-mediated glial gene activation." J Neurosci Res **69**(6): 848-60.
Gill, G., E. Pascal, et al. (1994). "A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation." Proc Natl Acad Sci U S A **91**(1): 192-6.
Goodrich, J. A., T. Hoey, et al. (1993). "Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB." Cell **75**(3): 519-30.
Gralla, J. D. (1996). "Global steps during initiation by RNA polymerase II." Methods Enzymol **273**: 99-110.
Grossmann, M., L. A. O'Reilly, et al. (2000). "The anti-apoptotic activities of Rel and RelA required during B-cell maturation involve the regulation of Bcl-2 expression." Embo J **19**(23): 6351-60.
Hampsey, M. (1998). "Molecular genetics of the RNA polymerase II general transcriptional machinery." Microbiol Mol Biol Rev **62**(2): 465-503.
Hansen, S. K., S. Takada, et al. (1997). "Transcription properties of a cell type-specific TATA-binding protein, TRF." Cell **91**(1): 71-83.
Helmann, J. D. and M. J. Chamberlin (1988). "Structure and function of bacterial sigma factors." Annu Rev Biochem **57**: 839-72.
Hermanson, O., K. Jepsen, et al. (2002). "N-CoR controls differentiation of neural stem cells into astrocytes." Nature **419**(6910): 934-9.
Hiller, M., X. Chen, et al. (2004). "Testis-specific TAF homologs collaborate to control a tissue-specific transcription program." Development **131**(21): 5297-308.
Hiller, M. A., T. Y. Lin, et al. (2001). "Developmental regulation of transcription by a tissue-specific TAF homolog." Genes Dev **15**(8): 1021-30.
Hoey, T., R. O. Weinzierl, et al. (1993). "Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators." Cell **72**(2):

247-60.
Hoffmann, A., C. M. Chiang, et al. (1996). "A histone octamer-like structure within TFIID." Nature **380**(6572): 356-9.
Holmes, M. C. and R. Tjian (2000). "Promoter-selective properties of the TBP-related factor TRF1." Science **288**(5467): 867-70.
Holstege, F. C., E. G. Jennings, et al. (1998). "Dissecting the regulatory circuitry of a eukaryotic genome." Cell **95**(5): 717-28.
Horikoshi, M., C. K. Wang, et al. (1989). "Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box." Nature **341**(6240): 299-303.
Huisinga, K. L. and B. F. Pugh (2004). "A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae." Mol Cell **13**(4): 573-85.
Isogai, Y. and R. Tjian (2003). "Targeting genes and transcription factors to segregated nuclear compartments." Curr Opin Cell Biol **15**(3): 296-303.
Jean, D., M. Harbison, et al. (1998). "CREB and its associated proteins act as survival factors for human melanoma cells." J Biol Chem **273**(38): 24884-90.
Kaltenbach, L., M. A. Horner, et al. (2000). "The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during C. elegans embryogenesis." Mol Cell **6**(3): 705-13.
Kargul, G. J., D. B. Dudekula, et al. (2001). "Verification and initial annotation of the NIA mouse 15K cDNA clone set." Nat Genet **28**(1): 17-8.
Kimura, H., Y. Tao, et al. (1999). "Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure." Mol Cell Biol **19**(8): 5383-92.
Kirschner, D. B., E. vom Baur, et al. (2002). "Distinct mutations in yeast TAF(II)25 differentially affect the composition of TFIID and SAGA complexes as well as global gene expression patterns." Mol Cell Biol **22**(9): 3178-93.
Koleske, A. J. and R. A. Young (1994). "An RNA polymerase II holoenzyme responsive to activators." Nature **368**(6470): 466-9.
Lee, T. I., H. C. Causton, et al. (2000). "Redundant roles for the TFIID and SAGA complexes in global transcription." Nature **405**(6787): 701-4.
Lemon, B. and R. Tjian (2000). "Orchestrated response: a symphony of transcription factors for gene control." Genes Dev **14**(20): 2551-69.
Lendahl, U., L. B. Zimmerman, et al. (1990). "CNS stem cells express a new class of intermediate filament protein." Cell **60**(4): 585-95.
Liu, Y., J. A. Ranish, et al. (2001). "Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes." J Biol Chem **276**(10): 7169-75.
Loriot, A., T. Boon, et al. (2003). "Five new human cancer-germline genes identified among 12 genes expressed in spermatogonia." Int J Cancer **105**(3): 371-6.
Lu, H., J. Lin, et al. (1994). "The regulation of p53-mediated transcription and the roles of hTAFII31 and mdm-2." Harvey Lect **90**: 81-93.
Maldonado, E. (1999). "Transcriptional functions of a new mammalian TATA-binding protein-related factor." J Biol Chem **274**(19): 12963-6.
Maldonado, E., R. Shiekhattar, et al. (1996). "A human RNA polymerase II complex associated with SRB and DNA-repair proteins." Nature **381**(6577): 86-9.
Mantamadiotis, T., T. Lemberger, et al. (2002). "Disruption of CREB function in brain leads to neurodegeneration." Nat Genet **31**(1): 47-54.
Martianov, I., G. M. Fimia, et al. (2001). "Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene." Mol Cell **7**(3): 509-15.
Matza, D., O. Wolstein, et al. (2001). "Invariant chain induces B cell maturation by activating a TAF(II)105-NF-kappaB-dependent transcription program." J Biol Chem **276**(29): 27203-6.
Mazzarelli, J. M., G. B. Atkins, et al. (1995). "The viral oncoproteins Ad5 E1A, HPV16 E7 and SV40 TAg bind a common region of the TBP-associated factor-110." Oncogene **11**(9): 1859-64.
Mazzarelli, J. M., G. Mengus, et al. (1997). "The transactivation domain of adenovirus

E1A interacts with the C terminus of human TAF(II)135." J Virol **71**(10): 7978-83.
Mengus, G., M. May, et al. (1997). "Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells." Genes Dev **11**(11): 1381-95.
Metsis, M., A. Brunkhorst, et al. (2001). "Cell-type-specific expression of the TFIID component TAF(II)135 in the nervous system." Exp Cell Res **269**(2): 214-21.
Mikolajczyk, M., J. Shi, et al. (2003). "The cyclin-dependent kinase 11(p46) isoform interacts with RanBPM." Biochem Biophys Res Commun **310**(1): 14-8.
Mohan, W. S., Jr., E. Scheer, et al. (2003). "TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice." Mol Cell Biol **23**(12): 4307-18.
Moore, P. A., J. Ozer, et al. (1999). "A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators." Mol Cell Biol **19**(11): 7610-20.
Morrison, S. J., S. E. Perez, et al. (2000). "Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells." Cell **101**(5): 499-510.
Muller, F. and L. Tora (2004). "The multicoloured world of promoter recognition complexes." Embo J **23**(1): 2-8.
Nakajima, T., C. Uchida, et al. (1997). "Analysis of a cAMP-responsive activator reveals a two-component mechanism for transcriptional induction via signal-dependent factors." Genes Dev **11**(6): 738-47.
Ninfa, A. J., L. J. Reitzer, et al. (1987). "Initiation of transcription at the bacterial glnAp2 promoter by purified E. coli components is facilitated by enhancers." Cell **50**(7): 1039-46.
Nooren, I. M. and J. M. Thornton (2003). "Structural characterisation and functional significance of transient protein-protein interactions." J Mol Biol **325**(5): 991-1018.
Ohbayashi, T., T. Kishimoto, et al. (1999). "Isolation of cDNA, chromosome mapping, and expression of the human TBP-like protein." Biochem Biophys Res Commun **255**(1): 137-42.
Palm, K., N. Belluardo, et al. (1998). "Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene." J Neurosci **18**(4): 1280-96.
Panchision, D. M. and R. D. McKay (2002). "The control of neural stem cells by morphogenic signals." Curr Opin Genet Dev **12**(4): 478-87.
Parvin, J. D. and R. A. Young (1998). "Regulatory targets in the RNA polymerase II holoenzyme." Curr Opin Genet Dev **8**(5): 565-70.
Perletti, L., J. C. Dantonel, et al. (1999). "The TATA-binding protein and its associated factors are differentially expressed in adult mouse tissues." J Biol Chem **274**(22): 15301-4.
Petty, K. J., Y. I. Krimkevich, et al. (1996). "A TATA binding protein-associated factor functions as a coactivator for thyroid hormone receptors." Mol Endocrinol **10**(12): 1632-45.
Pham, A. D., S. Muller, et al. (1999). "Mesoderm-determining transcription in Drosophila is alleviated by mutations in TAF(II)60 and TAF(II)110." Mech Dev **84**(1-2): 3-16.
Pointud, J. C., G. Mengus, et al. (2003). "The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation." J Cell Sci **116**(Pt 9): 1847-58.
Pugh, B. F. (2000). "Control of gene expression through regulation of the TATA-binding protein." Gene **255**(1): 1-14.
Rabenstein, M. D., S. Zhou, et al. (1999). "TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family." Proc Natl Acad Sci U S A **96**(9): 4791-6.
Rao, M. A., H. Cheng, et al. (2002). "RanBPM, a nuclear protein that interacts with and regulates transcriptional activity of androgen receptor and glucocorticoid receptor." J Biol Chem **277**(50): 48020-7.
Riccio, A., S. Ahn, et al. (1999). "Mediation by a CREB family transcription factor of NGF-dependent survival of sympathetic neurons." Science **286**(5448): 2358-61.

Roeder, R. G. (1996). "The role of general initiation factors in transcription by RNA polymerase II." Trends Biochem Sci **21**(9): 327-35.
Saluja, D., M. F. Vassallo, et al. (1998). "Distinct subdomains of human TAFII130 are required for interactions with glutamine-rich transcriptional activators." Mol Cell Biol **18**(10): 5734-43.
Sanders, S. L. and P. A. Weil (2000). "Identification of two novel TAF subunits of the yeast Saccharomyces cerevisiae TFIID complex." J Biol Chem **275**(18): 13895-900.
Selleck, W., R. Howley, et al. (2001). "A histone fold TAF octamer within the yeast TFIID transcriptional coactivator." Nat Struct Biol **8**(8): 695-700.
Shachar, I. and R. A. Flavell (1996). "Requirement for invariant chain in B cell maturation and function." Science **274**(5284): 106-8.
Shen, W. C., S. R. Bhaumik, et al. (2003). "Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly." Embo J **22**(13): 3395-402.
Shen, W. C. and M. R. Green (1997). "Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator." Cell **90**(4): 615-24.
Shi, Y. and C. Mello (1998). "A CBP/p300 homolog specifies multiple differentiation pathways in Caenorhabditis elegans." Genes Dev **12**(7): 943-55.
Shimohata, T., O. Onodera, et al. (2000). "Interaction of expanded polyglutamine stretches with nuclear transcription factors leads to aberrant transcriptional regulation in polyglutamine diseases." Neuropathology **20**(4): 326-33.
Silkov, A., O. Wolstein, et al. (2002). "Enhanced apoptosis of B and T lymphocytes in TAFII105 dominant-negative transgenic mice is linked to nuclear factor-kappa B." J Biol Chem **277**(20): 17821-9.
Sims, R. J., 3rd, R. Belotserkovskaya, et al. (2004). "Elongation by RNA polymerase II: the short and long of it." Genes Dev **18**(20): 2437-68.
Smale, S. T., M. C. Schmidt, et al. (1990). "Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID." Proc Natl Acad Sci U S A **87**(12): 4509-13.
Somers, J. P., J. A. DeLoia, et al. (1999). "Adenovirus-directed expression of a nonphosphorylatable mutant of CREB (cAMP response element-binding protein) adversely affects the survival, but not the differentiation, of rat granulosa cells." Mol Endocrinol **13**(8): 1364-72.
Struhl, K. (1987). "Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast." Cell **49**(3): 295-7.
Taatjes, D. J., A. M. Naar, et al. (2002). "Structure, function, and activator-induced conformations of the CRSP coactivator." Science **295**(5557): 1058-62.
Takizawa, T., K. Nakashima, et al. (2001). "DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain." Dev Cell **1**(6): 749-58.
Tanese, N., D. Saluja, et al. (1996). "Molecular cloning and analysis of two subunits of the human TFIID complex: hTAFII130 and hTAFII100." Proc Natl Acad Sci U S A **93**(24): 13611-6.
Teichmann, M., Z. Wang, et al. (1999). "Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells." Proc Natl Acad Sci U S A **96**(24): 13720-5.
Thuault, S., Y. G. Gangloff, et al. (2002). "Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain." J Biol Chem **277**(47): 45510-7.
Thut, C. J., J. L. Chen, et al. (1995). "p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60." Science **267**(5194): 100-4.
Tora, L. (2002). "A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription." Genes Dev **16**(6): 673-5.
Verrijzer, C. P., K. Yokomori, et al. (1994). "Drosophila TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA." Science **264**(5161): 933-41.
Verschure, P. J., I. van Der Kraan, et al. (1999). "Spatial relationship between

transcription sites and chromosome territories." J Cell Biol **147**(1): 13-24.
Walker, A. K., J. H. Rothman, et al. (2001). "Distinct requirements for C.elegans TAF(II)s in early embryonic transcription." Embo J **20**(18): 5269-79.
Wallen, A. A., D. S. Castro, et al. (2001). "Orphan nuclear receptor Nurr1 is essential for Ret expression in midbrain dopamine neurons and in the brain stem." Mol Cell Neurosci **18**(6): 649-63.
Wang, E. H., S. Zou, et al. (1997). "TAFII250-dependent transcription of cyclin A is directed by ATF activator proteins." Genes Dev **11**(20): 2658-69.
Wang, P. J. and D. C. Page (2002). "Functional substitution for TAF(II)250 by a retroposed homolog that is expressed in human spermatogenesis." Hum Mol Genet **11**(19): 2341-6.
Wang, Y., E. Marion Schneider, et al. (2002). "HIPK2 associates with RanBPM." Biochem Biophys Res Commun **297**(1): 148-53.
Wassarman, D. A., N. Aoyagi, et al. (2000). "TAF250 is required for multiple developmental events in Drosophila." Proc Natl Acad Sci U S A **97**(3): 1154-9.
Weis, L. and D. Reinberg (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**(6): 2973-84.
White-Cooper, H., D. Leroy, et al. (2000). "Transcription of meiotic cell cycle and terminal differentiation genes depends on a conserved chromatin associated protein, whose nuclear localisation is regulated." Development **127**(24): 5463-73.
Wolstein, O., A. Silkov, et al. (2000). "Specific interaction of TAFII105 with OCA-B is involved in activation of octamer-dependent transcription." J Biol Chem **275**(22): 16459-65.
Xie, X., T. Kokubo, et al. (1996). "Structural similarity between TAFs and the heterotetrameric core of the histone octamer." Nature **380**(6572): 316-22.
Yamit-Hezi, A., S. Nir, et al. (2000). "Interaction of TAFII105 with selected p65/RelA dimers is associated with activation of subset of NF-kappa B genes." J Biol Chem **275**(24): 18180-7.
Yao, T. P., S. P. Oh, et al. (1998). "Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300." Cell **93**(3): 361-72.
Zaman, Z., A. Z. Ansari, et al. (2001). "Interaction of a transcriptional repressor with the RNA polymerase II holoenzyme plays a crucial role in repression." Proc Natl Acad Sci U S A **98**(5): 2550-4.
Zhou, J., J. Zwicker, et al. (1998). "TAFII mutations disrupt Dorsal activation in the Drosophila embryo." Proc Natl Acad Sci U S A **95**(23): 13483-8.
Zou, Y., S. Lim, et al. (2003). "Serine/threonine kinase Mirk/Dyrk1B is an inhibitor of epithelial cell migration and is negatively regulated by the Met adaptor Ran-binding protein M." J Biol Chem **278**(49): 49573-81.

MATERIALS AND METHODS

I have used different protocols for similar methods in different papers. I present the methods by the paper they are used in.

Paper I:

Cloning of mouse TAFII135 cDNA.

Partial mouse TAFII135 cDNA was obtained by RT-PCR from total brain RNA using degenerate primers designed from the human TAFII135 protein sequence. The N-terminal degenerate primer covered amino acids 733 to 739 and the C-terminal primer covered amino acids 1005 to 1011 in the published sequence (Swiss-Prot Accession No. O00268). The obtained PCR products were gel purified and the fragment with the expected length was cloned and sequenced. The obtained partial cDNA was used for screening of a P0 mouse brain cDNA library (Stratagene) by conventional methods. The longest obtained cDNA clone was fully sequenced on both strands. The plasmid pCMVTAF4 was created by insertion of the cloned TAF4 cDNA into pRcCMV plasmid after addition of an in-frame ATG and Kozak sequence.

RNA extraction and Northern blot analysis.

Total RNA was isolated by the acid phenol extraction procedure (chomezynski 1987 anal biochem). Twenty-five micrograms of total RNA was run in each lane and fractionated on a 1.2% agarose–formaldehyde gel before transfer onto a nylon membrane (Hybond-N; Amersham). The amount and quality of transferred RNA were monitored by methylene blue staining of the filters before hybridization. The cDNA fragments were radiolabeled with [a-32P]dCTP, using the multiprobe DNA labeling system (Amersham) and then these fragments were used as probes. The blots were washed with high stringency and exposed to X-ray film for 1–10 days.

In situ hybridization.

In situ hybridization was performed by standard methods as described earlier (Metsis et al., 1993). Briefly, 14-mm sections were cut from fresh-frozen tissue, mounted on glass slides, postfixed with 4% paraformaldehyde/PBS for 20 min, deproteinated with 0.1 M HCl, and blocked with triethanolamin/acetic anhydride. Hybridization was performed for 16 h at 42°C in a humidified chamber with hybridization buffer (50% formamide, 20 mM Tris–HCl (pH 7.6), 1 mM EDTA, 0.3 M NaCl, 0.1 M dithiothreitol, 0.5 mg/ml yeast tRNA, 0.1 mg/ml poly(A) RNA (Sigma), 13 Denhardt's solution, and 10% dextran sulfate) containing 2.5 3 106 cpm/ml [35S]UTP-labeled riboprobe. After hybridization slides were RNase treated and washed in a dilution series of SSC with a final concentration of 0.13 SSC at 65°C. The slides were subsequently dehydrated and exposed to X-ray film or dipped into NTB-2 emulsion and exposed for 4 weeks. For visualization of cell nuclei, tissue was

counterstained with cresyl violet after the emulsion was developed.

Immunohistochemistry

For immunological detection, antibodies were created by immunization of rabbits with peptide NH2-PQSTMAPRPATPTG from the central part of the molecule corresponding to amino acids 482 to 495 of the human TAFII135 protein sequence (Swiss-Prot Accession No. O00268). Antibodies were purified by protein A–Sepharose chromatography followed by affinity chromatography using peptide–antigen immobilized on Affi-Gel 10 matrix. The 14-mm cryosections from fresh frozen tissue were mounted on glass slides and postfixed for 15 min with 4% paraformaldehyde in PBS. Before application of anti-TAFII135 antibodies, the tissue was blocked for 2 h with 2.5% Blocking Reagent (Roche) in TBS and then incubated with anti-TAFII135 antibodies overnight at 14°C in the same buffer. Immunoreactivity was detected using the Vectastatin ABC kit (Vector Laboratories).

Cell culture and CAT assays

PCC7 cells were grown in DMEM supplemented with 10% fetal calf serum. For differentiation, culture medium was supplemented with 1026 M all-*trans*-retinoic acid and 1 mM dibutyryl cyclic AMP. Transfection of cells and CAT assays were performed as described earlier [32] using [14C]chloramphenicol and acetyl-CoA as substrates. The products of the reaction were separated using TLC. The quantification of products was performed based on analysis of phosphoimager data.

Paper II

Cloning of mTAF4 cDNA´s

Mouse TAF4 cDNA (metsis 2001) was used to screen a neonatal (P0) BALBc whole brain cDNA library (Stratagene) by conventional methods. cDNA inserts from hybridization positive clones were subcloned and completely sequenced on both strands.

RT-PCR detection of alternatively spliced forms of TAF4 mRNA

Highly stringent primer pairs were designed using OLIGO 6.0 program (MBI, Cascade, USA) based on cDNA sequences. These primers and Superscript One-Step RT-PCR System (Life Technologies) were used to detect alternatively spliced mRNAs of mTAF4. PCR products with sizes different to the distance between primers in full-length mouse cDNA sequence were eluted, cloned into pCDNA-TOPO vector (Invitrogen) and sequenced.

Genomic cloning

Genomic DNA covering the TAF4 gene was isolated from a mouse BAC library by screening with the mixture of cDNA fragments obtained from cDNA library screening. Two independent BAC clones were obtained. Subsequent restriction site analysis revealed a common genomic region that was subcloned and entirely sequenced. This region was used as a landmark for identification of the exon-intron structure by PCR with Elongase (Life Sciences) using primers designed from the cDNA sequence. PCR products were either cloned into pTOPO TA 2.1 vector and sequenced using vector specific primers or directly sequenced using sequence specific primers after excision of PCR fragments from the agarose gel and using a gel purification kit (Qiagen).

RNase protection assay

RNase protection assay for the detection of different isoforms of TAF4 mRNA was performed with RPA II Kit (Ambion). For cRNA probe generation, selected regions from different TAF4 cDNA clones were obtained by PCR and subsequently cloned into pTOPO TA 2.1 vector (Invitrogen). For *in vitro* translation, plasmids were linearized with suitable restriction enzyme and translated with Riboprobe System (Promega) using T7 RNA polymerase. Protected RNA fragments were separated on denaturating polyacrylamide gels and analysed with a phosphoimager (Storm 840, Molecular Dynamics).

Detection of endogenous mTAF4 protein

For detection of endogenous mTAF4 protein, cell extracts (c17.2 neural stem cells) were incubated with antibody against human TAF4 (kindly provided by Dr. I. Davidson, IGBMC Strasbourg) for 1-2 hours at 4°C on a rotating wheel followed by 1 hour incubation with protein A/G agarose beads. Beads were subsequently centrifuged, washed, and the protein was eluted. The protein eluates were separated on 10% SDSPAGE gels. After electrophoresis, proteins were transferred to Hybond-N membrane (Amersham) and probed with the TAF4 antibody. Detection was performed using an ECL kit (Amersham).

Transient transfections

All mTAF4 derivates were cloned into pCDNA3 based TOPO3.1 eukaryotic expression plasmids (Invitrogen). All constructs were sequenced and *in vitro* translated using T7 polymerase and a rabbit reticulocyte system (Promega), and analyzed on SDS-PAGE gel to exclude clones with spontaneous frame-shift mutations. For transfections, cells were plated 24 hrs before transfection in 24-well plates at 2x104cell/cm2 in 10%FBS/DMEM. For all experiments Effectene transfection Reagent (Qiaqen) was used according to manufacturers recommendations. The total amount of DNA used was 2 µg per well. The total amount of DNA was adjusted by co-transfection of empty expression vector. The cells were harvested 48 h following transfection.

Luciferase assays

For functional analysis of mTAF4 alternative forms, PathDetect Trans-Reporting System (Stratagene) was used. For experiments on retinoic acid signal transduction RARGAL4DB and RXR-GAL4DB expression constructs were used (kindly provided by Dr.T. Perlmann, CMB and Ludwig Institute for Cancer Research, Karolinska Institutet, Stockholm). For CREB signal transduction pFA2-CREB plasmid was used (Stratagene). In all transfection experiments, pFR-luc plasmid was used as the reporter and pRL-CMV vector expressing *Renilla reniformis* luciferase was included as an internal standard. Transfection assays were performed in triplicate and analysed with Dual-Luciferase Reporter Assay System (Promega) using a Victor luminometer (Wallac). All transfection experiments were repeated at least three times with different DNA preparations.

Paper III

Neural stem cell culture

Isolation and culture of neural stem cells were performed essentially as described elsewhere (Johe et al., 1996; Hermanson et al., 2002). In brief, cortices from rats at embryonic day 15 were dissected and mechanically dispersed in a modified serumfree, HEPES-free N2-supplemented DMEM/F12 medium (Invitrogen). The primary cells were plated on dishes (35mm, 60mm, 100mm) coated with poly-L-ornithine and fibronectin (Sigma). The cells were treated with human recombinant basic FGF (FGF2) at 10 ng/ml every 24h and the N2 medium was replaced every 48h. All treatments and transfections were performed after the first passage when >>99% of the cells displayed nestin expression and typical neural stem cell morphology and <<1% of the cells expressed neuronal and glial differentation markers. To achieve astrocytic differentiation, the cells received ciliary neurotrophic factor (CNTF) at 10 ng/ml for 48h. For neuronal differentiation, the cells received platelet-derived growth factor-BB (PDGF-BB) at 20 ng/ml for 6-72h. All growth factors were obtained from R&D Systems. Inhibitors used included LY294002 (10-50μM, Cell Signaling) and PD98059 (10-20μM, Calbiochem).

Chromatin and whole-cell preparations

For whole cell lysates, cells from 10 cm dishes were scraped and lysed in NETN buffer with 1 x protease inhibitor cocktail (Complete_, Roche) for 15 min followed by centrifugation at 4°C for 20 min. Chromatin precipitation of NSC was performed essentially as described elsewhere (Hermanson, Jepsen et al. 2002). Briefly, NSC were crosslinked by 1% formaldehyde for 10 min at RT, followed by a 15 min incuation at 30°C in 100 mM Tris-HCl (pH9.4) and 10 mM DTT. The pellets were washed in buffer I (pH6.5; Triton X-100 (0.25%), EDTA (10mM), EGTA (0.5mM), HEPES (10 mM)) and buffer II (pH6.5; 200mM NaCl, EDTA (1mM), EGTA (0.5mM), HEPES (10 mM)), and resuspended in lysis buffer (1%SDS, 10mM EDTA, 50mM Tris-HCl, 1 x protease inhibitor cocktail), before sonicated and centrifuged. The soluble chromatin was retrieved and stored in 1% Triton X-100, 2mM EDTA,

150mM NaCl, 20 mM Tris-HCl, and 1 x protease inhibitor cocktail.

Immunocytochemistry, immunoblotting and antibodies

For immunocytochemistry, cells were fixed in the plates with formalin for 20 min at room temperature (RT), rinsed three times in PBS, and incubated in primary antibodies O/N at 4°C under agitation. The cells were then rinsed six times in PBS before incubation in secondary antibody for 1h at RT in the dark under agitation. After three additional rinses, the plates were coverslipped with Vectashield with DAPI. Protein electrophoresis and immunoblotting were performed with standard procedures as described elsewhere (Hermanson, Jepsen et al. 2002). Primary antibodies: mouse anti-TAF4 (1:500-2000; Pharmingen), goat anti-RanBPM (1:100-500; Mikolajczyk et al., 2003), mouse anti-Tuj1 (1:500; Covance), mouse anti-MAP2 (1:500; Sigma), rabbit anti-GFAP (1:500; DAKO), rabbit anti-Actin (1:2000; Sigma), chicken anti-GFP (1:1000; Chemicon). Secondary antibodies for immunocytochemistry were purchased from Molecular Probes and included donkeyanti-mouse, donkey-anti-goat, donkey-anti-rabbit, and donkey-anti-chicken conjugated to Alexa 488 and Alexa 594 and used at 1:500. Secondary antibodies for immunoblotting were purchased from Chemicon and included peroxidase-conjugated rabbit-anti-goat and rabbit-anti-mouse used at 1:2000.

DNA delivery and constructs

Transient transfections of NSC were performed essentially according to the supplier's recommendations (Effectene, Qiagen). Expression constructs used in this study: pCMV-TAF4 (Metsis, Brunkhorst et al. 2001) corresponding to aa 42-702 of the 702 amino acid annotated mouse protein (XP130764); pCMV-TAF4del harboring a deletion of the Cterminus (aa 338-702); pCMV-HA-RanBPM (Mikolajczyk, Shi et al. 2003); pCMVEGFP.

Yeast two-hybrid System

A 288 nucleotide domain was amplified using sense primer
(5′- CTCCAGCCCCCGGTCA-3′) and antisense primer
(5- TGCTTGTGCAGATACAGTAGA-3′) using mouse TAF4 as the template. The PCR product was sequenced and cloned into pAS2-1 vector in frame with the DNAbinding domain of the yeast GAL4 protein using PstI and BamHI. The resulting contruct was transformed into S. cerevisiae strain AH109 and used as bait in two hybrid screening of MATCHMAKER mouse brain cDNA library transformed into S. cerevisiae strain Y187 (Clontech Cat#638849). Screening procedures were performed according to the manufacturers protocol. Transformants were plated on –Leu–Trp–Ade–His medium and incubated at 30 CÅã for 3-5 days. Resulting clones were sequenced and GenBank database was searched using BLAST (National Center for Biotechnology Information, Bethesda, MD).

β-Galactosidase filter and liquid assay

For the filter assay single colonies were picked and transferred to a Whatman No. 5 filter paper, which was further incubated on a fresh plate for 2–3 days. The filters were frozen in liquid nitrogen and then layered over a second filter presoaked with Zbuffer (16.1 g/L Na2HPO4*7H2O, 5.5 g/L NaH2PO4*H2O, 0.75 g/L KCl, and 0.246 g/L MgSO4*7H2O) which contained 0.27 ml of 2-mercaptoethanol and 1.67 ml X-gal (20 mg/ml in dimethylformamide) per 100 ml. Incubation was done at 30 C for up to 12 h.

Immunohistochemistry

For immunohistochemistry (Wallen, Castro et al. 2001), slides were air-dried, washed in phosphate-buffered saline (PBS), and incubated with blocking solution (3% bovine serum albumin or 0.1% fetal calf serum, 0.3% Triton X-100 in PBS or 10 mM Hepes buffer) before overnight incubation with primary antibody diluted in blocking solution at 4°C. Following rinsing of slides, they were incubated with secondary antibody at 1:200 dilution (Alexa Fluor 488 and 594 IgG (Molecular Probes) for 1 h. After rinses in PBS, slides were mounted with Vectashield mounting medium (Vector).

In situ hybridization

In situ hybridization was performed by standard methods as described earlier (Metsis, Brunkhorst et al. 2001). Briefly, 14 mm sections were cut from fresh-frozen tissue,
mounted on glass slides, and postfixated with 4% paraformaldehyde for 20 min. Hybridization was performed for 16 hrs at 42°C with hybridization buffer containing 2.5x106cpm/ml [35S]UTP labeled riboprobe (Metsis, Brunkhorst et al. 2001). After hybridization, the slides were RNAse treated and washed in a series of SSC rinses at decreasing concentrations with a final concentration of 0.1X SSC at 65°C. The slides were subsequently dehydrated and exposed to X-ray film or dipped into NTB-2 emulsion and exposed for 4 weeks. For visualization of cell nuclei, tissue was counterstained with cresyl violet after development of the emulsion.

Paper IV

Construction of targeting vector and inactivation of TAF4

The murine TAF4 locus was isolated by screening an ES genomic DNA library with the TAF4 cDNA. Genomic DNA from positive phages was subcloned into the Not I site of plasmid pZERO (InvitroGen). The genomic locus was then sequenced using internal primers and automated DNA sequencing (Applied Biosystems). Using the cloned locus as template, PCR was used to generate a targeting vector in which a PGK promoter-driven hygromycin resistance gene flanked by LoxP sites was inserted downstream of exon 12 and a second loxP site inserted upstream of exon 11 along with an additional EcoRI site. This allows deletion of exons 11 and 12 by Cre treatment. All

constructions were verified by restriction enzyme digests and automated DNA sequencing. Genotyping of ES cell clones and mice was performed by Southern blot and digests. The TAF4lox/+ ES cell clone was electroporated with a vector expressing the Cre recombinase and 40 subclones were analysed by Southern Blot for the presence of the deleted allele. Several TAF4+/- ES clones were thus identified. Blastocyst injection and generation of the mutant mice from the TAF4lox/+ and TAF4-/+ ES cell clones was performed as previously described (Martianov, Fimia et al. 2001)

Isolation of embryonic fibroblasts

Taf4lox/- mice were crossed and the embryos isolated at E13.5. After dissection, the head was used for genotyping and the inner organs were removed before homogenisation and initiation of a primary cell culture. Fibroblasts from TAF4lox/- embryos were passed at high dilution until a population of immortalised cells was obtained. These cells were transfected with vectors expressing the Cre recombinase and the green fluorescent protein. 48 hours later, cells were sorted by FACS and individual GFP-expressing cells were seeded in 96 well plates. Clones were amplified and genotyped to yield more than 5-10 clones of each TAF4lox/- and TAF4-/- genotype. Cells were grown in Dulbecco's minimal essential media supplemented with glutamax and 10% fœtal calf serum.

Microarray preparations

Total RNA from the four imortalized fibroblast cell lines was isolated with RNeasy (Qiagen) and were reverse transcribed to produce cDNAs that were labelled with either Cy3 (control group, wt) or Cy5 (TAF -/-) In the present experiment, a 1/1 mixture of cells from the two wt samples was used as a reference. The total mRNA mixture from each group was adjusted to 20.0 µg/slide. The above target DNA was hybridized on NIA microarray chips. Glass arrays of a 15K mouse cDNA clone set developed at the NIA Intramural Research Program Laboratory of Genetics. The clones were derived from early embryonic cDNA libraries (Kargul, Dudekula et al. 2001). The labeling and hybridization was done at the IGBMC microarray core facility. Following hybridization and rinsing, the expression levels were quantified by the pixel intensity values (PIVs) for each spot (gene) and data processing such as the Cy5/Cy3 fluorescence ratio was performed using ImaGene™ 3.0 software (Bio-Discovery Inc., Los Angeles, CA, USA). Using this software, the PIV of each spot from each group was compared with the background PIV, and then the Cy5 (TAF-/-)/Cy3 intensity ratios (RT) were calculated. Spots with RT ≥ 2 or ≤ 0.5 were considered to represent genes that were differentially expressed, compared with the control group.