

Université de Montréal

**Études structurales d'interactions
protéine/protéine impliquées dans l'érythropoïèse**

par

Caroline Mas

Département de Biochimie

Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
En Biochimie

6 Août, 2010

© Caroline Mas, 2010

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée :
Études structurales d'interactions protéine/protéine impliquées dans l'érythropoïèse

Présentée par :
Caroline Mas

A été évaluée par un jury composé des personnes suivantes :

Muriel Aubry, président-rapporteur
James G. Omichinski, directeur de recherche
Éric Milot, membre du jury
Pierre Lavigne, examinateur externe
Michel Lafleur, représentant du doyen

Résumé

Le développement hématopoïétique est régulé par l'action combinée de facteurs de transcription lignée-spécifiques et de la machinerie transcriptionnelle de base, permettant ainsi l'expression de gènes en temps et lieu appropriés. Les travaux présentés dans cette thèse portent sur l'étude structurale et fonctionnelle d'interactions décisives pour la régulation de l'expression de gènes et impliquant des domaines de transactivation (TAD). En effet, les interactions faisant intervenir les TAD d'activateurs permettent de réguler l'activation de la transcription de façon spécifique.

La première étude présentée dans cette thèse relate l'identification et la caractérisation d'une nouvelle interaction entre deux facteurs de transcription : le facteur hématopoïétique GATA-1 et la protéine suppresseur de tumeur p53. En combinant des études *in vitro* par titrage calorimétrique en condition isotherme (ITC) et par spectroscopie RMN et des études *in vivo*, nous avons identifié et caractérisé cette nouvelle interaction. Il s'avère que le TAD de p53 et le domaine de liaison à l'ADN de GATA-1 sont les domaines minimaux requis pour la formation de ce complexe. L'inhibition de la voie p53 par GATA-1 s'est avérée être la conséquence majeure de cette interaction, permettant ainsi le maintien en vie des précurseurs érythrocytaires *via* l'inhibition de l'apoptose.

Un deuxième type d'interaction a fait l'objet d'études : l'interaction entre divers TAD et la machinerie transcriptionnelle de base, plus spécifiquement avec le Facteur général de Transcription IIIH (TFIIH). La structure des complexes constitués par la sous-unité Tfb1/p62 du facteur TFIIH en interaction avec le TAD viral de VP16 d'une part, et avec le TAD humain du facteur érythrocytaire « Erythroid Krüppel-like factor » (EKLF) d'autre part, ont été résolues par spectroscopie RMN. La structure du complexe Tfb1/VP16 a révélée que le mode de liaison de VP16 à Tfb1 est similaire au mode de liaison du TAD de p53 avec le même partenaire. En effet, les TAD de VP16 et de p53 forment tous deux une hélice- α de 9 résidus en interaction avec Tfb1. En dépit de partager avec p53 et VP16 le même site de liaison sur Tfb1/p62, la structure RMN du complexe EKLF/Tfb1 démontre que le mode d'interaction de ce TAD se distingue du mode de liaison canonique des activateurs transcriptionnels. Etonnamment, EKLF adopte un

mécanisme de liaison semblable au mécanisme de liaison du facteur général de transcription TFIIE α avec p62, leurs conformations demeurent étendues en interaction avec Tfb1/p62. En se basant sur nos données structurales, nous avons identifié un résidu dans le TAD d'EKLF décisif pour la formation du complexe EKLF/p62 : le Trp⁷³. La mutation de cet acide aminé perturbe son interaction avec Tfb1PH/p62PH et réduit significativement l'activité transcriptionnelle d'EKLF dans les érythrocytes.

Mots-clés : Facteur général de transcription, érythropoïèse, GATA-1, EKLF, p53, VP16, transcription, domaines de transactivation, résonance magnétique nucléaire, calorimétrie à titrage isotherme.

Abstract

Hematopoietic development is regulated through a combinatorial interplay between lineage-specific activators and the general transcription machinery that enables cell-specific patterns of gene expression. This thesis reports structural and functional studies of interactions involving the transactivation domains (TAD) of activator proteins and their role in hematopoietic development. Interactions between the TAD of activators and their partners play an important role in the transcriptional regulation of all genes including those regulating hematopoiesis.

The first section reports the identification and characterization of a novel interaction between the erythroid transcription factor GATA-1 and the tumor suppressor protein p53. Using a combination of isothermal titration calorimetry (ITC), NMR spectroscopy and *in vivo* studies, we identified and characterized the direct interaction between these two important transcription factors in an attempt to determine the role of this interaction in erythroid development. Based on our results, the TAD of p53 directly interacts with the DNA-binding domain of GATA-1 in a cell-type specific manner. Through this interaction, GATA-1 inhibits activation of select p53-regulated genes and we postulate that the inhibition of p53-dependent apoptotic pathways is essential for survival of erythroid precursor cells.

In the second section, we report on the interactions between two acidic TADs and the general transcription factor IIIH (TFIIH). The structure of the complexes formed by the Tfb1/p62 subunit of TFIIH (Tfb1PH/p62PH) and the acidic TAD of Herpes Simplex viral protein 16 (VP16) and the Erythroid Krüppel-like factor (EKLF) were determined by NMR spectroscopy. The structure of the Tfb1PH/VP16 complex demonstrated that a viral TAD has the ability to mimic the actions of the TAD from the human p53 with Tfb1PH/p62PH. The TADs of both VP16 and p53 adopt a 9-residue α -helix in complex with Tfb1PH/p62PH. Interestingly, the NMR structure of the EKLF/Tfb1PH complex demonstrated that despite sharing a common binding site with p53 and VP16 on Tfb1PH, the EKLF/Tfb1PH binding interface is distinctly different from the binding interfaces we previously observed with p53/Tfb1PH and VP16/Tfb1PH complexes. Surprisingly, EKLF adopted a similar binding mechanism as the

general transcription factor TFIIE α in interaction with p62PH as both interact in an extended conformation. Moreover, based on our structural data, we have identified Trp⁷³ as a key residue within the TAD of EKLF that is required for the formation of the EKLF/Tfb1PH complex. Mutations of Trp73 disrupted the binding to Tfb1PH/p62PH and significantly reduced the transcriptional activity of EKLF in red blood cells.

Keywords: General transcription factor, erythropoiesis, GATA-1, EKLF, p53, VP16 transcription, transactivation domains, nuclear magnetic resonance spectroscopy, isothermal titration calorimetry.

Table des matières

Résumé.....	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux.....	x
List des figures.....	xi
Liste des abréviations.....	xiv
Remerciements.....	xix
Avant propos	1
CHAPITRE I.....	4
INTRODUCTION	4
A - Transcription des ARN messagers chez les eucaryotes	5
1 - Formation du complexe de pré-initiation.....	5
2 - Le facteur général de transcription TFIID.....	8
a - TFIID et la régulation de l'expression de la β -globine	8
b - La sous-unité p62/Tfb1 de TFIID.....	10
3 - Activateurs et co-activateurs de la transcription.....	14
a - Les activateurs	14
b - Les co-activateurs	15
i - Complexes médiateurs	16
ii - Histones acétyltransférases	16
B - Facteurs de transcription impliqués dans la différenciation des érythrocytes.....	22
1 - Les érythrocytes.....	22
2 - Le Facteur de Transcription GATA-1	23
a - Domaines de GATA-1	23
b - Rôles de GATA-1	26
c - Interactions protéine/protéine.....	28
d - Maintien en vie des précurseurs des érythrocytes.....	32
3 - Erythroid Krüppel-like factor (EKLF).....	35
a - Domaines d'EKLF	35

b - Similitudes des TAD d'EKLF et de p53	35
c - Rôles d'EKLF	38
d - Interactions protéine/protéine	39
C - Désordres sanguins	40
1 - GATA et désordres sanguins	40
2 - EKLF et désordres sanguins	41
D - Objectifs de recherche et justification des méthodes employées	42
1 - Hypothèses et Objectifs de recherche	42
2 - Résonance magnétique nucléaire	44
a - Principe théorique de la RMN	44
b - Détermination de structures tridimensionnelles par spectroscopie RMN	48
c - Cartographie des régions d'interactions	51
3 - Calorimétrie à titrage isotherme	53
Préface au Chapitre II	58
CHAPITRE II	59
GATA-1 interagit avec et inhibe p53	59
Article 1: GATA-1 associates with and inhibits p53.	59
Préface au Chapitre III	93
CHAPITRE III	95
Structure RMN du complexe formé par la sous unité Tfb1 de TFIIH et le domaine de transactivation de VP16 :	
Similitudes entre VP16 et p53.	95
Article 2: NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16:	
structural similarities between VP16 and p53.	95
Préface au Chapitre IV	130
CHAPITRE IV	131

Caractérisation structurale et fonctionnelle d'un domaine de transactivation acide atypique d'EKLF	131
Article 3: Structural and Functional Characterization of an Atypical Acidic Transactivation Domain in the Erythroid-Kruppel-Like-Factor (EKLF)	131
CHAPITRE V	180
CONCLUSIONS ET PERSPECTIVES.....	180
A - Conclusions	181
B - Perspectives	184
1 - GATA-1 et ses partenaires	184
2 - Effet de modifications post-traductionnelles	185
3 - Mécanismes de liaison des TAD	186
4 - La transcription, une cible thérapeutique.....	187
C - Conclusions générales	189

Liste des tableaux

CHAPITRE III

Table 3. 1. Comparison of the K_d (nM) values for the binding of VP16 and Tfb1₁₋₁₁₅ using ITC. 121

Table 3. 2. Structural Statistics of the Tfb1/VP16C Complex^a 122

CHAPITRE IV

Table 4. 1. Comparison of the K_d (μ M) Values for the Binding of EKLFTAD2 with Tfb1PH, p62PH and domains of CBP/p300. 161

Table 4. 2. Structural Statistics of the Tfb1PH/EKLFTAD2 Complex^a 162

Table 4. 3 : Comparison of the K_d (μ M) Values for the Binding of EKLFTAD2 to Tfb1PH and Tfb1PH mutants. 163

Liste des figures

CHAPITRE I

Figure 1. 1 . L'hématopoïèse	3
Figure 1. 2 . Représentation schématique de l'assemblage séquentiel du complexe de pré-initiation (PIC) au promoteur	7
Figure 1. 3 . Représentation schématique du facteur général de transcription TFIID.....	9
Figure 1. 4 . Structures du domaine PH de p62/Tfb1 libres ou en complexe.....	12
Figure 1. 5 . Structures de différents domaines de CBP/p300 en complexe avec différents TAD de facteurs de transcription.....	19
Figure 1. 6 . Structure de GATA-1 _{CF+BA} lié à l'ADN.....	25
Figure 1. 7 . Structure et expression de la β -globine sur le gène globine « β -like ».....	28
Figure 1. 8 . Domaines fonctionnels de GATA-1.....	29
Figure 1. 9 . Érythropoïèse et énucléation.....	31
Figure 1. 10 . Structure et expression des gènes de globines « α -like » et « β -like » embryonnaire, fœtale et adulte.....	34
Figure 1. 11 . Domaines fonctionnels d'EKLF.....	37
Figure 1. 12 . Alignement de séquences du TAD1 (A) et du TAD2 (B) d'EKLF humain (hEKLf) avec les TAD respectifs de p53 humain (hp53).....	37
Figure 1. 13 . Étapes standard pour la détermination de structures par spectroscopie RMN hétéronucléaire.....	50
Figure 1. 14 . Corrélations entre les atomes ^{15}N et ^1H observées dans un spectre ^1H - ^{15}N HSQC hétéronucléaire à deux dimensions.....	51
Figure 1. 15 . Principe de la cartographie d'une interaction par spectroscopie RMN.....	52
Figure 1. 16 . Thermogramme d'une expérience d'ITC.....	54

CHAPITRE II

Figure 2. 1. The full-length p53TAD binds to the zinc-finger domains of GATA-1 in an <i>in vitro</i> pull-down assay.....	79
Figure 2. 2. The linker region, the carboxyl-terminal zinc finger and the basic arm of GATA-1 are sufficient for binding p53TAD in an <i>in vitro</i> pull-down assay.....	81
Figure 2. 3. GATA-1 and p53 interact <i>in vitro</i> in MEL cell extract..	83
Figure 2. 4. GATA-1, p53TAD and a GATA-1 DNA-binding site do not form a ternary complex	84
Figure 2. 5. GATA-1 can inhibit transactivation by p53 in 6C2 cells..	85
Figure 2. 6. The linker between the zinc fingers and the CF of GATA-1 are required for p53 inhibition. .	87
Figure 2. 7. p53 1-65 can compete with WT p53 for binding to GATA-1 and reduce the inhibition of p53.....	89
Supplementary Figure 2. 1. Amino Acids 232 to 256 of the linker between the two zinc fingers of GATA-1 are required for inhibition of p53.	90
Supplementary Figure 2. 2. GATA-1 ZFD and GATA-1 L +CF bind p53TAD with the same affinity..	92

CHAPITRE III

Figure 3. 1. VP16C bind to the PH domain of Tfb1.....	123
Figure 3. 2. Structure of the Tfb1/VP16C complex.....	124
Figure 3. 3. Structural details of the Tfb1 ₁₋₁₁₅ /VP16C interface.....	125
Figure 3. 4. Comparison of VP16C and p53TAD2 in complex with Tfb1.....	126
Supplementary Figure 3. 1. ITC analysis of VP16C binding to Tfb1 ₁₋₁₁₅	127
Supplementary Figure 3. 2. VP16C interacts with the PH domain of Tfb1 ₁₋₁₁₅	128
Supplementary Figure 3. 3. NMR titration of ¹⁵ N-labeled VP16C with Tfb1 ₁₋₁₁₅	129

CHAPITRE IV

Figure 4. 1 : EKLFTAD and EKLFTAD2 are homologous to p53TAD1 and p53TAD2.....	164
Figure 4. 2 : EKLFTAD2 and p53TAD2 share a common binding site on Tfb1PH/p62PH.....	165

Figure 4. 3 The KIX and IBiD domains of CBP interact with EKLFTAD2.....	167
Figure 4. 4: Structure of the Tfb1PH/EKLFTAD2 Complex.	169
Figure 4. 5: Electrostatic interactions in the Tfb1PH/EKLFTAD2 complex.....	171
Figure 4. 6: EKLFTAD2 and TFIIE α bind similarly to Tfb1PH/p62PH, but different then p53TAD2 and VP16C.....	173
Figure 4. 7: Trp73 is crucial for the binding and in vivo activity of the minimal EKLFTAD.....	175
Supplementary Figure 4. 1. EKLFTAD1 and EKLFTAD2 function as independent activation domains in yeast	177
Supplementary Figure 4. 2. EKLFTAD2 binds to Tfb1PH.	178

Liste des abréviations

A	Alanine
ACTR	Activator of Thyroid Receptor
ADN	acide désoxyribo nucléique
Ala	alanine
ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
ARC/DRIP	Activator-Recruited Co-factor / vitaminD-Receptor Interacting Proteins
Arg	Arginine
ARN	Acide ribonucléique
ARN pol II	ARN polymérase II
Asp	Acide aspartique
ATP	Adénosine triphosphate
BA	Basic Arm (Bras basique)
BAF	BRG1/brm-associated factor
Bcl	B-cell lymphoma 2
BRG1	Brahma-related gene 1
C	Cystéine
CBP/p300	CREB binding protein
cdk	Cyclin dependent kinase
CF	Carboxy Zinc Finger
CITED2	CBP/p300-Interacting Transactivator ED-rich 2
CLL	Chronic lymphocytic leukemia
CML	Chronic myelogenous leukemia
CREB	c-AMP response element-binding protein
CRSP/ PC2	Cofactor required for Sp1/ positive component 2
CSH	Cellules souches hématopoïétiques
CTD	Carboxy-terminal domain

D	Acide aspartique
DBD	DNA Binding Domain
E	Acide glutamique
E1A	Early region 1A
EKLF	Erythroid Krüppel-Like Factor
FasL	Fas ligant
FGT	Facteurs généraux de transcription
Fli-1	Friend leukemia virus integration 1
FOG1	Friend-Of-GATA.
F	Phénylalanine
Gln	Glutamine
Glu	Acide glutamique
HAT	Histones AcétylTransférases
HIF-1	Hypoxia-Induced Factor 1
HS	Hypersensibilité
Hsp70	Heat Shock Protein 70
HSQC	Heteronuclear Single Quantum Correlation
HSV	Herpes simplex virus
HTLV-1	Human T cell Leukemia Virus type 1
HTLV-2	Human T cell Leukemia Virus type 2
IBiD	IRB3 binding Domain
I	Isoleucine
Ile	Isoleucine
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KIX	Kinase-induced Interacting Domain
L	Leucine
LCR	Locus control region
Leu	Leucine
Lys	Lysine
M	Méthionine
MAT1	Ménage-à-Trois 1

Mdm2	Murine double minute 2
Mdm4	Murine double minute 4
MED	Mediator
Met	Méthionine
MLL	Mixed Lineage Leukemia protein
mM	Millimolaire mM
MOZ	Monocytic leukemia zinc finger
N	Asparagine
NER	Nucleotide Excision Repair
NF	Doigt de zinc amino-terminal
NFκB,	Nuclear factor-κB
NOE	Nuclear Overhauser Effect
NF-E2	Nuclear factor erythroid-derived 2
P/CAF	p300/CBP-associated factor
PC4	Positive Cofactor 4
PCR	Polymerase Chain Reaction (R ² action de polym ² risation en cha ² ne)
PDB	Protein data bank
Phe	Phenilalanine
PH	Pleckstrin homologie
PIC	Pre-Initiation Complex
pKID	Phosphorylated Kinase-Inducible Domain
ppm	Parti par million
P	Proline
Pro	Proline
pRB	Protéine du rétinoblastome
Q	Glutamine
R	Arginine
RMN	Résonance Magnétique Nucléaire
RMSD	Root Mean Square Deviation
RPA70	Replication protein A
S	Sérine
Ser	Sérine

SRC1	Steroid Receptor Coactivator 1
STAT1	Signal Transducer and Activator of Transcription 1
STAT2	Signal Transducer and Activator of Transcription 2
SWI/SNF	SWItching mating type/Sucrose Non Fermenting
TAD	Transactivation domain
TAF	TBP-Associated Factor
TAZ1	Transcription Adaptator putative Zinc finger 1
TAZ2	Transcription Adaptator putative Zinc finger 2
TBP	TATA box-Binding Protein
TFII	ARN polymerase II Transcription Factor
Thr	Thréonine
TRAP/SMCC	Thyroid hormone Receptor-Associated Proteins/SRB/mediator containing complex
Tris	Tris(hydroxyméthyl)méthane
T	Thréonine
Trp	Tryptophane
V	Valine
Val	Valine
VP16	Virion protein 16
W	Tryptophane
μM	Micromolaire
^{13}C	Carbone 13
^{15}N	Azote 15

À la mémoire de mon frère, Philippe.

Remerciements

J'aimerais en premier lieu témoigner ma reconnaissance envers mon directeur de recherche, Jim Omichinski. Je le remercie de m'avoir accueilli dans son laboratoire il y a maintenant presque cinq ans, sans même que nous nous soyons rencontré, je le remercie d'avoir guidé mon apprentissage de la recherche scientifique, et d'avoir été disponible pour faire, entre autre, de l'attribution de spectres avec moi. Ces années passées au Québec auront été des plus enrichissantes tant professionnellement que personnellement.

Merci à tous les membres des labos Legault-Omichinski, étudiants, postdocs, stagiaires et assistants de recherche avec qui j'ai eu la chance de travailler, vous avez tous contribué à créer une belle ambiance. Je tiens tout particulièrement à remercier Chantal Langlois et Vanessa Delfosse pour la relecture de se manuscrit. Merci Pascale, Paola, Chantal, Luca Riola, et Julien Lafrance-Vanasse pour leur disponibilité à discuter ou répondre à mes questions de RMN ou à enregistrer des spectres avec moi. Je tiens aussi à remercier mes collègues et amis, compagnons de randonnées, d'escalade, de voyage, de camping d'hiver... Patricia, Alex, Luca, Thomas, Sabrina et Julie se reconnaîtrons...que de bons souvenirs (à l'exception du camping d'hiver bien sur)! Merci aussi à Thom, Matt et Patrick mes stagiaires " mal traités ". Une petite pensée à Nicolas Girard (ou GB) pour les caricatures qu'il a fait de moi. Merci à Mathieu Coinçon pour ses blagues, pour les pauses café et merci à lui et à Sandrine de m'avoir fait l'honneur d'être la marraine de leur fils Gaël. Merci Chantal pour ton support moral, et pour m'avoir mise dans un avion en Août 2006.

Finalement, je voudrais aussi remercier infiniment ma famille. Mes parents, Marie-Claire et André pour leur soutien tout au long de mes études, pour toutes les cartes postales que ma mère m'a envoyées tout au long de ces cinq années passées loin d'eux et qui m'ont fait chaud au cœur... Merci aussi à mon frère Nicolas, et à Manu pour toutes ces photos et les vidéos des Pyrénées qui me rendaient parfois nostalgique....

Enfin je tiens à remercier mon copain, Colin, pour son amour et son affection. Il aura fait ce bout de chemin avec moi, et même à plusieurs milliers de kilomètres, il aura su me soutenir dans les moments difficiles.

Avant propos

Le mot « hématopoïèse » provient du grec ancien ‘aïma, aïmatos’ qui désigne le sang et ‘poiësis’ qui signifie création. Ce terme désigne les processus engendrant les différentes lignées cellulaires sanguines dont les fonctions varient de l’immunité au transport de l’oxygène. Au cours du développement des mammifères, le lieu de production des cellules sanguines varie ; les sites séquentiels de l’hématopoïèse sont : la vésicule vitelline, le foie et finalement la moelle osseuse [1-3]. Les **Cellules Souches Hématopoïétiques** (CSH) présentes dans ces tissus sont multipotentes, c’est-à-dire qu’elles ont la capacité de s’auto-renouveler et de se différencier exclusivement en diverses cellules sanguines. Les cellules sanguines matures produites ont une durée de vie limitée. On estime à 10^{11} le nombre de cellules hématopoïétiques produites quotidiennement chez l’adulte afin de maintenir l’homéostasie [4-6] ; ce chiffre pouvant rapidement augmenter en condition de stress physiologique.

Les facteurs de transcription hématopoïétiques orchestrent la différenciation des CSH en régulant la transcription des gènes afin que ceux-ci soient exprimés ou réprimés de manière spécifique au cours du développement de chaque type de cellules sanguines (Figure 1. 1 en page 3) [2, 7-9]. Ces protéines possèdent, entre autre, des domaines de liaison à l’ADN (DBD : DNA Binding Domain) capables de reconnaître des séquences d’ADN particulières, ciblant ainsi l’expression de gènes spécifiques. Elles comportent également des domaines de transactivation (TAD) leur conférant la faculté d’activer l’expression de gènes en participant à des réseaux complexes d’interactions avec des partenaires protéiques. L’altération de la production de cellules sanguines est généralement associée à des mutations retrouvées dans divers facteurs de transcription hématopoïétiques, causant ainsi le développement d’anémies ou de leucémies [7, 9]. Étant donné le rôle crucial de ces protéines dans la prolifération de certains cancers, la compréhension des interactions avec leurs partenaires s’avèrent essentielle pour la conception de nouveaux traitements thérapeutiques.

L’objectif de mon projet de recherche fut l’étude d’interactions protéine/protéine régissant le développement des globules rouges (érythrocytes). Au cours de mes études doctorales je me

suis donc principalement intéressée à l'étude structurale et fonctionnelle d'interactions faisant intervenir des TAD de facteurs de transcription avec différents partenaires protéiques. Une première partie introductive sera consacrée au mécanisme de transcription des ARN messagers par l'ARN polymérase II ainsi qu'aux rôles des activateurs et des co-activateurs de la transcription. Les fonctions régénées par les facteurs de transcriptions érythropoïétiques GATA-1 et EKLF (**E**rythroid **K**rüppel-**L**ike **F**actor) feront l'objet d'une deuxième partie. Suite à cela, une troisième partie sera consacrée aux désordres sanguins. Finalement, les objectifs de recherche et les méthodes utilisées seront exposés dans une dernière partie.

Tout au long du premier chapitre, une attention particulière sera portée sur les liens entre la machinerie transcriptionnelle et la régulation de l'érythropoïèse, et plus particulièrement sur la régulation de l'expression du gène de la β -globine. La β -globine est une protéine essentielle aux érythrocytes puisqu'elle est une composante de l'hémoglobine adulte. Constituée de deux sous-unités d' α -globine et de deux sous-unités de β -globine chez l'adulte, l'hémoglobine est une protéine tétramérique responsable du transport de l'oxygène dans les tissus.

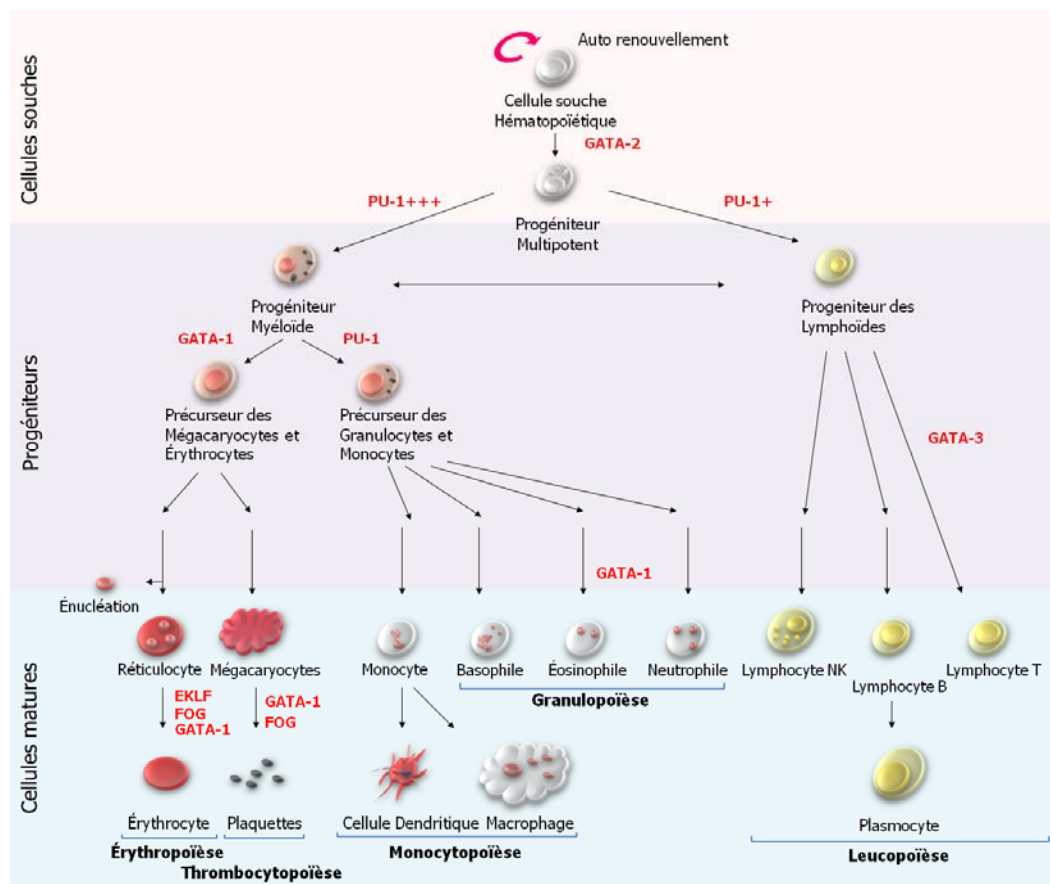


Figure 1. 1. L'hématopoïèse - L'hématopoïèse désigne le processus assurant le maintien des cellules souches hématopoïétiques, leur différenciation et leur maturation de manière à assurer le renouvellement de l'ensemble des lignées cellulaires sanguines. La différenciation des CSH en lymphocytes, granulocytes, monocytes, plaquettes (ou thrombocytes) et en globules rouges (ou érythrocytes) sont respectivement appelés : leucopoïèse, granulopoïèse, monocytopoïèse, thrombocytopoïèse, et érythropoïèse. Les facteurs de transcription qui seront discutés sont indiqués en rouge. Adapté de Orkin & Zon (2008) [2].

Chapitre I

Introduction

A - Transcription des ARN messagers chez les eucaryotes

1 - Formation du complexe de pré-initiation

Chez les eucaryotes, la transcription des différents types d'ARN est assurée par différentes ARN polymérases. Les ARN messagers, matrices de la production des protéines, sont transcrits par l'ARN polymérase II (ARN pol II). La régulation de la transcription de ces ARN se fait *via* l'action combinée de facteurs généraux de transcription, d'activateurs spécifiques et de co-activateurs. Le processus de transcription peut être divisé en trois étapes majeures :

- 1) Initiation : synthèse des premiers nucléotides,
- 2) Élongation : allongement de l'ARN messager,
- 3) Terminaison : relâchement de l'ARN messager et départ de l'ARN pol II.

Afin d'initier la transcription, l'ARN pol II a besoin de la présence des facteurs généraux de transcription (FGT) qui sont au nombre de 6 : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH (TFII pour **T**ranscription **F**actor de l'ARN polymérase **II**) (Figure 1. 2 en page 7). Le complexe ainsi formé est appelé complexe de pré-initiation ou PIC (**P**re-**I**nitiation **C**omplex). Les FGT et l'ARN pol II constituent la machinerie transcriptionnelle de base qui est la cible d'activateurs ou d'inhibiteurs régulant l'expression des gènes.

Selon le modèle d'assemblage séquentiel, la première étape de la formation du PIC est l'arrivée de TFIID (Figure 1. 2A) sur le promoteur [10, 11]. L'interaction de TFIID avec l'ADN se fait au niveau de sa sous-unité TBP (**T**ATA box-**B**inding **P**rotein) qui se lie à l'ADN de la boîte TATA. En plus de la TBP, le facteur TFIID contient des TAF (**T**BP-**A**ssociated **F**actor) [11]. Ceux-ci ont pour rôle de stabiliser l'interaction entre TFIID et le promoteur ainsi que d'interagir avec d'autres facteurs de transcription.

Une fois TFIID lié au promoteur, il mobilise les autres facteurs généraux de transcription en commençant par TFIIA (Figure 1. 2B). Ce facteur interagit avec la TBP, facilitant ainsi sa

liaison à la TATA-box et donc la formation du PIC. Le facteur TFIIB, impliqué dans la détermination du site d'initiation de la transcription, est ensuite incorporé. La polymérase associée à TFIIF (Figure 1. 2C) regagne alors au PIC en formation [12]. A son tour, TFIIF s'introduit dans le PIC et favorise l'enroulement de l'ADN autour de l'ARN pol II [13]. L'étape suivante est l'arrivée de TFIIE [14] qui se lie directement à la polymérase et qui est responsable du recrutement de TFIIH (Figure 1. 2D). Ce dernier est impliqué dans l'initiation et l'élongation de la transcription [15-17]. Il est l'unique facteur à posséder des activités enzymatiques : hélicase et kinase. Ces activités sont responsables de l'ouverture locale de la double hélice d'ADN au niveau du promoteur et de la phosphorylation du CTD (Carboxy-Terminal Domain) de l'ARN pol II. Lors de son arrivée sur un gène, le CTD de l'ARN pol II est hypophosphorylée puis est séquentiellement phosphorylé sur la serine 5 par la sous-unité cdk7 de TFIIH, et sur la serine 2 par la sous-unité cdk9 de P-TEFb [18]. La phosphorylation de la serine 5 par cdk7 semble être l'évènement clef de l'étape d'initiation puisqu'elle entraîne la dissociation du site d'initiation [19, 20] et déstabilise les interactions entre l'ARN pol II et les facteurs d'initiation (Figure 1. 2E) [21]. La phosphorylation de la serine 2 par cdk9, quant à elle, semble être nécessaire à la phase d'élongation de la transcription

Ce modèle d'assemblage du PIC est basé sur des études de reconstitution de la réaction de transcription *in vitro*. Il semble que le modèle de l'holoenzyme reflète mieux la réalité *in vivo*. Le modèle de l'holoenzyme suggère qu'une fois TFIID lié au promoteur, un complexe macromoléculaire comprenant l'ARN pol II, les FGT ainsi que des co-activateurs viendrait rejoindre le promoteur. La composition de l'holoenzyme pourrait être sujette à variation en fonction du lieu, de l'instant et des gènes à transcrire.

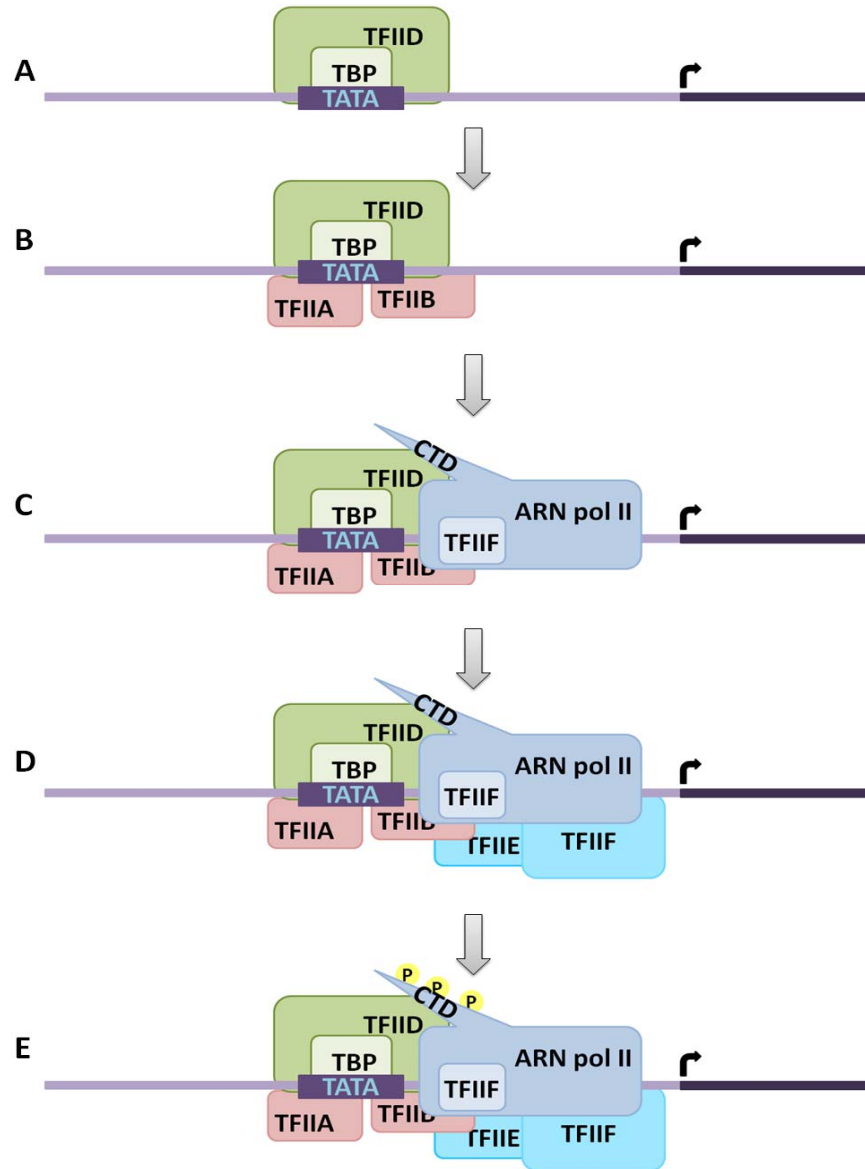


Figure 1. 2 . Représentation schématique de l'assemblage séquentiel du complexe de pré-initiation (PIC) au promoteur.

La phase d'initiation s'achève par le départ des facteurs d'initiation de la bulle de transcription (TFIIB, TFIIE, TFIIF ou encore TFIIH) et fait place à la phase d'élongation. Des facteurs d'élongation sont recrutés et l'ARN pol II s'échappe alors du promoteur pour transcrire le gène en aval et amorce l'addition successive de ribonucléotides. Des sites riches en adénines sur l'ADN sont reconnus comme étant des signaux de terminaison. La polymérase est alors relâchée et la transcription du gène terminée.

2 - Le facteur général de transcription TFIIH

Le facteur général de transcription TFIIH est connu pour être impliqué dans deux processus fondamentaux : la transcription ainsi que la réparation de l'ADN suite à des dommages (activité NER : Nucleotide Excision Repair) [22-25]. Le facteur TFIIH humain, ainsi que son homologue chez la levure, le facteur b, sont composés de dix sous-unités réparties en deux domaines :

- le cœur ou « core », composé des sous-unités XPB, p62, p52, p44, p34, et TTDA chez l'humain ; et respectivement de ses homologues chez la levure Ssl2, Tfb1, Tfb2, Ssl1, Tfb4, et Tfb5.
- le CAK (CDK-Activating Complex), composé de cdk7, cycline H et MAT1 chez l'humain ; et respectivement de ses homologues chez la levure Kin28, Ccl1 et Tfb3.

Le cœur et le domaine CAK interagissent au niveau de la sous-unité XPD, ou RAD3 chez la levure (Figure 1. 3 en page 9). Alors que les sous-unités XPB et XPD de TFIIH [26, 27] sont responsables de l'activité hélicase, l'activité kinase est portée par la sous-unité cdk7. [15, 16, 28, 29].

a - TFIIH et la régulation de l'expression de la β -globine

TFIIH joue un rôle non négligeable dans l'érythropoïèse. En effet, des mutations de sa sous-unité XPD ont été identifiées chez des patients atteints de β -thalassémie [30, 31]. Cette affection se caractérise par l'absence de production de β -globine. Deux chaînes de β -globine et

deux chaînes d' α -globine composent l'hémoglobine adulte. Bien que ces deux globines soient hautement exprimées durant la différenciation terminale des érythrocytes, l'expression de l' α -globine n'est pas altérée par les mutations de TFIIH. Par contre, plusieurs études ont démontré que TFIIH est un facteur important lors de la phase d'élongation de la transcription du gène de la β -globine [29]. TFIIH apparaît donc comme étant un élément indispensable à l'expression de cette protéine vitale.

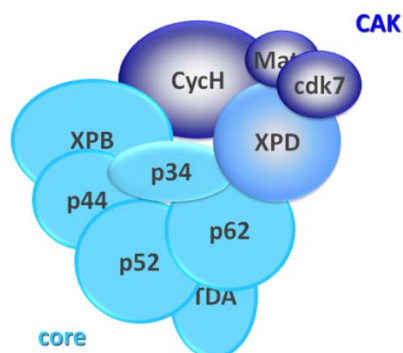


Figure 1. 3 . Représentation schématique du facteur général de transcription TFIIH. Les sous-unités composant le cœur sont représentées en turquoise, les sous-unités composant le CAK sont représentées en bleu foncé.

b - La sous-unité p62/Tfb1 de TFIIE

La sous-unité p62/Tfb1 (humain/levure) de TFIIE joue un rôle particulièrement important dans la régulation de la transcription. En effet, c'est en interagissant avec le domaine amino-terminal de p62/Tfb1 que TFIIE mobilise TFIIE dans le PIC [19, 20, 32]. De plus, p62/Tfb1 interagit directement avec les TAD acides d'un grand nombre d'activateurs dont notamment E2F-1, la sous-unité p53 de NF κ B, la protéine suppresseur de tumeur p53, ou encore la protéine virale VP16 du virus de l'herpes HSV (*herpes simplex virus*). VP16 est responsable de l'initiation de l'expression des gènes viraux dans les cellules nouvellement infectées par le HSV. Cette protéine régule la transcription *via* son interaction avec des facteurs de cellule hôte. [33-35]. Ces interactions corrélerent avec l'aptitude de ces activateurs à initier l'élongation de la transcription et dans de nombreux cas la région amino-terminale de p62/Tfb1 est requise pour cette interaction.

Plusieurs structures du domaine amino-terminal de p62/Tfb1, libres ou en complexe avec des partenaires, ont été résolues par spectroscopie de résonance magnétique nucléaire (RMN). La structure tertiaire adoptée par p62/Tfb1 étant homologue à la structure de la pleckstrine (protéine retrouvée dans les plaquettes) ce domaine est communément appelé domaine PH (PH : Pleckstrin Homology). Les domaines PH sont caractérisés par deux feuillets β antiparallèles perpendiculaires suivis d'une hélice α amphipatique qui vient se positionner sur les feuillets. Les boucles qui connectent les brins β sont quant à elles sujettes à variations [36].

Récemment, les structures de Tfb1 libre [37] et en complexe avec le TAD2 de p53 [38] ont été résolues dans notre laboratoire. Les études structurales visant à caractériser les interactions de p62/Tfb1 avec le TAD de p53 ont été menées sur Tfb1 plutôt que sur son homologue humain p62. En effet, Tfb1 est plus stable en solution que p62 et possède un domaine PH structurellement équivalent à celui de p62. Nos résultats montrent que les structures de Tfb1 sous forme libre et de Tfb1 en complexe avec p53 TAD2 sont très similaires. La liaison du TAD sur Tfb1 a lieu au niveau des brins β 5, β 6 et β 7 et de la boucle comprise entre les brins β 5 et β 6 [38] Figure 1. 4C & 1. 4D en page 12).

Le spectre RMN 2D ^1H - ^{15}N HSQC du TAD2 libre de p53 révèle l'absence de structures secondaires. En revanche, en complexe avec Tfb1, le TAD2 de p53 adopte une structure en hélice α amphipatique de neuf résidus. L'analyse structurale du complexe a mis en évidence une série d'interactions cation- π , sulfure- π et hydrophobes entre trois résidus hydrophobes de l'hélice de p53 (Ile⁵⁰, Trp⁵³ et Phe⁵⁴) et Tfb1. La chaîne latérale de la Phe⁵⁴ est alors positionnée dans une poche formée par les résidus Gln⁴⁹, Ala⁵⁰, Thr⁵¹, Met⁵⁹, Leu⁶⁰, et Arg⁶¹. Cette poche se situe entre les brins β 5 et β 6 de Tfb1. Le groupement aromatique de la Phe⁵⁴ participe ainsi à une interaction amino-aromatique avec la Gln⁴⁹ de Tfb1 et à une interaction cation- π avec l'Arg⁶¹ de Tfb1. Le groupement indole du Trp⁵³, inséré entre les chaînes latérales de la Met⁵⁹ et de la Lys⁵⁷ de Tfb1, participe à des interactions sulfure- π et cation- π avec ces deux résidus. L'Ile⁵⁰ est le troisième résidu impliqué dans la formation du complexe, il interagit avec Tfb1 *via* des contacts de van der Waals avec les méthionines en positions 59 et 88 de Tfb1. Par ailleurs, les résidus chargés positivement retrouvés à la surface de Tfb1 permettraient le positionnement de p53 en établissant des interactions électrostatiques avec les résidus Glu⁵¹ et Glu⁵⁶ de p53 [38].

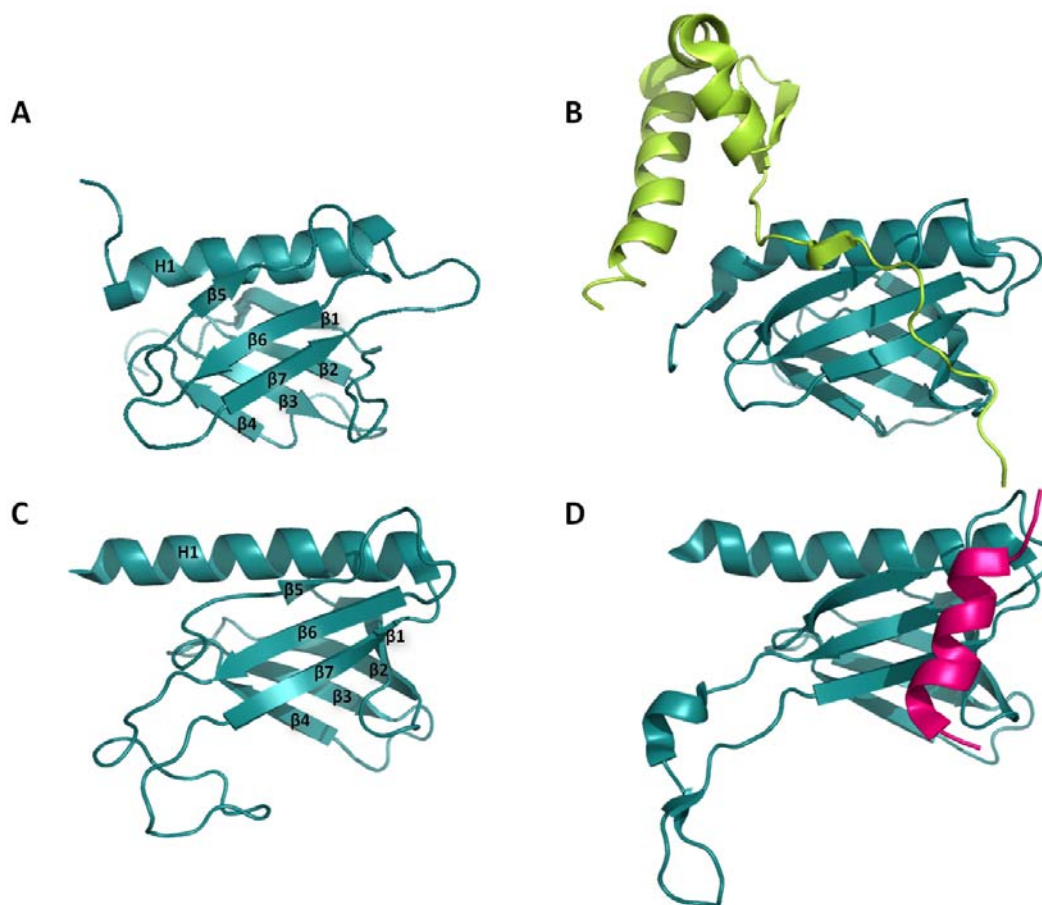


Figure 1. 4 . Structures du domaine PH de p62/Tfb1 libres ou en complexe. Structures résolues par spectroscopie RMN de (A) p62 libre [39], (B) p62 en complexe avec TFIIIE α (vert clair) [40], (C) Tfb1 libre [37] et (D) Tfb1 en complexe avec le TAD2 de p53 (rose) [38]. Les codes PDB sont respectivement : 1PFJ, 2RNR, 1Y50 et 2GS0. Les feuillets impliqués dans des interactions sont indiqués sur les structures de p62 et Tfb1 libres (A et C).

Récemment, plusieurs études structurales ont été menées sur le complexe formé par TFIIH et la sous-unité alpha de TFIIIE [40, 41]. Nécessaire à la transcription, cette interaction est responsable du recrutement de TFIIH dans le PIC et semble suffisante à l'induction d'un certain niveau de transcription [40, 41]. Ces travaux démontrent que TFIIH interagit avec TFIIIE α *via* sa sous-unité p62/Tfb1 et que TFIIIE α , le TAD de p53 ainsi que le TAD de VP16 partagent le même site de liaison sur p62 [37, 38, 40, 41]. La structure du complexe TFIIIE α /p62 révèle également que la région carboxy-terminale acide et hydrophobe de TFIIIE α (résidus 376-393) est en grande partie responsable de l'interaction avec Tfb1/p62. Le rôle crucial de deux résidus hydrophobes de TFIIIE α a été mis en évidence par des tests d'activité *in vivo* et des études de liaison *in vitro*. En effet, les mutations des résidus Phe³⁸⁷ et Val³⁹⁰ de TFIIIE α sont capables à elles seules d'inhiber l'interaction avec le domaine PH de p62. De fait, des mutations de ces résidus entraînent une baisse de l'activité transcriptionnelle *in vivo* [40, 41]. Finalement, tout comme la Phe⁴⁷ de p53 qui est positionnée dans une poche entre les brins β 5 et β 6 de Tfb1, la Phe³⁸⁷ de TFIIIE α vient se placer dans la poche équivalente formée entre les brins β 5 et β 6 de p62 [40].

Malgré ces similitudes, TFIIIE α utilise un mode de liaison entièrement différent de l'hélice amphipatique adoptée par l'activateur transcriptionnel p53. En fait, les structures de TFIIIE α libre ou en complexe sont similaires, excepté pour les résidus acides responsables de l'interaction. Sous la forme libre, ces acides aminés se retrouvent dans le segment amino-terminal non structuré de TFIIIE α . Cependant, suite à la formation du complexe, ils se stabilisent et forment un brin qui s'étend sur les brins β 5, β 6 et β 7 de p62 et établissent plusieurs contacts avec la boucle comprise entre les brins β 1 et β 2. Pour la première fois, on observe la présence de contacts de van der Waals et d'interactions hydrophobes entre TFIIIE α et l'extrémité carboxy-terminale de l'hélice H1 de p62 [40]. Il en résulte que la surface d'interaction entre ces deux partenaires s'avère plus importante que dans le cas de l'interaction Tfb1/p53 TAD2. TFIIIE α n'agit donc pas comme un activateur, son mode de liaison se distingue du mode de liaison canonique des activateurs transcriptionnels pour lesquels une région intrinsèquement désordonnée s'ordonne en hélice- α lors de l'interaction avec sa cible.

3 - Activateurs et co-activateurs de la transcription

La machinerie transcriptionnelle de base, décrite précédemment, est essentielle à la survie et à la différenciation d'une cellule. Cependant, le choix des gènes à exprimer en temps et en lieu appropriés se fait grâce à d'autres protéines : les activateurs de la transcription. Ces protéines ont la capacité de lier des séquences d'ADN spécifiques et de recruter des co-activateurs. Les co-activateurs, eux, ne sont pas prédisposés à lier l'ADN mais sont des composants essentiels de la modulation de la transcription.

a - Les activateurs

Tel que mentionné précédemment, les activateurs transcriptionnels comprennent deux domaines fonctionnels : un domaine de liaison à l'ADN (DBD) et un domaine d'activation de la transcription (TAD). Les DBD permettent aux activateurs de lier une séquence spécifique d'ADN au niveau du promoteur. Plusieurs motifs structuraux sont connus pour lier l'ADN : les domaines à doigts de zinc, les domaines hélice-tour-hélice ou encore les glissières à leucine. Comme leur nom l'indique, les TAD ont pour fonction de stimuler la transcription *via* leur interaction avec des partenaires protéiques. À ce jour, divers partenaires des TAD ont été identifiés *in vivo* et *in vitro*. Parmi ceux-ci on retrouve des facteurs généraux de la transcription (TBP, TFIIB, TAF31, TAF80, TAF9, TFIIF) [42-45], des composants du complexe médiateur (MED15, MED25) [46, 47], des histones acétyltransferases (CBP/p300) [48, 49] et des composants du complexe ATP-dépendent de remodelage du nucléosome (SWI/SNF : **SW**Itching mating type/**S**ucrose **N**on **F**ermenting) [50-52].

Les TAD sont traditionnellement classés en fonction de leur composition en acides aminés [53, 54]. On distingue les domaines riches en résidus acides, riches en glutamine, riches en proline, et enfin riches en sérine et thréonine. La classe des TAD acides se distingue des autres puisqu'elle est présente dans plusieurs oncogènes et virus cancérigènes tels que la protéine suppresseur de tumeur humaine p53, VP16, E2-F1, ou encore la sous-unité p65 de NFκB [35, 55, 56]. Comme leur nom l'indique, ces TAD comportent des séquences répétitives de résidus acides.

Étant donné leur composition particulière, la structure de ces TAD est connue pour être désordonnée sous forme libre. Les protéines possédant des domaines non structurés sont très répandues. En effet, plus de 30 % du génome des eucaryotes code pour des régions protéiques désordonnées de longueur supérieure à trente résidus [57, 58]. Ces domaines offrent des avantages fonctionnels tels que la possibilité d'interagir avec de multiples partenaires, ou encore facilitent l'accès à des modifications post-translotionnelles permettant ainsi de réguler leur activité. Habituellement, les TAD non structurés à l'état libre se structurent lorsqu'ils interagissent avec leurs cibles. Il est intéressant de constater que la majorité des protéines possédant des domaines intrinsèquement désordonnés sont impliquées dans le processus de transcription et que, majoritairement, plus une interaction entre deux partenaires est forte, plus cette interaction tolère un état natif désordonné [59].

Comme cela a été mentionné dans la section 2 -b -, plusieurs études RMN sur les TAD acides ont confirmé que ces domaines ne sont pas structurés lorsqu'ils sont libres, tandis qu'ils adoptent une structure secondaire ordonnée lorsqu'ils sont complexés à leurs partenaires. Le TAD2 de p53 en interaction avec Tfb1/p62 (Figure 1. 4 en page 12), les TAD1 et TAD2 de p53 en complexe avec MDM2 [60] et RPA70 [61] respectivement, le TAD de VP16 en interaction avec PC4 (Positive Cofactor 4) et TFIIB [62] ainsi que le TADC de VP16 en interaction avec TAF31 [63] sont quelques exemples de TAD acides qui adoptent une structure hélicale en interaction avec leur cible. Les mécanismes par lesquels les TAD interagissent avec leurs multiples partenaires sont néanmoins très peu connus dû, entre autres, au faible nombre de structures résolues à ce jour.

b - Les co-activateurs

Ne possédant pas la capacité de se lier à l'ADN, les co-activateurs de la transcription sont recrutés au promoteur par les activateurs pour y exercer diverses fonctions. On distingue deux classes de co-activateurs :

- *les co-activateurs interagissant avec la machinerie transcriptionnelle de base* tels que les complexes médiateurs qui recrutent et ajustent les activités de la machinerie transcriptionnelle,
- *les co-activateurs ayant un rôle dans la modification de la chromatine* tels que les complexes HAT (**H**istones **A**cétyl **T**ransférases) qui possèdent des activités enzymatiques modifiant les histones, ou le complexe SWI/SNF qui est capable de déplacer les nucléosomes [64].

i - Complexes médiateurs

Les gros complexes protéiques que sont les complexes médiateurs sont des adaptateurs entre les activateurs transcriptionnels spécifiques et la machinerie transcriptionnelle de base [65]. Différents complexes médiateurs ont été identifiés chez l'Homme tels que TRAP/SMCC (**T**hyroid **H**ormone **R**eceptor-**A**ssociated **P**roteins/**S**RB/**M**ediator **C**ontaining **C**omplex, CRSP/PC2 (**C**ofactor **R**equired for **S**p1/ **P**ositive **C**omponent **2**) ou encore ARC/DRIP (**A**ctivator-**R**ecruited **C**o-factor /vitamin**D**-**R**eceptor **I**nteracting **P**roteins) [66]. Ils diffèrent entre eux par le nombre et la nature de leurs composants. La diversité de leurs assemblages serait due au fait que les différents complexes médiateurs seraient spécifiques à certains activateurs ou inhibiteurs de la transcription [67].

Outre leur rôle dans le recrutement de la machinerie transcriptionnelle, les complexes médiateurs possèdent des sous-unités capables d'activer ou de réprimer l'activité transcriptionnelle tel que la kinase cdk8 qui est capable de phosphoryler le CTD de l'ARN pol II [68].

ii - Histones acétyltransférases

L'activation de la transcription de gènes est généralement associée à une hyper-acétylation des histones, ce qui permet une relaxation de la structure de la chromatine dans une région définie et facilite ainsi l'accès à l'ADN. Certains co-activateurs possèdent une activité enzymatique HAT permettant de faciliter l'accès au promoteur. Les protéines CBP (**C**REB (**c**-

AMP Response Element-Binding protein)-Binding Protein) et p300 sont des exemples de co-activateurs ayant une activité HAT [69, 70]. CBP/p300 sont des HAT homologues possédant des similarités de fonction et de structure [71]. Ces protéines sont largement exprimées et interagissent directement avec des FGT tels que TFIIB [72], la sous-unité TBP de TFIID [73] ainsi que l'ARN pol II [74], suggérant que ces protéines pourraient connecter la machinerie de base avec des activateurs spécifiques. Elles sont d'ailleurs considérées comme des régulateurs généraux de la transcription.

De plus, CBP et p300 interagissent et régulent l'activité d'un grand nombre de facteurs de transcription. Quatre domaines de CBP/p300 ont été identifiés comme étant impliqués dans des interactions avec des TAD d'activateurs : TAZ1, TAZ2 (Transcription Adaptator putative Zinc finger), KIX (Kinase-induced Interacting Domain) [75] et IBiD (IRB3 binding Domain) [76]. À ce jour, plusieurs structures tridimensionnelles impliquant des TAD et divers domaines de CBP/p300 ont été résolues par spectroscopie RMN (Figure 1. 5 en page 19).

Les structures du domaine KIX en complexe avec le TAD du facteur hématopoïétique c-Myb (Figure 1. 5A) [77] et en complexe avec le domaine pKID (Phosphorylated Kinase-Inducible Domain) du TAD de CREB (Figure 1. 5B) [78, 79] révèlent que ces deux protéines lient de façons différentes le même partenaire. Le facteur c-Myb adopte une structure en hélice- α courbée, tandis que CREB forme deux hélices- α (α A et α B). Malgré ces différences, c-Myb et l'hélice- α de CBP se lient à KIX dans le même sillon hydrophobe. Dans les deux cas, une leucine (Leu³⁰² de c-Myb et Leu¹⁴¹ de CBP) se retrouve à la même position dans le sillon hydrophobe de KIX et joue un rôle crucial pour l'interaction [77]. Par ailleurs, des études de cartographie d'interactions par RMN impliquant KIX et des TAD ont démontré que le site de liaison du TAD de MLL (Mixed Lineage Leukemia protein) est différent de celui des TAD de c-Myb et de CREB. De la même façon, il a été démontré que le TAD de c-Jun partage le même site de liaison sur KIX que le TAD de MLL [79]. Ces travaux montrent que différents domaines d'activation peuvent simultanément lier le domaine KIX de CBP (Figure 1. 5B) [79], illustrant le recrutement combinatoire de CBP par des activateurs multiples par l'intermédiaire du domaine de KIX. D'autre part, la cartographie de l'interaction du TAD de p53 sur KIX a mis en évidence que les

TAD1 et TAD2 de p53 lient respectivement KIX sur les sites de liaison de MLL/c-Jun et CREB/c-Myb [80].

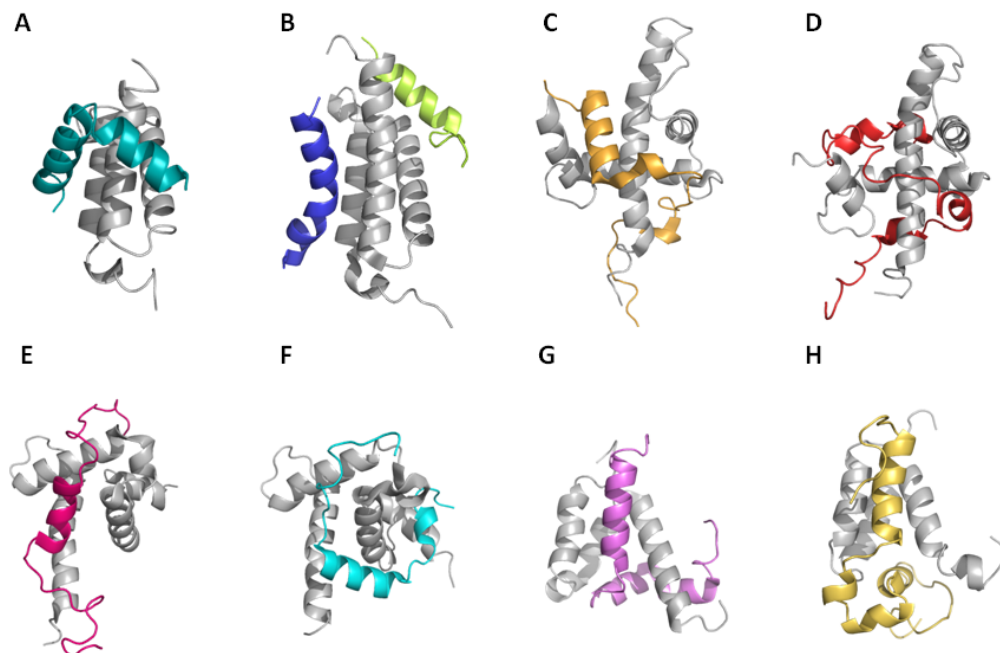


Figure 1. 5 . Structures de différents domaines de CBP/p300 en complexe avec différents TAD de facteurs de transcription. Les différents domaines de CBP/p300 sont représentés en gris, chaque TAD est représenté dans une couleur différente. Structures du domaine KIX en complexe avec (A) le TAD de c-Myb et en complexe avec (B) le TAD de CREB (bleu) et le TAD de MLL (vert). Structures du domaine TAZ1 en complexe avec (C) le TAD de CITED2 et en complexe avec (D) le TAD de STAT2. Structures du domaine TAZ2 en complexe avec (E) le TAD de p53 (F) et le TAD de STAT1. Structures du domaine IBiD en complexe avec (G) ACTR et avec (H) le TAD de SRC1. Les codes PDB sont respectivement : 1KDX, 2AGH, 1R8U, 2KA4, 2K8F, 2KA6, 1KBH et 2C52 [79, 81-85].

Les structures du domaine TAZ1 en complexes avec divers TAD acides ont permis de comparer les modes de liaison à ce domaine. Les structures des TAD de CITED2 (**CBP/p300-Interacting Transactivator ED-rich 2**) (Figure 1. 5C) [84], STAT2 (**S**ignal **T**ransducer and **A**ctivator of **T**ranscription **2**) (Figure 1. 5D) [81] et HIF-1 α (**H**ypoxia-**I**nduced **F**actor **1**) [86] occupent tous le même sillon sur TAZ1, s'enroulant autour de celui-ci. De leur côté, les TAD de CITED2 et de HIF-1 α adoptent tous deux des structures secondaires en hélices- α bien définies [84, 86]. Par contre, pour la première fois, on observe une conformation majoritairement étendue dans le cas du TAD de STAT2 (**S**ignal **T**ransducer and **A**ctivator of **T**ranscription **2**) [81].

D'autre part, les structures des TAD acides de STAT1 (**S**ignal **T**ransducer and **A**ctivator of **T**ranscription **1**) (Figure 1. 5E) [81], de p53 (Figure 1. 5F) [85] et de la protéine virale E1A [87] ont été résolues en complexes avec TAZ2. Ces divers travaux démontrent que ces TAD lient TAZ2 sur le même site de liaison, au niveau du croisement des trois hélices amino-terminales. Cependant, ces TAD adoptent des structures en hélices- α de longueur et d'orientation différentes. Il est aussi intéressant de constater que les structures adoptées par le TAD1 de p53 en interaction avec TAZ2 et avec MDM2 sont quasiment identiques. Effectivement, la liaison du TAD1 de p53 à MDM2 induit la formation d'une hélice- α allant des résidus 18 à 25, et sa liaison à TAZ2 induit une hélice- α s'étendant des résidus 16 à 25. Dans les deux cas, deux résidus aromatiques, Phe¹⁹ et Trp²², jouent un rôle majeur dans l'interaction [60, 85].

Finalement, deux structures du domaine IBiD en complexe avec des TAD de protéines membres de la famille p160 ont été résolues. Les protéines de cette famille sont des co-activateurs des récepteurs nucléaires capables d'activer la transcription grâce, en partie, à leur capacité à recruter CBP/p300. Les structures des TAD d'ACTR (**A**ctivator of **T**hyroid **R**eceptor) [88] et de SRC1 (**S**teroid **R**eceptor **C**oactivator **1**) [82] liés au domaine IBiD montrent que ces domaines adoptent eux aussi une structure hélicale. De façon similaire, l'hélice amino-terminale de ces TAD se lie dans le sillon hydrophobe entre les hélices H1 et H2 d'IBiD tandis que le mode de liaison de leurs domaines carboxy-terminaux diffère de part leur structure secondaire et de par leur positionnement sur IBiD (Figure 1. 5G et 5H) [82].

Il apparaît que, même s'ils partagent généralement les mêmes sites de liaison sur des partenaires communs et s'ils semblent former majoritairement des structures en hélice- α , différents TAD adoptent des modes de liaison distincts avec une même protéine cible. Étonnamment, la structure nouvellement résolue du domaine TAZ2 de CBP en complexe avec le TAD acide de STAT1 est la première structure à démontrer qu'un TAD peut interagir spécifiquement avec sa cible sous une forme majoritairement non structurée [81].

Depuis plusieurs années, beaucoup d'effort ont été fournis pour comprendre l'interaction entre les co-activateurs généraux CBP/p300 et le TAD de la protéine suppresseur de tumeur p53 [89]. L'activité apoptotique de p53 est régulée, entre autre, *via* son interaction avec l'ubiquitine-ligase MDM2 ainsi qu'avec CBP/p300. De nombreuses études ont mis en évidence une interaction directe entre le TAD de p53 et les quatre domaines KIX, I β DiD, TAZ1 et TAZ2 [90, 91]. Il a été suggéré que MDM2/p53 et CBP/p300 sont capables de former un complexe ternaire. Selon cette hypothèse, p53, sous forme tétramérique, interagirait avec quatre molécules de MDM2 *via* ses TAD1 et avec quatre domaines de CBP/p300 *via* ses TAD2. La formation de ce complexe entraînerait l'ubiquitination de p53 et sa dégradation. La phosphorylation de la Ser¹⁸ du TAD1 de p53 inhibe son interaction avec MDM2 et parallèlement augmente l'affinité de ce TAD pour les divers domaines de CBP/p300, déclenchant ainsi l'acétylation et donc l'activation de p53 [90].

Par ailleurs, CBP et p300 interagissent directement et acétylent divers facteurs de transcription hématopoïétiques (GATA-1 [92, 93], GATA-2 [94], PU.1 [95], EKLF [49], NF-E2 [96]) impliqués virtuellement dans la formation de toutes les lignées sanguines. Cette modification post-transcriptionnelle corrèle avec l'activation de l'expression des gènes cibles de ces activateurs. En outre, l'absence de CBP conduit à une diminution du nombre de cellules sanguines progénitrices, suggérant un rôle dans la prolifération ou la survie de ces cellules [75]. De ce fait, CBP/p300 se révèlent être des co-activateurs indispensables à l'hématopoïèse puisqu'ils sont requis pour l'auto-renouvellement des CSH et pour la différenciation des cellules sanguines. Malgré de nombreux travaux sur les mécanismes d'interaction de CBP/p300 avec

leurs partenaires, des études additionnelles sont indispensables afin de caractériser les interactions spécifiques impliquant CBP/p300 et les facteurs de transcription érythropoïétiques.

B - Facteurs de transcription impliqués dans la différenciation des érythrocytes

1 - Les érythrocytes

L'érythropoïèse définit l'ensemble des processus de production des érythrocytes dans la moelle osseuse à partir de cellules souches hématopoïétiques multipotentes. Ce processus peut être stimulé par un facteur de croissance, l'érythropoïétine. Chez l'Homme la durée de vie des globules rouges est comprise entre 110 et 120 jours. L'érythropoïèse comporte différents stades de différenciation en commençant par le stade proérythroblaste, auquel font suite l'érythroblaste basophile, l'érythroblaste polychromatophile, l'érythroblaste acidophile, puis, suite à la perte du noyau et des organelles (telles que les mitochondries), vient le stade réticulocyte. Ces cellules possèdent encore quelques ribosomes et quelques mitochondries. De plus, elles ont la capacité de quitter la moelle osseuse pour être libérées dans le sang circulant. Suite à la perte des ribosomes et des mitochondries restantes, le réticulocyte devient un érythrocyte. Le rôle de l'énucléation des érythrocytes serait de libérer de l'espace pour laisser place à la grande quantité d'hémoglobine présente dans ces cellules.

Les globules rouges ont un rôle capital dans l'organisme puisqu'ils sont chargés de transporter l'oxygène des poumons vers les tissus. En effet, l'hémoglobine présente en grande quantité dans ces cellules se combine à l'oxygène au niveau des poumons pour former de l'oxyhémoglobine. Cette combinaison permet de véhiculer le dioxygène jusqu'au niveau des tissus où l'oxyhémoglobine sera transformée en carboxyhémoglobine. Les globules rouges ont une forme de disque biconcave particulière. Cette morphologie, apte à la déformation, leur permet de s'adapter à la taille variable des vaisseaux qu'ils parcourent et facilite les échanges

transmembranaires. Le vieillissement de ces cellules s'accompagne de changements au niveau de leurs membranes plasmiques. Ces modifications sont identifiées par des phagocytes qui vont induire la phagocytose au niveau de la rate.

Depuis plusieurs années, des études ont permis de mieux comprendre l'hématopoïèse. De nombreux activateurs de la transcription régulant la différenciation des CSH en différentes lignées ont été identifiés et étudiés. Les fonctions régentées par ces divers facteurs ne peuvent s'expliquer par une action indépendante de ceux-ci. Les connaissances actuelles confluent vers le concept d'un contrôle transcriptionnel dépendant de divers facteurs membres de complexes multi-protéiques, dont la composition varie en fonction du temps et du lieu, et où des modifications post-translationnelles ajoutent un niveau de régulation supplémentaire.

Les travaux présentés dans cette thèse portent essentiellement sur les facteurs de transcription érythrocytaires GATA-1 et EKLF. L'état des connaissances actuelles sur le mécanisme de la régulation de l'expression des gènes par ces activateurs transcriptionnels sera présenté dans cette section. Nous allons aborder le rôle d'EKLF en mettant tout particulièrement l'accent sur son rôle dans la régulation du gène de la β -globine. Mais auparavant, le rôle de GATA-1 dans la différenciation des globules rouge et dans le maintien en vie de ces cellules sera discuté.

2 - Le Facteur de Transcription GATA-1

a - Domaines de GATA-1

Le facteur de transcription GATA-1 est le membre fondateur de la famille GATA. Ce groupe de facteurs de transcription doit son nom à sa capacité à lier une séquence d'ADN consensus A/TGATAA/G [97, 98]. Cette famille peut être divisée en deux sous-familles, hématopoïétique et non-hématopoïétique. GATA-1, GATA-2 et GATA-3 sont exprimés principalement lors de l'hématopoïèse. Plus particulièrement, GATA-1 est crucial pour la formation des globules rouges, GATA-2 jouant quant à lui un rôle dans l'auto-renouvellement des cellules progénitrices et GATA-3 étant impliqué dans la différenciation des lymphocytes.

GATA-4, GATA-5 et GATA-6 sont exprimés lors du développement de divers tissus tels que les intestins, les poumons et le cœur.

Le site de liaison à l'ADN de ces facteurs est hautement conservé, il est composé de deux domaines adjacents de liaison du zinc de type Cys-X₂-Cys-X₁₇-Cys-X₂-Cys [99]. Ces deux domaines se distinguent par leur position amino- et carboxy-terminale (NF et CF respectivement). Le domaine CF est essentiel à l'activité de GATA-1 puisqu'il est responsable de sa liaison au site d'ADN consensus [100]. Adjacent au CF, se trouve le bras basique BA (**B**asic **A**rm). La structure de GATA-1_{CF+BA} en complexe avec l'ADN a révélé le rôle déterminant du BA lors de l'interaction avec l'ADN [101] (Figure 1. 6 en page 25). Néanmoins, le CF est aussi impliqué dans des interactions avec des partenaires protéiques comme les facteurs de transcriptions PU.1, EKLF, CBP ou avec lui-même [102-106]. Le domaine NF, quant à lui, est capable de lier l'ADN GATA avec une plus faible affinité que le CF, tout en contribuant à la stabilisation de la liaison sur des sites d'ADN comprenant plusieurs motifs GATA [107]. Il est lui aussi connu pour son implication dans des interactions protéine/protéine en particulier avec la protéine FOG1 (**F**riend-**O**f-**G**ATA).

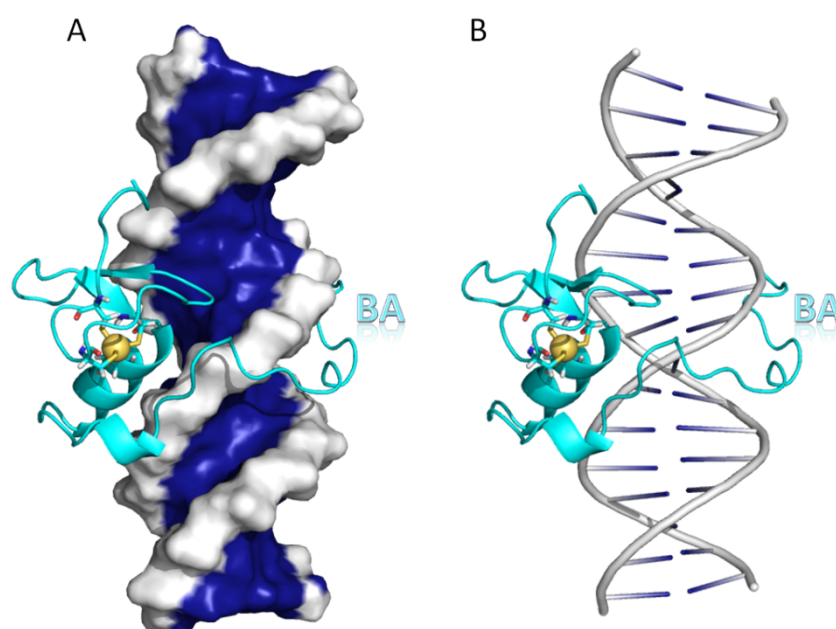


Figure 1. 6. Structure de GATA-1_{CF+BA} lié à l'ADN. GATA-1_{CF+BA} est représenté en ruban turquoise, l'ADN GATA est représenté successivement en surface (A) et en ruban (B). Le bras basique est annoté BA. Code PDB : 2GAT [101].

b - Rôles de GATA-1

Une quantité suffisante de globules rouges est requise afin de transporter l'oxygène au niveau des tissus, mais un trop grand nombre de ces cellules peut causer une hypertension artérielle. Ainsi, la formation des globules doit aussi bien être régulée positivement que négativement. Il s'avère que le facteur de transcription GATA-1 joue un rôle important dans les mécanismes de survie, de mort et de maturation des érythroblastes. En effet, ce facteur de transcription contrôle la production et la survie de ces cellules en régulant l'arrêt du cycle cellulaire et en bloquant l'apoptose dans les cellules progénitrices. Des études de différenciation *in vitro* ont confirmé l'arrêt du développement des cellules au stade pré-érythroblaste en l'absence de GATA-1 et ont montré que ce précurseur meurt par apoptose [108]. De plus, GATA-1 est impliqué dans le contrôle négatif de la production de globules rouges. Bien que ce mécanisme soit encore méconnu, il a été montré que l'activation des récepteurs de mort Fas par leur ligand (FasL) conduit au clivage sélectif du facteur de transcription GATA-1 par la caspase-8, conduisant ainsi à la mort de ces cellules [109]. Il semblerait ainsi que les décisions de vie, de mort et de maturation convergent tous vers une même cible : GATA-1.

Le facteur de transcription GATA-1 est exprimé dans plusieurs lignées cellulaires sanguines (mastocytes, éosinophiles, mégacaryocytes, érythrocytes) avec toutefois un très haut niveau d'expression dans les globules rouges [110]. GATA-1 peut tout aussi bien jouer le rôle d'activateur que de répresseur de la transcription afin de diriger la différenciation des progéniteurs vers la lignée érythrocytaire [105, 111-114]. Il est en outre capable de s'autoréguler. Chez des souris déficientes en GATA-1, les érythrocytes sont les seules cellules à ne pas se développer normalement et l'embryon meurt d'une sévère anémie au bout de 10.5 à 11.5 jours [115]. Parallèlement, la surexpression de GATA-1 dans des progéniteurs myéloïdes ou lymphoïdes a démontré que GATA-1 induit l'engagement vers la voie érythrocytaire et inhibe donc la différenciation vers les autres lignées [116].

L'importance de GATA-1 dans l'érythropoïèse vient en partie du fait que ce facteur active la production des protéines essentielles aux globules rouges [117, 118], tel que l'hémoglobine adulte [119]. L'expression de la β - et de l' α -globine, protéines composant l'hémoglobine, est régulée par GATA-1 [103, 119-121]. Un rapprochement du promoteur distal et du promoteur proximal du gène de la β -globine est requis pour l'activation de ce gène. GATA-1, *via* sa liaison sur la région locus contrôle (LCR) du promoteur distal, est responsable de ce rapprochement [119, 122] (Figure 1. 7 en page 28 et Figure 1. 10 en page 34). Par ailleurs, l'expression de l' α -spectrine [123] (protéine structurale responsable de la forme concave des globules rouges), du récepteur de l'érythropoïétine (responsable de l'orientation vers la production de globules rouges) [124], ou encore de plusieurs enzymes impliquées dans la synthèse de l'hème de l'hémoglobine sont sous le contrôle transcriptionnel de GATA-1 [125-127].

L'inhibition des facteurs de transcription PU.1 et GATA-2 sont deux exemples illustrant la répression de gènes par GATA-1. Il a été mentionné auparavant que GATA-2 est indispensable au renouvellement des précurseurs, mais lorsque la différenciation progresse ce facteur est inhibé puis est graduellement remplacé par GATA-1 [119]. L'absence de GATA-1 induit la surexpression de GATA-2, suggérant que GATA-1 réprime l'expression de GATA-2. Il a en effet été démontré que la liaison de GATA-1 sur une séquence d'ADN en amont du promoteur proximal de GATA-2 bloque le recrutement de l'ARN pol II, entraînant ainsi l'inhibition de l'expression de GATA-2 [111].

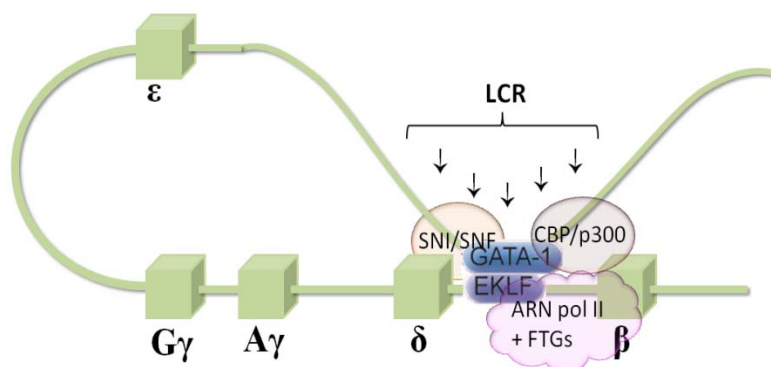


Figure 1. 7 . Structure et expression de la β -globine sur le gène globine « β -like ». Rapprochement du LCR à proximité du promoteur proximal du gène de la β -globine grâce aux facteurs de transcription spécifiques et aux co-activateurs, favorisant ainsi la formation d'une boucle d'ADN.

c - Interactions protéine/protéine

La protéine GATA-1 est non seulement capable de lier l'ADN, mais est également impliquée dans des interactions protéine/protéine qui modulent son activité (Figure 1. 8 en page 29). Il a été montré que, durant l'érythropoïèse et la thrombocytopoïèse, l'interaction entre GATA-1 et FOG-1 est nécessaire à l'activation de certains gènes [128-131] La synergie de ces deux activateurs permet de moduler l'activité de GATA-1 en fonction du promoteur [132-135].

GATA-1 est aussi connue pour interagir avec les facteurs de transcription de la famille Krüppel : EKLF et Sp1 [104, 121]. Sur plusieurs promoteurs, les séquences consensus d'ADN de GATA-1 et des membres de la famille Krüppel se retrouvent à proximité suggérant une coopérativité entre ces protéines [136]. En effet, l'interaction entre les facteurs GATA-1 et EKLF

a été mise en évidence et semble être directement corrélée à l'augmentation de l'activité transcriptionnelle de GATA-1 [104]. Récemment, des études ont révélé le rôle indiscutable d'EKLF dans la formation d'un complexe comprenant le LCR ainsi que le promoteur proximal du gène de la β -globine [103, 120]. L'interaction entre EKLF et GATA-1 semble être déterminante afin de promouvoir le rapprochement des séquences activatrices et des LCRs à proximité des promoteurs, favorisant ainsi la formation de boucles d'ADN (Figure 1. 7 en page 28).

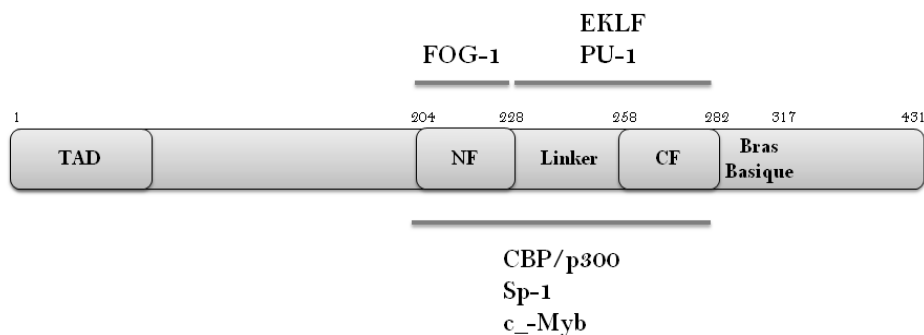


Figure 1. 8 . Domaines fonctionnels de GATA-1. Diagramme montrant la localisation sur le DBD (NF + linker + CF) de GATA-1 d'interactions essentielles au développement hématopoïétique. Les zones d'interaction sont représentées par les barres grises.

Une des étapes fondamentales de l'érythropoïèse est le point où le précurseur des myéloïdes décide de s'engager dans la voie des précurseurs des mégacaryocytes et des érythrocytes, soit des précurseurs des granulocytes et des monocytes. La différenciation vers une lignée particulière dépend de la quantité de GATA-1 et de PU.1. Lors de cette étape, l'expression de GATA-1 et de PU.1 est mutuellement exclusive, ces deux protéines interagissent et sont antagonistes l'une de l'autre [105, 137]. GATA-1 inhibe la différenciation vers la lignée myéloïde en empêchant l'interaction de PU.1 avec son partenaire c-Jun [105, 112]. Quant à PU.1, il inhibe l'érythropoïèse *via* son interaction avec GATA-1 au niveau des gènes cibles de celui-ci. En coopération avec pRB, PU.1 recrute alors un complexe de répression au promoteur [138-141].

En plus de réguler ses gènes cibles en coopérant avec divers facteurs de transcription, GATA-1 est la cible de co-activateurs capables de contrôler son activité, tels CBP/p300 [142]. L'interaction entre ces partenaires entraîne l'acétylation de GATA-1 au niveau de son DBD. Cette modification s'accompagne de l'augmentation de l'activité transactivationnelle de GATA-1 et semble être importante *in vivo* puisque des mutations des résidus acétylés altèrent la capacité de GATA-1 à rétablir la différenciation en érythrocytes d'une lignée cellulaire déficiente en GATA-1 [143].

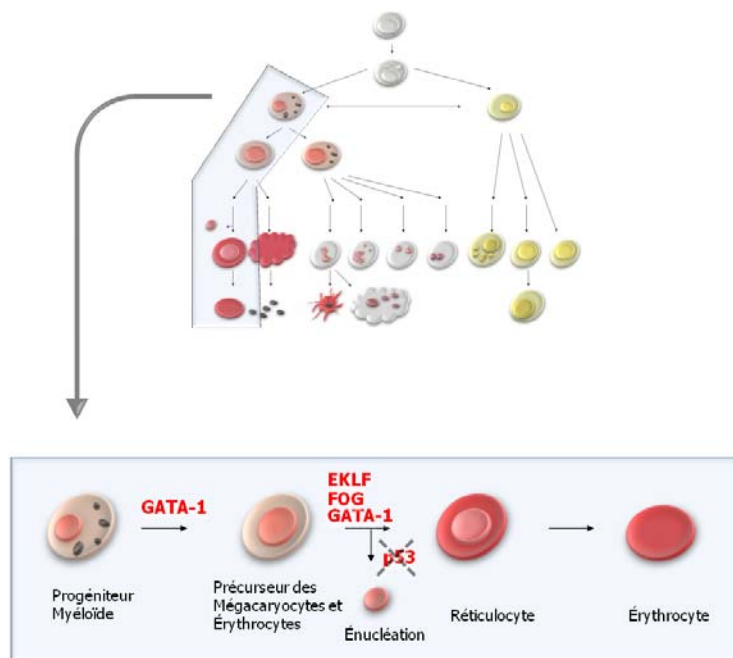


Figure 1. 9 . Érythropoïèse et énucléation. Différentes étapes de la différenciation en érythrocytes et facteurs de transcription impliqués.

d - Maintien en vie des précurseurs des érythrocytes

Durant la différenciation, la mort cellulaire doit absolument être réprimée afin d'assurer le renouvellement des cellules. La protéine p53, qualifiée de « protéine suppresseur de tumeur », est un facteur de transcription dont l'activation conduit à l'arrêt de la croissance cellulaire et à l'apoptose. La protéine p53 est activée par des dommages à l'ADN ainsi que par des instabilités chromosomiques afin d'éviter la formation de cancers [144]. Des changements morphologiques caractérisent la différenciation terminale des érythrocytes. En effet, la condensation de la chromatine, qui s'accompagne de dommages à l'ADN, précède l'expulsion du noyau (énucléation) (Figure 1. 9 en page 31). Ces modifications sont analogues à celles qui déclenchent l'apoptose dans les cellules non-sanguines et devraient, par conséquent, induire la mort programmée des cellules *via* l'activation de la protéine p53. Pourtant les cellules ne meurent pas et p53 demeure inactive durant l'érythropoïèse.

Le rôle primordial de GATA-1 dans le maintien en vie des globules rouges a été très tôt mis en évidence. Des études effectuées sur des proérythroblastes n'exprimant pas GATA-1 ont montré que les cellules entraient en apoptose sans accumulation de p53 [108], montrant ainsi que GATA-1, en plus de réguler les gènes érythrocytaires, est garant du maintien en vie de ces cellules. Cependant, le mécanisme par lequel GATA-1 réprime l'apoptose reste très peu connu. Plusieurs mécanismes impliquant le facteur GATA-1 et inhibant la mort des érythroblastes ont été identifiés.

Entre autres, le contrôle de l'apoptose par GATA-1 passe par l'induction de l'expression d'un membre de la famille Bcl-2, Bcl-x_L un variant d'épissage de Bcl-x, inhibiteur de l'apoptose p53-dépendante [145]. La délétion de Bcl-x_L induit un blocage de la maturation et une induction de l'apoptose des érythroblastes [146]. Curieusement, ces observations s'avèrent être analogues aux observations faites en l'absence de GATA-1. En outre, un fait intéressant a été observé au cours de l'énucléation : la caspase-3, protéase effectrice de l'apoptose, clive sélectivement les protéines impliquées dans le maintien de l'intégrité du noyau (lamine B) et de la chromatine (acinus) sans pour autant induire le clivage de GATA-1 ni la mort cellulaire. C'est la protéine

Hsp70 qui, au cours de la différenciation terminale, protège spécifiquement GATA-1 de la protéolyse médiée par la caspase-3, inhibant ainsi l'apoptose [109, 147, 148].

Par ailleurs, Mdm2 et Mdm4, inhibiteurs connus de p53, sont des acteurs importants du maintien en vie des cellules au cours de l'érythropoïèse. Il a été montré que Mdm2 est requis pour inhiber p53 lors de l'érythropoïèse primitive (chez l'embryon), alors que Mdm4 est requis lors de l'érythropoïèse définitive adulte, démontrant que p53 est présent mais inhibé dans les érythrocytes [149]. Malgré les nombreuses études effectuées, le mécanisme par lequel les érythrocytes sont maintenus en vie au cours de la différenciation terminale reste énigmatique.

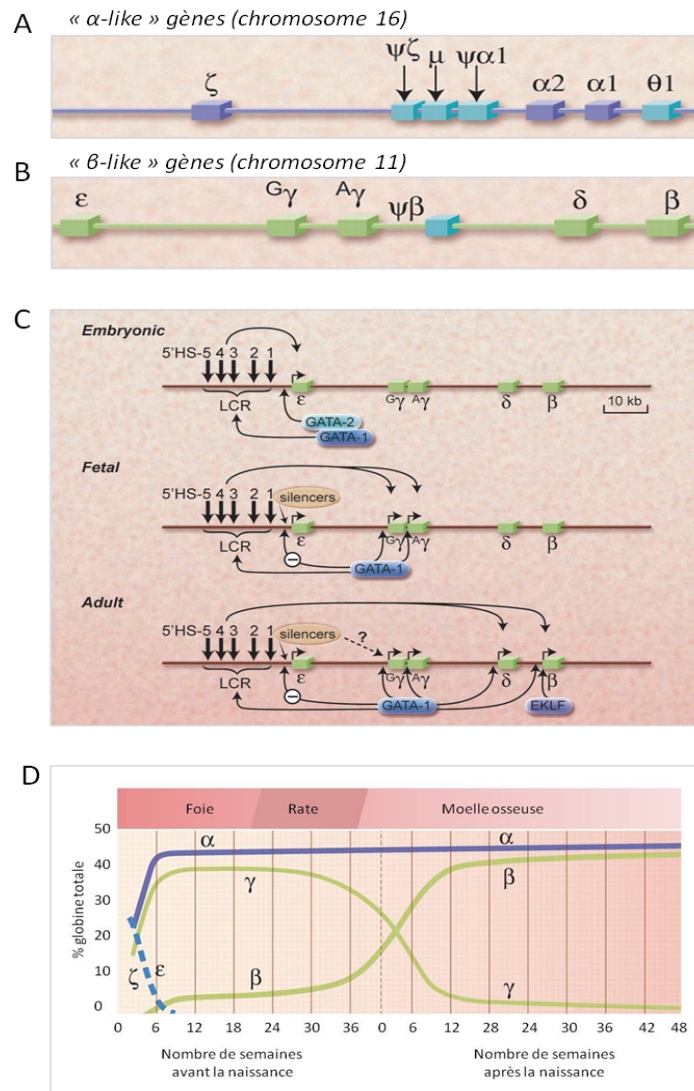


Figure 1. 10 . Structure et expression des gènes de globines « α -like » et « β -like » embryonnaire, fœtale et adulte. (A) Structure des gènes de globine « α -like ». (B) Structure des gènes de globine « β -like ». (C) Implication des facteurs GATA-1, GATA-2 et EKLF dans la régulation des gènes de globine « β -like » au cours du développement. LCR indique la région locus contrôlé. HS désigne les sites d'hypersensibilité à la DNase. (D) Variations de l'expression des différents gènes de globine et lieux de production des érythrocytes chez l'humain au cours du développement (Adapté de Schechter (2008) [150]).

3 - Erytroid Krüppel-like factor (EKLF)

a - Domaines d'EKLF

La famille des facteurs de type Krüppel doit son nom à l'homologie que possède le domaine de liaison de ces protéines à l'ADN avec la protéine Krüppel de *Drosophila melanogaster*. Les membres de cette famille possèdent trois doigts de zinc de type Cys₂/His₂ situés en position carboxy-terminale et séparés par des « linker » H/C conservés (H désigne la dernière histidine du doigt de zinc *i* et C la première cystéine du doigt de zinc *i+1*). Les doigts de zinc de cette famille ont une affinité pour les séquences d'ADN riches en G/C et les sites CACCC [151, 152]. Le promoteur du gène humain de la β-globine possède une séquence d'ADN CACCC située sur le promoteur proximal, à la position -90, reconnue par le DBD d'EKLF [153, 154] (Figure 1. 10C en page 34). Ce facteur de transcription est un élément clef dans le développement érythropoïétique puisqu'il est responsable du haut niveau d'expression du gène de la β-globine. L'expression d'EKLF est, entre autre, régulée par GATA-1 [155]. Son implication dans la régulation du gène de la β-globine en fait l'un des facteurs de transcription érythropoïétique les plus étudiés.

La protéine EKLF possède un domaine d'activation en amont de son site de liaison à l'ADN. Plusieurs travaux ont permis de localiser un domaine de transactivation acide minimal en position amino-terminale comprenant les cent premiers résidus (TAD A) [153, 156, 157]. Plus récemment, il a été montré que la région riche en prolines comprise entre les résidus 140 à 225 (TAD B) (Figure 1. 11 en page 37) est, elle aussi, capable d'activer indépendamment la transcription du gène de la β-globine [157]. Ces divers travaux indiquent qu'EKLF possède au moins deux domaines de transactivation dont la fonction est redondante.

b - Similitudes des TAD d'EKLF et de p53

La composition riche en acides aminés acides du domaine de transactivation TAD A d'EKLF classe celui-ci dans la catégorie des TAD acides. Ce domaine a été identifié comme étant le domaine minimum d'activation de la transcription[153]. Il est intéressant de constater que

le TAD d'EKLF est extrêmement similaire, tant au point de vue de la séquence que de la fonction, à celui de la protéine suppresseur de tumeur p53. Tout comme le TAD de p53, le TAD d'EKLF possède de nombreux résidus acides et hydrophobes. De plus, ces deux activateurs de la transcription peuvent être divisés en deux sous-domaines qui possèdent des rôles distincts et capables d'activer indépendamment la transcription. En effet, le TAD de p53 se divise en deux sous-domaines : le TAD1 allant des résidus 1 à 40 et le TAD2 incluant les résidus 40 à 73. Le premier sous-domaine d'EKLF est quant à lui situé entre les résidus 1 à 40 et le deuxième entre les résidus 50 à 90 (Figure 1. 11 en page 37).

Les TAD1 et TAD2 de p53 sont notamment connus pour interagir respectivement avec MDM2 et Tfb1. Lors de leur interaction avec leurs partenaires respectifs, ces deux TAD adoptent une structure secondaire en hélice- α amphipatique (résidus soulignés dans la Figure 1. 12 en page 37) dont la face hydrophobe est responsable de l'interaction avec la protéine cible. L'alignement des séquences des TAD1 et TAD2 de p53, avec les TAD1 et TAD2 d'EKLF respectivement, révèle qu'à chacune des trois positions hydrophobes clefs des TAD1 et TAD2 de p53 correspond un résidu hydrophobe dans EKLF. De plus, à plusieurs sites de phosphorylation de p53 correspondent des sites potentiels de phosphorylation d'EKLF. Ces observations mettent en lumière la similitude de ces domaines de transactivation.

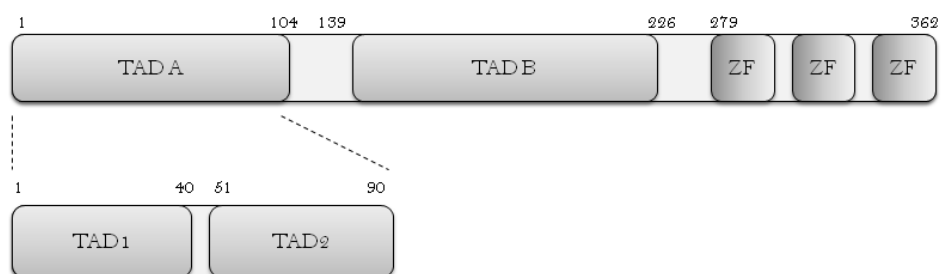


Figure 1. 11. Domaines fonctionnels d'EKLF. Diagramme représentant les domaines d'EKLF : TAD représente un domaine de transactivation et ZF défini un doigt de zinc.

A	hEKLf TAD1	19	P F P D T Q D D F L K W W R S E E A Q D M	39
	hp53 TAD1	11	E P P L S Q E T F S D L W K L L P E N N V	31
B	hEKLf TAD2	64	D D E R G A D A T W D L D L L L T N F S G	84
	hp53 TAD2	39	A M D D L M L S P D D I E Q W F T E D P G	59

Figure 1. 12. Alignement de séquences du TAD1 (A) et du TAD2 (B) d'EKLF humain (hEKLf) avec les TAD respectifs de p53 humain (hp53). Les résidus connus pour former une hélice sont soulignés en noir. Les résidus hydrophobes impliqués dans les interactions sont surlignés en gris. Les sites potentiels de phosphorylation sont surlignés en jaune.

c - Rôles d'EKLF

Chez les mammifères, l'expression du groupe de gènes globines « β -like » est régulée spécifiquement en fonction du stade du développement. Chez l'humain, les premières cellules hématopoïétiques formées dans la vésicule vitelline expriment la ξ -globine embryonnaire. Vient ensuite le stade fœtal où le foie devient le site de formation des cellules sanguines. La γ -globine fœtale remplace alors la ξ -globine. Puis, lors du passage au stade adulte, la γ -globine est graduellement remplacée par la β -globine adulte. La moelle osseuse devient dès lors le site principal de production des cellules sanguines [2] (Figure 1. 10D en page 34). La régulation de ce groupement de gènes est complexe et se produit *via* la compétition des promoteurs proximaux des différentes globines pour l'interaction avec la région LCR. Cela permettrait une activité transcriptionnelle élevée d'un seul membre à un temps approprié du développement (Figure 1. 7 en page 28) [2, 158, 159].

Le facteur EKLF joue un rôle décisif lors du passage du stade fœtal au stade adulte, c'est-à-dire quand la β -globine vient remplacer la γ -globine. En effet, la délétion d'EKLF chez la souris provoque la diminution des transcrits de β -globine uniquement. Les souris apparaissent normales au cours du stade embryonnaire mais meurent d'anémie au cours de l'érythropoïèse fœtale [160]. Elles présentent un déséquilibre dans la composition de leur hémoglobine, phénotype typique des β -thalassémies. Ces observations confortent le rôle d'EKLF dans le remplacement de la γ -globine par la β -globine lors de la transition fœtal/adulte chez l'humain [161, 162]. L'absence d'EKLF est associée à une réduction de l'hypersensibilité à la DNaseI au niveau du promoteur de la β -globine et de la région HS3 du LCR, suggérant un rôle important d'EKLF dans le contrôle de la structure de la chromatine (Figure 1. 10C en page 34). Accréditant cette hypothèse, il a été démontré qu'EKLF interagit avec les acétyltransférases CBP/p300 [49, 163, 164] et des complexes de remodelage de la chromatine dépendante de l'ATP tel que le complexe SWI/SNF [164, 165]. L'interaction avec CBP/p300 a pour conséquence l'acétylation d'EKLF au niveau de son domaine riche en prolines, ce qui a pour effet d'augmenter son activité transcriptionnelle [49, 163]. Le complexe SWI/SNF facilite quant à lui la formation de sites d'hypersensibilité à la DNaseI dépendant d'EKLF *in vitro*. Ce résultat corrèle avec le fait que les

sous-unités BRG1 et BAF170 de SWI/SNF sont associées au promoteur de la β -globine *in vivo* [166-168]. La stimulation de l'interaction BRG1/EKLF suite à l'acétylation d'EKLF suggère qu'il existe un lien potentiel entre l'acétylation d'EKLF et le recrutement du complexe SWI/SNF. Le déroulement des événements serait donc le suivant : EKLF lierait CBP/p300 entraînant l'acétylation des histones et d'EKLF. La modification d'EKLF stimulerait alors le recrutement de SWI/SNF qui lui, remanierait la chromatine et entrainerait alors l'activation de la transcription [169] (Figure 1. 7 en page 28).

Non seulement EKLF contrôle l'expression de la β -globine, mais il contribue également au blocage de la différenciation en mégacaryocytes [170]. Les précurseurs communs des érythrocytes et des mégacaryocytes ne disposent plus que de deux choix de différenciation possibles : érythrocytes ou mégacaryocytes. Le choix de produire des érythrocytes est guidé par la présence d'EKLF. En effet, EKLF bloque la différenciation des mégacaryocytes au bénéfice de la différenciation des érythrocytes (Figure 1. 1 en page 3) *via* l'inhibition de Fli-1, un activateur indispensable des mégacaryocytes [138, 155].

d - Interactions protéine/protéine

La régulation des divers gènes cibles d'EKLF est modulée par des interactions avec de nombreux partenaires protéiques tels que les facteurs de transcription GATA-1 [103, 104] et Fli-1 [171], mais aussi avec des co-activateurs de la transcription tels que CBP/p300 [163, 164], P/CAF [49] ou encore avec des facteurs généraux de transcription telle que la sous-unité TAF9 de TFIID [172]. En effet, au niveau du locus de la β -globine, il a été récemment démontré que la sous-unité TAF9 du facteur général de transcription TFIID interagit avec une séquence d'ADN en aval du promoteur ainsi qu'avec le facteur de transcription érythropoïétique EKLF. Cette dernière interaction est impliquée dans le recrutement de TFIID au promoteur. La présence de TAF9 s'avère être indispensable pour l'expression de la β -globine [172]. En revanche, l'absence de TAF9 au promoteur du gène codant pour l'AHSP (α -Hemoglobin-Stabilizing Protein), gène également régulé par EKLF, n'affecte pas l'activation de la transcription [172]. Le recrutement spécifique des TAF sur certains promoteurs met en évidence la mise en place de mécanismes de

régulation propres à chaque gène [163]. Les interactions protéine/protéine impliquant EKLf sont régulées par des modifications post-traductionnelles au niveau de son domaine de transactivation telles que la phosphorylation [153, 173, 174], l'ubiquitination [175], la sumoylation [176] ou encore l'acétylation [49, 163].

C - Désordres sanguins

Comme dans tous les tissus, des cancers et des malformations peuvent apparaître dans le sang. Il existe différents types de cancers du sang : les leucémies aiguës lymphoblastique (ALL) ou myéloblastique, (AML) qui sont caractérisées par la prolifération rapide de cellules immatures de la moelle osseuse, et les leucémies lymphoïde (CLL) ou myéloïde chroniques (CML), où les cellules cancéreuses sont plus matures. Certains virus, tel HTLV-1 et HTLV-2 (**H**uman **T** cell **L**eukemia **V**irus type 1 et 2) sont à l'origine de l'apparition d'ALL. Il existe également un grand nombre de malformations sanguines telles que la thrombocytopénie qui se traduit par une diminution du nombre de plaquettes. Par ailleurs, de nombreuses maladies sanguines sont reliées à une hémoglobine non fonctionnelle. La drépanocytose, également appelée sicklémie ou anémie falciforme, est une maladie héréditaire caractérisée par une mutation de l'hémoglobine qui favorise son agrégation en filament causant une déformation des érythrocytes qui deviennent inaptes à circuler dans les vaisseaux sanguins. De plus les β - et α -thalassémies sont des maladies répandues caractérisées par une chute de la production de β - et d' α -globine respectivement, rendant les globules rouges inaptes à transporter l'oxygène. La plupart de ces maladies sont causées par des mutations altérant des interactions protéine/protéine ou protéine/ADN primordiales pour la différenciation des cellules hématopoïétiques.

1 - GATA et désordres sanguins

Récemment, des formes mutantes de GATA-1 et GATA-2 ont été identifiées chez l'humain et corrént directement avec l'apparition de désordres sanguins tels que des

thalassémies, des thrombocytopénies ou encore des leucémies megakarioblastiques aiguës [177-182]. Des mutations dans le doigt de zinc amino-terminal, perturbant son interaction avec l'ADN ou avec FOG, ont été identifiés chez des patients atteints de thrombocytopénies ou d'anémies [183, 184]. De plus, chez les personnes atteintes du syndrome de Down, aussi nommé trisomie 21, l'apparition de leucémie est initiée par des mutations de GATA-1 [185-187]. Ces mutations perturbent des interactions auxquelles GATA-1 participe et sont directement responsables de l'apparition de maladies.

Par ailleurs, la perturbation de la balance entre les facteurs de transcription PU.1 et GATA-1 peut être la cause de l'apparition de leucémies. Une mutation sur la séquence activatrice distale du gène de PU.1 a été identifiée chez des patients souffrant de leucémie myéloïde aiguë (AML) [188], démontrant que la répression de l'expression de PU.1 perturbe la balance entre la production des précurseurs des mégacaryocytes et des érythrocytes et des précurseurs des granulocytes et des monocytes, et conduit à l'apparition de cancers du sang.

2 - EKLF et désordres sanguins

La présence d'EKLF est essentielle à la production de globules rouges. Son absence, de même que des mutations sur son site de liaison au promoteur de la β -globine, provoquent l'apparition de β -thalassémie due à une diminution de la production de β -globine dans les érythrocytes [23-25]. En effet, des mutations du site d'ADN consensus CACCC du gène de la β -globine ont été identifiées chez des patients et ont pour conséquence l'abolition de la liaison d'EKLF au promoteur, abrogeant ainsi l'activation du gène [189, 190]. Les connaissances actuelles sur le dispositif complexe de contrôle de la transcription par EKLF restent incomplètes. Malgré de nombreuses études, aucune information structurale concernant le mode d'activation des gènes cibles d'EKLF ne sont à notre disposition.

L'implication des facteurs GATA-1 et EKLF dans l'apparition de malformations sanguines justifie les études structurales et fonctionnelles des complexes auxquels ces protéines participent. Les informations fournies par de telles études sont des outils précieux pour

approfondir notre compréhension des mécanismes détaillés de formation de cancers et pour contribuer au développement de nouveaux traitements contre les anémies et les leucémies.

D - Objectifs de recherche et justification des méthodes employées

1 - Hypothèses et Objectifs de recherche

La régulation de la différenciation des cellules souches hématopoïétiques en diverses lignées sanguines passe par un contrôle précis de la transcription de gènes cibles en temps et lieu appropriés. La fine régulation de la transcription des gènes est accomplie par l'action combinée de co-activateurs, de FGT et d'activateurs de la transcription. Les TAD des activateurs occupent une place importante dans ce mécanisme de régulation puisqu'ils ont la capacité d'activer la transcription *via* leur interaction avec de multiples partenaires. Ces interactions sont responsables du recrutement de co-activateurs ou de FGT au niveau du promoteur, accélérant ainsi l'initiation et l'élongation. Étant présente chez plusieurs oncogènes, la classe des TAD acides se distingue des autres catégories de TAD. Malgré quelques études structurales de ces domaines en complexe avec différentes protéines, les mécanismes par lesquels les TAD interagissent avec leurs multiples partenaires restent néanmoins très peu connus.

Afin de vérifier l'hypothèse selon laquelle divers TAD acides pourraient interagir avec un partenaire commun en utilisant des modes de reconnaissance analogues, nous avons entamé des études structurales de ces domaines en interaction avec des partenaires. Suggestions : Les trois objectifs principaux de cette thèse ont porté sur l'étude de :

1) L'interaction du TAD de p53 avec GATA-1:

Au cours de l'hématopoïèse, le facteur GATA-1 est connu pour interagir et inhiber PU.1 *via* une interaction avec le TAD acide de ce dernier. Étant donné le rôle anti-apoptotique primordial de GATA-1 dans les érythrocytes, le rôle pro-apoptotique de p53, les similitude entre les TAD acides de PU.1 et de p53 et puisque les voies de GATA-1 et de p53 possèdent des

partenaires en commun, nous nous sommes demandé : Est-ce qu'une interaction entre GATA-1 et le TAD de p53 pourrait jouer un rôle dans le maintien en vie des globules rouges ? Et si oui, quels sont les domaines requis pour cette interaction?

2) L'interaction du TAD viral de VP16 avec TFIID :

Une partie des travaux relatés dans cette thèse, portent sur l'étude d'interaction de TAD acides avec le facteur général de transcription TFIID. Étant donné que le TAD de VP16 est connu pour interagir avec la sous-unité Tfb1/p62 de TFIID et puisque celui-ci possède des similarités de séquence avec le TAD de p53, nous nous sommes demandé : Est-ce que le TAD viral de VP16 interagit avec Tfb1/p62 de façon similaire au TAD cellulaire de p53 ?

3) L'interaction du TAD de EKLF avec TFIID :

Étant donné que le TAD du facteur de transcription érythropoïétique EKLF possède une séquence similaire aux TAD de p53 et de VP16, nous avons émis l'hypothèse que le TAD d'EKLF pouvait lui aussi interagir avec Tfb1/p62 de la même manière que les TAD de VP16 et de p53.

Étant donné l'implication des facteurs de transcription GATA-1 et EKLF dans l'érythropoïèse et dans l'apparition de malformations sanguines, la caractérisation détaillée des interactions auxquelles ces protéines participent s'avèrent être des outils précieux pour la conception rationnelle de traitements contre les anémies et les leucémies. C'est donc dans le but d'améliorer notre compréhension de la régulation des gènes au cours de la différenciation des érythrocytes et d'approfondir nos connaissances sur le mode d'action des TAD acides que nous avons entrepris des études structurales et fonctionnelles de plusieurs interactions impliquant notamment les TAD acides de la protéine suppresseur de tumeur p53, de l'activateur viral VP16 et du facteur de transcription érythropoïétique EKLF en complexe avec divers partenaires tel que le facteur de transcription érythropoïétique GATA-1 et le facteur général de transcription TFIID.

2 - Résonance magnétique nucléaire

Deux techniques permettent de résoudre des structures tridimensionnelles de macromolécules biologiques à résolution atomiques : la radiocristallographie et la spectroscopie de Résonance Magnétique Nucléaire (RMN). La méthode basée sur la diffraction des rayons X est celle qui a fourni le plus grand nombre de structures tridimensionnelles. La cristallographie offre l'avantage de ne pas être limitée par la taille de la molécule étudiée mais nécessite néanmoins l'obtention de cristaux ayant un fort pouvoir diffractant (arrangement périodique et ordonné de la molécule dans l'espace). Ce processus étant aléatoire, il représente habituellement une étape limitante. De plus, étant donné que la présence de domaines flexibles ou intrinsèquement désordonnés empêche la formation de cristaux, l'étude de nombreuses protéines ne peut être effectuée par radiocristallographie.

La spectroscopie RMN, quant à elle, est limitée par la taille des macromolécules étudiées mais offre l'avantage de pouvoir étudier les protéines en solution. Un autre atout de la RMN est la possibilité de travailler avec des domaines non structurés. Il est intéressant de constater que la majorité des structures impliquant des domaines de transactivation, tels que les complexes impliquant des domaines de CBP/p300 présentés dans la section **B-3-b-ii**, ont été résolues par spectroscopie RMN [38, 40, 79, 81-85]. L'utilisation de la spectroscopie RMN s'est imposée pour l'étude des facteurs de transcription présentés dans ce manuscrit, étant donné que les TAD sont connus pour être non structurés à l'état libre et puisque la majeure partie des protéines impliquées dans la transcription possèdent des domaines intrinsèquement désordonnés.

a - Principe théorique de la RMN

La résonance magnétique nucléaire utilise les propriétés magnétiques des atomes. Les noyaux atomiques sont caractérisés par leur mouvement cinétique \vec{P} ainsi que par un moment magnétique $\vec{\mu}$ reliés par l'équation :

$$\vec{\mu} = \gamma \cdot \vec{P}$$

Ou

$$P = I(I+1) \frac{h}{2\pi}$$

Avec :

γ : Rapport gyromagnétique

I : Nombre quantique de spin nucléaire

h : Constante de Planck

Le rapport gyromagnétique γ est dépendant du noyau considéré. Le nombre quantique de spin nucléaire I dépend du noyau. Pour un noyau A_ZX , où A est la masse atomique, et Z le nombre atomique :

- si A est pair et Z est pair $I = 0$
- si A est pair et Z est impair $I = \text{entier}$
- si A est impair et Z est impair $I = \text{demi-entier}$

Les atomes que l'on souhaite observer en RMN sont l'hydrogène, le carbone et l'azote. Cependant, les noyaux dont le moment magnétique μ est égal à zéro ne sont pas observables en RMN. C'est à dire que les noyaux dont $I = 0$ sont inactifs en RMN, tel le Carbone ${}^{12}_6C$. De plus, les noyaux observées étant ceux possédant un spin nucléaire de $\frac{1}{2}$, l'Azote ${}^{14}_7N$ n'est lui aussi pas observé en RMN. En revanche, tout comme le proton 1_1H , l'azote ${}^{15}_7N$ et le carbone ${}^{13}_6C$, possèdent un spin de $\frac{1}{2}$. Étant donné la faible abondance naturelle de ces isotopes (0,37 % pour l'azote ${}^{15}_7N$ et 1,1 % pour le carbone ${}^{13}_6C$), il est nécessaire de marquer uniformément les protéines. Pour cela, les cultures bactériennes permettant de surproduire la protéine étudiée sont effectuées dans un milieu minimum dont la seule source d'azote est du chlorure d'ammonium ${}^{15}NH_4Cl$ et dont la seule source de carbone est du glucose ${}^{13}C_6H_{12}O_6$.

Noyau	I	Abondance Naturelle
1_1H	$\frac{1}{2}$	99,98%
${}^{12}_6C$	0	98,9%
${}^{13}_6C$	$\frac{1}{2}$	1,1%
${}^{14}_7N$	1	99,63%
${}^{15}_7N$	$\frac{1}{2}$	0,37%

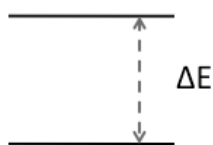
Les moments magnétiques $\dot{\mu}$ des noyaux sont capable d'interagir avec un champ magnétique extérieur. Le principe de la spectroscopie RMN consiste à placer la macromolécule d'intérêt dans un champ magnétique intense. Les spins des noyaux atomiques composant la macromolécule s'orientent alors le long de l'axe principal du champ magnétique. Dans le champ magnétique \dot{B}_0 d'un spectromètre RMN, le moment magnétique des noyaux $\dot{\mu}$ et \dot{B}_0 interagissent. Cette interaction est caractérisée par une énergie :

$$E = \mu \cdot B_0$$

Un noyau de spin $\frac{1}{2}$ possède deux niveaux d'énergie possibles :

$$\text{Niveau } \beta : E = -\frac{1}{2} \gamma \frac{h}{2\pi} B_0$$

$$\text{Niveau } \alpha : E = +\frac{1}{2} \gamma \frac{h}{2\pi} B_0$$



La différence de population entre ces deux niveaux appelé aussi énergie de transition du moment magnétique (ΔE) résulte en une magnétisation macroscopique le long de l'axe du champ magnétique \dot{B}_0 . ΔE peut être définie en fonction de la fréquence de l'onde électromagnétique ν_0 qui permet la résonance :

$$\Delta E = \gamma \frac{h}{2\pi} B_0$$

$$\Delta E = h\nu_0 \quad \text{avec} \quad \nu_0 = \gamma \frac{B_0}{2\pi}$$

En présence du champ \dot{B}_0 la magnétisation d'un noyau $\dot{\mu}$ tourne autour de ce champ à une fréquence angulaire ω_0 appelé précession définie par :

$$\omega_0 = 2\pi\nu_0 = \gamma B_0$$

Il est ensuite possible de perturber l'état d'équilibre de ces spins grâce à des impulsions électromagnétiques et de mesurer leur retour à l'équilibre. Suite à l'application d'un champ magnétique \dot{B}_1 à l'échantillon, le noyau va passer d'un niveau d'énergie à un autre. Dès la fin de l'excitation, il va y avoir un retour à l'état d'équilibre à la fréquence ν . Ce phénomène de relaxation, appelé résonance, est détectable. Le signal RMN ainsi obtenu contient la contribution de l'ensemble des atomes de la macromolécule. La transformée de Fourier de ce signal permet d'obtenir les différentes fréquences de résonance des spins observés. Chaque atome d'une molécule résonne à une fréquence particulière. Selon le type de spectre enregistré, des corrélations entre divers atomes vont être observées et vont permettre d'identifier les résonances des atomes.

b - Détermination de structures tridimensionnelles par spectroscopie RMN

La détermination d'une structure par RMN hétéronucléaire comporte plusieurs étapes (Figure 1. 13 en page 50). L'acquisition de différents types de spectres en deux et trois dimensions est nécessaire à la détermination de structures par RMN. L'analyse des spectres RMN peut être divisée en deux étapes :

1) la première étape consiste à établir la liste des déplacements chimiques de tous les atomes observables. Pour cela nous utilisons des spectres à deux et trois dimensions corrélant les atomes à travers les liaisons chimiques tels que les spectres à deux dimensions:

- ^1H - ^{15}N HSQC
- ^1H - ^{13}C CT-HSQC

ou encore les spectres a trois dimensions :

- ^{15}N -edited TOCSY-HSQC
- HNCACB
- (HB)CBCA(CO)NNH
- HNCO
- H(CCO)NNH-TOCSY
- C(CO)NNH-TOSCY

2) la seconde étape consiste à analyser des expériences NOE (Nuclear Overhauser Effect) qui permettent de déterminer la majorité des contraintes structurales. L'effet NOE permet d'obtenir des corrélations résultant de la proximité spatiale des différents noyaux et permet donc d'identifier les atomes voisins dans l'espace d'un atome particulier. L'intensité de l'effet NOE diminue en fonction de $\frac{1}{r^6}$ (ou r est la distance entre deux atomes). Pour cela on enregistre les spectres a trois dimensions tels que :

- ^{15}N -edited NOESY-HSQC

- ^{13}C -edited HMQC-NOESY

De plus, afin de résoudre la structure d'un complexe protéique par RMN, des contraintes intermoléculaires sont obtenues par l'enregistrement de spectres ^{13}C -filtered-edited NOESY.

Des contraintes de distance sont dérivées de l'intensité des signaux observés dans les expériences NOE. Ces contraintes de distances sont alors utilisées pour calculer un ensemble de structures satisfaisant les contraintes imposées.

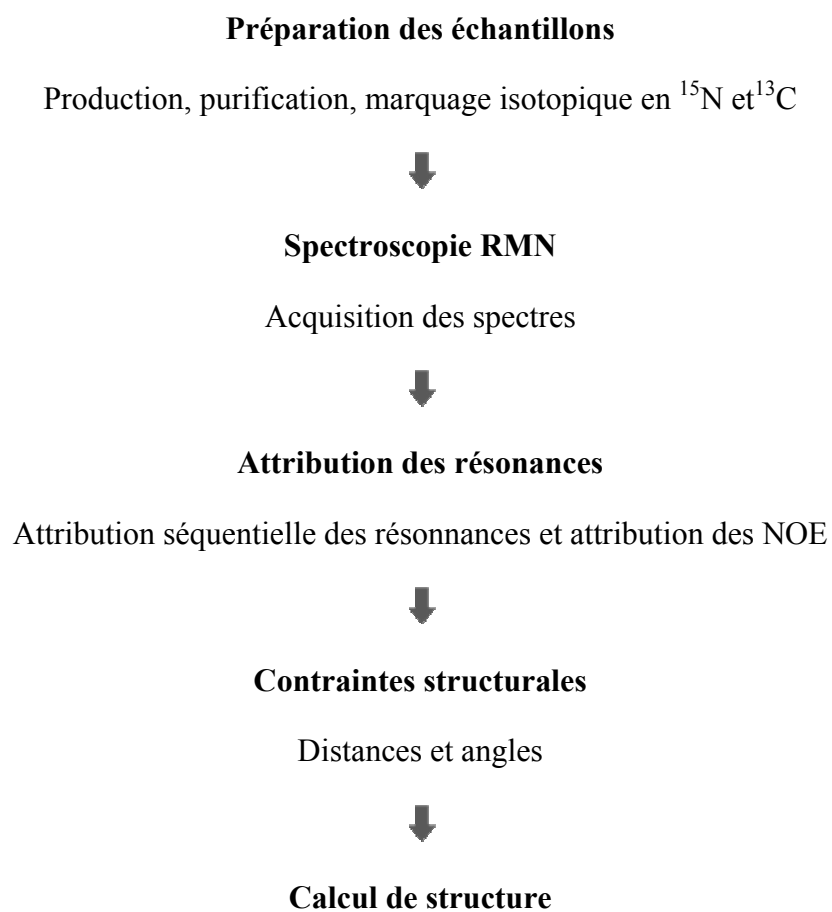


Figure 1. 13 . Étapes standard pour la détermination de structures par spectroscopie RMN hétéronucléaire.

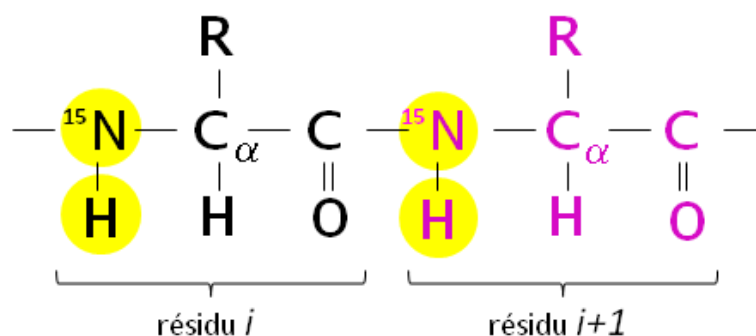


Figure 1. 14 . Corrélations entre les atomes ^{15}N et ^1H observées dans un spectre ^1H - ^{15}N HSQC hétéronucléaire à deux dimensions. Un signal est observé pour chaque liaison amide présente dans la protéine.

c - Cartographie des régions d'interactions

La RMN hétéronucléaire s'avère être une méthode particulièrement efficace pour cartographier les zones d'interactions entre macromolécules. Pour cela, nous utilisons des spectres à deux dimensions ^1H - ^{15}N HSQC (**H**eteronuclear **S**ingle **Q**uantum **C**oherence). Ces spectres possèdent une dimension ^1H et une dimension ^{15}N . Ils présentent un signal pour chaque liaison amide de la protéine, c'est-à-dire que pour chaque acide aminé un signal va être observé (sauf pour la proline) (Figure 1. 14 ci-dessus). L'enregistrement de spectres ^1H - ^{15}N HSQC d'une protéine uniformément marquée en ^{15}N en absence et en présence de son partenaire non marqué en ^{15}N sont nécessaires.

La position de chaque pic sur le spectre dépend de l'environnement chimique de l'acide aminé auquel il correspond. Lors d'une interaction avec un partenaire, l'environnement chimique des acides aminés impliqués dans l'interaction va être modifié, les signaux de ces résidus vont donc se déplacer. Il est donc possible d'identifier les acides aminés participant à l'interaction (Figure 1. 15 en page 52).

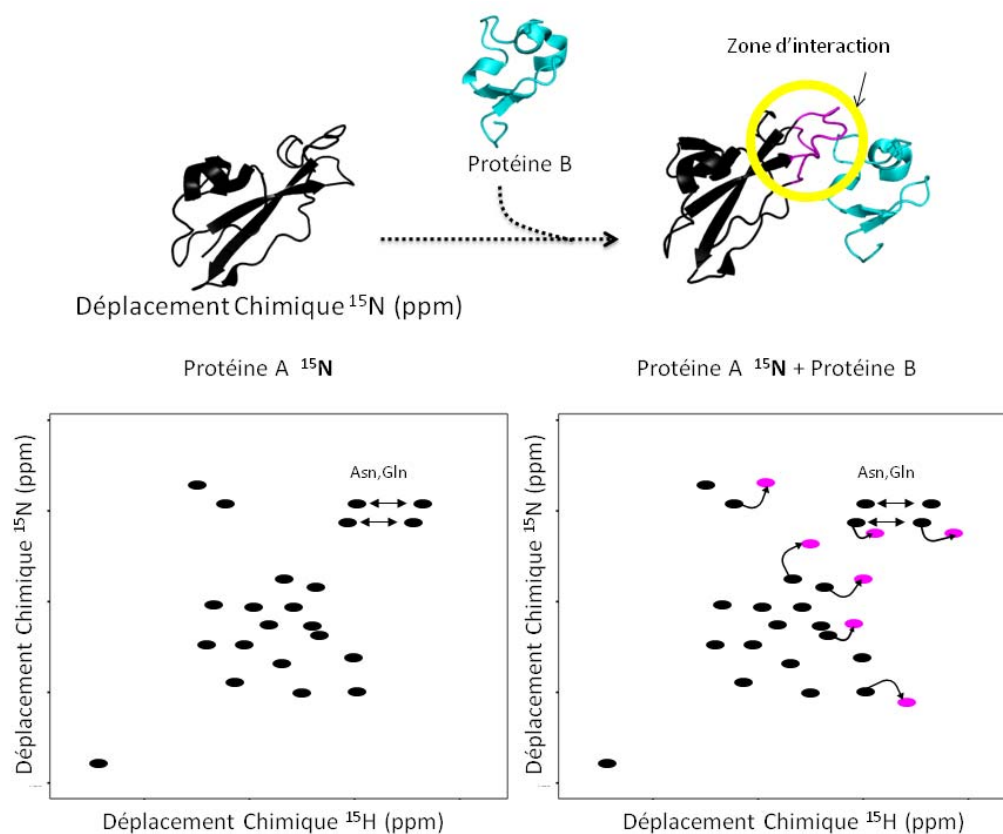


Figure 1. 15 . Principe de la cartographie d'une interaction par spectroscopie RMN. Des spectres 2D ^1H - ^{15}N HSQC ^1H - ^{15}N HSQC d'une protéine A (marquée ^{15}N) sont enregistrés pour la forme libre et pour la forme complexée à une protéine B (non marquée). Il y a modification des déplacements chimiques des résidus impliqués dans l'interaction suite à l'ajout de la protéine B. L'identification des résidus correspondant aux signaux modifiés permet de cartographier la zone d'interaction

3 - Calorimétrie à titrage isotherme

Afin de mesurer les constantes d'affinités régissant l'interaction entre différents partenaires ainsi que d'étudier l'effet de mutations ponctuelles sur ces interactions, nous avons utilisé la calorimétrie à titrage isotherme (ITC : Isothermal Titration Calorimetry). L'ITC est une technique thermodynamique qui permet de caractériser l'interaction entre deux macromolécules biologiques. Une interaction entre deux partenaires peut être exothermique ou endothermique. Lors d'une expérience d'ITC la chaleur émise ou absorbée par la formation d'un complexe entre deux molécules est mesurée et permet de déterminer la constante d'association (K_a), la stœchiométrie (n) ainsi que le profil thermodynamique de la réaction qui comprend l'enthalpie (ΔH) et l'entropie (ΔS). Le domaine des constantes d'association directement mesurables s'étend de 10^2 à 10^9 M^{-1} .

Lors d'un titrage ITC, réalisé à température constante, le « ligand » contenu dans une seringue est progressivement injecté dans une solution de la « macromolécule » contenue dans une cellule. Après chaque ajout d'une petite quantité de ligand, l'énergie libérée ou absorbée est mesurée et comparée à celle d'une cellule de référence. Ces deux cellules, la cellule de référence ainsi que la cellule contenant l'échantillon, sont placés dans un compartiment adiabatique. La différence de chaleur entre ces deux compartiments est exprimée comme la puissance électrique requise pour maintenir une température constante entre la cellule contenant l'échantillon et la cellule de référence.

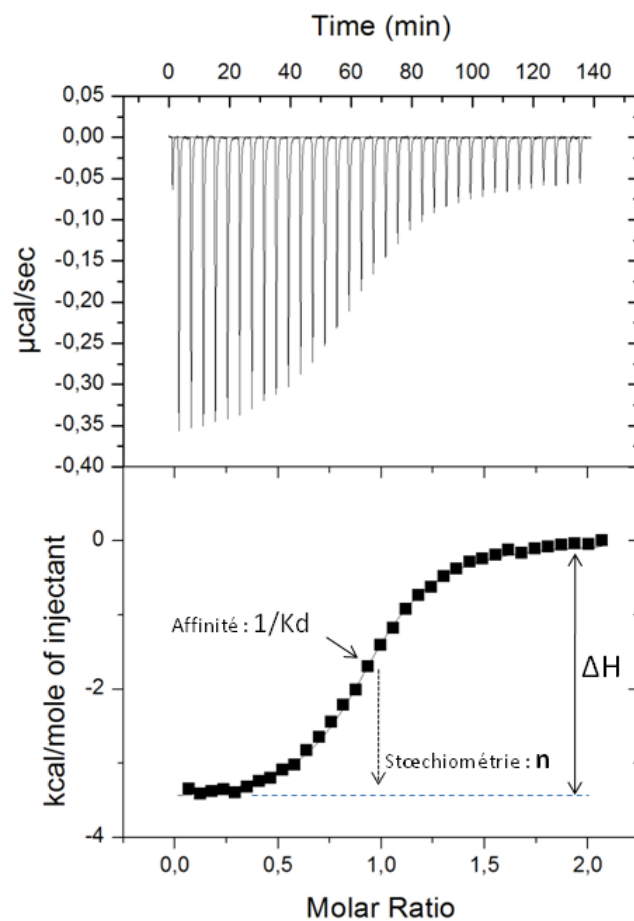


Figure 1. 16 . Thermogramme d'une experience d'ITC.

Lors des premières injections (Figure 1. 16 en page 54), tout le ligand injecté se lie à la macromolécule. Au fur et à mesure que la molécule contenue dans la cellule est saturée en ligand, la quantité de chaleur générée diminue jusqu'à atteindre une valeur correspondant au bruit de fond (chaleur de dilution).

Les données expérimentales ont été traitées avec le programme MicoCAL Origin 7.0. Les études d'interactions effectuées lors des travaux présentés dans cette thèse coïncident avec un modèle de site de liaison unique. Selon le manuel « ITC Data Analysis in Origin » de MicoCAL, les équations (1) et (2) qui suivent sont utilisées pour corréler les données expérimentales. Dans ces équations :

K_a = Constante de liaison

n = Nombre de site de liaison

V_0 = Volume actif de la cellule

dV_i = Volume d'une injection

M_t = Concentration de la macromolécule dans V_0

X_t = Concentration du ligand dans V_0

ΔH = Enthalpie de liaison

La chaleur totale Q de l'échantillon contenu dans le volume de la cellule V_0 est définie par l'équation (1):

$$Q = \frac{nM_t\Delta HV_0}{2} \left[1 + \frac{X_t}{nM_t} + \frac{1}{nK_aM_t} - \sqrt{\left(1 + \frac{X_t}{nM_t} + \frac{1}{nK_aM_t} \right)^2 - \frac{4X_t}{nM_t}} \right]$$

La chaleur $\Delta Q(i)$ dégagée ou absorbée suite à la $i^{\text{ème}}$ injection est définie par l'équation :

$$\Delta Q(i) = Q(i) + \frac{dV_i}{V_0} \times \left[\frac{Q(i) + Q(1-i)}{2} \right] - Q(i-1)$$

Connaissant les concentrations initiales (M_t et X_t), la constante d'affinité (K_a), l'enthalpie de liaison (ΔH) et la stœchiométrie (n), il est possible de calculer la chaleur impliquée dans chaque injection. Le processus de corrélation des données expérimentales implique plusieurs étapes :

- 1) Estimation des paramètres thermodynamiques (K_a , ΔH et n) par le programme MicroCAL Origin 7.0
- 2) Calcul de $\Delta Q(i)$ pour chaque injection et comparaison avec la valeur mesurée pour l'injection correspondante
- 3) Ajustement des valeurs K_a , ΔH et n
- 4) Itération de la procédure jusqu'à ce qu'on n'observe plus d'amélioration significative de l'ajustement des données expérimentales avec les paramètres thermodynamiques (K_a , ΔH et n).

La position du point d'inflexion de la courbe correspond au rapport molaire n entre ligand et la macromolécule. Ainsi, n définit la stœchiométrie de la réaction. Dans le cas où $n=1$, la pente de la courbe au point d'inflexion représente la constante d'association K_a . L'enthalpie ΔH de la réaction correspond à la différence de kcal/mol libérées ou absorbées entre le début (quand toutes les molécules de ligand se lient à la macromolécule) et la fin du titrage (quand la macromolécule est saturée en ligand et qu'on observe juste la chaleur de dilution) (Figure 1. 16 en page 54)

[191]. À partir de ces mesures, il est possible de calculer l'énergie libre de la réaction ΔG ainsi que l'entropie ΔS , grâce aux relations suivantes :

$$\Delta G = -RT \ln K_a$$

$$\Delta G = \Delta H - T\Delta S$$

où R , la constante des gaz parfaits, vaut $8,314472 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$.

Une expérience d'ITC peut être caractérisée par la valeur c qui est le produit de la constante d'association (K_a), de la concentration de la macromolécule ($[M]$) et de la stœchiométrie de la réaction (n).

$$c = K_a [M] n$$

Pour une détermination précise de la constante d'association, la valeur de c doit être comprise entre 1 et 1000. En effet, pour une trop grande valeur de c , la saturation de la macromolécule est trop rapide et il n'y a pas assez de points à l'inflexion pour déterminer une constante d'affinité avec précision. De même, pour des faibles valeurs de c la courbe perd sa forme sigmoïdale caractéristique et se rapproche de la linéarité. Le point d'inflexion ne peut donc pas être déterminé.

L'ITC permet également de déterminer la capacité calorifique (ΔC_p) associée à la formation d'un complexe. L'analyse à plusieurs températures donne la variation de la capacité calorifique selon la relation :

$$\Delta C_p = \frac{d\Delta H}{dT}$$

Préface au Chapitre II

Non seulement GATA-1 harmonise l'expression des gènes érythrocytaires, mais ce facteur de transcription contrôle aussi la différenciation et la survie des érythrocytes notamment *via* le contrôle des gènes conduisant à l'apoptose, affiliés à la voie de p53. Malgré les nombreuses études effectués, le mécanisme par lequel les érythrocytes sont maintenus en vie au cours de la différenciation terminale reste peu connu. En collaboration avec le Dr. Cecelia Trainor (National Institutes of Health, Bethesda, MD), nous avons identifié et caractérisé une interaction entre les facteurs de transcription GATA-1 et p53. L'analyse de mutants tronqués nous a permis de déterminer les domaines minimum requis pour cette interaction : le DBD de GATA-1 lie le TAD de p53. Les études de liaison *in vitro* nous ont révélé que l'affinité de cette interaction est de l'ordre du nanomolaire. Dans des cellules érythrocytaires, la conséquence principale de cette interaction est l'inhibition de l'activité de p53 par GATA-1. À la lumière de ces résultats, nous suggérons que cette inhibition de la voie p53 par GATA-1 peut être essentielle pour le développement et la survie des globules rouges.

Les résultats de ce chapitre sont exposés sous la forme d'un article scientifique publié dans la revue *Blood* en juillet 2009. Lors de ces travaux, Patrick Archambault a aidé à la production et à la purification de protéines. Paola Di Lello a été impliquée dans l'acquisition de spectres RMN. J'ai procédé à la production et la purification d'échantillons protéiques non-marqués et marqués. Puis j'ai réalisé les études par spectroscopie RMN ainsi que les expériences d'ITC. Les études *in vivo* ont été effectuées par notre collaboratrice, le Dr. Cecelia Trainor. Ayant participé aux aspects expérimentaux et théoriques de cette étude, ainsi qu'à la rédaction du manuscrit, ma contribution à ces travaux fût significative. C'est pourquoi je suis co-premier auteur sur cet article.

Chapitre II

GATA-1 interagit avec et inhibe p53

Article 1: GATA-1 associates with and inhibits p53.

Cecelia D. Trainor*, Caroline Mas*, Patrick Archambault, Paola Di Lello and James G. Omichinski.

Reprinted with permission from : Cecelia D Trainor, Caroline Mas, Patrick Archambault, Paola Di Lello and James G. Omichinski.

This research was originally published in *Blood*. Cecelia D Trainor*, Caroline Mas*, Patrick Archambault, Paola Di Lello and James G. Omichinski. GATA-1 associates with and inhibits p53

Reproduced with permission from *Blood*. 2009 Jul 2;114(1):165-73© the American Society of Hematology

GATA-1 associates with and inhibits p53.

Cecelia D Trainor^{1*}, Caroline Mas^{2*}, Patrick Archambault², Paola Di Lello² and James G. Omichinski².

¹Laboratory of Molecular Biology, NIDDK, National Institutes of Health Building 5, Room 208, Bethesda, MD 20892, U.S.A. ²Département de Biochimie, Université de Montréal C.P. 6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada.

* These two authors contributed equally to this publication

Correspondence: Cecelia D. Trainor, Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Building 5, Room 208, Bethesda, MD 20892 and James G. Omichinski, Département de Biochimie, Université de Montréal C.P. 6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada.

Abstract

In addition to orchestrating the expression of all erythroid-specific genes, GATA-1 controls the growth, differentiation and survival of the erythroid lineage through the regulation of genes that manipulate the cell cycle and apoptosis. The stages of mammalian erythropoiesis include global gene inactivation, nuclear condensation and enucleation to yield circulating erythrocytes, and some of the genes whose expression are altered by GATA-1 during this process are members of the p53 pathway. Here, we demonstrate a specific *in vitro* interaction between the transactivation domain of p53 (p53TAD) and a segment of the GATA-1 DNA-binding domain (DBD) that includes the carboxyl-terminal zinc-finger domain. We also show by immunoprecipitation that the native GATA-1 and p53 interact in erythroid cells and that activation of p53-responsive promoters in an erythroid cell line can be inhibited by the over-expression of GATA-1. Mutational analysis reveals that GATA-1 inhibition of p53 minimally requires the segment of the GATA-1 DBD that interacts with p53TAD. This inhibition is reciprocal, as the activation of a GATA-1 responsive promoter can be inhibited by p53. Based on these findings, we conclude that inhibition of the p53 pathway by GATA-1 may be essential for erythroid cell development and survival.

Introduction

GATA-1 is the founding member of the GATA factor family of DNA binding proteins [1], and is critical to the development of the erythroid [2, 3], megakaryocyte [4], mast [5], and eosinophil [6] cell lineages. It is multifunctional and acts as a global regulator of erythroid-specific genes [7]. It is both an activator and repressor of transcription [8, 9] and participates in long-range gene regulation through the β -globin locus control region (LCR) [10, 11] and the c-kit locus [12]. GATA-1 alters chromatin structure; it is a substrate for acetylation [13, 14] sumolation [15] and ubiquitination [16], and recruits the corresponding activities to chromatin.

The GATA-1 DNA-binding domain (DBD) consists of two zinc-finger domains (C-X₂-C-X₁₇-C-X₂-C) and is highly conserved among the six vertebrate GATA factors. The carboxyl-terminal zinc-finger (GATA-1 CF) constitutes the primary DBD and is capable of independent high-affinity binding [9] to the consensus motif A/TGATAA/G [17]. The amino-terminal zinc-finger domain (GATA-1 NF) binds independently to a related sequence containing a GATC core with lower affinity [18, 19], and stabilizes binding to DNA with multiple GATA sites [9, 20, 21]. In addition to their role in DNA binding, both zinc fingers participate in protein-protein interactions [22, 23] with multiple partners, including Friend-of-GATA-1 (FOG-1) [23] and PU.1 [24]

GATA-1 is required for generation of erythroid cells [2, 3] and GATA-1 negative embryonic stem (ES) cells have been used to demonstrate that definitive erythroid cells lacking GATA-1 are blocked at the proerythroblast stage of development [25]. Unlike wild-type cells, GATA-1 null definitive erythroid progenitors die of apoptosis, leading to the conclusion that GATA-1 controls the survival of erythroid cells in addition to contributing to growth and differentiation. GATA-1 stimulates the expression of Bcl-xL [26], an anti-apoptotic Bcl-2 family member that is critical to the erythroid lineage [27] and Bcl-xL likely contributes to increased cell survival seen in the presence of GATA-1. This was confirmed in studies with the *GATA-1.05* mouse [28] where it was shown that 5% of the normal level of GATA-1 is sufficient to stop the apoptosis of day 11 definitive erythroid precursors and to increase the expression of Bcl-xL [29]. The observed apoptosis in the absence of GATA-1 occurs without the accumulation of the tumor suppressor protein p53, through an undetermined mechanism [25, 29].

The process of terminal erythroid differentiation is one that might induce the p53 pathway since DNA breakage accompanies nuclear condensation and enucleation. p53 is activated by chromosome instability and by DNA damage, however the p53 pathway remains largely inactive during erythropoiesis even though apoptosis is an integral part of the process [30]. p53 accumulates at low levels during nuclear apoptosis in the normoblast stage of erythropoiesis just prior to enucleation, but surprisingly the cells survive [31]. A modified p53 response rather than the classical pathway is observed, concomitant with a reduction of caspase 3/7, the ultimate agent of apoptosis. Recently, erythroid conditional knock-outs of the inhibitory p53 partner protein genes Mdm-2 and Mdm-4 in wild type and in p53^{-/-} mice have shown unequivocally that a competent p53 pathway is present and is induced in erythroid cells in the absence of these proteins [32]. Mdm-2 and -4 are clearly involved in preventing the p53 response, but there are likely other factors since a specialized p53 response is observed (in normoblasts).

The last few cell divisions are critical in erythroid differentiation, because a controlled exit from the cell cycle without induction of apoptosis is essential. GATA-1 regulates cell cycle progression in a concentration and context dependent manner, and can both block and accelerate the cell cycle (reviewed in) [33]. The restoration of GATA-1 to the GATA-1 null erythroid cell line, G1E, altered the expression of many cell cycle regulators in a manner consistent with the induction of cell cycle arrest [33]; at least one of these, c-myc, was shown to be directly inhibited by GATA-1 [33]. The expression of the p21 gene, which causes cell cycle withdrawal, was stimulated by GATA-1 in a murine erythroleukemia cell line (MEL) [34]. p53 also induces cell cycle arrest in part by regulating the same genes. It directly inhibits the transcription of the proliferation inducing oncogene, c-myc [35], and activates the p21 gene, a primary component of the p53 response. In other instances, GATA-1 and p53 appear to have opposing functions, as GATA-1 stimulates the anti-apoptotic Bcl-xL gene, but p53 induces several pro-apoptotic Bcl-2 family members such as Bax [36] and Noxa [37] and all function in the same intrinsic apoptotic pathway. In addition, the tumor suppressor RB interacts with both the p53-Mdm-2 complex [38] and with GATA-1 [39]. Thus several proteins that are prominent in GATA-1 function are also participants in the p53 pathway. The observation that the p53 transactivation domain (TAD) has similar amino acid composition (high in acidic and hydrophobic residues) to the TAD of PU.1, which interacts with the GATA-1 DBD resulting in inhibition of GATA-1 [24], caused us to

consider that a similar mechanism might explain the interplay between the GATA-1 and p53 pathways.

Here, we demonstrate a specific interaction between the TAD of p53 (p53TAD) and a segment of GATA-1 that includes the linker between the zinc fingers and the carboxyl-terminal zinc-finger domain (GATA-1 L+CF). We also show that the native proteins interact and can be co-immunoprecipitated from MEL cell extract. Furthermore, the activation of p53-responsive promoters in an erythroid cell line are inhibited by over-expression of GATA-1 but not by the expression of a GATA-1 mutant that lacks a region required for binding p53. The inhibition is reciprocal, as the activation of a GATA-1 responsive promoter can be inhibited by p53.

Materials and Methods

Clones.

The p53 Glutathione-S-Transferase (GST)-fusion constructs were provided by Dr. Rong Li. The GATA-1 GST-fusion constructs were prepared by inserting hGATA-1 fragments into the pGEX-5X (GE Healthcare) vector. The GATA-1₁₉₉₋₃₁₇ and GATA-1₂₅₂₋₃₁₇ expression constructs were prepared by inserting the fragments of hGATA-1 into the pet11d (Novagen) vector. Mutations were made by site-directed mutagenesis and sequenced.

GST-fusion proteins.

The GST-GATA-1 fragments were expressed in *E. coli* strain BL21(DE3) (Novagen). The cells were lysed, centrifuged at 100,000 g and incubated with glutathione (GSH)-sepharose resin (GE Healthcare). GST-GATA-1 fragments were eluted off the resin with GSH (Sigma, MO). The proteins were dialyzed into storage buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl and 20 % glycerol), and rebound to fresh GSH-sepharose for binding studies.

Protein purification.

The p53 [40] and GATA-1 [21] fragments were purified as previously described.

Binding Assays.

GST-pull downs: 1 μM of GST or GST-GATA-1 fragment was incubated with the resin in 0.5X Superdex. Purified p53 fragments were then (0.01 μM -1 μM) incubated for 1 hour. Following washing, the bound protein was eluted and loaded onto a 15% acrylamide gel and separated by electrophoresis in Tris-Glycine. The proteins were then transferred to an Immobilon-P membrane (Millipore) and the membrane incubated overnight in TBS/Blotto. The membrane was then incubated with an anti-p53 primary antibody (p53 DO-1 or Pab 1801; Santa Cruz) and with a secondary antibody (goat-antimouse IgG-HRP; Santa Cruz). The protein was detected through chemiluminescence using ECL Plus (GE Healthcare). The concentrations of GST-fusion proteins were determined from A_{280} and the membrane was Coomassie stained to ensure equal inputs.

Isothermal Titration Calorimetry binding experiments: ITC titrations were performed essentially as described [40] (For more details see supplementary materials and methods).

Immunoprecipitations.

MEL cell nuclei were isolated [41] and nuclear extract was made as described [41, 42]. 320 μg of extract and 40 μl of Protein G Plus agarose precleared with anti-mouse IgG (Pierce 31168) was used for each IP. Samples were rotated at 4°C for 4 hours followed by centrifugation at 1000xg at 4°C. The agarose pellets were washed in PBS and suspended in SDS Page loading buffer with 0.1% βME and heated to 100°C for 10 min. GATA-1 antibodies (N6) sc265 and (M20) sc1234; p53 antibodies pab 240, 246 and FL393; cyclin B1 (GNS1) antibody, GAPDH sc47724, anti rat IgG HRP sc2032 and the protein G plus agarose were obtained from Santa Cruz Biotechnologies. The anti-rat IgG2a Clone R2A-2 antibody was obtained from Sigma. Denatured protein samples were electrophoresed and blotted with Tris glycine gels using PVDF membrane. 3 % milk in PBS was used for blocking and antibody binding and blots were washed with PBS with 0.1 % Tween 20. Blots were developed with SuperSignal West Dura (Pierce).

DNA Constructs.

FR 7Luc, a GATA responsive reporter gene, was made by inserting a blunt end 470 bp HpaII to XmnI restriction fragment from the chicken folate receptor gene promoter into the SmaI site of pGL3 basic. This fragment contains two GATA sites. [43] The GATA-1 mutant CM was made by site directed mutagenesis. Amino acids KKR₍₂₄₅₋₂₄₇₎ and K_(314 to 316) of human GATA-1 were converted to alanine and sequenced.

Cell Culture and Transfections.

MEL cells were grown in IMDM supplemented with 10% FBS and 1% PS. 6C2 cells were grown in α MEM with 10 % FBS and 2 % chicken serum, 0.5 mM β ME, 1 mM HEPES and 1 % PS. Cells were maintained at 37°C and 5 % CO₂. 6C2 cells were transfected with Effectene (Qiagen) according to manufacturers instructions and harvested 18-24 hours later. 5×10^6 cells per sample were utilized with 150 η g of control plasmid (CMV Renilla Luciferase), 500 η g of reporter plasmid p53 Luc (Stratagene) or pBAX Luc and 250-500 η g of CMV p53 (in pLPC), while 0.5-3 μ g of GATA-1 expression vectors or control vectors were used. The Dual Luciferase Assay Kit (Promega) was used to process and analyze the samples according to manufacturer's instructions. Luciferase activity was determined with a Victor Multilabel Counter from Perkin Elmer. QT6 cells were transfected with Lipofectamine or Lipofectamine 2000 (Invitrogen). The Dual Luciferase Assay was utilized. All transfections were normalized to CMV renilla luciferase, by setting the highest value for the positive control to 100% and adjusting other values in proportion.

NMR Studies.

NMR spectra were collected at 300 K on a Varian 600 MHz spectrometer. The GATA L + CF / DNA complex consisted of a 1:1 molar ratio of ¹⁵N-labeled GATA L + CF and unlabeled double stranded DNA (5' GTTGCAGATAAACATT) at 0.5 mM as previously described [44]. For studies of the p53TAD, the sample consisted of ¹⁵N-labeled p53 at 0.5 mM (pH=6.5, 12.5 mM NaCl, 90% H₂O/10% D₂O). For binding of the GATA L + CF / DNA complex with p53TAD, we performed two titrations. The first titration consisted of a 0.5 mM GATA L +CF (¹⁵N-labeled)/DNA complex and p53TAD (unlabeled) was added in incremental amounts to a 1:1

ratio. The second titration consisted of 0.5 mM of labeled p53TAD and the complex of GATA L +CF (unlabeled)/DNA (unlabeled) was added in incremental amounts to a 1:1 ratio.

Results

The zinc-finger domains of GATA-1 bind the full length TAD of p53 *in vitro*.

The TAD of p53 (p53TAD) is located within its amino terminus (residues 1-73). p53TAD contains two subdomains, p53TAD1 (residues 1-40) and p53TAD2 (residues 40-73), that are each capable of independently activating transcription (Figure 2. 1A). Given the similarity between the TADs of p53 and PU.1, we tested if the zinc-finger domains of GATA-1 interacted with the TAD of p53. A GST-fusion protein containing the two zinc-finger domains of GATA-1 (GATA-1 ZFD; GATA-1₍₁₉₉₋₃₁₇₎) was incubated with increasing concentrations of purified fragments of p53 corresponding to p53TAD (p53₍₁₋₇₃₎) and p53TAD1 (p53₍₁₋₄₀₎) in a GST pull-down assay. GST-GATA-1 ZFD bound to p53TAD but not to p53TAD1 and GST alone did not interact with either p53TAD or p53TAD1 (Figure 2. 1B). Next, we tested whether a fragment of p53 containing p53TAD2 (p53₍₂₀₋₇₃₎) interacted with the GST-GATA-1 ZFD. Again, GST-GATA-1 ZFD bound to p53TAD but not to the p53TAD2 subdomain (Figure 2. 1C). Based on the GST pull-down assay, both subdomains of the TAD of p53 (p53TAD) are necessary and sufficient for interaction with the ZFDs of GATA-1 (GATA-1 ZFD).

Identification of the minimal p53-interacting region of GATA-1

After establishing that both of the subdomains of the TAD of p53 (p53TAD) were required for interaction with the zinc-finger domains of GATA-1 (GATA-1 ZFD), we tested the ability of the individual zinc-finger domains of GATA-1 to interact with p53TAD. Three GST-fusion proteins containing different combinations of the zinc-fingers domains of GATA-1 were tested with p53TAD (Figure 2. 2A). The three constructs included: 1) The amino-terminal zinc finger plus the linker region (GATA-1 NF + L; GATA-1₍₂₀₀₋₂₅₁₎); 2) The linker region plus the complete carboxyl-terminal zinc finger domain (GATA-1 L + CF; GATA-1₍₂₂₈₋₃₁₇₎); 3) The carboxyl-terminal zinc finger (GATA-CF; GATA-1₍₂₅₂₋₃₁₇₎). First, GST-GATA-1 NF + L was incubated with p53TAD and the binding compared to GST-GATA-1 ZFD. This experiment

revealed that the GATA-1 NF + L was not sufficient for the interaction with p53TAD (Figure 2. 2B). Next, we tested the GATA-1 L + CF fragment. Incubation of p53TAD with GST-GATA-1 L + CF resulted in similar binding to p53TAD as observed for GATA-1 ZFD (Figure 2. 2C). Lastly, GATA-CF (GATA-1₍₂₅₂₋₃₁₇₎) was tested for its ability to bind to p53TAD. The results demonstrated that GATA-1 CF was not sufficient for binding p53TAD (Figure 2. 2D). Based on the results in the pull-down assay, the binding of p53TAD to GATA-1 minimally requires the linker region and the carboxyl-terminal zinc finger of GATA-1 (GATA-1 L +CF).

GATA-1 and p53 interact in erythroid cells.

To determine whether the native p53 and GATA-1 proteins physically associate with one another in erythroid cells, immunoprecipitations were carried out with MEL cell nuclear extract and antibodies against GATA-1, p53, cyclin B1 and IgG. The immunoprecipitates were analyzed by Western blotting, and the antibodies themselves were run in adjacent lanes of the gels in addition to molecular weight markers. There was a single GATA-1 band in the MEL cell extract that migrated at 50 kd with both GATA-1 antibodies (N6 and M20), and the antibodies ran at slightly higher molecular weights. GATA-1 is precipitated with both the GATA-1 and p53 antibodies, but not with anti IgG2 (Figure 2. 3A). The immunoprecipitations were repeated using anti-cyclin B1 antibody followed by anti-rat IgG2 as the negative control, and probing the blots with an antibody reactive to a different epitope of GATA-1 (M20 Santa Cruz, data not shown) or an HRP anti-p53 antibody (Figure 2. 3B). GATA-1 and p53 again co-immunoprecipitate, strongly suggesting that they associate with one another in MEL cells.

GATA-1, p53 and GATA DNA do not form a ternary complex *in vitro*.

The binding site for the p53TAD on GATA-1 overlaps the minimal DNA binding domain of GATA-1. To test if p53, DNA and GATA-1 L +CF are capable of forming a ternary complex, we performed two sets of NMR experiments. To a sample containing ¹⁵N-labeled GATA-1 L + CF (0.5 mM), we added a stoichiometric amount of unlabeled AGATAA containing DNA (0.5 mM) and recorded a 2D ¹H -¹⁵N HSQC spectra (Figure 2. 4A, black signals). To the sample containing the ¹⁵N-GATA-1 L + CF/ DNA complex, we then added unlabeled p53TAD (0.5 mM) and we observed that there were no differences in the ¹H and ¹⁵N resonances of ¹⁵N-GATA-1 L +

CF (Figure 2. 4A, red signals). This result indicates that under these experimental conditions, there does not appear to be a ternary complex involving GATA-1, DNA and p53. To further test for the presence of a ternary complex, we performed a second set of NMR experiments. First, we recorded a 2D ^1H - ^{15}N HSQC spectra of a sample containing ^{15}N -labeled p53TAD (0.5 mM) (Figure 2. 4B, black signals). To the sample containing the ^{15}N -p53TAD, we then added the unlabeled GATA-1 L +CF / DNA complex (0.5 mM) and again we did not observe any differences in the ^1H and ^{15}N resonances of ^{15}N -p53TAD (Figure 2. 4B, red signals). These results further support the conclusion that under the NMR conditions, there does not appear to be a ternary complex involving GATA-1, DNA and p53. These experiments suggest that p53TAD and the AGATAA DNA compete for binding to GATA-1 L + CF and that GATA-1 L + CF binds tighter to the DNA than to the p53TAD. This is consistent with the fact that GATA-1 L +CF binding to DNA is approximately 10-fold tighter than to the p53TAD.

GATA-1 and p53 interfere with one another.

To determine whether the p53 and GATA-1 proteins can influence the activity of one another, transactivation assays were performed. A p53 responsive reporter gene (p53 Luc) can be activated in 6C2 cells by the expression of CMV p53 (Figure 2. 5A, Bars 1 and 2). The co-expression of RSV-Cat or the inclusion of empty expression vector have little affect on p53 activation, but the addition of RSV hGATA-1 inhibits this activation in a dose dependent manner (Figure 5A-B). Protein was prepared from a portion of the transfected cells and subjected to Western blotting with anti-p53 antibody to show that the inhibition of p53 activity seen in these experiments is not due to repression of p53 expression by GATA-1 (Figure 2. 5C). We observed high levels of p53 even with the highest concentration of GATA-1 expression vector, and GAPDH levels are comparable among all samples. Similar results were obtained in K562 cells, and p53 can also inhibit the activation of a GATA-1 responsive reporter gene in transactivation assays (data not shown). Under similar conditions, GATA-1 did not inhibit the dexamethasone-induced activation of a GRE-responsive reporter gene by the glucocorticoid receptor indicating that the inhibition is specific (data not shown). Furthermore, addition of erythropoietin to K562 cells caused a slight decrease in p53Luc activation that correlated with a slight increase in endogenous GATA-1 levels (Supplementary Figure 2. 1) and this result is consistent with the results seen following the addition of RSV hGATA-1.

While inhibition in the transactivation assay is reciprocal between the p53 and GATA-1 proteins, we hypothesize that p53 inhibition of GATA-1 does not serve an important biological function in red cell development. GATA-1 levels far exceed p53 levels at all stages of red cell development, and p53 null mice have normal erythroid cells [45, 46], although possible compensation by p53 homologs p63 and p73 has not been ruled out. Consequently only the inhibition of p53 by GATA-1 has been studied further.

To determine which regions of GATA-1 are required for the inhibition of p53, a series of deletion mutants of mouse GATA-1 in pEF-BOS (Figure 2. 6A) were tested in the transactivation assay in 6C2 cells. Wild-type GATA-1 and a construct missing the amino-terminal zinc-finger domain (Δ NF) inhibit p53 activity to similar extents (Figure 2. 6B). In contrast, constructs missing the carboxyl-terminal zinc finger (Δ CF) or the amino-terminal finger and the linker region (Δ NFL) do not (Figure 2. 6B) indicating that the carboxyl-terminal zinc finger and the linker between the two zinc fingers are sufficient for the inhibition of p53 activity, but the amino-terminal zinc-finger domain is not required. The Δ NFL mutant is capable of activation of FR7Luc, a GATA-1 responsive reporter gene, in QT6 fibroblasts (and in 6C2 cells, data not shown) whereas the Δ CF mutant is not (Figure 2. 6C). This is consistent with the fact that the Δ CF mutant lacks the major DNA-binding domain of GATA-1. The activation levels seen with the Δ NFL mutant are similar to wild type GATA-1 indicating that it accumulates stably in cells. Thus the Δ NFL mutant can transactivate a GATA responsive reporter (belowC) and bind to DNA (data not shown) and yet does not inhibit transactivation by p53.

The amino-terminal zinc finger of GATA-1 is not required for p53 inhibition, but is critical to other GATA-1 functions, so we restored it to the construct in two forms. Amino acids 196 to 232 were restored to the Δ NFL mutant to generate the Δ NL mutant, and this protein does not inhibit p53 (Supplementary Figure 2. 2A-B). A construct restoring 196 to 234 to the Δ NFL mutant to generate the Δ NL2 mutant (Supplementary Figure 2. 2A) was also tested and this mutant inhibits p53 but less efficiently than Δ NFL (data not shown) suggesting that amino acids 233 and 234 may be involved but are not sufficient for interaction with p53. Both the Δ NL and Δ NL2 mutants can transactivate a GATA responsive reporter (Supplementary Figure 2. 2C) and bind to the GATA cofactor, FOG-1 (data not shown).

Plasmids with the promoters of p53 responsive genes driving luciferase reporter genes were obtained and assayed in 6C2 cells. Consistent with our results above, both the BAX and p21 promoters are partially inhibited by wild type GATA-1, but not by the Δ NFL mutant (data not shown). The observation that the Δ NF mutant inhibits p53 while the Δ NFL mutant does not demonstrates a strict requirement for the linker between the GATA-1 zinc-finger domains for p53 inhibition and these results correlate perfectly with results of our GST-pull down assay (above). The linker region and the carboxyl-terminal zinc finger are required both for p53TAD binding and inhibition of p53 activity.

GATA-1 L + CF and p53TAD interact strongly and specifically *in vitro*.

To compare the binding affinity of the two subdomains within p53TAD for the zinc-finger domains of GATA-1, we determined dissociation constants (K_D) by isothermal titration calorimetry (ITC) (Supplementary Figure 2. 2). By ITC, p53TAD (p53₍₁₋₇₃₎) bound GATA-1 ZFD (GATA-1₍₁₉₉₋₃₁₇₎) with apparent K_D values of 77 ± 16 nM and GATA-1 L + CF (GATA-1₍₂₂₈₋₃₁₇₎) with an apparent K_D of 67 ± 7 nM (Supplementary Figure 2. 2). In contrast, we were not able to see any interaction between either p53TAD1 (p53₍₁₋₄₀₎) or p53TAD2 (p53₍₄₀₋₇₃₎) with either GATA-1 L + CF ((Supplementary Figure 2. 2) or GATA-1 ZBD by ITC. The ITC results are consistent with our GST-pull down and transactivation assays. Taken together the results indicate that the minimal region of GATA-1 required for the interaction with p53TAD and for the inhibition of p53 activity is the linker region plus the carboxyl-terminal zinc-finger domain.

The inhibition of p53 activity by GATA-1 requires the interaction of p53TAD with GATA-1.

A competition assay was developed to examine the mechanism of inhibition of p53 by GATA-1. Two truncation mutants of p53 producing p53₁₋₆₅ or p53₁₋₄₀ were constructed in CMVp53. These truncation mutants that no longer activate p53-responsive reporters (data not shown), were expressed along with wild-type p53 and GATA-1 in 6C2 cells. In the presence of wild type GATA-1, p53₁₋₆₅ increased the expression of p53 Luc to 80 % of the levels seen with p53 alone, while p53₁₋₄₀ did not (Figure 2. 7 bars 1-6). A GATA-1 mutant (GATA-1 CM, Material and Methods) that does not inhibit p53 (data not shown) was used as a control and

neither p53 fragment affected the activation of p53Luc by p53 in the presence of this mutant GATA-1 (Figure 2. 7 bars 1, 7, and 8). These p53 fragments have no effect on the activity of p53 in the absence of GATA-1, and GATA-1 does not affect the activity of p53 Luc in the absence of p53 (data not shown). Taken together these results indicate that p53₁₋₆₅ competes with wild-type p53 for GATA-1, and this reduces the inhibitory affect of GATA-1; p53₁₋₄₀, which does not bind to GATA-1, did not compete. Thus the interaction of p53₁₋₆₅ with GATA-1 is essential for the inhibition of p53.

Discussion

The observation that GATA-1 interacts strongly with p53 *in vitro* caused us to consider that GATA-1 directly influences p53 and that there may be interdependency between these two regulatory networks in erythroid cells. We have shown that GATA-1 is a potent inhibitor of p53 transactivation in erythroid precursor cells and that this inhibition requires a region of the GATA-1 DNA-binding domain critical to a high affinity interaction with the TAD of p53. This inhibition is competed by p53 fragments that bind to GATA-1 but not by ones that do not. Mutants of GATA-1 that do not inhibit p53 transactivation can still bind to DNA and to FOG-1, and transactivate GATA responsive reporter genes like native GATA-1.

The region of the GATA-1 DBD required for p53 inhibition correlates with the region required for binding to the p53TAD. Residues 228 to 317 of GATA-1 are sufficient for maximal binding to p53 and significantly lower binding is seen when residues 228-251 from the linker of GATA-1 are deleted (Figure 2. 2D). In the transactivation assay, the GATA-1 Δ NF mutant (removal of residues 196 to 233) still inhibits p53. This mutant removes the cysteine residues of the amino-terminal zinc finger but retains most of the linker region including several residues that are part of the helix associated with the amino-terminal zinc finger. In contrast, the GATA-1 Δ NL and Δ NL2 mutants, which contain the amino-terminal zinc-binding domain but lack either residues 234-254 or residues 236-254 from the linker, do not inhibit p53 activity efficiently. Residues within the linker between the two zinc fingers of GATA-1 are critical for both the *in vitro* binding and the inhibition of p53 transactivation.

All GATA-1 DNA binding mutants tested thus far are defective in the inhibition of p53 transactivation. This is consistent with the requirement for the complete carboxyl-terminal zinc finger of GATA-1 (the major DBD) for p53 binding, and does not necessarily indicate that DNA binding by GATA-1 is required for the inhibition of p53. Our NMR studies indicate that the p53TAD does not disturb the GATA-1/DNA complex *in vitro*, suggesting that a ternary complex of the three molecules does not occur (Figure 2. 4). This is consistent with the fact that the affinity of the GATA-1 DNA interaction is 10-fold higher than that of GATA-1 and p53TAD, and there is overlap between the regions of GATA-1 required for its interaction with DNA and with p53. We do not know if a ternary complex forms at p53 DNA-binding sites, but this is plausible since the p53 DBD is not required for the interaction with GATA-1. It is important to remember that GATA-1 is present at very high concentration in erythroid cells and could either sequester p53 from its targets, or bind to p53 at its targets and inactivate it. Further experiments are required to determine if GATA-1 and p53 are able to form a ternary complex on a p53 DNA-binding site.

While the inhibition of p53 by GATA-1 has been seen in two types of erythroid cells, we have not observed it in non-erythroid cells and thus cell-type specific modifications of either GATA-1 or p53 may be important. Alternately, the cell type specificity may be the result of a significantly different balance between the concentrations of the two proteins in other cell types. Both proteins can be toxic at high concentrations, and it is difficult to express GATA-1 at the normally high levels that occur in red blood cells in non-hematopoietic cells. High levels of GATA-1 may be required to prevent cell death induced by ectopically expressed p53, and this may be unattainable except in an erythroid cell. In this case, the inhibition of p53 by GATA-1 would be an erythroid cell specific event.

GATA-1 has been linked to leukemias in mice and humans. GATA *1.05* knockdown mice, producing 5% of normal levels of GATA-1, develop erythroleukemia at 3-5 months of age, and B-cell leukemia at 1 year; *1.05* fetal erythroid cells are unable to differentiate due to lack of GATA-1, but escape the apoptosis seen in GATA-1 null cells and are hyperproliferative. [47] After birth, secondary mutations in this unchecked cell population lead to overt transformation and disease; the inhibition of p53 by GATA-1 observed here could facilitate this cell expansion.

In humans GATA-1 mutations are associated with transient myeloproliferative disorder (TMD) and with acute megakaryoblastic leukemia (AMKL) in Down's Syndrome (DS): both diseases occur frequently in DS patients. [22] TMD and AMKL patients produce GATA-1s, a shorter form of GATA-1 lacking the N-terminal transactivation domain (NTAD) that is translated from an internal ATG: GATA-1s is a critical factor in these diseases. DS TMD is usually a transient disease but 20-30% of cases develop into DS AMKL derived from the original clone of diseased cells with identical mutations in GATA-1. [22] In several experimental systems, substitution of GATA-1s for the native protein has resulted in transient hyperproliferation of fetal megakaryocyte progenitors, but leukemia was not observed consistent with the requirement for trisomy 21. [22, 47] It is thought that during DS fetal development, mutations producing GATA-1s occur at high incidence leading to expansion of megakaryocyte progenitors to generate TMD. These cells undergo secondary mutation leading to transformation and AMKL. [22] The p53 tumor suppressor pathway is evidently unable to prevent the accumulation of these cells, but p53 mutations are observed in approximately 21 % of cases and are thus not prerequisite for DS AMKL [48]. GATA-1s has the complete DNA-binding domain and should interact with p53 as native GATA-1. It would be interesting to investigate whether the interaction between GATA-1 and p53 described here prevents the p53 response, allowing the expansion of this cell type. The interaction could also be relevant in preventing the activation of the p53 pathway in other leukemias that express hematopoietic GATA factors. Targeting the interaction in diseased cells may be therapeutic.

Inhibition in the transactivation assay is reciprocal between the p53 and GATA-1 proteins, and our competition assay (Figure 2. 7) shows the p53TAD is involved. We propose that only the inhibition of p53 by GATA-1 is functionally important, and may be needed to prevent the induction of the p53 pathway and to avoid apoptosis and/or senescence during red cell development. To accomplish this GATA-1 regulates p53 and several proteins usually controlled by p53 to cause cell cycle withdrawal in order to facilitate enucleation and the formation of functional erythrocytes.

There is precedent for the participation of GATA-1 in this type of regulation, although the mechanism may not be identical. The association of PU.1 with GATA-1 through the DBDs of both proteins, and the PU.1 TAD also results in a reciprocal inhibition of activity, which specifies

the choice of either the myeloerythroid or myelolymphoid lineages [24, 42, 49, 50]. Each transcription factor drives cells towards a specific hematopoietic potential, while simultaneously inactivating the master regulator that specifies alternate lineages. The similarity of the TADs of p53 and PU.1 suggests that both the transactivation and the inhibitory functions of this domain have been conserved between the two proteins.

Acknowledgements

This work was supported by Canadian Institutes of Health Research grant MOP-135604 (J.G.O.) and by NIDDK, NIH. The authors thank Gary Felsenfeld for advice and support. We thank Dr. Masayuki Yamamoto for mouse GATA-1 expression clones, and Dr. Sam Lee for reporter plasmids. We thank Stoney Simons for a glucocorticoid receptor expression vector and reporter gene. We thank Gerardo Ferbeyre for helpful discussions and the p53 expression vectors. We thank Connie Noguchi for advice and discussions.

Authorship and Conflict of interest Statements: **C.T.**- designed research, performed research, contributed vital new reagents, analyzed data and wrote the paper. **C.M.**- designed research, performed research, contributed vital new reagents, analyzed data and wrote the paper. **P.A.**- performed research and contributed vital new reagents. **P.D.**- designed research and performed research. **J.G.O.**- designed research, performed research, contributed vital new reagents, analyzed data and wrote the paper. The authors declare no competing financial interests.

References

1. Tsai, S.F., et al., Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature*, 1989. 339(6224): p. 446-51.
2. Fujiwara, Y., et al., Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci U S A*, 1996. 93(22): p. 12355-8.
3. Pevny, L., et al., Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, 1991. 349(6306): p. 257-60.

4. Shivdasani, R.A., et al., A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *Embo J*, 1997. 16(13): p. 3965-73.
5. Migliaccio, A.R., et al., GATA-1 as a regulator of mast cell differentiation revealed by the phenotype of the GATA-1^{low} mouse mutant. *J Exp Med*, 2003. 197(3): p. 281-96.
6. Yu, C., et al., Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage in vivo. *J Exp Med*, 2002. 195(11): p. 1387-95.
7. Welch, J.J., et al., Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood*, 2004. 104(10): p. 3136-47.
8. Letting, D.L., et al., Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc Natl Acad Sci U S A*, 2004. 101(2): p. 476-81.
9. Martin, D.I. and S.H. Orkin, Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes Dev*, 1990. 4(11): p. 1886-98.
10. Pal, S., et al., Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc Natl Acad Sci U S A*, 2004. 101(4): p. 980-5.
11. Vakoc, C.R., et al., Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell*, 2005. 17(3): p. 453-62.
12. Jing, H., et al., Exchange of GATA Factors Mediates Transitions in Looped Chromatin Organization at a Developmentally Regulated Gene Locus. *Mol Cell*, 2008. 29(2): p. 232-42.
13. Boyes, J., et al., Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, 1998. 396(6711): p. 594-8.
14. Hung, H.L., et al., CREB-Binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites. *Mol Cell Biol*, 1999. 19(5): p. 3496-505.
15. Collavin, L., et al., Modification of the erythroid transcription factor GATA-1 by SUMO-1. *Proc Natl Acad Sci U S A*, 2004. 101(24): p. 8870-5.
16. Hernandez-Hernandez, A., et al., Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *Embo J*, 2006. 25(14): p. 3264-74.
17. Evans, T., M. Reitman, and G. Felsenfeld, An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc Natl Acad Sci U S A*, 1988. 85(16): p. 5976-80.
18. Newton, A., J. Mackay, and M. Crossley, The N-terminal zinc finger of the erythroid transcription factor GATA-1 binds GATC motifs in DNA. *J Biol Chem*, 2001. 276(38): p. 35794-801.
19. Ghirlando, R. and C.D. Trainor, Determinants of GATA-1 binding to DNA: the role of non-finger residues. *J Biol Chem*, 2003. 278(46): p. 45620-8.

20. Tsai, S.F., E. Strauss, and S.H. Orkin, Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes Dev*, 1991. 5(6): p. 919-31.
21. Trainor, C.D., et al., A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction. *Mol Cell Biol*, 1996. 16(5): p. 2238-47.
22. Crispino, J.D., GATA1 in normal and malignant hematopoiesis. *Semin Cell Dev Biol*, 2005. 16(1): p. 137-47.
23. Tsang, A.P., et al., FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*, 1997. 90(1): p. 109-19.
24. Zhang, P., et al., Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci U S A*, 1999. 96(15): p. 8705-10.
25. Weiss, M.J. and S.H. Orkin, Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc Natl Acad Sci U S A*, 1995. 92(21): p. 9623-7.
26. Gregory, T., et al., GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood*, 1999. 94(1): p. 87-96.
27. Motoyama, N., et al., Massive cell death of immature hematopoietic cells and neurons in Bcl-x-deficient mice. *Science*, 1995. 267(5203): p. 1506-10.
28. Takahashi, S., et al., Role of GATA-1 in proliferation and differentiation of definitive erythroid and megakaryocytic cells in vivo. *Blood*, 1998. 92(2): p. 434-42.
29. Pan, X., et al., Graded levels of GATA-1 expression modulate survival, proliferation, and differentiation of erythroid progenitors. *J Biol Chem*, 2005. 280(23): p. 22385-94.
30. Testa, A., et al., Deletion of a negatively acting sequence in a chimeric GATA-1 enhancer-long terminal repeat greatly increases retrovirally mediated erythroid expression. *J Biol Chem*, 2004. 279(11): p. 10523-31.
31. Peller, S., et al., The onset of p53-dependent apoptosis plays a role in terminal differentiation of human normoblasts. *Oncogene*, 2003. 22(30): p. 4648-55.
32. Maetens, M., et al., Distinct roles of Mdm2 and Mdm4 in red cell production. *Blood*, 2007. 109(6): p. 2630-3.
33. Rylski, M., et al., GATA-1-mediated proliferation arrest during erythroid maturation. *Mol Cell Biol*, 2003. 23(14): p. 5031-42.
34. Choe, K.S., et al., Reversal of tumorigenicity and the block to differentiation in erythroleukemia cells by GATA-1. *Cancer Res*, 2003. 63(19): p. 6363-9.
35. Ho, J.S., et al., p53-Dependent transcriptional repression of c-myc is required for G1 cell cycle arrest. *Mol Cell Biol*, 2005. 25(17): p. 7423-31.
36. Miyashita, T. and J.C. Reed, Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell*, 1995. 80(2): p. 293-9.

37. Oda, E., et al., Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science*, 2000. 288(5468): p. 1053-8.
38. Xiao, Z.X., et al., Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature*, 1995. 375(6533): p. 694-8.
39. Rekhtman, N., et al., PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation. *Mol Cell Biol*, 2003. 23(21): p. 7460-74.
40. Di Lello, P., et al., Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. *Mol Cell*, 2006. 22(6): p. 731-40.
41. Andrews, N.C. and D.V. Faller, A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res*, 1991. 19(9): p. 2499.
42. Rekhtman, N., et al., Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev*, 1999. 13(11): p. 1398-411.
43. Prioleau, M.N., et al., An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. *Embo J*, 1999. 18(14): p. 4035-48.
44. Omichinski, J.G., et al., NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science*, 1993. 261(5120): p. 438-46.
45. Donehower, L.A., et al., Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, 1992. 356(6366): p. 215-21.
46. Metz, T., A.W. Harris, and J.M. Adams, Absence of p53 allows direct immortalization of hematopoietic cells by the myc and raf oncogenes. *Cell*, 1995. 82(1): p. 29-36.
47. Shimizu, R., J.D. Engel, and M. Yamamoto, GATA1-related leukaemias. *Nat Rev Cancer*, 2008. 8(4): p. 279-87.
48. Malinge, S., S. Izraeli, and J.D. Crispino, Insights into the manifestations, outcomes and mechanisms of leukemogenesis in Down syndrome. *Blood*, 2009.
49. Nerlov, C., et al., GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood*, 2000. 95(8): p. 2543-51.
50. Arinobu, Y., et al., Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. *Cell Stem Cell*, 2007. 1(4): p. 416-27.

Figures

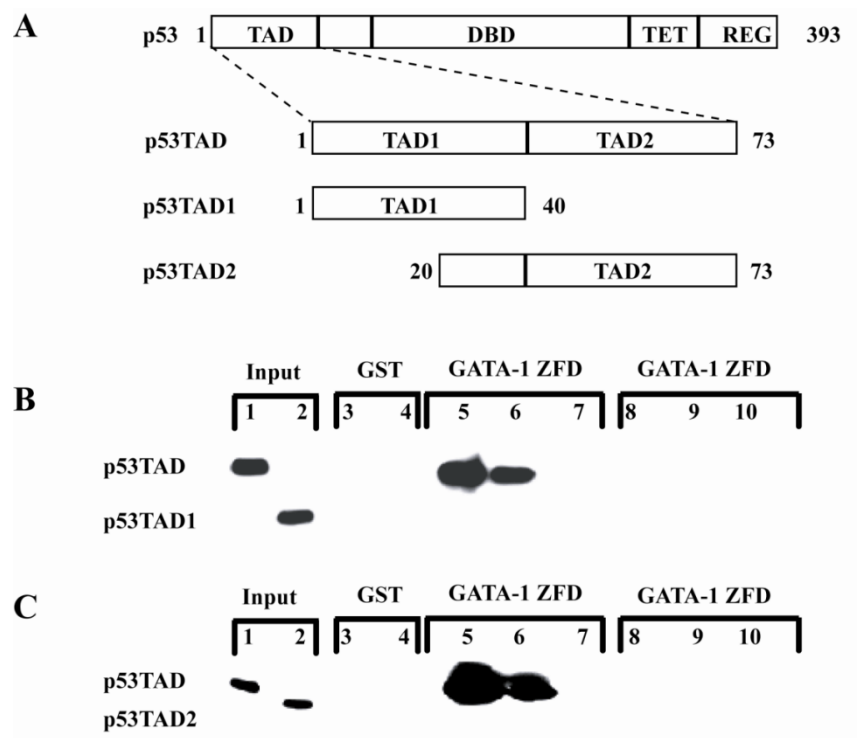


Figure 2. 1. The full-length p53TAD binds to the zinc-finger domains of GATA-1 in an *in vitro* pull-down assay. **A.** Schematic of the TAD of p53 demonstrating position of the full-length p53TAD and the location of the two independent subdomains; p53TAD1 and p53TAD2. **B.** Binding of p53TAD and p53TAD1 to GST-GATA-1 ZFD (GATA-1₍₁₉₉₋₃₁₇₎). Varying concentrations (1.0, 0.1, 0.01 μ M) of p53TAD (lanes 5-7) and p53TAD1 (lanes 8-10) were incubated with 1 μ M GST-GATA-1 ZFD. In the GST lanes, 1 μ M of purified p53TAD (lane 3) and p53TAD1 (lane 4) were incubated with 1 μ M GST as a control. The input lanes are p53TAD (lane 1) and p53TAD1 (lane 2) at 0.5%. Bound protein was detected with an anti-p53 antibody DO-1. **C.** Binding of p53TAD and p53TAD2 to GATA-1 ZFD. Varying concentrations (1.0, 0.1, 0.01 μ M) of p53TAD (lanes 5-7) and p53TAD2

(lanes 8-10) were incubated with 1 μ M GST-GATA-1 ZFD. In the GST lanes, 1 μ M of purified p53TAD (lane 3) and p53TAD2 (lane 4) were incubated with 1 μ M GST as a control. The input lanes are p53TAD (lane 1) and p53TAD2 (lane 2) at 0.5%. Bound protein was detected with an anti-p53 antibody Pab 1801.

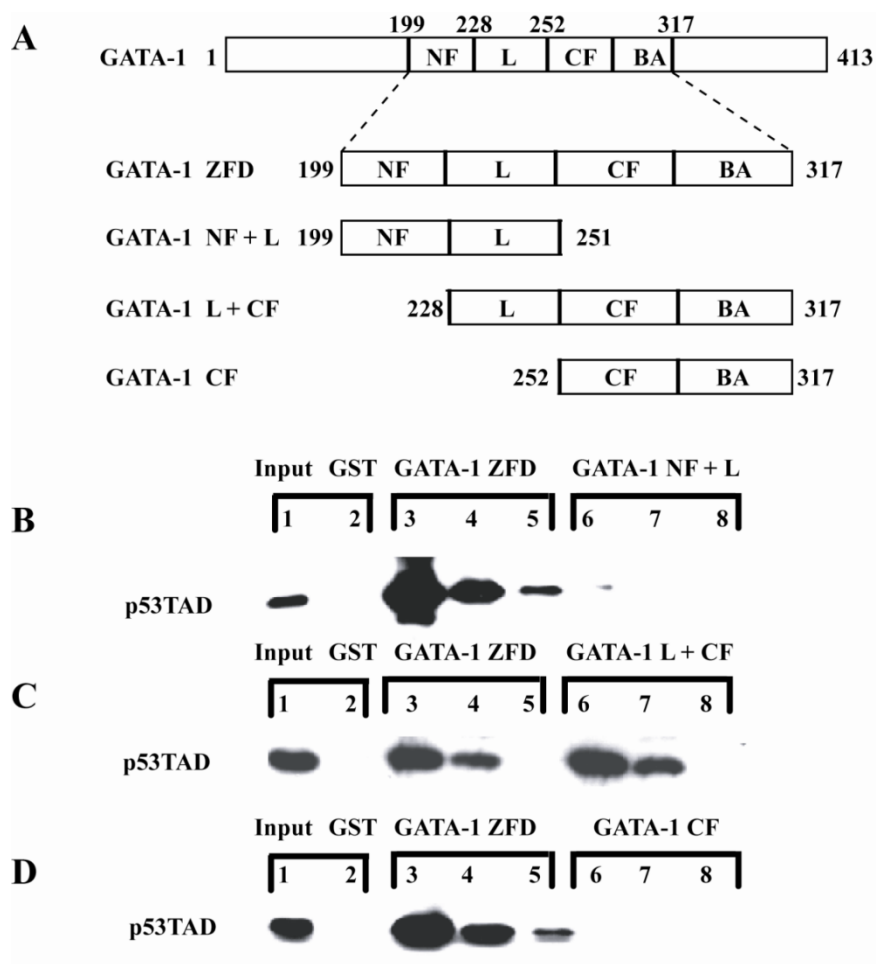


Figure 2. 2. The linker region, the carboxyl-terminal zinc finger and the basic arm of GATA-1 are sufficient for binding p53TAD in an *in vitro* pull-down assay. **A.** Schematic of GATA-1 highlighting the zinc-finger domains (GATA-1 ZFD) and the various deletion mutants of GATA-1 ZFD used as GST-fusion proteins. **B.** Comparative binding of GST-GATA-1 ZFD and GST-GATA-1 NF + L (GST-GATA-1₍₂₀₀₋₂₅₁₎) to p53TAD. Varying concentrations (1.0, 0.1, 0.01 μ M) of p53TAD were incubated with either 1 μ M GST-GATA-1 ZFD (lanes 3-5) or GST-GATA-1 NF + L (lanes 6-8). In the GST lane, 1 μ M of purified p53TAD was incubated with 1 μ M GST as a control (lane 2). The input lane is 0.5% p53TAD (lane 1). **C.** Comparative binding of GST-GATA-1 ZFD and GST-

GATA-1 L + CF (GST-GATA-1₍₂₂₈₋₃₁₇₎) to p53TAD. Varying concentrations (1.0, 0.1, 0.01 μ M) of p53TAD were incubated with either 1 μ M GST-GATA-1 ZFD (lanes 3-5) or GATA-1 L + CF (lanes 6-8). In the GST lane, 1 μ M of purified p53TAD was incubated with 1 μ M GST as a control (lane 2). The input lane is 0.5% p53TAD (lane 1). D. Comparative binding of GST-GATA-1 ZFD and GST-GATA-1 CF (GST-GATA-1₍₂₅₂₋₃₁₇₎) to p53TAD. Varying concentrations (1.0, 0.1, 0.01 μ M) of p53TAD were incubated with either 1 μ M GST-GATA-1 ZFD (lanes 3-5) or GATA-1 CF + BA (lanes 6-8). In the GST lane, 1 μ M of purified p53TAD was incubated with 1 μ M GST as a control (lane 2). The input lane is 0.5% p53TAD (lane 1). In all experiments, bound protein was detected with an anti-p53 antibody DO-1.

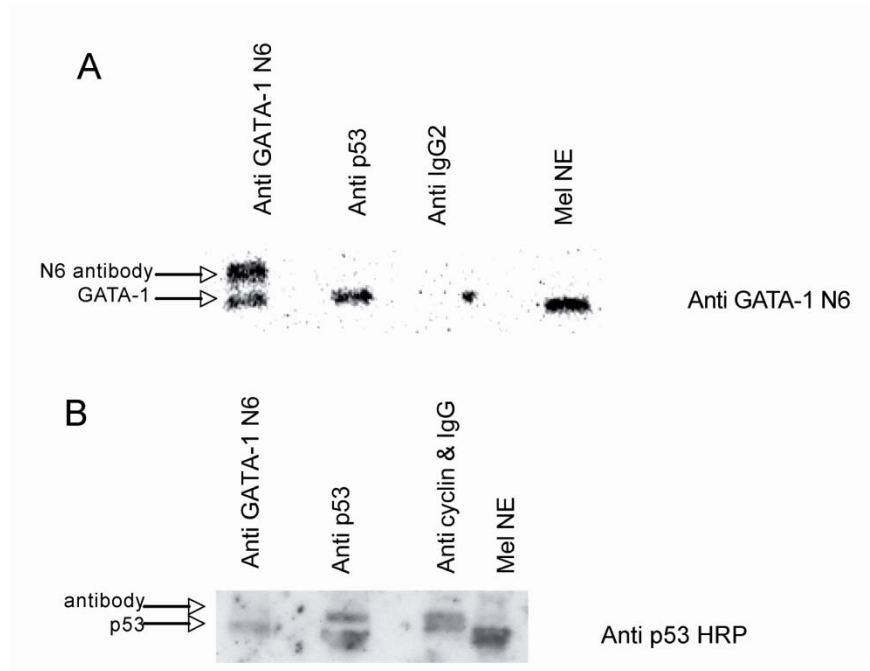


Figure 2. 3. GATA-1 and p53 interact *in vitro* in MEL cell extract. A. 320 μ g of MEL nuclear extract was incubated with Lane 1, 4 μ l anti GATA-1 N6 and 1 μ l anti rat IgG2a; lane 2, 4 μ l of anti p53 (PAB 240) and 1 μ l anti-mouse IgG; lane 3, 5 μ l anti rat IgG₂ for 1.5 hours at 4°C with rotation. 40 μ l of Protein G Plus agarose was added with 4 additional hours of rotation at 4°C; lane 4, MEL nuclear extract. Western blots are shown using GATA -1 N6 antibody. **B.** Lane 1, 8 μ l anti GATA-1 N6 and 1 μ l anti rat IgG2a; lane 2. 8 μ l of anti p53 (PAB 246) and 1 μ l anti-mouse IgG; lane 3, 8 μ l anti-cyclin B1(GNS1) and 1 μ l anti-mouse IgG for 1.5 hours at 4°C with rotation. 40 μ l of Protein G Plus agarose was added with 4 more hours of rotation at 4°C; lane 4, MEL nuclear extract. A western blot with HRP conjugated anti-p53 is shown.

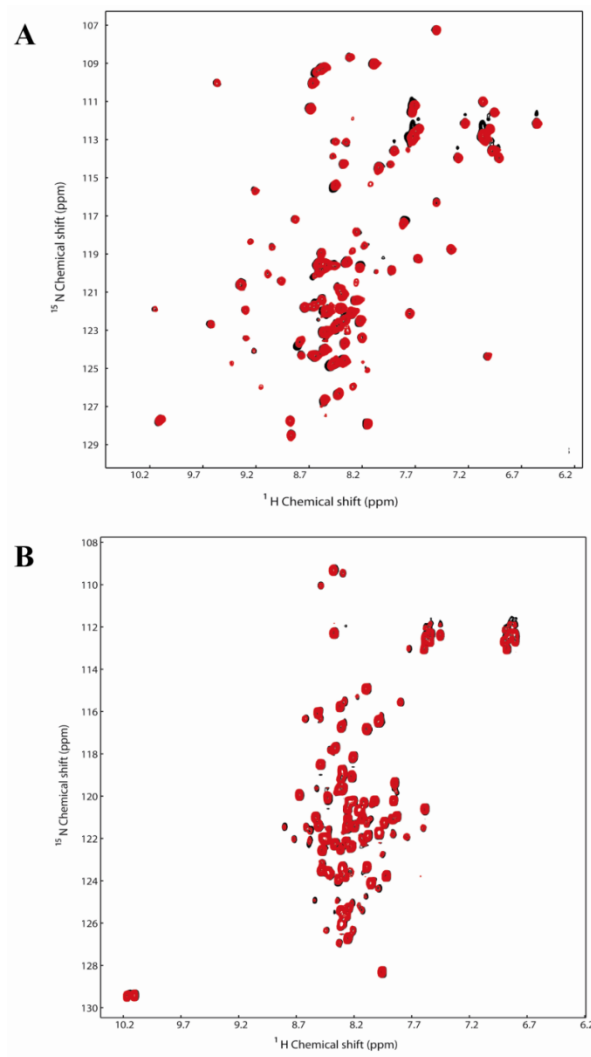


Figure 2. 4. GATA-1, p53TAD and a GATA-1 DNA-binding site do not form a ternary complex. (A) Overlay of the 2D ^1H - ^{15}N HSQC spectra for a 0.5 mM sample of ^{15}N -labeled GATA-1 L +CF/DNA complex (black) and in the presence of 0.5mM unlabeled p53TAD (red). (B) Overlay of the 2D ^1H - ^{15}N HSQC spectra for a 0.5 mM sample of ^{15}N -labeled p53TAD in the free form (black) and in the presence of 0.5 mM unlabeled GATA-1 L +CF/DNA complex (red).

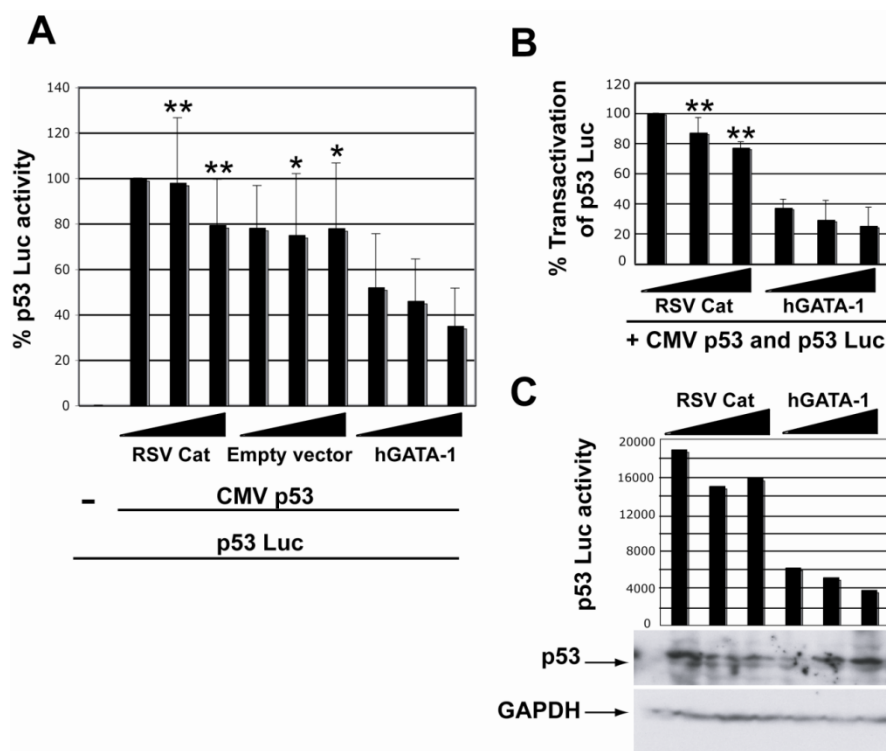


Figure 2. 5. GATA-1 can inhibit transactivation by p53 in 6C2 cells. **A.** All samples were transfected with 150 ng of CMV renilla luciferase and 0.5 µg of p53 Luc reporter plasmid; and 0.25 µg of CMV p53 was added to all but the first sample. Increasing amounts of RSV Cat, Empty RSV expression vector or RSV hGATA-1 were added (1, 2 and 3 µg of DNA). The dual luciferase reporter assay was performed on aliquots of the samples and firefly luciferase activity normalized to renilla luciferase is plotted. The error bars indicate the standard deviations and n=3. Means of the two highest DNA concentrations of RSV Cat and empty vector were compared to means for equal weights of hGATA-1 by Student's t test and * p < 0.05, ** p < 0.005. **B.** RSV Cat and hGATA-1 were compared in six transfections. Results were normalized to renilla luciferase activity. The firefly luciferase value for 1 µg of RSV Cat was then set to 100 % and all other values were normalized to this. The error bars indicate the standard deviation and n = 6. We compared means for equal weights of the two highest DNA concentrations of the RSV Cat and hGATA-1 plasmids to test for statistical significance by Student's t test and ** p < 0.005. **C.** A comparable transfection was performed adding CMV renilla

luciferase, p53 Luc reporter plasmid and CMV p53 to all samples and co-transfecting increasing amounts of RSV Cat or RSV hGATA-1. The cells were divided at harvest and half were used for the dual luciferase reporter assay and half for Western blotting with HRP conjugated anti-p53 (upper) and with anti-GAPDH antibody (lower).

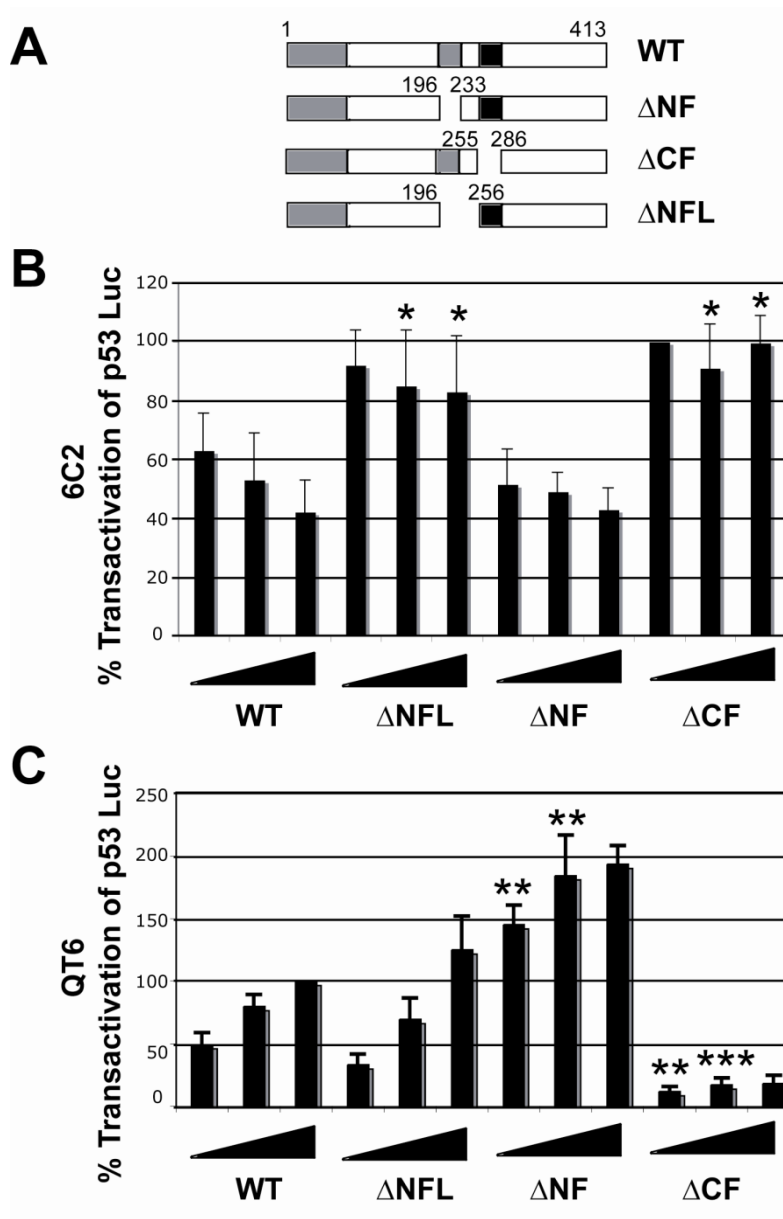


Figure 2. 6. The linker between the zinc fingers and the CF of GATA-1 are required for p53 inhibition. A. Schematic of deletion mutants of mouse GATA-1. B. Transfections into 6C2 cells were performed as described in the figure 5 legend. 150 ng of CMV renilla luciferase, 0.5 μg of p53

Luc reporter plasmid, and 0.25 μg of CMV p53 were present in all samples, while varying amounts of mouse WT GATA-1 expression vector or mutants were added as indicated (1, 2 or 3 μg). After normalization to renilla luciferase activity, the firefly luciferase value for 1 μg of ΔCF was set to 100% and the values for all other samples were correspondingly adjusted. The error bars indicate the standard deviation and the experiments were performed between 4 and 7 times. We compared means for equal weights of the two highest DNA concentrations of ΔNFL , ΔNF and ΔCF to WT by Student's t test and * $p < 0.05$. C. CMV renilla luciferase and 1 μg FR7Luc, a GATA-1 responsive reporter plasmid, were transfected into QT6 fibroblasts with increasing amounts of mouse GATA-1 expression plasmid as indicated (0.5, 1 or 2 μg). Normalization and error bars are as in B with values for 2 μg of wild type GATA-1 set to 100 %. The means for equal weights of the two lowest concentrations of wild type were compared to those of all other constructs by Student's t test and ** $p < 0.005$, *** $p < 0.0005$.

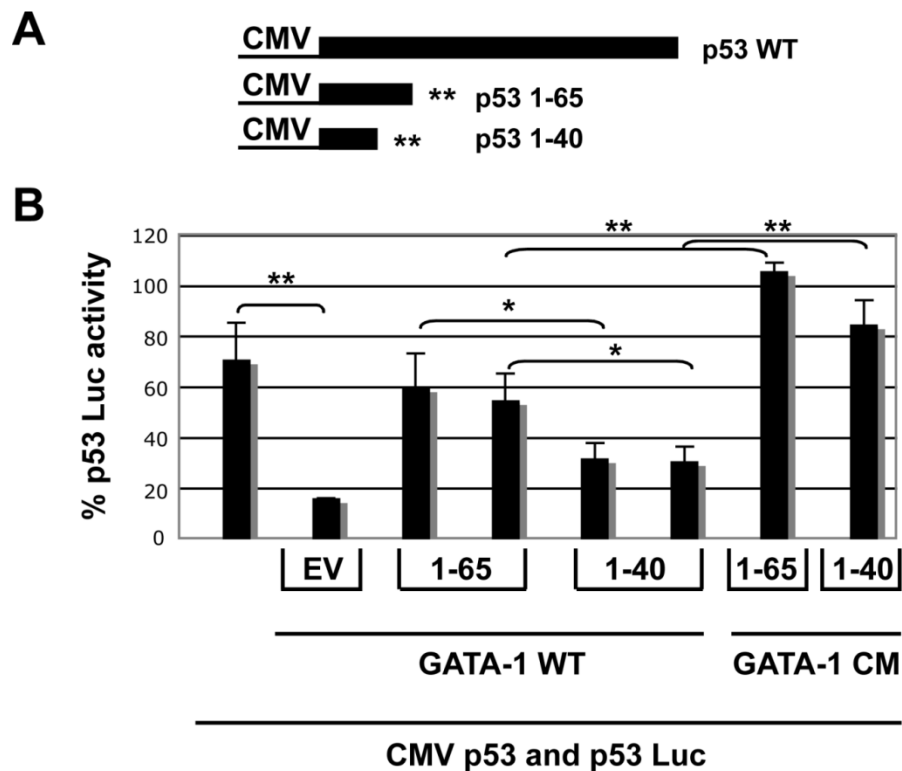
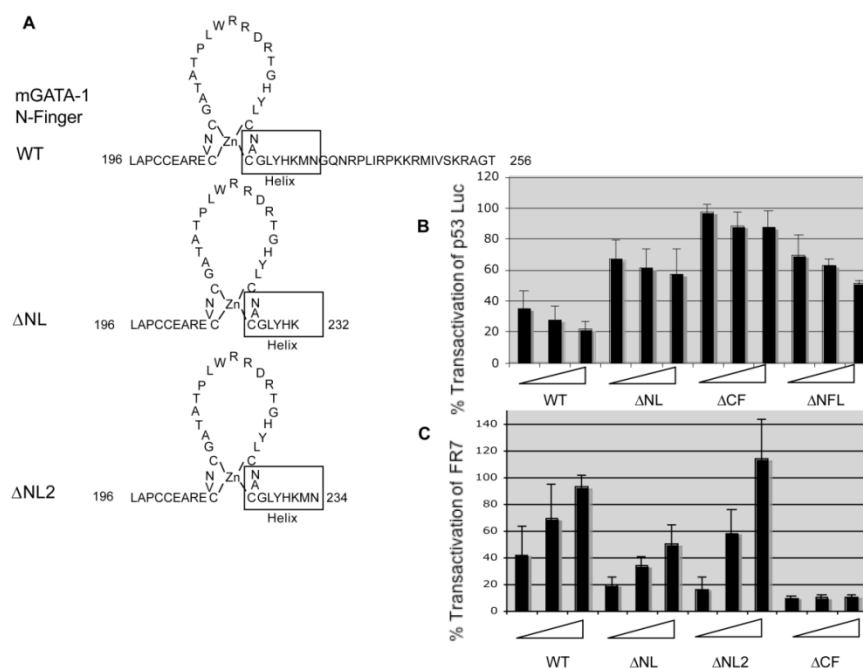


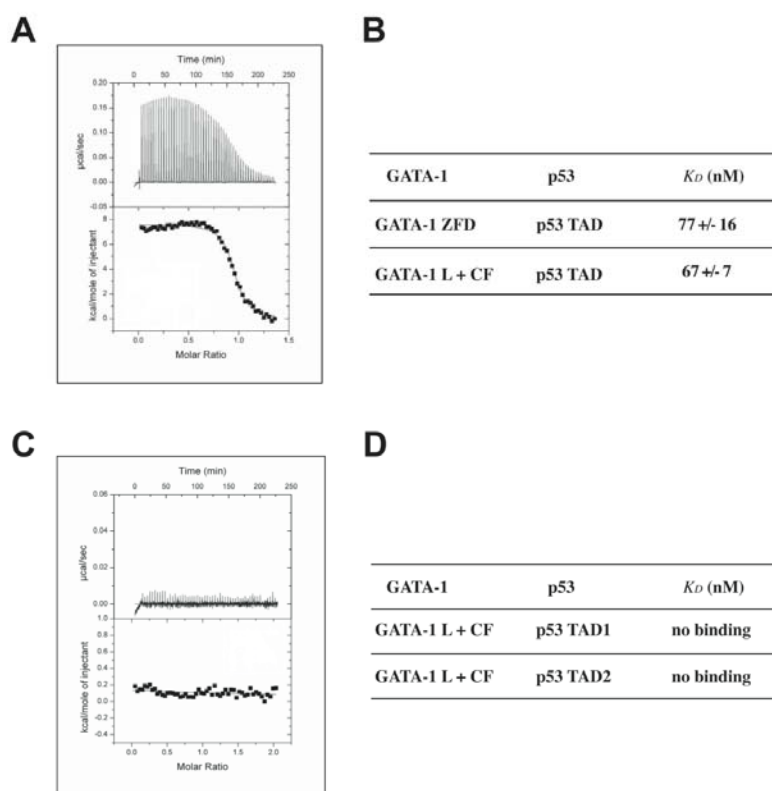
Figure 2. 7. p53 1-65 can compete with WT p53 for binding to GATA-1 and reduce the inhibition of p53. A. A schematic of constructs made by inserting translation termination signals in the CMV p53 expression vector to generate peptides 1 to 65 and 1 to 40 is shown. B. A transfection assay in which CMV renilla luciferase, p53 Luc reporter plasmid, and CMV p53 were present in all samples and expression plasmids for mGATA-1 were added to samples 2 to 6 along with 6 μ g of empty vector (bar 2), 5 or 6 μ g CMV p53 1- 65 (bar 3 and 4) or CMVp53 1 to 40 (bar 5 and 6) as indicated. For the last two samples GATA-1CM, a mutant which does not interfere with transactivation by p53, was used instead of WT GATA-1 and 6 μ g of CMV 1 to 65 (bar 7) or 6 μ g of CMV 1 to 40 (bar 8) were added. Normalization is as described in the Figure 6 legend with the average of the 5 μ g GATA-1 CM 65 and 40 samples set to 100%. Results are the average of 3 experiments and error bars indicate the standard deviation. Means of equal weights

Supplementary Figures



Supplementary Figure 2. 1. Amino Acids 232 to 256 of the linker between the two zinc fingers of GATA-1 are required for inhibition of p53. **A.** A Schematic of the N-finger and linker of GATA-1 (amino acids 196 to 256) that was removed to generate Δ NFL (top, Figure 5A) and of the region of the N-finger that was returned to Δ NFL to generate Δ NL (middle) and Δ NL2 (bottom) is shown. **B.** A transfection into 6C2 cells was performed as above. CMV renilla luciferase, p53 Luc reporter plasmid, and CMV p53 were present in all samples and expression plasmids for GATA-1 and mutants were added as indicated (1, 2, and 3 μ g). A sample in which empty vector was added was set at 100% and all values were normalized accordingly. Values are averages of six experiments and the error bars indicate the standard deviations. The differences between the means of equal weights of WT vs. all other constructs are significant, $p < 0.005$. **C.** CMV renilla luciferase and 1 μ g FR7Luc, a GATA-1 responsive reporter plasmid, were transfected into QT6 fibroblasts with increasing amounts of mouse GATA-1, Δ NL, Δ NL2 and Δ CF expression plasmids as indicated (0.5, 1 or 2 μ g). Normalization and error bars are as in B and values are the average of 3 experiments.

On average, the differences between the mean of WT vs. Δ NL and Δ NL2 are not significant, although results are variable. The means of all three are significantly different from that of Δ CF, $p < 0.02$.



Supplementary Figure 2. 2. GATA-1 ZFD and GATA-1 L +CF bind p53TAD with the same affinity. A. A typical ITC thermogram obtained by successive additions of p53TAD to GATA-1 L + CF. B. A typical ITC thermogram obtained by successive additions of p53TAD1 to GATA-1 L + CF. C. Dissociation constants (K_d) for the binding of GATA-1 fragments to p53 fragments as determined by ITC. The ITC curves were fit using a single-binding site to solve for the stoichiometry (N), the dissociation constant (K_d) the change in enthalpy (ΔH). The experimental conditions are as described in the experimental procedures. All injections fit the single binding site mechanism with 1:1 stoichiometry.

Préface au Chapitre III

De précédents travaux ont démontré que les TAD de la protéine virale VP16 et de la protéine p53 interagissent directement avec la sous-unité p62/Tfb1 du facteur général de transcription TFIID, et que cette liaison corrèle avec leur aptitude à activer la transcription. Sur la base de ces travaux, nous avons entrepris une étude structurale et fonctionnelle du complexe Tfb1/VP16 TADc. En utilisant la spectroscopie RMN ainsi que le titrage calorimétrique en conditions isothermes (ITC) combiné à de la mutagenèse dirigée, nous avons analysé le mode de liaison de ces deux partenaires ainsi que la contribution relative de résidus présents à l'interface. Nous avons ainsi identifié une région comprise dans le TADc de VP16 présentant une similarité de séquence avec le TAD2 de p53 et dont l'affinité pour Tfb1 est de l'ordre du nanomolaire. Nous avons également déterminé la structure tridimensionnelle de ce complexe par spectroscopie RMN. La structure révèle que p53 TAD2 et VP16 TADc adoptent un mode de liaison similaire à Tfb1. Tout comme p53 TAD2, le TADc de VP16 adopte une structure secondaire hélicale sur 9 résidus en complexe avec Tfb1. Par ailleurs, plusieurs résidus de cette hélice sont impliqués dans des interactions hydrophobes, cation- π et ioniques. Cependant, des différences entre les complexes Tfb1/p53 et Tfb1/VP16 illustrent et expliquent la façon dont les domaines de transactivation peuvent sélectivement activer la transcription selon leur état de phosphorylation.

Les résultats de ce chapitre sont exposés sous la forme d'un article scientifique ayant été publié dans la revue *Journal of the American Chemical Society* en juillet 2008. Au cours de ce projet, j'ai travaillé en étroite collaboration avec Chantal Langlois. Paola Di Lello a été impliquée dans l'acquisition de spectres RMN. Chantal Langlois et moi avons été impliquées au niveau de la production et la purification des protéines marquées et non marquées nécessaires à l'acquisition des spectres RMN. L'attribution séquentielle ainsi que l'attribution des spectres NOE (Nuclear Overhauser Effect) ainsi que les calculs de structure ont été fait en coopération avec Chantal Langlois. J'ai participé à l'écriture des parties « Matériel et Méthodes » et

« Résultats ». Ayant été impliquée dans les aspects de la recherche expérimentale et théorique de cette étude, ma contribution à ces travaux est majeure. *Journal of American Chemical Society* n'acceptant pas les co-premiers auteurs, je suis deuxième auteur sur cet article.

Chapitre III

Structure RMN du complexe formé par la sous unité Tfb1 de TFIID et le domaine de transactivation de VP16 : Similitudes entre VP16 et p53.

Article 2: NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: structural similarities between VP16 and p53.

This research was originally published in JACS. Langlois C, Mas C, Di Lello P, Jenkins LM, Legault P, Omichinski JG. NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: structural similarities between VP16 and p53. JACS 2008, *130*, 10596–10604 10.102

Reprinted with permission from : Langlois C, Mas C, Di Lello P, Jenkins LM, Legault P and Omichinski JG.

Reproduced with permission from J. AM. CHEM. SOC. 2008, *130*, 10596–10604 10.102 ©2008 American Chemical Society

**NMR Structure of the complex between the Tfb1
subunit of TFIIH and the activation domain of VP16: structural
similarities between VP16 and p53**

Chantal Langlois¹, Caroline Mas¹, Paola Di Lello¹, Lisa M. Miller Jenkins², Pascale Legault¹ and James G. Omichinski^{1*}

¹Département de Biochimie, Université de Montréal C.P. 6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada. ²Laboratory of Cell Biology, NCI, National Institutes of Health, 37 Convent Drive, Bethesda, MD 20892-4256, USA.

* Corresponding Author:

Mailing address: Université de Montréal, Département de Biochimie, C.P. 6128, Succursale Centre-Ville, Montréal, QC H3C 3J7, Canada.

Abstract

The Herpes Simplex Virion Protein 16 (VP16) activates transcription through a series of protein/protein interactions involving its highly acidic transactivation domain (TAD). The acidic TAD of VP16 (VP16TAD) has been shown to interact with several partner proteins both in vitro and in vivo, and many of these VP16 partners also bind the acidic TAD of the mammalian tumour suppressor protein p53. For example, the TADs of VP16 and p53 (p53TAD) both interact directly with the p62/Tfb1 (human/yeast) subunit of TFIID, and this interaction correlates with their ability to activate both the initiation and elongation phase of transcription. In this manuscript, we use NMR spectroscopy, isothermal titration calorimetry (ITC) and site-directed mutagenesis studies to characterize the interaction between the VP16TAD and Tfb1. We identify a region within the carboxyl-terminal subdomain of the VP16TAD (VP16C) that has sequence similarity with p53TAD2 and binds Tfb1 with nanomolar affinity. We determine an NMR structure of a Tfb1/VP16C complex, which represents the first high-resolution structure of the VP16TAD in complex with a target protein. The structure demonstrates that like p53TAD2, VP16C forms a 9-residue α -helix in complex with Tfb1. Comparison of the VP16/Tfb1 and p53/Tfb1 structures clearly demonstrates how the viral activator VP16C and p53TAD2 shares numerous aspects of binding to Tfb1. Despite the similarities, important differences are observed between the p53TAD2/Tfb1 and VP16C/Tfb1 complexes, and these differences demonstrate how selected activators such as p53 depend on phosphorylation events to selectively regulate transcription.

Introduction

Activators function by stimulating transcription through protein/protein interactions involving their transactivation domain (TAD) [1, 2]. The current model is that activators function by recruiting a number of different transcriptional regulatory factors including nucleosome-remodeling complexes, the mediator complex and several general transcription factors (TFIIB, TBP, TFIID) to enhance the rate of transcription by affecting nucleosome assembly/disassembly, pre-initiation complex formation, promoter clearance and/or the rates of elongation [3-17]. Originally, transcription activators were classified based on the presence of specific amino acids within their TAD, and these included the acidic-rich (aspartic and glutamic acid), the glutamine-rich, the proline-rich and the serine/threonine-rich activators. Given their occurrence in a number of crucial transcriptional regulatory proteins, the most extensively studied activators are those that contain acidic TADs and two highly investigated acidic activators are the human tumour suppressor protein p53 and the Herpes Simplex Virion (HSV) protein 16 (VP16) [18, 19].

VP16 functions to stimulate transcription of viral immediate early genes in HSV-infected cells, whereas p53 induces the expression of many human target genes whose products regulate numerous functions, including cell-growth arrest and apoptosis [20, 21]. The acidic TADs of VP16 (residues 412-490) and p53 (residues 1-72) share many similarities related to both their functions and amino acid composition [18-20, 22, 23]. The TAD of both VP16 and p53 can be subdivided into two subdomains, and each subdomain is capable of independently activating transcription when tethered to a DNA-binding domain. In the case of VP16, the subdomains are referred to as the amino subdomain (VP16N or VP16₄₁₂₋₄₅₆) and the carboxyl subdomain (VP16C or VP16₄₅₆₋₄₉₀) [6, 22, 24-26], whereas for p53 they are referred to as the first TAD (p53TAD1 or p53₁₋₄₀) and the second TAD (p53TAD2 or p53₄₀₋₇₂) [27]. Mutagenesis studies have clearly demonstrated that the activity of VP16 and p53 not depends only on acidic residues, but also on key hydrophobic and aromatic amino acids within their respective TADs [22, 26, 28-33]. Importantly, both the p53 and the VP16 TADs have been shown to interact directly with several

of the same target proteins both *in vitro* and *in vivo*, and these targets include the TATA-binding protein (TBP), the CREB-binding protein (CBP), the general transcription factor IIB (TFIIB), TBP-associated factor (TAF) TAF_{II}31 and the p62/Tfb1 (human/yeast) subunit of the general transcription factor IIIH (TFIIH) [6, 13, 17, 34-39].

Given the high percentage of aspartic acid and glutamic acid residues within their sequences, acidic TADs are generally thought to be devoid of regular secondary structure elements in their free state and have been referred to as “acid blobs” or “negative noodles” [40]. NMR studies of the TADs of p53 and VP16 in their free state support this view [41-43]. In some cases, it has been shown that acidic TADs become more ordered when bound to their target proteins [31, 33, 44-49], and that the intrinsically unstructured nature of TADs helps them mediate multiple protein/protein interactions with different partners [50]. Three high-resolution structures have been determined for complexes containing one of the two subdomains of the p53TAD: a p53TAD1/MDM2 complex [45], a p53TAD2/RPA70 (replication protein A 70) complex [47] and a p53TAD2/Tfb1 (Tfb1 subunit of TFIIH) complex [49]. In all three complexes, the TAD of p53 adopts a short α -helical conformation in the presence of its partner protein.

Despite the fact that the VP16TAD has been extensively studied and utilized as a model activation domain to examine activation mechanisms and protein interactions in a wide number of organisms including mammals [51], yeast [52] and plants [53], there is currently no high-resolution structure of a protein/protein complex containing the VP16TAD. Previous NMR studies have examined the structure of the TAD of VP16 in complex with the TAF_{II}31 [31], the human co-factor protein PC4 and TFIIB [33]. In the first study, VP16C (VP16₄₅₆₋₄₉₀) complexed to TAF_{II}31 was predicted to form an α -helix between residues 472 and 483 [31]. In the second study, the full-length TAD of VP16 (VP16₄₁₂₋₄₉₀) in complex with either PC4 or TFIIB forms two α -helices [33]. The first helix was between residues 435 and 450 within VP16N and the second helix was between residues 465 and 485 in VP16C. In this study, models of VP16N and VP16C in complex with either PC4 or TFIIB were proposed based on docking approaches in combination with NMR chemical shift mapping. In both studies, high-resolution structures of the

complexes were not obtained [31, 33]. Therefore, the current mechanism by which the VP16TAD recognizes its target partners at the atomic level is poorly understood.

It has previously been demonstrated that the TADs of VP16 and p53 directly bind the p62/Tfb1 subunit of TFIID and that their binding to TFIID correlates directly with their ability to activate both the initiation and elongation phase of transcription [8, 54]. In this manuscript, we have used Isothermal Titration Calorimetry (ITC), NMR spectroscopy and site-directed mutagenesis experiments to define the molecular basis of the interaction between the TAD of VP16 and the Tfb1 subunit of TFIID. We demonstrated by ITC that VP16C binds to Tfb1 with higher affinity than VP16N. We determined by NMR spectroscopy a high-resolution structure of a complex between Tfb1 and VP16C. This structure demonstrates that like p53, VP16C forms a 9-residue α -helix when bound to Tfb1 and that several residues within the α -helix make important hydrophobic, cation- π and ionic interactions with Tfb1. Structural comparison of the Tfb1/VP16 complex with our previously published Tfb1/p53 complex [49] indicates that the two complexes are strikingly similar. The two structures explain how the TAD of VP16 can mimic mammalian activators such as p53. In addition, there are important differences between the Tfb1/p53 and Tfb1/VP16 complexes that help explain how activation domains can selectively regulate transcription, and these differences appear to be related to the phosphorylation state of p53. Such detailed structural information is absolutely essential to our efforts to design molecules that mimic transcription activators such as p53 and VP16.

Experimental Procedures

Cloning of recombinant proteins.

Tfb1₁₋₁₁₅ and related mutants were constructed as GST-fusion proteins as previously described [49]. Dr. Steven Triezenberg generously provided the clones expressing VP16₄₁₂₋₄₅₆ (VP16N) and VP16₄₅₆₋₄₉₀ (VP16C) as GST-fusion proteins. Site directed mutagenesis of VP16C and addition of a tyrosine residue at the carboxyl-terminal position of the VP16N segment were carried out with the QuickChange II site-directed mutagenesis kit (Stratagene).

Protein expression and purification.

Tfb1₁₋₁₁₅ and its mutant were purified as previously described [49, 55]. VP16N, VP16C and VP16C mutants were expressed as GST-fusion protein in *E. coli* host strain TOPP2 and bound to GSH-resin (General Electric). The resin bound protein was incubated overnight with thrombin (Calbiochem). Following cleavage, the supernatant was purified using a Q-Sepharose High Performance column (General Electric). Uniformly (>98%) ¹⁵N-labeled and ¹⁵N/¹³C-labeled proteins were prepared in minimal media containing ¹⁵NH₄Cl and/or ¹³C₆-glucose as the sole nitrogen and carbon sources⁵⁵.

Isothermal titration calorimetry studies.

The ITC titration experiments were performed as previously described [56] in 20 mM TRIS at pH 7.5. The concentrations of the injected proteins were determined from A₂₈₀. All titrations fit the single-binding site mechanism with 1:1 stoichiometry.

NMR samples.

For the NMR structural studies of the Tfb1₁₋₁₁₅/VP16C complex we used four samples. One sample contained 1.0 mM of ¹⁵N-Tfb1₁₋₁₁₅ in 10 mM sodium phosphate (pH 6.5), 1 mM EDTA and 90% H₂O/10% D₂O; unlabeled VP16C was added to a final ratio of 1:1. The second sample consisted of 1.0 mM of ¹⁵N/¹³C- Tfb1₁₋₁₁₅ in 10 mM sodium phosphate (pH 6.5), 1 mM EDTA and 90% H₂O/10% D₂O to which unlabeled VP16C was added to a final ratio of 1:1. The third sample contained 1.0 mM of ¹⁵N-VP16C in 10 mM sodium phosphate (pH 6.5), 1 mM EDTA and 90% H₂O/10% D₂O; unlabeled Tfb1₁₋₁₁₅ was added to a final ratio of 1:1. The fourth sample consisted of 1.0 mM of ¹⁵N/¹³C- VP16C in 10 mM sodium phosphate (pH 6.5), 1 mM EDTA and 90% H₂O/10% D₂O to which unlabeled Tfb1₁₋₁₁₅ was added to a final ratio of 1:1. For experiments in D₂O, samples two and four were lyophilized and resuspended in 100% D₂O.

For the NMR mapping studies of the Tfb1₁₋₁₁₅/VP16C complex we used two samples. One sample contained 0.5 mM of ¹⁵N-Tfb1₁₋₁₁₅ in 10 mM sodium phosphate (pH 6.5), 1 mM

EDTA and 90% H₂O/10% D₂O; unlabeled VP16C was added to a final ratio of 1:1. The second sample consisted of 0.5 mM of ¹⁵N- VP16C in 10 mM sodium phosphate (pH 6.5), 1 mM EDTA and 90% H₂O/10% D₂O to which unlabeled Tfb1₁₋₁₁₅ was added to a final ratio of 1:1.

NMR spectroscopy.

The NMR experiments were carried out at 300 K on Varian Unity Inova 500 MHz and 600 MHz spectrometers. The backbone and aliphatic side chain resonances (¹H, ¹⁵N and ¹³C) were assigned using a combination of triple resonance experiments [HNCO, HNCACB, (HB)CBCA(CO)NNH, C(CO)NNH, H(CCO)NNH and HCCH-COSY]. Interproton distance restraints were derived from ¹⁵N-edited NOESY-HSQC and ¹³C-edited HMQC-NOESY spectra ($\tau_m=90$ ms). The NMR data were processed with NMRPipe/NMRDraw [57] and analyzed with NMRView [58].

Structure calculations.

The NOE-derived distance restraints were divided into three classes defined as strong (1.8-2.8 Å), medium (1.8-3.4 Å) and weak (1.8-5.0 Å). Backbone dihedral angles were derived with the program TALOS [59]. Following exchange in D₂O, slow-exchanging amide protons were identified by recording a 2D ¹H-¹⁵N HSQC spectrum of Tfb1₁₋₁₁₅. Hydrogen bonds were used as distance restraints only for the slowly exchanging amide protons in elements of secondary structure and after the initial rounds of calculations had revealed the protein fold. The structures of VP16C and the PH domain of Tfb1₁₋₁₁₅ were calculated using the program CNS, with a combination of torsion angle and Cartesian dynamics [60] and starting from two extended structures with standard geometry. The quality of the structures was assessed using PROCHECK-NMR [61] and MOLMOL [62]. All of the figures representing the structures were generated with the program PyMol (<http://www.pymol.org>).

Results

VP16C binds the PH domain of Tfb1 with higher affinity than VP16N.

We have previously shown by chemical shift mapping studies that VP16N binds to the amino-terminal PH domain of Tfb1 (Tfb1₁₋₁₁₅) and that the binding site for VP16N on Tfb1 is virtually identical to the binding site for p53TAD2 on Tfb1 [49, 55]. Since both VP16N and VP16C are each capable of independently activating transcription and VP16C has been shown to interact with a number of transcriptional regulatory proteins including TBP, PC4,TFIIB, TAF9 and TAF_{II}31 [24, 31, 33, 63], we investigated whether or not VP16C was also able to interact with Tfb1. To check this, we determined the dissociation constants (K_d) for the complexes formed between VP16N and Tfb1₁₋₁₁₅ and between VP16C and Tfb1₁₋₁₁₅ by ITC (Supplementary Figure 3. 1). By this method, VP16C bound to Tfb1 with an apparent K_d of 360 ± 40 nM and VP16N bound to Tfb1 with an apparent K_d of 1000 ± 100 nM (Table 1). These results suggested that both subdomains of the VP16TAD were capable of interacting with Tfb1, and the K_d obtained for the Tfb1/VP16C complex is very similar to the K_d observed between Tfb1 and the p53TAD2 (390 nM) [49]. Our VP16C binding to Tfb1 (360 nM), as measured by ITC, is higher affinity than what has been previously reported for VP16C binding to TFIIB (3 μ M), PC4 (14 μ M), TAF9 (73 μ M) and humanTAF_{II}31(100 μ M) but lower affinity than VP16C binding to yeast TBP (44 nM) [31, 33, 63].

VP16C binds to the same region of the PH domain of Tfb1 as VP16N

To define the binding site for VP16C on Tfb1, NMR chemical shift mapping studies were performed. Addition of VP16C (VP16₄₅₆₋₄₉₀) to ¹⁵N-labeled Tfb1₁₋₁₁₅ produced changes in ¹H and ¹⁵N chemical shifts for several signals of Tfb1₁₋₁₁₅ in the 2D ¹H-¹⁵N HSQC spectra (Figure 3. 1). The residues displaying significant chemical-shift changes ($\Delta\delta > 0.1$ ppm; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) are clustered on a positively charged surface that includes strands β 5, β 6 and β 7 of the PH domain when mapped onto the NMR solution structures of free Tfb1 (Supplementary Figure 3. 2). The VP16C binding site on Tfb1 is virtually identical to the binding sites for VP16N [55].

This result indicates that VP16C and VP16N share a common binding site on Tfb1 and this binding site is virtually identical to the binding site for p53TAD2

Structure determination of the Tfb1/VP16C complex.

Given the similar affinity of VP16C for Tfb1₁₋₁₁₅ and that VP16C shares a common binding site on Tfb1 with p53TAD2, we pursued NMR structural studies of a complex containing Tfb1₁₋₁₁₅ and VP16C. The NMR structure of a Tfb1/VP16C complex would not only provides us with the first high-resolution structure of the TAD of VP16 bound to a target protein, but it would allows us to make direct structural comparison between the TADs of p53 and VP16 bound to an identical target (Tfb1 in this case) [49].

The three-dimensional structure of the Tfb1₁₋₁₁₅/VP16C complex was calculated using 1568 NOE-derived distance restraints, 36 hydrogen-bond restraints and 158 dihedral angle restraints. A total of 109 structures were calculated, and 100 of them satisfied the experimental constraints with no NOE violation greater than 0.2 Å and no backbone dihedral angle violation greater than 2° (Table 2). The structure of the Tfb1₁₋₁₁₅/VP16C complex is well defined by the NMR data (Figure 3. 2). The 20 lowest-energy structures are characterized by good backbone geometry, no significant restraint violation and low pairwise rmsd values (Figure 3. 2A and Table 3. 2).

Structure of Tfb1 and VP16C in the Tfb1/VP16C complex.

The structure of Tfb1 in the Tfb1/VP16 complex is very similar to what was seen for the free form [55] and for the Tfb1/p53 complex [49]. Tfb1 contains a PH fold, which consists of two perpendicular anti-parallel β-sheets arranged in a β-sandwich and flanked on one side by a long α-helix (Figure 3. 2). This indicates that Tfb1 does not undergo any significant structural change between the free and the VP16C-bound forms.

The free VP16C is very flexible in solution and it does not appear to possess any regular elements of secondary structure [31, 33, 41, 43]. Binding of VP16C to targets such as TAF_{II}31,

TFIIB and PC4 induces formation of an α -helical structure within VP16C [31, 33]. In the case of TAF_{II}31, the α helix is located between residues 472 and 483, whereas with TFIIB and PC4 the α -helix is located between residues 465 and 485 [31, 33]. These helices were inferred from NMR studies based on NOE intensities. In the Tfb1/VP16C complex, chemical shift mapping studies demonstrate that addition of unlabeled Tfb1 to ¹⁵N-labeled VP16C induced changes in the ¹H and ¹⁵N chemical shifts of several signals in the 2D ¹H-¹⁵N HSQC spectra of VP16C (Supplementary Figure 3. 3). As was observed for p53 in complex with Tfb1 [49], VP16C forms a short 9-residue α -helix between residues 472 and 480 (Figure 3. 3B). Except for this 9-residue helix and a few flanking amino acids, VP16C is largely unstructured in complex with Tfb1.

Important interactions at the interface of the Tfb1/VP16C complex.

The binding interface of the Tfb1/VP16C complex comprises strands β 5, β 6 and β 7 and the loop connecting β 5 and β 6 of Tfb1 and the 9-residue α -helix of VP16C (Asp⁴⁷²- Thr⁴⁸⁰). The VP16C helix is amphipathic and its hydrophobic face interacts with Tfb1. Three hydrophobic residues (Phe⁴⁷⁵, Met⁴⁷⁸ and Phe⁴⁷⁹) within the VP16C helix contribute a series of crucial contacts with Tfb1 (Figure 3. 3A). The aromatic side chain of Phe⁴⁷⁵ and Phe⁴⁷⁹ fill two adjacent shallow pockets on the surface of Tfb1, where they form a complex network of interactions with Gln⁴⁹, Met⁵⁹, Arg⁶¹ and Met⁸⁸ of Tfb1 (Figure 3. 3). The aromatic ring of Phe⁴⁷⁵ is positioned to form a cation- π interaction with Arg⁶¹ [64] and a sulfur- π interaction [65] with Met⁸⁸ on Tfb1 (Figure 3. 3B). In a similar manner to what was seen with Phe⁵⁴ of p53 in the p53/Tfb1 complex, Phe⁴⁷⁹ of VP16C binds in a pocket formed by Gln⁴⁹, Ala⁵⁰, Thr⁵¹, Met⁵⁹, Leu⁶⁰ and Arg⁶¹ of Tfb1 (Figure 3. 3B-C). In this pocket, the aromatic ring of Phe⁴⁷⁹ is positioned to participate in an amino-aromatic interaction with Gln⁴⁹ [66], a sulfur- π interaction with Met⁵⁹ [65] and a cation- π interaction with Arg⁶¹ of Tfb1 (Figure 3. 3C) [64]. The third important hydrophobic residue in VP16C is Met⁴⁷⁸. Met⁴⁷⁸ is inserted in a narrow cleft on the surface of Tfb1 formed by Lys⁵⁷ and Met⁵⁹, where it makes van der Waals contacts with the side chains of Lys⁵⁷ and Met⁵⁹ of Tfb1. Like the p53 binding site [49], the VP16C binding site on Tfb1 is characterized by a positively-charged surface formed by Lys⁴⁷, Arg⁶¹, and Arg⁸⁶ that helps to position the negatively charged

VP16C. In addition, one specific ionic interaction is observed involving Glu⁴⁷⁶ of VP16C and Arg⁶¹ of Tfb1 ($d_{N\zeta-O\epsilon 2} = 2.1 \text{ \AA}$ in the minimized average structure) (Figure 3. 3D).

ITC and mutational studies of the interaction between VP16C and Tfb1

In earlier studies, we demonstrated via ITC and mutational analysis that Gln⁴⁹ on Tfb1 makes key contributions to the binding of Tfb1 to p53 [49]. In our Tfb1/VP16C structure, Gln⁴⁹ is positioned to make crucial interactions with Phe⁴⁷⁹ of VP16C, and Gln⁴⁹ is a key member of a complex interaction network on the surface of Tfb1. Therefore, we were interested in determining if mutation of Gln⁴⁹ to alanine [Tfb1(Q49A)] would significantly alter the binding of Tfb1 to VP16C, as was observed for Tfb1 binding to p53TAD2. This mutant was selected because it is positioned on the surface of Tfb1 and we have previously demonstrated that this mutation does not significantly alter the structure of the PH domain of Tfb1 [49]. ITC experiments of VP16C with Tfb1(Q49A) demonstrated that Gln⁴⁹ is important for the formation of the Tfb1/VP16C complex as we were unable to detect binding for this mutant (Table 1). This dramatic effect of the Q49A mutation on the binding of Tfb1 to VP16C is identical to what was observed with the Tfb1/p53 complex.

To assess the contribution of residues in the VP16C helix to Tfb1 binding, we performed site-directed mutagenesis at the three hydrophobic positions that make important contributions to the Tfb1/VP16C interface and we examined the binding of Tfb1 to the VP16C mutants by ITC (Table 1). In this assay, we were unable to detect binding to the VP16C(F475P), VP16C(M478P) and VP16C(F479P) mutants and these results indicate that these three VP16C mutants have K_d 's $\geq 100 \mu\text{M}$ (Table 1). This is consistent with the fact that proline would significantly impair formation of the helix in VP16C. Replacing either Phe⁴⁷⁵ or Phe⁴⁷⁹ by alanine had a less dramatic effect [24, 63]. As measured by the ITC method, VP16C(F475A) bound to Tfb1 with an apparent K_d of $1000 \pm 200 \text{ nM}$ and VP16C (F479A) bound to Tfb1 with an apparent K_d of $900 \pm 100 \text{ nM}$ (Table 1). These results are in agreement with our structural studies indicating that these three hydrophobic residues are important for formation of the interface between Tfb1 and VP16. In addition, these results are consistent with the fact that VP16C(F475P) and VP16C(F479P) are

almost completely devoid of transcriptional activation activity whereas VP16C(F475A) and VP16C(F479A) retain approximately 20-25% of their transcriptional activity [63].

Previously, we have shown that binding of p53TAD2 to Tfb1/p62 is significantly enhanced by diphosphorylation of p53 at both Ser⁴⁶ and Thr⁵⁵ and that the enhanced binding of each phosphorylation was additive [49]. Mutational analysis in combination with the structure of the Tfb1/p53 complex indicated that the increase in binding due to phosphorylation of Ser⁴⁶ is the result of an ionic interaction with Lys¹¹ of Tfb1, whereas the increase in binding due to phosphorylation of Thr⁵⁵ is the result of an ionic interaction with Arg⁶¹ of Tfb1 [49]. Thr⁵⁵ of p53 aligns with Thr⁴⁸⁰ in VP16 so this residue is conserved between p53 and VP16 (Fig. 4A). In contrast, Ser⁴⁶ of p53 aligns with Ala⁴⁷¹ of VP16C, thus the interaction between phosphorylated Ser⁴⁶ of p53 and Lys¹¹ of Tfb1 does not appear to be conserved in the Tfb1/VP16 complex (Figure 3. 4A). Interestingly, Asp⁴⁷² is the residue adjacent to Ala⁴⁷¹ in VP16C, and in a subset of our NMR structures of the Tfb1/VP16C complex it appears that this Asp⁴⁷² could form an ionic interaction with Lys¹¹. Therefore, this interaction would be somewhat analogous to the interaction observed with the phosphorylated Ser⁴⁶ in p53. In order to determine if an ionic interaction between Lys¹¹ of Tfb1 and Asp⁴⁷² of VP16C is contributing to the binding of VP16C to Tfb1, we mutated Asp⁴⁷² of VP16C to alanine [VP16C(D472A)] and we mutated Lys¹¹ of Tfb1 to glutamic acid [Tfb1(K11E)]. ITC experiments of Tfb1 with VP16C(D472A) demonstrated that Asp⁴⁷² is important for the formation of the Tfb1/VP16C complex as this mutant decreased binding by a factor of ~ 5 when compared to the wild-type VP16C (Table 1). Surprisingly, ITC experiments with VP16C and the Tfb1(K11E) mutant demonstrated that Tfb1(K11E) actually had a slightly higher affinity for VP16C than wild-type Tfb1 (Table 1). These results indicate that Lys¹¹ of Tfb1 does not make a significant contribution to VP16C binding and that there is no apparent ionic interaction between Asp⁴⁷² of VP16C and Lys¹¹ of Tfb1. The role of the negatively charged Asp⁴⁷² in the first position of the helix of VP16C in the Tfb1/VP16C complex may be for the stabilization of the α -helix through interaction with the helix dipole [67, 68]. This is supported by experiments with model peptides that have shown that aspartic acid at the first position of a helix stabilizes the helix [69, 70].

Comparison of the structures of the Tfb1/VP16C and Tfb1/p53 TAD2 complexes

The TADs of VP16 and p53 share several common binding partners, and the residues of p53 TAD2 that form the Tfb1 interface display sequence homology with VP16C (Figure 3. 4A). Thus, one would anticipate that the Tfb1/p53TAD2 and Tfb1/VP16C complexes would display many common features, and this is in fact the case. First, the structure of the PH domain of Tfb1 is essentially unchanged in the two complexes. In addition, one observes the formation of a 9-residue α -helix in the TAD of both p53 and VP16C upon complex formation with Tfb1. The VP16C and p53TAD2 helices both possess a face that contains three hydrophobic amino acid residues that are essential for binding to Tfb1. Superposition of the hydrophobic residues from p53TAD2 and VP16C, clearly demonstrates that they are located in similar positions on the α helices in the two complexes (Figure 3. 4B). The hydrophobic residues also formed similar types of interactions with Tfb1 in the two complexes even though these residues are not strictly conserved. VP16C is clearly mimicking p53TAD2 in making several types of contacts with Tfb1 such as hydrophobic, amino-aromatic, cation- π interaction and sulfur- π interactions (Figure 3. 4B). In addition, the ionic interaction involving Glu⁴⁷⁶ of VP16C with Arg⁶¹ of Tfb1 in the Tfb1/VP16 complex directly mimics the interaction between Glu⁵¹ of p53 and Arg⁶¹ of Tfb1 (Figure 3. 4C). Despite the similarities between the Tfb1/p53 and Tfb1/VP16C complexes, there are some important differences that occur at the amino-terminal end of the helices. In the case of p53, there is a proline residue in the first position of the helix (Pro⁴⁷) that is preceded by a serine residue (Ser⁴⁶). Ser⁴⁶ of p53 has been shown to be phosphorylated *in vivo* by the p38 mitogen activated protein kinase (MAPK) and the homeodomain interacting protein kinase (HIPK) [71], and we have previously demonstrated that phosphorylation of Ser⁴⁶ increases the binding affinity of p53TAD2 to Tfb1/p62 *in vitro*. In VP16C, there is an aspartic acid residue in the first position of the helix (Asp⁴⁷²). Our initial assumption was that Asp⁴⁷² of VP16 was serving a similar function as the phosphorylated Ser⁴⁶ in p53. However, mutational analysis clearly demonstrated this is not the case. In addition, it is clear that Ser⁴⁶ of p53 and Asp⁴⁷² of VP16 are oriented very differently with respect to each other (Figure 3. 4C). The different orientations would preclude them from participating in similar interactions with Tfb1. As we proposed earlier, Asp⁴⁷² may

stabilize the VP16 helix through interactions with the helix dipole and thereby stabilize its complex with Tfb1. In contrast, the phosphorylated Ser⁴⁶ of p53 forms an ionic interaction with Tfb1/p62 that is dependent on a signaling pathway involving either MAPK or HIPK [71].

Discussion

In this manuscript, we have examined the interaction of the Tfb1 subunit of TFIID with the TAD of VP16 at the atomic level. We demonstrate that the VP16C subdomain within the VP16TAD binds with higher affinity to the PH domain of Tfb1 than does VP16N. NMR structural studies demonstrate that VP16C is disordered in the free state, and forms a 9-residue α -helix involving residues Asp⁴⁷² to Thr⁴⁸⁰ in complex with Tfb1. The high-resolution structure of the Tfb1/VP16C complex demonstrates that the VP16C helix contains three hydrophobic residues (Phe⁴⁷⁵, Met⁴⁷⁸ and Phe⁴⁷⁹) that make several crucial contacts with Tfb1 via cation- π , sulfur- π , amino-aromatic and hydrophobic interactions. In addition, two acidic residues (Asp⁴⁷² and Glu⁴⁷⁶) in the VP16C helix also contribute significantly to the stability of the complex. Glu⁴⁷⁶ of VP16C forms an ionic interaction with Arg⁶¹ of Tfb1, whereas Asp⁴⁷² likely stabilizes the helix dipole of VP16C. The structure and interactions of the VP16C helix in the Tfb1/VP16C complex are remarkably similar to what we previously observed for the 9-residue helix of p53 in our NMR structure of a Tfb1/p53TAD2 complex. It is clear that VP16 mimics p53 in many interactions with Tfb1 [49]. Despite the remarkable similarities between the Tfb1/p53TAD2 complex and the Tfb1/VP16C complex, there are some important differences. The differences between the two complexes are at the amino-terminal end of the respective α -helices in VP16C and p53TAD2. For p53, phosphorylation of Ser⁴⁶ significantly enhances binding of the p53TAD2 to Tfb1/p62 through formation of an ionic interaction with Lys¹¹ in a loop between β 1 and β 2 of the PH domain of Tfb1. In the case of VP16, there is no equivalent phosphorylation site present in the sequence, and the closest acidic residue in VP16C (Asp⁴⁷²) that could mimic the phosphoserine of p53 does not interact with Lys¹¹ of Tfb1.

The Tfb1/VP16C structure supports many of the previous mutagenesis studies examining the role of VP16C in VP16-dependent activation [8, 24, 31, 33, 54, 63, 72]. These studies demonstrated that mutating the three hydrophobic residues (Phe⁴⁷⁵, Met⁴⁷⁸ and Phe⁴⁷⁹) in the VP16C helix to proline almost completely eliminates VP16 activation of target genes but that mutation of these residues to alanine or other hydrophobic residues such as leucine has a significantly less dramatic effect on activity. Based on the Tfb1/VP16 complex structure, it is evident that mutation of these three hydrophobic residues to proline should significantly alter the interaction of VP16 with Tfb1 by disrupting formation of the α helix, and the results from our ITC experiments confirm this. In addition, the F475A and F479A mutants decrease the binding of VP16C to Tfb1 by only 2.8-fold and 2.5-fold respectively and the decrease in binding observed with these two mutations is consistent with the fact that these mutants retain 25% of their activity in vivo [72]. It is also similar to what has been previously observed with TBP where the F475A mutant decreased VP16C binding to TBP by 2.3-fold and the F479A mutant decreased VP16C binding to TBP by 2.7-fold [72]. Given the fact that alanine is a common amino acid in α helices, it appears that the F475A and F479A mutants do not disrupt formation of the helix in VP16C and this leads to a much smaller effect on VP16C binding to Tfb1 and transcriptional activity than the proline mutants. In comparing VP16C/Tfb1 complex with the p53TAD2/Tfb1 complex, Phe⁴⁷⁵ of VP16C occupies the same position in the helix as Ile⁵⁰ of p53. Phe⁴⁷⁵ forms a cation- π interaction with Arg⁶¹ and a sulfur- π interaction with Met⁸⁸ on Tfb1 (Figure 3B) and Ile⁵⁰ forms van der Waals interactions with Met⁵⁹ and Met⁸⁸ of Tfb1 [49]. In the case of VP16C (F475A), the methyl group of alanine could position itself in a similar manner as one of the methyl groups of Ile⁵⁰ and participate in van der Waals interactions with Tfb1. This would partially compensate for the interactions lost with the aromatic ring of Phe⁴⁷⁵ and this would explain why only a small decrease in binding is observed with the VP16C (F475A) mutant. A similar argument could be used to explain why the F475L mutant of VP16C has similar activity as wild-type VP16C in vivo [72]. Likewise, Phe⁴⁷⁹ occupies the same position in the helix as Phe⁵⁴ of p53. These two residues are identical and their side chain fills a shallow pocket formed by Gln⁴⁹, Ala⁵⁰, Thr⁵¹, Met⁵⁹, Leu⁶⁰ and Arg⁶¹ of Tfb1. Specifically, the aromatic ring participates in cation- π interactions with Arg⁶¹ and amino-aromatic interactions with Gln⁴⁹ of Tfb1 in both complexes. It is clear that the

alanine mutant cannot participate in similar interactions with Tfb1 but it could participate in van der Waals interactions with Ala⁵⁰, Thr⁵¹, Met⁵⁹ or Leu⁶⁰ of Tfb1. The van der Waals interactions would partially compensate for the interactions observed with the aromatic ring. Again as suggested for Phe⁴⁷⁵, a series of van der waals interactions would explain why the VP16C (F479L) mutant has similar activity to the wild-type VP16C [72].

The structure of the Tfb1/VP16 structure also indicates that Phe⁴⁷³ within VP16C does not play a critical role in the formation of the interface with Tfb1 (Figure 3. 3A). Mutagenesis studies indicate that although Phe⁴⁷³ is important for transcriptional activation by VP16C, equivalent substitutions at Phe⁴⁷⁵ and Phe⁴⁷⁹ have more pronounced effects [24, 63]. Based on the structure of the Tfb1/VP16 complex, it appears that Phe⁴⁷³ plays a secondary role by stabilizing the formation of the helix or that it is important for other interactions during the transcription process. This is supported by the fact that mutation of Phe⁴⁷³ to proline completely abolishes transcriptional activity whereas mutation of Phe⁴⁷³ to alanine has very little effect [72]. These results are consistent with the fact that the proline mutation would block formation of the helix and the alanine mutation would allow for formation of the helix.

In addition to hydrophobic residues, mutational analysis indicates that acidic residues are important for VP16C-dependent activation albeit less critical than hydrophobic residues [22, 24, 63]. The structure of the complex indicates that many of the acidic residues of VP16C probably play a more general role by contributing to binding through a non-specific electrostatic effect with the basic binding site on Tfb1. This is consistent with mutagenesis studies indicating that for several cases multiple acidic residues must be mutated before one observes a decrease in VP16C-dependent activation [22]. Our structure of the complex does demonstrate that there is one crucial ionic interaction involving Glu⁴⁷⁶ of VP16C and this is consistent with the fact that mutation of Glu⁴⁷⁶ to alanine results in a 55% decrease in VP16C activity in a yeast-based transcriptional activation assay [63]. The role of Asp⁴⁷² is more complex. Our results indicate that Asp⁴⁷² of VP16C plays a crucial role in binding to Tfb1 although no direct ionic interaction is formed. In our NMR studies, we observe significant chemical shift changes in Asp⁴⁷² indicating that it is near the interface in the VP16C/Tfb1 complex. Similar chemical shift changes for Asp⁴⁷² have

been observed in previous NMR studies examining VP16C binding to TFIIB, PC4 and humanTAF_{II}31 [31, 33]. In addition, our ITC studies clearly demonstrate that alanine substitution at Asp⁴⁷² significantly reduces VP16C binding to Tfb1. We attribute this to the fact that Asp⁴⁷² is stabilizing the formation of the helix through interactions with the dipole moment. However, previous studies indicate that the D472A mutation does not alter the transactivation capacity of VP16C [24]. There are several possible explanations for this observation including the possibility that the D472A mutation is having a stabilizing effect on VP16C turnover and this helps compensate for the loss in binding activity. However, additional studies are required to understand more clearly the role of Asp⁴⁷² in VP16C binding to Tfb1.

It is interesting to compare our Tfb1/VP16C NMR structure with the results of other NMR studies examining the interactions of the VP16TAD with partner proteins. Two other studies have used NMR spectroscopy to examine the interactions of the VP16 TAD with target proteins [31, 33]. The first study examined binding between VP16C and TAF_{II}31 [31]. This study predicted that a 12-residue helix formed between residues Asp⁴⁷² and Leu⁴⁸³ in VP16C when in complex with TAF_{II}31. The second study [33] examined interactions between the full VP16TAD and TFIIB and PC4. In the second study, a 21-residue helix was predicted to form between residues Pro⁴⁶⁵ and Ile⁴⁸⁵ of VP16C when in complex with either TFIIB or PC4. As was seen in our Tfb1/VP16C complex, each of these studies provided strong NMR experimental evidence for the formation of an α -helix in VP16C upon complex formation with the partner proteins. In both studies, the residues predicted to form the helix overlap with the VP16C residues that interact with Tfb1 in our Tfb1/VP16C complex. Consistent with our results, mutational analysis of VP16C indicated that Phe⁴⁷⁹ was extremely crucial for the interaction with TAF_{II}31, TFIIB and PC4 [31, 33]. The difference between our results and results from previous studies is that the helix in the Tfb1/VP16C complex is slightly shorter than the helices from the three other complexes. Together, this suggests that VP16C might use similar interactions when binding to TAF_{II}31, TFIIB and PC4 and that a binding pocket for Phe⁴⁷⁹ is clearly a crucial determinant for interaction with VP16C.

The comparison of the Tfb1/p53 and Tfb1/VP16 structures is an important step towards developing a better understanding for acidic TADs binding to target proteins and in particular acidic TADs binding to the Tfb1/p62 subunit of TFIID. Such detailed structural information is absolutely essential to our efforts to design molecules that mimic transcription activators such as p53 and VP16. For example, it is now clear that the VP16C interacting helix aligns much better with the p53TAD2 helix as opposed to p53TAD1 helix as earlier reports suggested [31, 32]. Based on the structures of the two Tfb1 complexes, it is now possible to develop several simple rules that would begin to describe the requirements for a 9-residue helix binding to Tfb1/p62 subunit of TFIID. The rules would start with position 3 through 9 in the α -helix and these 7 positions make a significant number of interactions found at the complex interface. The simplified sequence for position 3 through 9 would be D/E Φ EQ Φ Φ T, where Φ is a hydrophobic residue with a slight preference for phenylalanine. Although the glutamine at position 6 is conserved between VP16 and p53, it does not contribute significantly to the interaction surface. Therefore this residue may not be strictly conserved in other acidic TADs that bind to p62/Tfb1. The threonine at position 9 (Thr⁵⁵ in p53 and Thr⁴⁸⁰ in VP16C) does not contribute significantly to binding in the case of p53 unless it is phosphorylated. To date there is no evidence that Thr⁴⁸⁰ of VP16 is phosphorylated, but it would be in position to make similar contributions as the phosphorylated Thr⁵⁵ does in the p53/p62 and p53/Tfb1 complexes (Figure 3. 4C). The requirements for the first two positions of the helix are not well defined by the two structures. Not surprisingly, the Tfb1/p62 subunit has been shown to interact with a number of additional transcriptional activators that contain acidic TADs, including the p65 subunit of NF κ B [73], E2F1 [74], EBNA2 [75, 76] and CIITA [77]. In each case, the TAD within the proteins contains a number of hydrophobic amino acids in addition to being highly acidic and therefore these domains are likely binding sites for p62/Tfb1. Structures of Tfb1/p62 in complex with other acidic TADs are needed to better define the requirements for the nine-residue helix.

The structure of the Tfb1/VP16C complex is the first high-resolution structure of the TAD of VP16 in a protein/protein complex. The structure sheds light on the role of the general transcription machinery in VP16 transcriptional activation. The structure of the Tfb1/VP16C

complex in combination with our previous NMR structure of the Tfb1/p53 complex [49] enables us to make a detailed comparison between a mammalian activator (p53) and a viral activator (VP16) on a common member of the general transcription machinery (Tfb1). Our results clearly demonstrate how the viral activator VP16 is very similar to the mammalian activator p53 when in complex with Tfb1. In addition, the structures of the Tfb1/VP16 and Tfb1/p53 complexes demonstrate how simple changes in the sequence could alter the requirement for post-translational events such as phosphorylation. It is interesting to point out that most of the acidic TADS studied to date are intrinsically unstructured in the free state, but form α -helical domains in complex with their target proteins as is seen with p53 and VP16 in complex with Tfb1 [31, 33, 41-43, 49, 78]. This observation, in combination with the fact that phosphorylation events often take place in unstructured regions of proteins [79], suggests that phosphorylation events within acidic TADS can play an important role in regulating their interactions with partner proteins. Future structural studies with other mammalian and viral activators bound to Tfb1/p62 are needed to better understand this form of regulation is in the case of TFIIF. In addition, more detailed structure-function studies are needed to further refine the precise requirements for acidic TADS binding to Tfb1/p62 and other target proteins.

Coordinates. The atomic coordinates of the Tfb1/VP16C complex have been deposited in the Protein Data Bank, www.rcsb.org (2k2u).

Acknowledgements

We thank Jimmy Dikeakos for critical reading of this manuscript. We thank Steven Triezenberg for providing the GST-VP16C and GST-VP16N clones. This work was supported by NCIC grant 016302 with funds from the Terry Fox Foundation, (J.G.O.). C.L. is recipient of a Postdoctoral Fellowship from the FQRS. P.D.L is recipient of a Postdoctoral Fellowship from the CIHR. P.L. is a Canadian research chair in structural biology of RNA.

References

1. Ptashne, M. and A.A.F. Gann, *Transcriptional activation by recruitment*. Nature, 1997. **386**: p. 569-577.
2. Ptashne, M. and A.A.F. Gann, *Activators and targets*. Nature, 1990. **346**: p. 329-331.
3. Prochasson, P., et al., *Targeting activity is required for SWI/SNF function in vivo and is accomplished through two partially redundant activator-interaction domains*. Mol. Cell, 2003. **12**(4): p. 983-990.
4. Gutierrez, J.L., et al., *Activation domains drive nucleosome eviction by SWI/SNF*. EMBO J., 2007. **26**(3): p. 730-740.
5. Bryant, G.O. and M. Ptashne, *Independent recruitment in vivo by Gal4 of two complexes required for transcription*. Mol. Cell, 2003. **11**: p. 1301-1309.
6. Goodrich, J.A., et al., *Drosophila TAF₄₀ interacts with both a VP16 activation domain and the basal transcription factor TFIIB*. Cell, 1993. **75**: p. 519-530.
7. Brown, S.A., et al., *Transcription activation domains stimulate initiation and elongation at different times and via different residues*. EMBO J., 1998. **17**: p. 3146-3154.
8. Blau, J., et al., *Three functional classes of transcriptional activation domains*. Mol. Cell. Biol., 1996. **16**: p. 2044-2055.
9. Bentley, D., *Regulation of transcriptional elongation by RNA polymerase II*. Curr. Opin. Genet. Dev., 1995. **5**: p. 210-16.
10. Brown, C.E., et al., *Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit*. Science, 2001. **292**(5525): p. 2333-2337.
11. Ito, M., et al., *Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators*. Mol. Cell, 1999. **3**(3): p. 361-370.
12. Black, J.C., et al., *A mechanism for coordinating chromatin modification and preinitiation complex assembly*. Mol. Cell, 2006. **23**(6): p. 809-818.
13. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcription components as determined by protein:protein cross-linking in vivo*. J. Biol. Chem., 2002. **277**: p. 46043-46050.
14. Kuras, L. and K. Struhl, *Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme*. Nature, 1999. **399**(6736): p. 609-613.
15. Yang, F., et al., *An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis*. Nature, 2006. **442**(7103): p. 700-704.

16. Yang, F., et al., *The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator*. Proc. Natl. Acad. Sci. USA, 2004. **101**(8): p. 2339-2344.
17. Tansey, W.P., et al., *Multiple regions of TBP participate in the response to transcriptional activators in vivo*. Genes & Dev., 1994. **8**(22): p. 2756-69.
18. Fields, S. and S.K. Jang, *Presence of a potent transcription activating sequence in the p53 protein*. Science, 1990. **249**(4972): p. 1046-9.
19. Triezenberg, S.J., R.C. Kingsbury, and S.L. McKnight, *Functional dissection of VP16, the transactivator of herpes simplex virus immediate early gene expression*. Genes & Dev., 1988. **2**: p. 718-729.
20. Levine, A.J., *p53, the cellular gatekeeper for growth and division*. Cell, 1997. **88**: p. 323-331.
21. Wysocka, J. and W. Herr, *The herpes simplex virus VP16-induced complex: the makings of a regulatory switch*. Trends Biochem. Sci., 2003. **28**(6): p. 294-304.
22. Cress, W.D. and S.J. Triezenberg, *Critical structural elements of the VP16 transcriptional activation domain*. Science, 1991. **251**: p. 87-90.
23. Das, G., C.S. Hinkley, and W. Herr, *Basal promoter elements as a selective determinant of transcriptional activator function*. Nature (London), 1995. **374**(6523): p. 657-60.
24. Sullivan, S.M., et al., *Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus*. Nucleic Acid Res., 1998. **26**: p. 4487-4496.
25. Mitchell, P.J. and R. Tjian, *Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins*. Science, 1989. **245**: p. 371-378.
26. Regier, J.L., F. Shen, and S.J. Triezenberg, *Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator*. Proc. Natl. Acad. Sci. USA, 1993. **90**: p. 883-887.
27. Candau, R., et al., *Two tandem and independent sub-activation domains in the amino terminus of p53 require the adaptor complex for activity*. Oncogene, 1997. **12**: p. 807-816.
28. Hope, I.A., S. Mahadevan, and K. Struhl, *Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein*. Nature, 1988. **333**: p. 635-640.
29. Gill, G., I. Sadowski, and M. Ptashne, *Mutations that increase the activity of a transcriptional activator in yeast and mammalian cells*. Proc. Natl. Acad. Sci. USA, 1990. **87**: p. 2127-2131.
30. Drysdale, C.M., et al., *The transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids*. Mol. Cell. Biol., 1995. **15**: p. 120-1233.

31. Uesugi, M., et al., *Induced ζ -helix in the VP16 activation domain upon binding to a human TAF*. Science, 1997. **277**: p. 1310-1313.
32. Uesugi, M. and G.L. Verdine, *The ζ -helical FXXFF motif in p53: TAF interaction and discrimination by MDM2*. Proc. Natl. Acad. Sci. USA, 1999. **96**(26): p. 14801-14806.
33. Jonker, H.R.A., et al., *Structural properties of the promiscuous VP16 activation domain*. Biochemistry, 2005. **44**(3): p. 827-839.
34. Stringer, K.F., C.J. Ingles, and J. Greenblatt, *Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID*. Nature, 1990. **345**: p. 783-786.
35. Ingles, C.J., et al., *Reduced binding of TFIID to transcriptionally comprised mutants of VP16*. Nature, 1991. **351**: p. 588-590.
36. Neeley, K.E., et al., *Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays*. Mol. Cell, 1999. **4**: p. 649-655.
37. Neeley, K.E., et al., *Transcription activator interactions with multiple SWI/SNF subunits*. Mol. Cell. Biol., 2002. **22**: p. 1615-1625.
38. Yudkovsky, N., et al., *Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators*. Genes & Dev., 1999. **13**: p. 2369-2374.
39. Chang, J., et al., *Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein*. J. Biol. Chem., 1995. **270**(42): p. 25014-19.
40. Sigler, P.B., *Transcription activation. Acid blobs and negative noodles*. Nature, 1988. **333**: p. 210-212.
41. O'Hare, P. and G. Williams, *Structural studies of the acidic transactivation domain of the Vmw65 protein of herpes simplex virus using proton NMR*. Biochemistry, 1992. **31**: p. 4150-4156.
42. Lee, H., et al., *Local structural elements in the mostly unstructured transcriptional activation domain of human p53*. J. Biol. Chem., 2000. **275**: p. 29426-29432.
43. Donaldson, L. and J.P. Capone, *Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1*. J. Biol. Chem., 1992. **267**: p. 1411-1414.
44. Shen, F., et al., *Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors*. J. Biol. Chem., 1996. **271**: p. 4827-4837.
45. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain*. Science, 1996. **274**: p. 948-953.

46. Lee, C., et al., *Structural basis for the recognition of the E2F transactivation domain by the retinoblastoma tumour suppressor*. *Genes & Dev.*, 2002. **16**: p. 3199-3212.
47. Bochkareva, E., et al., *Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A*. *Proc. Natl. Acad. Sci. USA*, 2005. **102**(43): p. 15412-15417.
48. Chi, S.-W., et al., *Structural details on mdm2-p53 interaction*. *J. Biol. Chem.*, 2005. **280**(46): p. 38795-38802.
49. Di Lello, P., et al., *Structure of the Tfb1/p53 complex: insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53*. *Mol. Cell*, 2006. **22**(6): p. 731-740.
50. Dyson, H.J. and P.E. Wright, *Coupling of folding and binding for unstructured proteins*. *Curr. Opin. Struct. Biol.*, 2002. **12**(1): p. 54-60.
51. Uhlmann, T., et al., *The VP16 activation domain establishes an active Mediator lacking CDK8 in vivo*. *J. Biol. Chem.*, 2007. **282**(4): p. 2163-2173.
52. Taddei, A., et al., *Nuclear pore association confers optimal expression levels for an inducible yeast gene*. *Nature*, 2006. **441**(7094): p. 774-778.
53. Moore, I., M. Samalova, and S. Kurup, *Transactivated and chemically inducible gene expression in plants*. *Plant J.*, 2006. **45**(4): p. 651-683.
54. Xiao, H., et al., *Binding of basal transcription factor TFIIH to the acidic activation domain of VP16 and p53*. *Mol. Cell. Biol.*, 1994. **14**: p. 7013-7024.
55. Di Lello, P., et al., *NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites*. *Biochemistry*, 2005. **44**(21): p. 7678-7686.
56. Houtman, J.C.D., et al., *Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences*. *Biochemistry*, 2004. **43**(14): p. 4170-4178.
57. Delaglio, F., et al., *NMRPipe: A multidimensional spectral processing system based on UNIX pipes*. *J. Biomol. NMR*, 1995. **6**(3): p. 277-93.
58. Johnson, B.A. and R.A. Blevins, *NMRView: A computer program for the visualization and analysis of NMR data*. *J. Biomol. NMR*, 1994. **4**: p. 603-614.
59. Cornilescu, G., F. Delaglio, and A. Bax, *Protein Backbone angle restraints from searching a database for chemical shift and sequence homology*. *J. Biomol. NMR*, 1999. **13**: p. 289-302.
60. Brunger, A.T., et al., *Crystallography and NMR system (CNS): A new software system for macromolecular structure determination*. *Acta. Cryst.*, 1998. **D54**: p. 905-921.
61. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR*. *J. Biomol. NMR*, 1996. **8**(4): p. 477-486.

62. Koradi, R., M. Billeter, and K. Wüthrich, *MOLMOL: A program for display and analysis of macromolecular structures*. J. Mol. Graphics, 1996. **14**: p. 51-55.
63. Nedialkov, Y.A. and S.J. Triezenberg, *Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region*. Arch. Biochem. and Biophys., 2004. **425**(1): p. 77-86.
64. Gallivan, J.P. and D.A. Dougherty, *Cation- π interactions in structural biology*. Proc. Natl. Acad. Sci. USA, 1999. **96**(17): p. 9459-9464.
65. Reid, K.S.C., P.F. Lindley, and J.M. Thornton, *Sulfur-aromatic interactions in proteins*. FEBS Letters, 1985. **190**(2): p. 209-13.
66. Burley, S.K. and G.A. Petsko, *Amino-aromatic interactions in proteins*. FEBS Let., 1986. **203**: p. 139-143.
67. Wada, A., *The ζ -helix as an electric macro-dipole*. Adv. Biophys., 1976. **9**: p. 1-63.
68. Hol, W.G.L., P.T.v. Duijnen, and H.J.C. Berendsen, *The ζ -helix dipole and the properties of proteins*. Nature, 1978. **273**: p. 443-446.
69. Huyghues-Despointes, B.M.P., J.M. Scholtz, and R.L. Baldwin, *Effect of a single aspartate on helix stability at different positions in a neutral alanine-based peptide*. Prot. Sci., 1993. **2**(10): p. 1604-11.
70. Scholtz, J.M., et al., *The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide*. Biochemistry, 1993. **32**(37): p. 9668-76.
71. Bode, A.M. and Z. Dong, *Post-translational modification of p53 in tumorigenesis*. Nature Rev. Cancer, 2004. **4**(10): p. 793-805.
72. Nedialkov, Y.A., D.D. Shooltz, and S.J. Triezenberg, *Purification and protein interaction assays of the VP16C transcription activation domain*. Meth. Enzymology, 2003. **370**(RNA Polymerases and Associated Factors, Part C): p. 522-535.
73. Kim, Y.K., et al., *Recruitment of TFIIH to the HIV LTR is a rate-limiting step in the emergence of HIV from latency*. EMBO J., 2006. **25**(15): p. 3596-3604.
74. Pearson, A. and J. Greenblatt, *Modular organization of the E2F1 activation domain and its interaction with general transcription factors TBP and TFIIH*. Oncogene, 1997. **15**: p. 2643-2658.
75. Wang, B.Q. and Z.F. Burton, *Functional domains of human RAP74 including a masked polymerase binding domain*. J. Biol. Chem., 1995. **270**: p. 27035-27044.
76. Tong, X., et al., *The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIIB, TAF40, and RPA70 but not with TATA-binding protein*. J. Virol., 1995. **69**(1): p. 585-8.

77. Mahanta, S.K., et al., *Transactivation by CIITA, the type II bare lymphocyte syndrome-associated factor, requires participation of multiple regions of the TATA box binding protein*. Proc. Natl. Acad. Sci. USA, 1997. **94**(12): p. 6324-6329.
78. Dahlman-Wright, K., et al., *Structural characterization of a minimal functional transactivation domain from glucocorticoid receptor*. Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 1699-1703.
79. Iakoucheva, L.M., et al., *The importance of intrinsic disorder for protein phosphorylation*. Nucleic Acids Res., 2004. **32**(3): p. 1037-1049.

Tables

Table 3. 1: Comparison of the K_d (nM) values for the binding of VP16 and Tfb1₁₋₁₁₅ using ITC.		
Tfb1 ₁₋₁₁₅	VP16N	1000 ± 100
Tfb1 ₁₋₁₁₅	VP16C	360 ± 40
Tfb1 ₁₋₁₁₅	VP16C (D472A)	2000 ± 300
Tfb1 ₁₋₁₁₅	VP16C (F475P)	NB*
Tfb1 ₁₋₁₁₅	VP16C (F475A)	1000 ± 200
Tfb1 ₁₋₁₁₅	VP16C (F479P)	NB
Tfb1 ₁₋₁₁₅	VP16C (F479A)	900 ± 100
Tfb1 ₁₋₁₁₅	VP16C (M478P)	NB*
Tfb1 ₁₋₁₁₅ (Q49A)	VP16C	NB*
Tfb1 ₁₋₁₁₅ (K11E)	VP16C	250 ± 30
*NB , no binding detected $K_d \geq 100 \mu\text{M}$.		

Table 3. 2. Structural Statistics of the Tfb1/VP16C Complex^a

Restraints used for the Structure Calculations	
Total number of NOE distances restraints	1568
Short-range (intraresidue)	578
Medium-range ($ i-j \leq 4$)	640
Long-range	309
Intermolecular	41
Hydrogen bond	36
Number of dihedral angle restraints (φ, ψ)	158
Structural Statistics	
Rms deviations from idealized geometry	
Bonds (Å)	0.0025 ± 0.00006
Angles (deg)	0.3918 ± 0.0057
Impropers (deg)	0.2753 ± 0.132
Rms deviations from distance restraints (Å)	0.0215 ± 0.0005
Rms deviations from dihedral restraints (deg)	0.4438 ± 0.0357
Ramachandran statistics (%) ^b	
Residues in most favoured regions	82.7
Residues in additional allowed regions	15.8
Residues in generously allowed regions	1.4
Residues in disallowed regions	0.2
Coordinate Precision	
Atomic pairwise rmsd (Å) ^c	
Tfb1/VP16C complex	
Backbone atoms (C', C ^α , N)	0.74 ± 0.16
All heavy atoms	1.42 ± 0.19
Tfb1 alone	
Backbone atoms (C', C ^α , N)	0.62 ± 0.12
All heavy atoms	1.31 ± 0.17
VP16C alone	
Backbone atoms (C', C ^α , N)	0.43 ± 0.23
All heavy atoms	1.57 ± 0.37

^a The 20 conformers with the lowest energy were selected for statistical analysis. Due to the absence of medium-range, long-range, and intermolecular NOEs involving residues 456-468 and 483-490 of VP16C, these amino acids were not included in the structure calculations.

^b Based on PROCHECK-NMR analysis.

^c Only residues 4-64 and 85-112 of Tfb1 and residues 470-482 of VP16C were used for rmsd calculations. Residues at the N-terminus (1-3), at the C-terminal (113-115) and in the flexible loop (65-84) of Tfb1, as well as residues at the N-terminus (456-468) and at the C-terminus (484-490) of VP16C, were not included in the calculation.

Figures

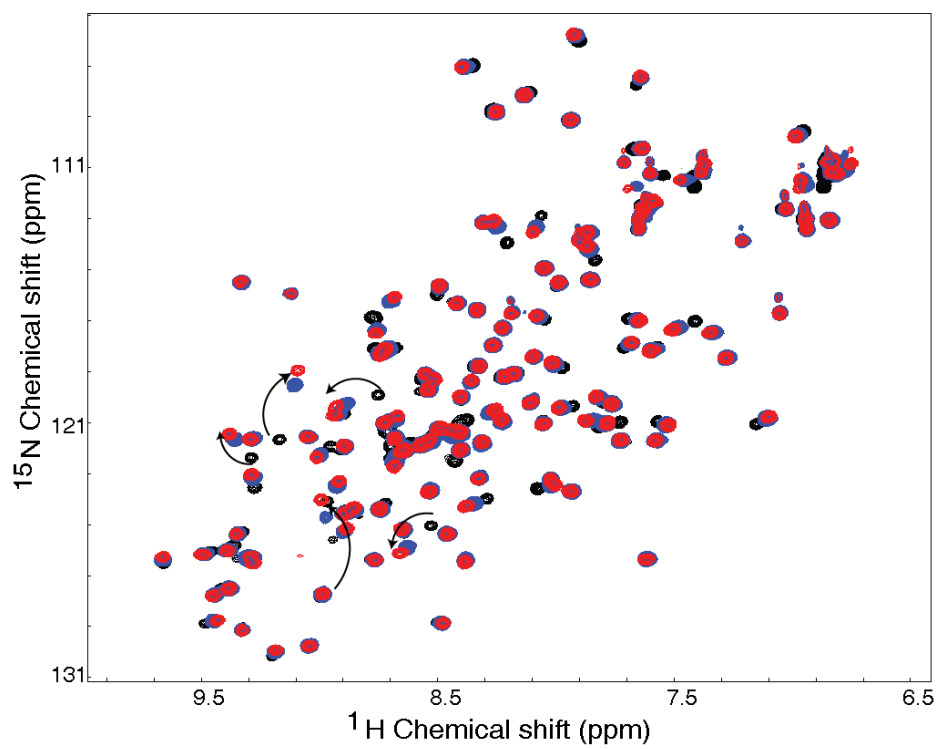


Figure 3. 1. VP16C bind to the PH domain of Tfb1. Overlay of the 2D ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Tfb1₁₋₁₁₅ in its free form (black), in the presence of 0.75 equivalent (blue) and 1 equivalent (red) of VP16C.

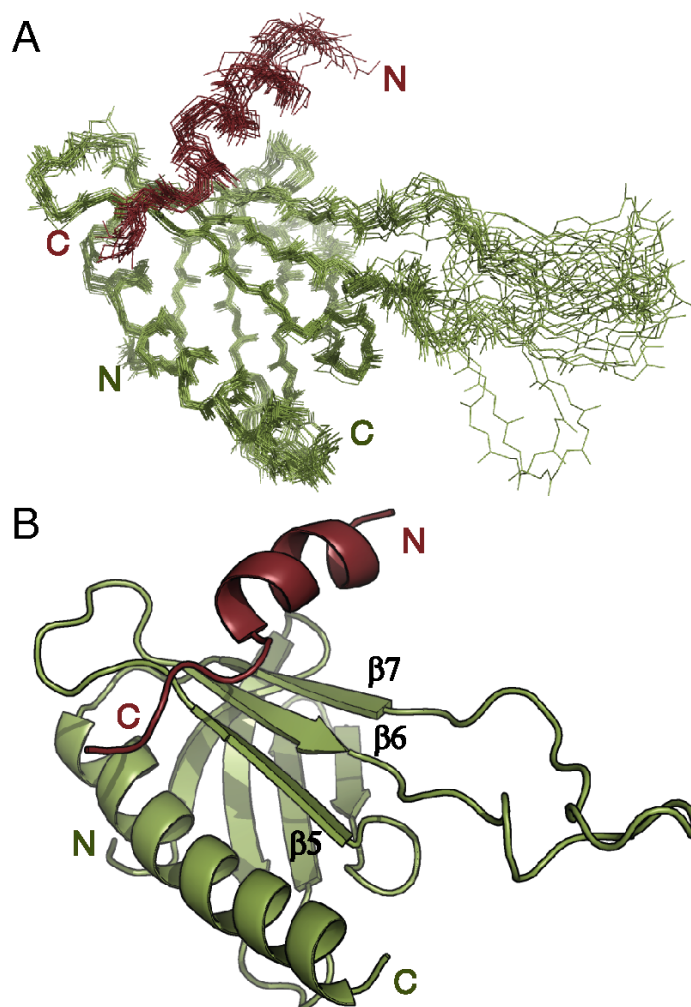


Figure 3. 2. Structure of the Tfb1/VP16C complex. (A) Overlay of the 20 lowest-energy structures of the complex between Tfb1₁₋₁₁₅ (green) and VP16C (red). The structures were superimposed using the backbone atoms C', C^α and N of residues 4-65 and 85-112 of Tfb1 and residues 470-482 of VP16C. (B) Ribbon representation of the lowest-energy structure of the Tfb1₁₋₁₁₅/VP16C complex.

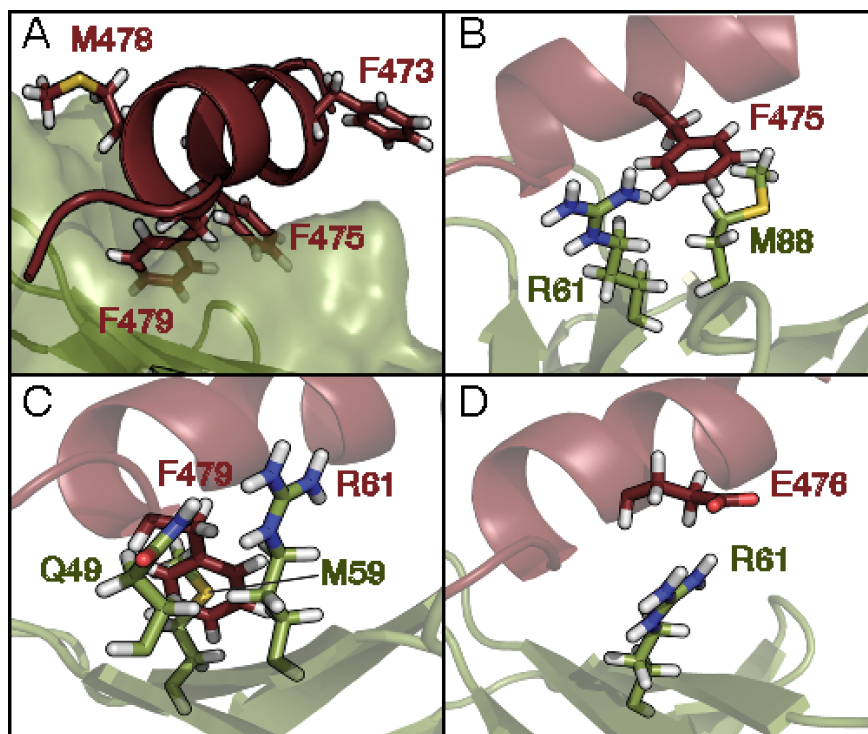


Figure 3. 3. Structural details of the Tfb1₁₋₁₁₅/VP16C interface. (A) The 3D structure of Tfb1 is shown as a ribbon within the transparent molecular surface (green), and the helix of VP16C is represented as a ribbon (red). (B-D) The backbone of Tfb1 and VP16C are shown as ribbons (green and red, respectively). (B) Phe⁴⁷⁵ of VP16C forms cation- π and sulfur- π interactions with Arg⁶¹ and Met⁸⁸ of Tfb1, respectively (C) Phe⁴⁷⁹ of VP16C forms cation- π , amino-aromatic and sulfur- π interactions with Arg⁶¹, Gln⁴⁹ and Met⁵⁸ of Tfb1, respectively. (D) Arg⁶¹ of Tfb1 forms an ionic interaction with Glu⁴⁷⁶ of VP16C.

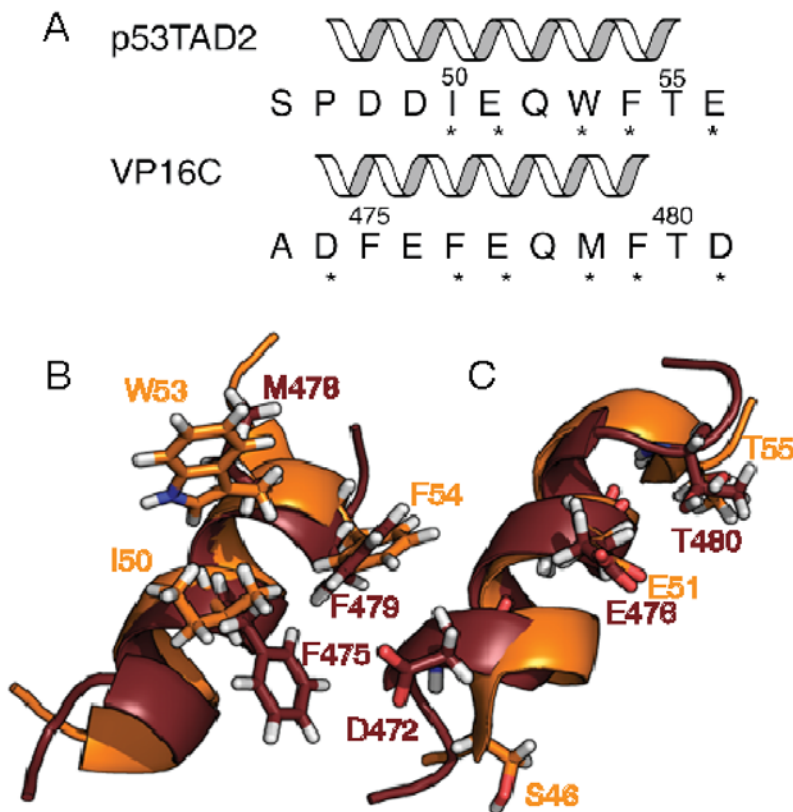
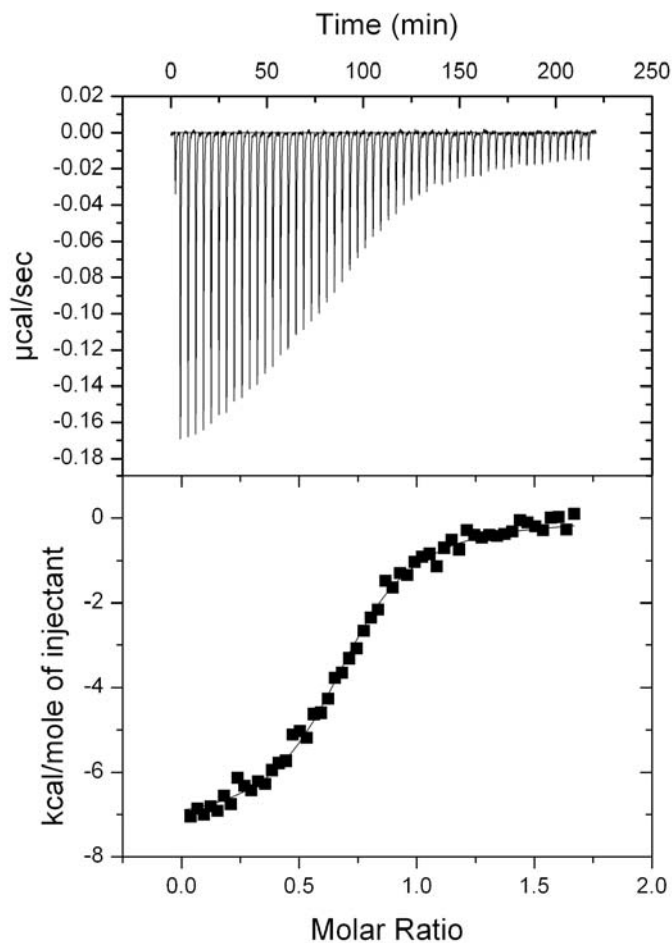
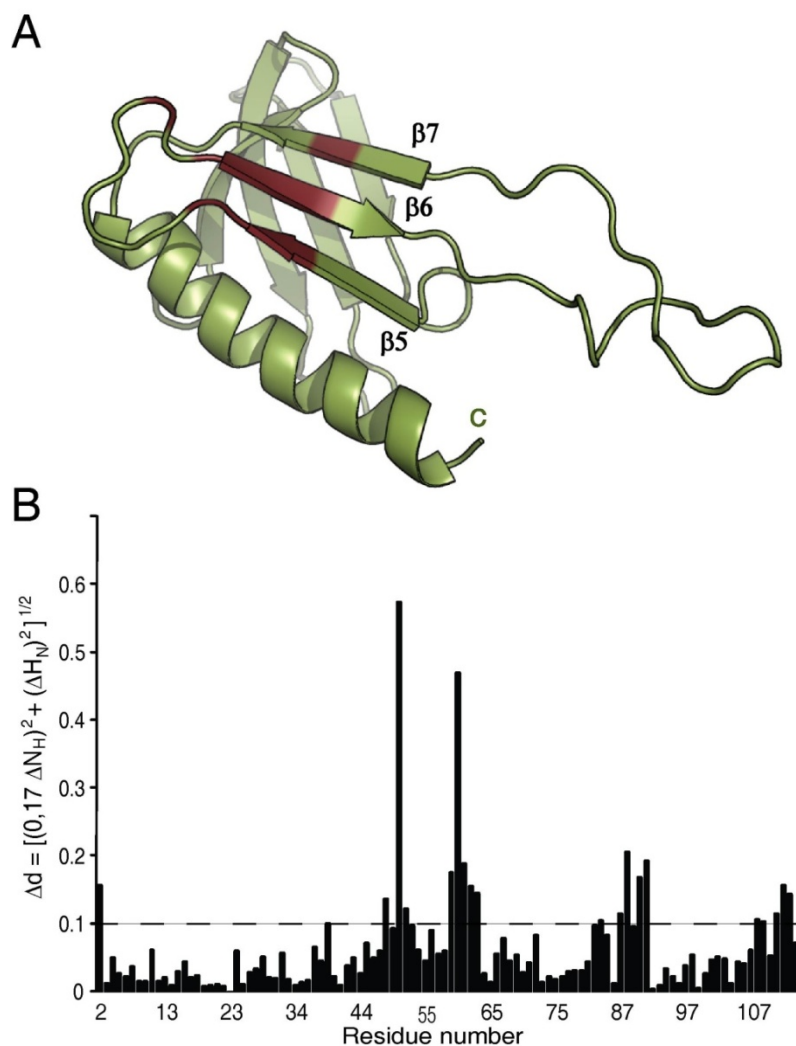


Figure 3. 4. Comparison of VP16C and p53TAD2 in complex with Tfb1. (A) Sequences of human p53TAD2 and Herpes Simplex Virus VP16C are aligned based on the 9-residue α helix with residue numbering and secondary structure elements represented above each protein sequences. An asterisk below a residue indicates that the amino acids show intermolecular NOE(s) with Tfb1 in its respective complex. (B-C) Overlay of residues 45-57 of the p53TAD2 (in orange) and residues 469-482 of VP16C (in red). (B) The three hydrophobic residues from the helix of VP16C (Phe⁴⁷⁵, Phe⁴⁷⁹ and Met⁴⁷⁸) at the complex interface are aligned with the three hydrophobic amino acids from the helix of p53TAD2 (Ile⁵⁰, Phe⁵⁴ and Trp⁵³) at the complex interface. (C) Glu⁴⁷⁶ and Thr⁴⁸⁰ of VP16C and Glu⁵¹ and Thr⁵⁵ of p53 are in similar positions in the two structures. In contrast, Asp⁴⁷² of VP16C and Ser⁴⁶ of p53TAD2 are positioned on different faces of the helices.

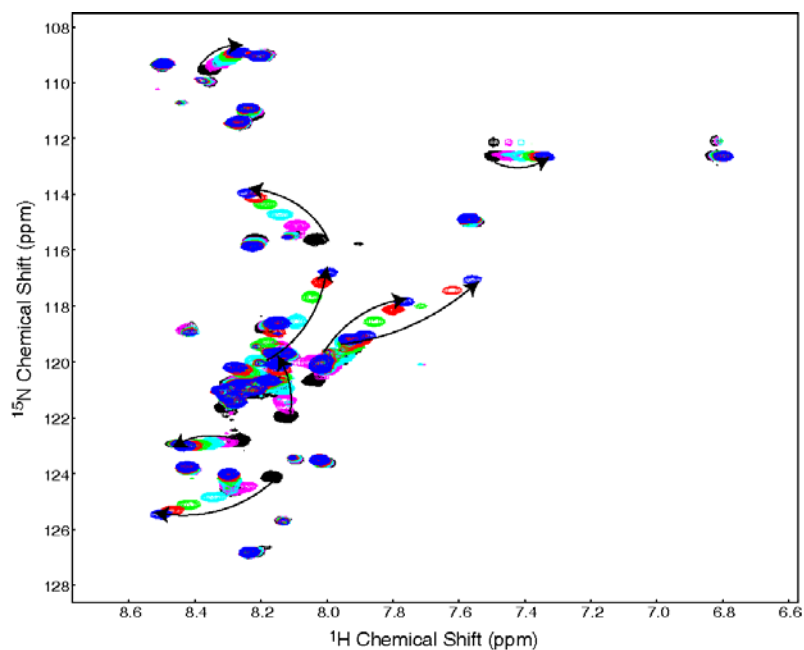
Supplementary Figures



Supplementary Figure 3. 1. ITC analysis of VP16C binding to Tfb₁₋₁₁₅. Thermogram obtained by successive additions of VP16C to Tfb₁₋₁₁₅. The curve was fit using a single-binding site to solve for the stoichiometry (N), the dissociation constant (K_d) the change in enthalpy (ΔH). The experimental conditions are as described in the experimental procedures and the dissociation constants are reported in Table 1. All injections fit the single binding site mechanism with 1:1 stoichiometry.



Supplementary Figure 3. 2. VP16C interacts with the PH domain of Tfb1₁₋₁₁₅. (A) Ribbon model of the free Tfb1₁₋₁₁₅ structure⁵⁵. The residues showing significant chemical shift changes ($\Delta\delta > 0.1$ ppm; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) in the 2D ¹H-¹⁵N HSQC spectra of Tfb1₁₋₁₁₅ in the presence of 1 equivalent of VP16C are shown in red. (B) Histogram of the variations in the chemical shifts ($\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) observed between the 2D ¹H-¹⁵N HSQC spectra of free ¹⁵N-labeled Tfb1₁₋₁₁₅ and that of Tfb1 in the presence of 1 equivalent of VP16C.



Supplementary Figure 3. 3. NMR titration of ^{15}N -labeled VP16C with Tfb1₁₋₁₁₅. Overlay of the 2D ^1H - ^{15}N HSQC spectra for ^{15}N -labeled VP16C in its free form (black), and in presence of 0.2 (pink), 0.4 (aqua), 0.6 (green), 0.8 (red) and 1 equivalent (blue) of unlabeled Tfb1₁₋₁₁₅.

Préface au Chapitre IV

Le facteur de transcription érythropoïétique EKLF ainsi que le facteur général de transcription TFIID sont des acteurs principaux de l'activation du gène de la β -globine. L'activateur EKLF possède un TAD similaire, tant au point de vue de sa fonction que de sa composition en acide aminés, au TAD de la protéine suppresseur de tumeur p53. Par conséquent, nous avons voulu savoir si le TAD d'EKLF pouvait interagir avec la sous-unité p62/Tfb1 de TFIID. En combinant des études de liaison *in vitro* par ITC et par spectroscopie RMN, nous avons déterminé la structure tridimensionnelle de ce complexe et analysé le mode de liaison de ces deux partenaires. La structure révèle que le mode de liaison d'EKLF à Tfb1 se distingue du mode de liaison canonique des activateurs transcriptionnels qui, habituellement, adoptent une structure secondaire en hélice- α , tels p53 TAD2 et VP16 TADC. En effet, la conformation d'EKLF en interaction avec Tfb1 est majoritairement étendue. Curieusement, ce mode de liaison est similaire au mode de liaison du facteur général TFIIE α à la sous unité p62. Ces résultats nous ont permis d'identifier un acide aminé essentiel à la transcription et à l'interaction p62/EKLF : le Trp73. La mutation de ce seul résidu abroge l'interaction entre ces deux partenaires et abolit l'activation de la transcription dans les érythrocytes.

Les résultats de ce chapitre sont exposés sous la forme d'un article scientifique prêt à être soumis dans la revue *Molecular Cell*. Lors de ces études, Mathieu Lussier-Price a été impliqué respectivement dans la production et la purification de protéines et dans la réalisation d'expériences d'ITC. Geneviève Arsenault a réalisé des essais *in vivo*. Paola Di Lello a été impliquée dans l'acquisition de spectres RMN. Pour ma part, j'ai réalisé les clonages pour les études *in vivo* et *in vitro*, la production et la purification des protéines marquées et non marquées nécessaires à l'acquisition des spectres RMN, l'attribution séquentielle, l'attribution des spectres NOE ainsi que les calculs de structure. J'ai également réalisé la majorité des expériences d'ITC. De plus, j'ai largement participé à la rédaction de ce manuscrit. C'est pourquoi je suis premier auteur de cet article.

Chapitre IV

Caractérisation structurale et fonctionnelle d'un domaine de transactivation acide atypique d'EKLF

Article 3: Structural and Functional Characterization of an Atypical Acidic Transactivation Domain in the Erythroid-Kruppel-Like-Factor (EKLF).

Caroline Mas, Mathieu Lussier-Price, Thomas Morse, Geneviève Arsenault, Paola Di Lello, Shefali Soni, James J. Bieker and James G. Omichinski¹

Reprinted with permission from : Mathieu Lussier-Price, Thomas Morse, Geneviève Arsenault, Paola Di Lello, Shefali Soni, James J. Bieker and James G. Omichinski¹

**Structural and Functional Characterization of an Atypical Acidic
Transactivation Domain in the Erythroid-Krüppel-Like-Factor
(EKLF)**

Caroline Mas¹, Mathieu Lussier-Price¹, Thomas Morse¹, Geneviève Arsenault¹, Paola Di Lello¹,
Shefali Soni², James J. Bieker² and James G. Omichinski^{1*}

¹Département de Biochimie, Université de Montréal C.P. 6128 Succursale Centre-Ville,
Montréal, QC H3C 3J7 Canada, ²Mount Sinai School of Medicine, Department of
Developmental and Regenerative Biology, Box 1020, One Gustave Levy Place, New York, NY
10029, U.S.A.

*Correspondence: James G. Omichinski, Département de Biochimie, Université de Montréal C.P.
6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada.

Abstract

Erythroid Kruppel-Like Factor (EKLF) plays an important role in erythroid development by stimulating expression of the β -globin gene. We have examined the details by which the minimal transactivation domain (TAD) of EKLF (EKLFTAD) interacts with transcriptional regulators to control gene expression. We demonstrate that EKLFTAD displays homology to the p53TAD and like the p53TAD, EKLFTAD can be divided into two subdomains (EKLFTAD1 and EKLFTAD2). Sequence analysis indicates that EKLFTAD2 is conserved in several other KLF proteins and that EKLFTAD2 directly binds the amino-terminal PH domain of the Tfb1/p62 subunit of TFIID (Tfb1PH/p62PH) and four domains of CBP/p300. Structural characterization of the EKLFTAD2/Tfb1PH complex demonstrates that EKLFTAD2 binds in an extended conformation, which is in contrast to the α -helical conformation seen with p53TAD2 and VP16 in complex with the same Tfb1PH. These studies provide the first mechanistic details as to how the minimal EKLFTAD functions as well as insights into potential targets of the TADs of other KLF family members. In addition, they demonstrate that not only have acidic TADs evolved so that they bind using different conformations on a common target, but that coupling folding and binding is not a strict requirement for their ability to bind multiple partners.

Introduction

The Krüppel-like factors (KLF) are a family of transcriptional regulatory proteins that function in a number of tissue-specific processes including hematopoietic development (KLF1), lung development (KLF2), stem cell differentiation (KLF4), and cellular proliferation (KLF5) [1, 2]. All KLF family members contain a homologous DNA-binding domain (DBD) consisting of three zinc-finger domains (Cys₂His₂) separated by a conserved linker sequence that helps distinguish them from the closely related Sp1 and Krox zinc-finger families [1, 2]. The DBD is positioned at the carboxyl-terminal end of the protein and enables KLF proteins to bind to “GT-box” or “CACCC” sites within promoter regions [1, 2]. In contrast to their highly homologous DBD, KLF proteins vary considerably at the amino-terminal end of the protein, and the KLF family can be divided into subgroups based on the types of domains present in this region [1, 2]. The amino-terminal regions of KLF proteins contain a number of different types of transcriptional regulatory domains including acidic transactivation domains (TADs), proline-rich TADs, Sin-3 interacting repressor domains (SIDs) and CtBP2 interacting repressor domains (PVLS/T motifs) [2]. These domains enable KLFs to regulate their target gene specificity through interactions with other transcriptional co-regulators.

In addition to being the founding member of the KLF family, Erythroid Krüppel-Like Factor (EKLF/KLF1) is a transcriptional activator that plays a key role in erythroid development [3-5]. EKLF stimulates expression of the adult β -globin gene and helps enable the switch from the fetal γ -globin to the adult β -globin [6-11]. EKLF specifically binds to a CAACC element in the β -globin promoter [1, 3, 12] and either the absence of EKLF or mutations in the β -globin CAACC element leads to severe β -thalassemia due to decreased β -globin production [6, 7, 13]. In addition, EKLF plays a role in generating an open chromatin structure on the adult β -globin gene through protein-protein interactions with a number of transcriptional regulatory factors including general transcription factors, chromatin remodeling complexes and histone acetyl transferases (HATs) [14-21].

The human EKLf protein contains 362 amino acids that can be divided into two functional domains (numbered as per the human sequence): the characteristic carboxyl-terminal DBD (residues 292-362) and an amino-terminal TAD (residues 1-271) [3, 22]. Like the DBD, the TAD plays an important role in EKLf's ability to activate β -globin gene expression [21-24]. Initial experiments demonstrated that the first one hundred amino acids constituted the minimal TAD of EKLf (EKLfTAD) [22]. Subsequent experiments demonstrated that the first forty amino acids within this minimal EKLfTAD play a role in recruiting additional cofactors that are required for β -globin activation and that residues between forty and one hundred are essential for β -globin gene activation [22]. The minimal EKLfTAD is highly acidic (pI = 3.8) and the combination of acidic and hydrophobic amino acids within this domain is similar to that found in the TAD of other potent transcriptional activators including the tumor suppressor protein p53 and the Herpes Simplex viral protein 16 (VP16) [25-30]. The presence of two critical regions within the minimal EKLfTAD is also consistent with the fact that the TAD of both p53 and VP16 can be divided into two subdomains each capable of independently activating transcription [29, 31-34]. In the case of the p53TAD, the two subdomains (p53TAD1 and p53TAD2) have been shown to participate in distinct interactions with general transcription factors (TFIID, TFIIB and TFIIF), the mediator complex, MDM2 and with various domains of the HAT CBP/p300 [35-38].

As is the case with other proteins containing acidic TADs including p53 [39-42], posttranslational modification of residues within the minimal EKLfTAD can significantly alter the functional properties of EKLf. EKLf has been shown to be phosphorylated at T15 and T23. In the case of T23, the phosphorylation is regulated by casein kinase 2 (CK2) and mutation of T23 to alanine significantly decreases the ability of EKLf to activate β -globin gene expression [24]. In addition, EKLf is sumoylated at K54 within the minimal EKLfTAD and sumoylation of EKLf is required for EKLf-mediated repression of megakaryocyte formation [43]. Despite the importance of the minimal EKLfTAD, mechanistic details by which it interacts with other proteins to regulate EKLf function are unknown. However, like other acidic TAD, it is presumed that EKLfTAD participates in protein-protein interactions with transcriptional regulatory factors

such as general transcription factors, chromatin remodeling complexes and histone acetyl transferases (HATs) [14, 16-18, 21].

In this article, we use isothermal titration calorimetry (ITC), NMR spectroscopy and in vivo activation experiments to define the molecular basis of the interactions between the minimal EKLFTAD (residue 1-90 of human EKLF [22]) and transcriptional regulatory factors required for β -globin gene activation. We demonstrate that EKLFTAD displays homology to the p53TAD and like the p53TAD, EKLFTAD can be divided into two subdomains (EKLFTAD1 and EKLFTAD2). Sequence analysis indicates that EKLFTAD2 is conserved in several other KLF proteins, and EKLFTAD2 directly binds with the amino-terminal PH domain of the Tfb1/p62 subunit of TFIIH (Tfb1PH/p62PH) and four domains of CBP/p300. We also demonstrate that EKLFTAD2 binds to a similar region on p62PH/Tfb1PH as p53TAD2 and that EKLFTAD2 competes with p53TAD2 for binding. NMR structural analysis of an EKLFTAD2/Tfb1PH complex demonstrates that EKLFTAD2 binds Tfb1PH in a unique manner that is distinct from that previously determined for the TAD of either p53 or VP16 [40, 44]. Our studies suggest that acidic TADs have evolved so that they are able to bind a common target using different mechanisms and that their targets have also co-evolved so that they are able to recognize different structural features in their cognate TADs. In addition, our results indicate that coupled folding and binding may not be a strict requirement for the ability of acidic TADs to bind multiple partners as well as provide insights into how several other KLF proteins may function in activating transcription.

Methods

Cloning of Recombinant Proteins.

EKLFTAD (residues 1-90), EKLFTAD1 (residues 1-40), EKLFTAD2 (residues 51-90) were cloned into the pGEX-2T vector generating GST-fusion proteins starting with the human EKLF cDNA. The KIX domain (residues 586-672) of human CBP was expressed as a his-tagged

fusion protein using the pET15b vector (provided by Dr. Alanna Schepartz, Yale University, New Haven, CT). The IbiD domain (residues 2065-2115) of CBP was cloned in the pGEX-2T vector generating GST-fusion protein starting with the human CBP cDNA. The TAZ1 domain (residues 345-439) of CBP was cloned in the pET11b vector starting with the cDNA of the mutated fragment from mouse CBP (provided by Dr. Timothy Osborne, University of California, Irvine, CA). The TAZ2 domain (residues 1723-1812)/C1738A, C1746A, C1789A, C1790A) of p300 was cloned into the pET11d vector starting with the cDNA of the mutated fragment from human p300 (provided by Dr. Ettore Appella, National Institutes of Health, Bethesda, MD). The GST-Tfb1PH (residues 1-115), and GST-p62PH (residues 1-108) were cloned as previously described [45]. QuickChange II site-directed mutagenesis kit (Startagene) was used to carry out site-directed mutagenesis of EKLFTAD2, Tfb1PH and p62PH. The GST-TFII α CTD was cloned as previously described [46]. The pEG202 vector was used to generate the LexA-fusion proteins; plasmids expressing the different fusion proteins were constructed by inserting the EcoRI-BamHI digested PCR products, which were generated by using the GST plasmids described previously for TFII α [46]. All the coding sequences were verified by DNA sequencing.

Protein Expression and Purification.

TFII α CTD, Tfb1PH, p62PH and related mutants were purified as previously described [45, 46]. EKLFTAD, EKLFTAD1, EKLFTAD2 and the CBP/p300 IbiD domain were expressed as GST-fusion proteins in *Escherichia Coli* host strain TOPP2 (EKLFTAD1, EKLFTAD2, IbiD domain) or *Escherichia Coli* host strain Rosetta2 (EKLFTAD) and bound to GSH resin, as described [45]. The resin bound proteins were incubated 2h with 100 units of thrombin (Calbiochem) to remove the GST-tag. Following the cleavage reaction, the supernatants containing the proteins were dialyzed overnight into 5% aqueous acetic acid. The proteins were further purified by reverse-phase HPLC over a C4-column (Vydac), lyophilized and stored at -80 °C. The his-tagged KIX domain of CBP was expressed in *Escherichia Coli* host strain BL21(DE3). The cells were grown at 37 °C, and protein expression was induced for 4 h with 0.7

mM IPTG at 37 °C. The cells were harvested by centrifugation, resuspended in buffer A [20 mM Phosphate Buffer pH 7.4, 50 mM imidazole, 0.5 M NaCl], lysed by passage through a French press, and centrifuged at 100,000g for 45 min. The supernatant was incubated for 1 h at 4 °C with chelating sepharose FF resin (GE Healthcare) charged with nickel. Following incubation, the resin was washed with buffer A, and the CBP-KIX protein was eluted with buffer B [20 mM Phosphate Buffer -HCl (pH 7.4), 500 mM imidazol, 0.5 M NaCl]. The eluted protein was dialyzed overnight into 10% aqueous acetic acid, further purified by reverse-phase HPLC over a C4-column (Vydac), lyophilized and stored at -80 °C.

The TAZ1 and TAZ2 domains of CBP were expressed in *Escherichia Coli* host strain BL21(DE3). The cells were grown at 37 °C, and protein expression was induced for 4 h with 0.7 mM IPTG at 37 °C. Cells were harvested, resuspended in buffer A [26 mM Tris Buffer pH 7.4, 1mM DTT], lysed by passage through a French press, and centrifuged at 15,000g for 20 min. The pellets from the centrifugation were then resuspended in buffer A with 6M guanidinium HCl, heated for 15 min at 50°C and centrifuged at 100,000g for 30 min. The supernatants from this centrifugation were then dialyzed overnight into 5% aqueous acetic acid containing 2mM DTT. The proteins were further purified by reverse-phase HPLC over a C4-column (Vydac), lyophilized and stored at -80 °C. For NMR studies, uniformly (>98%) ¹⁵N-labeled and ¹⁵N/¹³C-labeled proteins were prepared in minimal media containing ¹⁵NH₄Cl and/or ¹³C₆-glucose as the only nitrogen and carbon sources.

Isothermal Titration Calorimetry binding experiments.

ITC titrations were performed at 25 °C as previously described [40]. Proteins were dialyzed into 20 mM Tris buffer pH 7.5 except for KIX, TAZ1 and TAZ2, which were dialyzed into 20 mM sodium phosphate buffer pH 6.4. In the case of the TAZ1 and TAZ2 domains, the proteins were first refolded with 2.2 equivalent of ZnCl₂ using a previously published procedure [47]. The protein concentrations were determined from A₂₈₀. In all experiments, the CBP/p300 domains, Tfb1PH and p62PH proteins were in the stirred cell and increasing concentrations of

the EKLFTAD proteins were added incrementally. All injections fit a single-binding site mechanism with 1:1 stoichiometry.

NMR samples.

For structural studies of the Tfb1PH/EKLFTAD2 complex, the samples contained 1mM ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labeled Tfb1PH and EKLFTAD2 was added to a final ratio of 1:2. These studies were performed in 20 mM sodium phosphate buffer pH 6.5, 1mM EDTA, 1mM DTT (NMR buffer) with either 90% $\text{H}_2\text{O}/10\%\text{D}_2\text{O}$ or 99.9% D_2O . For the labeled EKLFTAD experiments, the samples contained 0.6 mM ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labeled EKLFTAD2, and Tfb1PH was added to a final ratio of 1:2 in NMR buffer.

For the NMR chemical shift titration studies of EKLFTADs with Tfb1PH, either unlabeled EKLFTAD1 or unlabeled EKLFTAD2 was added to a sample containing 0.5 mM of ^{15}N -Tfb1PH in NMR buffer to a final ratio of 2:1. For the NMR chemical shift titration studies of EKLFTAD2 on p62PH, unlabeled EKLFTAD2 was added to a final ratio of 1:1.5 to a sample containing 0.4 mM of ^{15}N -p62PH in NMR buffer.

For the NMR competition experiment, a sample containing 0.5 mM ^{15}N -labeled p53TAD2 (residues 40-73) in NMR buffer was used. To this sample, unlabeled Tfb1PH was added to a final concentration of 0.4 mM. In a second addition, unlabeled EKLFTAD2 was added to a final concentration of 0.5 mM.

NMR spectroscopy.

NMR spectra were collected at 300 K on Varian Unity Inova 500, 600 and 800 MHz spectrometers. The sequential assignment of the backbone and aliphatic side chain resonances (^1H , ^{15}N and ^{13}C) was performed with triple resonance experiments as reported for free Tfb1PH/p53TAD2 complex. Interproton distance restraints were derived from 3D ^{15}N -edited NOESY-HSQC and ^{13}C -edited HMQC-NOESY spectra ($\tau_m=90\text{ms}$). Intermolecular distance restraints were obtained from 3D $^{15}\text{N}/^{13}\text{C}$ {F1}-filtered, {F3}-edited NOESY experiments

($\tau_m=90$ ms). The NMR data were processed with NMRPipe/NMRDraw [48] and analyzed with NMRViewJ [49].

Structures calculation.

The NOE-derived distance restraints were divided into four classes defined as strong (1.8-2.8 Å), medium (1.8-3.4 Å), weak (1.8-5.0 Å) and very weak (3.3-6.0 Å). Backbone dihedral angles were derived with the program TALOS[50]. Slowly exchanged amide protons were identified by recording a 2D ^1H - ^{15}N HSQC spectrum of the Tfb1PH/EKLFTAD2 complex exchanged in D_2O . Hydrogen bonds were used as distance restraints only for the slowly exchanging amide protons in Tfb1PH. The structures of EKLFTAD2 and Tfb1PH were calculated using the program CNS [51], with a combination of torsion angle and Cartesian dynamics [52] and starting from the two extended structures (one for Tfb1PH and one for EKLFTAD2) with standard geometry. The quality of structures was assessed using PROCHEK-NMR [53] and MOLMOL [54]. The figures were generated with the program PyMol (<http://www.pymol.org>).

Transactivation assays in yeast.

All yeast strains were grown in synthetic complete media (SC; 0,67% yeast nitrogen base w/o amino acids, 2% glucose and amino acids drop-out mix) lacking the selection markers uracil and histidine. The EGY48 (Mat alpha leu2-3 his3-11,15 trp1-1 ura3-1 6lexAops-LEU2) strain was transformed with the LexA operator-Lac-Z fusion plasmid pSH18-34 combined with either pEG202NLS (pEG202 derivative with SV40 nuclear localization sequence) as a negative control, pSH17-4 (Activation domain of GAL4 cloned into pEG202 backbone) as a positive control or pEG202NLS with LexA fused to the activation domains to be tested. All yeast transformations were performed using the modified lithium acetate protocol [55]. Liquid β -galactosidase assays were performed as previously described [56]. Three different colonies of each strain were tested in three independent experiments and relative activities were determined based on the definition that 1 unit of β -galactosidase hydrolyses 1 μmol of CPRG to D-galactose and chlorophenol red per minute per cell [57]. The results are presented as the mean of the percentage of the β -

galactosidase units of the tested LexA-fusion proteins on the β -galactosidase units of the LexA-GAL4 positive control \pm standard of the mean [58]. Western-blot analyses were performed with an antibody directed at the LexA protein (Santa Cruz, sc-7544; 1:1000 dilution) to verify equivalent expression of the LexA-fusion proteins.

Transactivation assays in K562 Blood Cells.

Luciferase assays in K562 blood cells were performed as described previously [22]. The minimal EKLFTAD (residues 2-106 of mouse EKLF) was fused to the DNA-binding domain of GAL4 [22] to create the GAL4-fusion proteins. The W73 are based on the numbering of the equivalent tryptophan residue in humans. Western blot analysis with an antibody directed at the GAL4 protein to verify expression levels of the GAL-4 fusion protein.

Results

The Minimal EKLFTAD Resembles the TAD of p53.

Based on mutational analysis and sequence comparison with the TAD of p53 (Figure 4. 1A) the minimum EKLFTAD (residues 1-90 of human EKLF) can be divided in two subdomains EKLFTAD1 (residues 1-40 of EKLF) and EKLFTAD2 (residues 51-90 of EKLF) [22]. To further test the roles of these two subdomains in regulating EKLF function, we tested whether or not they are each capable of activating transcription when artificially tethered to a heterologous DNA-binding domain as previously observed with p53TAD1 (residues 1-40 of p53) and p53TAD2 (residues 40-73 of p53). To test this, EKLFTAD1, EKLFTAD2, p53TAD1 and p53TAD2 were fused to the DNA-binding domain of LexA (LexA-DBD) in order to compare their relative abilities to activate transcription in yeast. Similar to other studies in yeast [26], p53TAD1 and p53TAD2 are both able to stimulate transcription of a *lacZ* reporter gene when

fused to the LexA-DBD and the two subdomains of p53TAD activate transcription at levels just slightly lower (about 81% and 71% respectively) than the LexA-GAL4TAD positive control (Supplementary Figure 4. 1). Likewise, EKLFTAD1 and EKLFTAD2 activate transcription when fused to the LexA-DBD at levels 74% and 60% relative to the LexA-GAL4 control (Supplementary Figure 4. 1). These experiments demonstrate that EKLFTAD1 and EKLFTAD2 are both able to independently activate transcription in the yeast system at levels similar to p53TAD1 and p53TAD2 when tethered to the LexA-DBD.

An EKLFTAD2 Like Sequence in Other KLF Proteins

The KLF family can be divided into several groups based on either the presence of specific domains in their amino-terminal region or based on phylogenetic analysis [1, 2]. In an attempt to identify other KLF members with acidic TADs related to either EKLFTAD1 or EKLFTAD2, we analyzed the amino acid sequences of all known human and mouse KLF proteins. Our searches failed to identify any other KLF proteins that contain a region with significant homology to the region of EKLFTAD1 that aligns with the region of p53TAD1 that forms a helix when bound to MDM2. However, we did identify four other KLF proteins (KLF2, KLF4, KLF5 and KLF15) that contain a region with significant homology to the region of EKLFTAD2 that resembles the region of p53TAD2 that forms a helix when bound to Tfb1PH (Figure 4. 1B). This result is consistent with the fact that EKLF, KLF2, KLF4, KLF5 and KLF15 are members of group 3 and group 4 of the human KLF proteins based on the phylogenetic analysis [1]. In fact, these five KLF proteins constitute all known members of group 3 and group 4 with the exception of KLF14. This indicates that the EKLFTAD2 motif is conserved in other KLF proteins and suggests that this domain in KLF2, KLF4, KLF5, and KLF15 may bind to many of the same transcriptional regulatory factors as EKLFTAD2.

EKLFTAD1 and EKLFTAD2 Bind Tfb1PH/p62PH.

Several proteins containing acidic TADs including p53 and VP16 have been shown to directly interact with the Tfb1/p62 (yeast/human) subunit of TFIID and this interaction appears to

correlate with their ability to stimulate both the initiation and elongation phase of transcription [59]. In addition, recent studies have demonstrated that TFIID is one of the rate-limiting factors during the elongation phase of β -globin gene expression [60-62] and this is consistent with the fact that patients with specific mutations in TFIID suffer from β -thalassemia due to decreased β -globin production [63]. Together, these studies have demonstrated that TFIID plays a crucial role in β -globin gene expression and thus erythroid development. Recently, we have shown that the TADs of p53 and VP16 specifically interact with the amino-terminal PH domain of Tfb1/p62 (Tfb1PH/p62PH) [40, 44]. In the case of the TAD of p53, the interaction is specific for p53TAD2 [40] whereas for the TAD of VP16, both subdomains (VP16C and VP16N) have been shown to interact with Tfb1PH/p62PH [44]. Given the sequence homology between the TADs of p53 and EKL (Figure 4. 1A), we examined whether or not EKLFTAD1 and EKLFTAD2 are able to interact with Tfb1PH/p62PH. In order to quantitatively determine the dissociation constants [64], isothermal titration calorimetry (ITC) studies were performed using highly purified EKLFTAD1, EKLFTAD2, Tfb1PH and p62PH. The ITC experiments demonstrate that both p62PH and Tfb1PH are able to bind EKLFTAD2 with similar K_{ds} ($1.1 \pm 0.1 \mu\text{M}$ and $1.0 \pm 0.2 \mu\text{M}$ respectively) (Table 4. 1). A weaker affinity is observed for Tfb1PH with EKLFTAD1 ($11 \pm 1 \mu\text{M}$) by ITC. The K_{ds} for EKLFTAD2 are comparable to those previously determined for the interaction of p62PH and Tfb1PH with p53TAD2 ($3.2 \pm 0.6 \mu\text{M}$ and $0.39 \pm 0.07 \mu\text{M}$ respectively) [40] as well as for the Tfb1PH/VP16C interaction ($0.36 \pm 0.04 \mu\text{M}$) and Tfb1PH/VP16N interaction ($1.0 \pm 0.1 \mu\text{M}$) [44].

EKLFTAD2 and p53TAD2 Share a Common Binding Site

Given the sequence similarities between EKLFTAD2 and p53TAD2, we attempted to define the binding site for EKLFTAD2 on both Tfb1PH and p62PH using NMR chemical shift titration studies. Additions of unlabeled-EKLFTAD2 to either ^{15}N -Tfb1PH or ^{15}N -p62PH resulted in significant changes in ^1H and ^{15}N chemical shifts for several signals in both ^1H - ^{15}N HSQC spectra (Supplementary Figure 4. 2A). When mapped onto the structure of either Tfb1PH (Figure 4. 2A) or p62PH (Figure 4. 2B), the residues exhibiting significant chemical shift

changes upon addition of EKLFTAD2 are located in strands $\beta 5$, $\beta 6$, $\beta 7$ and the H1 helix. The chemical shift perturbations induced by EKLFTAD2 in both Tfb1PH and p62PH are very similar to the chemical shift changes we previously observed with p53TAD2 binding to Tfb1PH and p62PH where significant changes also occurred in strands $\beta 5$, $\beta 6$, $\beta 7$ [40]. These results suggest that EKLFTAD2 and p53TAD2 share a common binding site on Tfb1PH/p62PH.

EKLFTAD2 and p53TAD2 Compete for Binding to Tfb1PH

Given the similarity of the p53TAD2 and EKLFTAD2 binding sites on Tfb1PH/p62PH, we next used an NMR displacement experiment to test whether or not p53TAD2 and EKLFTAD2 compete for binding to Tfb1PH/p62PH. To a sample containing ^{15}N -p53TAD2 (0.5 mM), we added a substoichiometric concentration of unlabeled Tfb1PH (0.4 mM) and as expected significant changes in ^1H and ^{15}N chemical shifts in the ^1H - ^{15}N HSQC spectra of p53TAD2 are observed (Figure 4. 2C). We subsequently added an equimolar amount of unlabeled-EKLFTAD2 (0.5mM) to the ^{15}N -p53TAD2/Tfb1PH complex and we observed that the ^1H and ^{15}N resonances of p53TAD2 which had shifted upon formation of the p53TAD2/Tfb1PH complex moved back towards the values characteristic of the free p53TAD2 (Figure 4. 2D). These results demonstrate that p53TAD2 and EKLFTAD2 compete for a common binding site on Tfb1PH and their binding to Tfb1PH is mutually exclusive.

EKLFTAD2 and p53TAD2 Target Similar Domains of CBP/p300.

Since we had shown that EKLFTAD2 and p53TAD2 compete for a common binding site on Tfb1PH/p62PH, we were interested to test whether or not EKLFTAD2 shares other common binding targets with p53TAD2. The p53TAD has been shown to directly interact with the highly homologous histone acetyltransferases CBP and p300 (CBP/p300) and acetylation of p53 by CBP/p300 is essential for p53-dependent activation [65, 66]. The interaction with CBP/p300 occurs with both subdomains of the TAD of p53 and p53TAD2 has been shown to interact with at least four distinct domains of CBP/p300 [37, 67-69]. The four domains include the TAZ1/CH1 domain, the KIX domain, the TAZ2/CH3 domain and the IBID domain.

Given the fact that EKLF has been shown to interact with CBP/p300 and acetylation of EKLF by CBP is crucial for activation of β -globin gene expression in red blood cells [16-18], we tested if EKLFTAD2 was also capable of binding to the same four domains of CBP that have been previously shown to interact with p53TAD2 [67, 69]. To determine the dissociation constant of these interactions, ITC studies were performed with EKLFTAD2 and the TAZ1, KIX, IBiD and TAZ2 domains of CBP/p300 (Table 4. 1). As is the case for p53TAD2 [67, 69], all four domains of CBP bound to EKLFTAD2 and they exhibited K_d 's ranging between $0.74 \pm 0.08 \mu\text{M}$ (TAZ2) and $4.2 \pm 2.0 \mu\text{M}$ (IBiD).

In order to further analyze the interactions of EKLFTAD2 with CBP/p300, we performed NMR titration experiments with the KIX and IBiD domain of CBP/p300 (Figure 4. 3). To a sample containing ^{15}N -labeled EKLFTAD2 (0.5 mM), we added either unlabeled-KIX domain or unlabeled-IBiD domain. Significant changes in ^1H and ^{15}N chemical shifts in the ^1H - ^{15}N HSQC spectra are observed upon addition of either the KIX (Figure 4. 3A) or the IBiD domain (Figure 4. 3B). These results indicate that similar to p53TAD2 [67, 69], EKLFTAD2 is capable of interacting with multiple domains of CBP/p300, with micro-molar affinity.

Structure Determination of the Tfb1PH/EKLFTAD2 Complex

Since EKLFTAD2 competes for a common binding site on Tfb1PH/p62PH with p53TAD2, we pursued NMR structural studies of a complex containing Tfb1PH and EKLFTAD2. Tfb1PH [56] was chosen as a model system for structural NMR studies, because it is more stable than p62PH (human) in solution and Tfb1PH and p62PH are structurally and functionally homologous. In addition, we have previously determined the structures of a p53TAD2/Tfb1PH complex and VP16C/Tfb1PH complex, and the structures of these two complexes will be very valuable for comparison.

The three-dimensional structures of the Tfb1PH/EKLFTAD2 complex were calculated using 1686 NOE-derived distance restraints, 36 hydrogen-bond restraints, and 116 dihedral angle restraints. A total of 127 structures were calculated, of which 100 satisfied all the experimental

constraints with no NOE violation greater than 0.2 Å and no backbone dihedral angle violation greater than 2° (Table 4. 2). The structure of the Tfb1PH/EKLFTAD2 complex is well defined by the NMR data. The 20 lowest-energy structures (Figure 4. 4A) are characterized by good backbone geometry, no significant restraint violation and low pair wise rmsd values (Table 4. 2).

Structure of Tfb1PH and EKLFTAD2 in Complex.

The structure of Tfb1PH in complex with EKLFTAD2 is comparable to that of free Tfb1PH [45] indicating that the structure of Tfb1PH does not undergo significant structural changes between the free and complex form (Figure 4. 4A-B). Tfb1PH in complex with EKLFTAD2 consists of a typical PH domain fold with two perpendicular antiparallel-β sheets arranged in a β sandwich (β1-β7) and flanked on one side by a long α helix (H1). Significant changes in the ¹H and ¹⁵N chemical shifts of ¹⁵N-EKLFTAD2 in the ¹H-¹⁵N HSQC spectra upon addition of Tfb1PH (Supplementary Figure 4. 2B), however, EKLFTAD2 binds Tfb1PH in a predominantly extended conformation devoid of any regular secondary structural elements (Figure 4. 4A-B). In complex with EKLFTAD2, β strands β5-β7 and the H1 α-helix of Tfb1PH form an extensive interface with residues between Ala⁶⁹ and Thr⁸⁰ of EKLFTAD2 (Figure 4. 4C-D).

Tfb1PH/EKLFTAD2 Interface.

EKLFTAD2 binds in a predominately extended conformation to a contiguous shallow groove formed by β-strands β5, β6 and β7 and the H1 helix of Tfb1PH. The shallow groove of Tfb1PH is surrounded by positively charged residues (Lys⁴⁷, Lys⁵⁷, Arg⁶¹, Arg⁸⁶, Lys⁹⁷, Lys¹⁰¹ and Lys¹¹²) that position the negatively charged EKLFTAD2 (Figure 4. 5A-D) so that it can participate in a series of van der Waals contacts with hydrophobic pockets dispersed along the entire length of the groove (Figure 4. 4C). The interface of the complex begins with Ala⁶⁹ of EKLFTAD2 at one end of the groove and terminates with Thr⁸⁰ of EKLFTAD2 at the other end of the groove.

Numerous contacts are formed between Tfb1PH and Ala⁶⁹, Ala⁷¹, Thr⁷², Trp⁷³, Leu⁷⁵, Leu⁷⁷, Leu⁷⁹ and Thr⁸⁰ of EKLFTAD2 (Figure 4. 4C-D). The key residue of EKLFTAD2 along the interface with Tfb1PH appears to be Trp⁷³, which inserts into a pocket formed by Gln49, Ala50, Thr51, Met⁵⁹, Leu⁶⁰, Arg⁶¹ and Met⁸⁸ of Tfb1PH. This pocket is centrally located within the groove and in this pocket Trp⁷³ forms a cation- π with Arg⁶¹ and an amino- π interaction with Gln⁴⁹ (Figure 4. 4D). Other important van der Waals contacts are generated by Ala⁷¹ of EKLFTAD2 inserting into a narrow cleft formed by residues in strand β 5 (Thr⁵¹), strand β 6 (Met⁵⁹) and the loop between strands β 5 and β 6 (Pro52 and Lys57) of Tfb1PH at one end of the groove and Leu⁷⁹ of EKLFTAD2 inserting the shallow pocket formed by residues in strand β 5 (Leu⁴⁸) and in helix H1 (Gln¹⁰⁵, Ile¹⁰⁸ and Ser¹⁰⁹) of Tfb1PH at the other end of the groove.

In addition to the extensive network of interactions involving hydrophobic residues, there are two possible salt bridges between acidic residues of EKLFTAD2 (Figure 4. 5C-D) and basic amino acids of Tfb1PH (Figure 4. 5A-B). One is formed between Asp76 of EKLFTAD2 and Lys101 in the H1 helix of Tfb1PH ($d_{O\delta 2-N\zeta} = 4.3 \text{ \AA}$ in the minimized average structure), and the second one is formed between Asp⁷⁰ of EKLFTAD2 and Lys⁵⁷ of Tfb1PH.

Comparison of p53TAD2, VP16C and EKLFTAD2 in complex with Tfb1PH

In previous studies, we demonstrated that binding to Tfb1PH induced the formation of a 9-residue α -helix in both p53TAD2 and VP16C [40, 44]. In sharp contrast, EKLFTAD2 binds Tfb1PH in a predominantly extended conformation devoid of any classical secondary structural elements. This is despite the fact that EKLFTAD2, p53TAD2 and VP16C have very similar amino acid compositions (acidic and hydrophobic residues). In both the p53TAD2/Tfb1PH and the VP16C/Tfb1PH complexes [40, 44], three hydrophobic residues on one face of the α -helix make crucial contacts with Tfb1PH. Interestingly, in both cases one of the three hydrophobic residues (Phe⁵⁴ of p53TAD2 and Phe⁴⁷⁹ of VP16C) is inserted into the same pocket on Tfb1PH as Trp73 of EKLFTAD2 (Figure 4. 6A, D and E). In this pocket, the aromatic ring of the phenylalanine residues of p53TAD2 and VP16C participates in virtually identical interactions with Tfb1PH as observed for Trp⁷³ in EKLFTAD2. Surprisingly, this is the only distinct

similarity between the EKLFTAD2/Tfb1PH complex and either the p53TAD2/Tfb1PH or VP16C/Tfb1PH complex along the binding interface. By binding to Tfb1PH in an extended conformation, EKLFTAD2 contributes a significantly larger interface ($\approx 700 \text{ \AA}^2$) than either p53TAD2 ($\approx 380 \text{ \AA}^2$) or VP16C ($\approx 500 \text{ \AA}^2$) in complex with Tfb1PH.

In all three complexes, the positively charged residues on Tfb1PH appear to serve two functions. First, they position the negatively charged TADs so that they can participate in van der Waals contacts with hydrophobic pockets on the surface of Tfb1PH. Second, they participate in ionic interactions with acidic residues in the TADs. As we have previously observed for both the p53TAD2/Tfb1PH and VP16C/Tfb1PH complexes [40, 44], EKLFTAD2 can form two possible ionic interactions with Tfb1PH based on our NMR structures. The Lys⁵⁷/Asp⁷⁰ (Tfb1PH/EKLFTAD2) ion pair is similar to the interaction involving Lys⁵⁷ in both the p53TAD2/Tfb1PH and VP16C/Tfb1PH complexes. In contrast, the Lys¹⁰¹/Asp⁷⁶ ion pair appears to be distinct to the EKLFTAD2/Tfb1PH complex and results from the fact that EKLFTAD2 participates in interactions with the H1 helix of Tfb1PH, and such interactions are not observed with either the p53TAD2 or VP16C complex where the binding interface is much smaller.

Mutational Analysis of Tfb1PH Binding Interface

In previous studies, we identified five mutants of Tfb1PH that significantly disrupted binding to the p53TAD2 domain [40]. The five mutants [Tfb1PH (Q49A), Tfb1PH (K57E), Tfb1PH (M59A), Tfb1PH (R61E), and Tfb1PH (M88A)] are all located on the surface of Tfb1PH and these mutations do not alter the structure of the PH domain fold [40]. We were unable to detect binding to p53TAD2 with four of the mutants (Q49A, K57E, M59A, R61E) by ITC (K_d values $\geq 10 \text{ mM}$) and the M88A mutation decreased binding by over 10-fold ($K_d = 4.8 \text{ }\mu\text{M}$) [40]. To assess the relative contributions of these same residues on the surface of Tfb1PH to the formation of the Tfb1PH/EKLFTAD2 complex, we determined the K_d values between each of the five Tfb1PH mutants and EKLFTAD2 by ITC (Table 4. 3). As was the case with p53TAD2 [40], there is no binding detected with either the Tfb1PH (K57E) or the Tfb1PH (R61E) mutants (K_d values $\geq 100 \text{ }\mu\text{M}$). These results demonstrate that Arg61 and Lys57 of Tfb1PH also make

crucial contributions to the binding to EKLFTAD2. In contrast, the Tfb1PH (Q49A), Tfb1PH (M59A) and the Tfb1PH (M88A) mutants bound to EKLFTAD2 with K_d s very similar to the wild-type Tfb1PH. These results are further evidence that the Tfb1PH/ EKLFTAD2 and Tfb1PH/p53TAD2 structures are distinct and indicate that the interactions stabilizing the two complexes are quite different despite the fact that these two acidic TADs compete for a common binding site on Tfb1PH/p62PH.

EKLFTAD2/Tfb1PH Complex Resembles the TFII α CTD/p62PH Complex

TFIIH is recruited to the pre-initiation complex (PIC) through an interaction between Tfb1PH/p62PH and the acidic carboxyl-terminal domain of the alpha subunit of the general transcription factor TFII α (TFII α CTD) [70]. The structure of TFII α CTD (residues 336-439 in human TFII α) can be divided into two regions; the amino-terminal region (residues 336-394) is highly acidic and devoid of any regular secondary structural elements, whereas the carboxyl-terminal region (residues 395-439) folds into a novel $\beta\beta\alpha\alpha$ topology [46]. Recently, a structure of TFII α CTD in complex with p62PH was determined [71]. The structure demonstrated that residues from both regions of TFII α CTD contribute to the binding interface with p62PH. In addition, ITC studies demonstrated that the acidic amino-terminal region predominantly contributes to the binding energy. In complex with p62PH, the acidic amino-terminal region of TFII α CTD makes multiple hydrophobic contacts along the groove formed by strands $\beta 5$, $\beta 6$, and $\beta 7$ and the H1 helix of p62PH. TFII α CTD binds p62PH in an extended conformation and the structure of the TFII α CTD/p62PH complex is very similar to the structure of the EKLFTAD2/Tfb1PH complex (Figure 4. 6B-C) [71]. The key interactions involve a phenylalanine residue of TFII α CTD (Phe³⁸⁷). Like Trp⁷³ in EKLFTAD2, Phe³⁸⁷ from TFII α CTD forms amino-aromatic interactions and cation- π interaction with the homologous residues in p62PH. In addition, TFII α CTD forms ionic interactions with two lysine residues of p62PH (Lys⁶² and Lys⁹⁷) that are in equivalent positions to the two lysine residues in Tfb1PH (Lys⁵⁷ and Lys¹⁰¹) that form ionic interactions with EKLFTAD2. Overall, the EKLFTAD2/Tfb1PH complex is similar to the TFII α CTD/p62PH complex and suggests that

the acidic TAD of EKLF can mimic the interactions of the general transcription factor TFIIE with p62PH/Tfb1PH.

Trp73 is Crucial for EKLFTAD Binding and Activation in Red Blood Cells

We have previously shown that upon binding to Tfb1PH, both p53TAD2 and VP16C form 9-residue α -helices that constitute the binding interfaces [40, 44]. The structures of the p53TAD2/Tfb1PH and VP16/Tfb1PH complex demonstrating the importance of folding on binding are consistent with the fact that proline mutants within VP16C led to a decreased affinity for Tfb1PH and a decrease in activation potential [32, 34, 44]. In EKLFTAD2, Leu⁷⁵ is one of the hydrophobic residues located along the binding interface with Tfb1PH, but the equivalent position in the mouse EKLFTAD2 is a proline residue (Figure 4. 1A). To test the effect of a proline substitution at Leu⁷⁵ in human EKLFTAD2, we prepared a L75P mutant and determined its affinity for Tfb1PH/p62PH by ITC. The ITC experiments demonstrate that the EKLFTAD2 (L75P) mutant binds to p62PH with an apparent K_d of $0.8 \pm 0.2 \mu\text{M}$ (Figure 4. 7A-B) and this value is virtually identical to that observed with the wild-type EKLFTAD2. This result is consistent with the fact that EKLFTAD2 binds Tfb1PH/p62PH as an extended structure as opposed to forming an α -helix as seen with p53TAD2 and VP16C and this interaction is conserved in the mouse.

Analysis of the structure of the EKLFTAD2/Tfb1PH complex indicates that Trp⁷³ of EKLFTAD2 is a key component of the binding interface. To test the role of Trp⁷³ on the activity of the minimal EKLFTAD, we prepared both an alanine and a proline mutant of Trp⁷³ and tested their affect using an in vivo activation assay. The assay is performed in K562 erythroid cells transfected with the minimal EKLFTAD (residues 2-106 of mouse EKLF) fused to the DNA-binding domain of GAL4 [22]. In this system, both the EKLFTAD (W73P) mutant and the EKLFTAD (W73A) mutant significantly reduce the ability of EKLFTAD to activate transcription of a GAL-TK- β -globin-reporter gene relative to the wild-type EKLFTAD (Figure 4. 7C). In addition, ITC analysis demonstrates that the EKLFTAD2 (W73P) mutant is unable to bind to either p62PH or the IBiD domain of CBP/p300 (Figure 4. 7B). These results support the

conclusion that Trp⁷³ of EKLFTAD2 forms important interactions at the interface with transcriptional regulatory factors such as CBP/p300 and TFIID and that these interactions are crucial for the *in vivo* activity of EKLF.

Discussion

The structures of EKLFTAD2, p53TAD2 [40] and VP16C [44] bound to Tfb1PH allow us to make direct comparisons between the three similar acidic TADs bound to a common target. Acidic TADs are typically disordered in the unbound form [72-74], and adopt an α -helical conformation in complex with their partner proteins [40, 44, 75-80]. Based on these structural studies, it has been suggested that acidic TADs have the flexibility to mediate interactions with multiple partners due to their ability to couple folding with binding. The transition of p53TAD2 and VP16C from a disordered state to a helical conformation when bound to Tfb1PH is consistent with other structural studies with acidic TADs [38, 40, 44, 75-82]. However, we determined that EKLFTAD2 binds to Tfb1PH in an extended conformation despite sharing sequence and functional homology with other acidic TADs and this demonstrates how acidic TADs can recognize identical target proteins using different structural conformation. In addition, the extended structure observed in the EKLFTAD2/Tfb1PH complex indicates that coupling folding and binding is not a strict requirement for acidic TADs interacting with multiple partners.

The structures of p53TAD2, VP16C, and EKLFTAD2 in complex with Tfb1PH/p62PH also suggests that not only have acidic TADs evolved so that they are able to bind using different conformations, but that their targets like Tfb1PH have evolved to recognize different structural features in their partners. This adaptability of both components of transcriptional regulatory complexes is supported by the observation that transcription factors contain a high percentage of intrinsically unstructured regions in their sequences in comparison to other proteins [83]. This is also supported by previous work showing that TFIIE α CTD binds the homologous p62PH in a virtually identical conformation along the same binding interface where EKLFTAD2 binds to Tfb1PH, and that TFIIE α CTD competes with p53TAD2 and EKLFTAD2 for binding to

Tfb1PH/p62PH [46, 71]. In addition, a somewhat similar situation is observed when comparing the structures of the acidic TADs of STAT2, CITED2, and HIF-1 α bound to the TAZ1 domain of CBP [75, 80, 81]. These three acidic TADs bind to the same surface of TAZ1, but they bind using two different orientations (N to C and C to N) and by utilizing different combinations of helical segments. Despite binding through three distinct mechanisms they differ from EKLFTAD2 in that all three of these acidic TADs require the formation of an alpha helical conformation through a coupled folding and binding mechanism [75, 80, 81].

The interactions of EKLFTAD with TFIIH and CBP provide detailed information regarding the role these two transcriptional regulatory factors have in β -globin gene expression and red blood cell development. The interaction of EKLFTAD with Tfb1PH/p62PH is consistent with the fact that TFIIH is required for β -globin expression and that mutations in TFIIH lead to β -thalassemia [60-62]. The structure of the Tfb1PH/EKLFTAD2 complex demonstrated that Trp⁷³ is crucial for this interaction and is supported by our in vivo results demonstrating that mutations of Trp⁷³ leads to significantly lower levels of transcriptional activity. The results demonstrating that EKLFTAD2 binds the same four subdomains of CBP as p53TAD2 is in agreement with studies demonstrating that acetylation of EKLF by CBP is required for activation of β -globin gene expression in red blood cells [16-18]. As was the case for binding to Tfb1PH/p62PH, Trp⁷³ is important for EKLFTAD2 interaction with CBP and this is also supported by the in vivo results with the Trp⁷³ mutations. These results also suggest that p53TAD2 and EKLFTAD2 could share other common binding partners despite the fact that they may bind these targets using significantly different mechanisms.

In addition to the minimal TAD, EKLF contains a second domain between residues 140-225 that is sufficient for activation of β -globin gene expression [84, 85]. Mutational analysis demonstrated that the minimal EKLFTAD and the domain between residues 140-225 appear to serve redundant functions in β -globin gene expression [84], although the mechanistic details by which this region of EKLF functions is not clearly understood [85]. The primary sequences of these two domains are very different, with the minimal EKLFTAD being very acidic (pI=3.8)

and the region between residues 140-225 being very basic (pI=9.7) with only two acidic residues and no tryptophans. However, the redundant functions suggest that the region between residues 140-225 of EKLF and the minimal EKLFTAD recruit the same transcriptional regulatory factors. Thus, future studies are needed to determine if this region also directly binds Tfb1/p62 and CBP/p300 as does the minimal EKLFTAD.

Sequence comparisons indicate that a domain similar to EKLFTAD2 is also present in a number of other KLF family members and specifically in KLF members belonging to group 3 and group 4 based on phylogenetic analysis (Figure 4. 1B) [1]. This is supported by studies demonstrating that KLF2, KLF4, KLF5 and KLF15 all interact directly with CBP/p300 through a domain in their sequence that includes the region that is homologous to EKLFTAD2 [86] [87-89]. This strongly suggests that KLF2, KLF4, KLF5 and KLF15 will bind to a number of the same transcriptional regulatory factors as EKLF, including possibly the Tfb1/p62 subunit of TFIIH. It is interesting to note that based on our alignment (Figure 4. 1B), Trp⁷³ does not appear to be conserved in any of the other KLF factors, and in the case of KLF5 and KLF15 the conserved regions appear more homologous to the regions of p53TAD2 and VP16C than to the region of EKLFTAD2 that binds Tfb1PH. Additional structural and functional studies will be needed to determine whether these homologous regions of the other KLF factors bind to their target proteins in an extended conformation like EKLFTAD2 or if they couple folding and binding as typically seen with acidic TADs such as p53 and VP16.

Acknowledgements

This work was supported by a grant from the CIHR to JGO (MOP-74739) and by a USPHS grant to JJB (NIH DK46865). The microcalorimeter for ITC experiments was purchased with funds from the NSERC. The 800 MHz NMR experiments were recorded at the Québec/Eastern Canada High Field NMR Facility, supported by grants from the CFI, NSERC and the Québec Ministère de la Recherche en Science et Technologie.

References

1. Bieker, J.J., *Kruppel-like factors: Three fingers in many pies*. J. Biol. Chem., 2001. **276**: p. 34355-34358.
2. Kaczynski, J., T. Cook, and R. Urrutia, *Sp1- and Kruppel-like transcription factors*. Genome Biol, 2003. **4**(2): p. 206.
3. Miller, I.J. and J.J. Bieker, *A novel, erythroid cell-specific murine transcription factor that binds the CACCC element and is related to the Kruppel family of nuclear proteins*. Molecular and Cellular Biology, 1993. **13**(5): p. 2776-2786.
4. Bieker, J.J. and C.M. Southwood, *The erythroid kruppel-like factor transactivation domain is a critical component for cell-specific inducibility of a beta-globin promoter*. Molecular and Cellular Biology, 1995. **15**(2): p. 852-860.
5. Hodge, D., et al., *A global role for EKLF in definitive and primitive erythropoiesis*. Blood, 2006. **107**(8): p. 3359-70.
6. Perkins, A.C., A.H. Sharpe, and S.H. Orkin, *Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF*. Nature, 1995. **375**(6529): p. 318-22.
7. Nuez, B., et al., *Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene*. Nature, 1995. **375**(6529): p. 316-8.
8. Wijgerde, M., et al., *The role of EKLF in human beta-globin gene competition*. Genes & Development, 1996. **10**(22): p. 2894-2902.
9. Perkins, A.C., K.M.L. Gaensler, and S.H. Orkin, *Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(22): p. 12267-12271.
10. Zhou, D.W., et al., *Differential binding of erythroid Kruppel-like factor to embryonic/fetal globin gene promoters during development*. Journal of Biological Chemistry, 2006. **281**(23): p. 16052-16057.
11. Im, H., et al., *Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(47): p. 17065-17070.
12. Lee, J.S., et al., *Erythroid Kruppel-like factor is recruited to the CACCC box in the beta-globin promoter but not to the CACCC box in the gamma-globin promoter: The role of the neighboring promoter elements*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(6): p. 2468-2473.

13. Feng, W.C., C.M. Southwood, and J.J. Bieker, *Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor*. J Biol Chem, 1994. **269**(2): p. 1493-500.
14. Armstrong, J.A., J.J. Bieker, and B.M. Emerson, *A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro*. Cell, 1998. **95**(1): p. 93-104.
15. Chen, X.Y. and J.J. Bieker, *Stage-specific repression by the EKLF transcriptional activator*. Molecular and Cellular Biology, 2004. **24**(23): p. 10416-10424.
16. Zhang, W. and J.J. Bieker, *Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases*. Proc Natl Acad Sci U S A, 1998. **95**(17): p. 9855-60.
17. Zhang, W., et al., *Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex*. Mol Cell Biol, 2001. **21**(7): p. 2413-22.
18. Sengupta, T., et al., *Acetylation of EKLF is essential for epigenetic modification and transcriptional activation of the beta-globin locus*. Molecular and Cellular Biology, 2008. **28**(20): p. 6160-6170.
19. Bottardi, S., et al., *Lineage-specific activators affect beta-globin locus chromatin in multipotent hematopoietic progenitors*. Embo Journal, 2006. **25**(15): p. 3586-3595.
20. Drissen, R., et al., *The active spatial organization of the beta-globin locus requires the transcription factor EKLF*. Genes & Development, 2004. **18**(20): p. 2485-2490.
21. Sengupta, T., et al., *Distinct modes of gene regulation by a cell-specific transcriptional activator*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4213-8.
22. Chen, X. and J.J. Bieker, *Erythroid Krueppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions*, in *EMBO J*. 1996. p. 5888-5896.
23. Bieker, J.J. and X. Chen, *Modification of critical regions in the EKLF transactivation domain plays a determining role in its activity*. Blood, 1995. **86**(10): p. 976-976.
24. Ouyang, L., X. Chen, and J.J. Bieker, *Regulation of erythroid Kruppel-like factor (EKLF) transcriptional activity by phosphorylation of a protein kinase casein kinase II site within its interaction domain*. J Biol Chem, 1998. **273**(36): p. 23019-25.
25. Fields, S. and S.K. Jang, *Presence of a potent transcription activating sequence in the p53 protein*. Science, 1990. **249**(4972): p. 1046-9.
26. Candau, R., et al., *Two tandem and independent sub-activation domains in the amino terminus of p53 require the adaptor complex for activity*. Oncogene, 1997. **12**: p. 807-816.

27. Unger, T., et al., *p53: A transdominant regulator of transcription whose function is ablated by mutations occurring in human cancer*. EMBO Journal, 1992. **11**(4): p. 1383-90.
28. Das, G., C.S. Hinkley, and W. Herr, *Basal promoter elements as a selective determinant of transcriptional activator function*. Nature (London), 1995. **374**(6523): p. 657-60.
29. Cress, W.D. and S.J. Triezenberg, *Critical structural elements of the VP16 transcriptional activation domain*. Science, 1991. **251**: p. 87-90.
30. Triezenberg, S.J., R.C. Kingsbury, and S.L. McKnight, *Functional dissection of VP16, the transactivator of herpes simplex virus immediate early gene expression*. Genes & Dev., 1988. **2**: p. 718-729.
31. Goodrich, J.A., et al., *Drosophila TAF_{II}40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB*. Cell, 1993. **75**: p. 519-530.
32. Sullivan, S.M., et al., *Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus*. Nucleic Acid Res., 1998. **26**: p. 4487-4496.
33. Mitchell, P.J. and R. Tjian, *Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins*. Science, 1989. **245**: p. 371-378.
34. Regier, J.L., F. Shen, and S.J. Triezenberg, *Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator*. Proc. Natl. Acad. Sci. USA, 1993. **90**: p. 883-887.
35. Abramova, N.A., et al., *Interaction between replication protein A and p53 is disrupted after UV damage in DNA repair-dependent manner*. Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 7186-7191.
36. Chang, J., et al., *Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein*. J. Biol. Chem., 1995. **270**(42): p. 25014-19.
37. Van Orden, K., et al., *Binding of p53 to the KIX domain of CREB binding protein. A potential link to human T-cell leukemia virus, type I-associated leukemogenesis*. Journal of Biological Chemistry, 1999. **274**(37): p. 26321-26328.
38. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain*. Science, 1996. **274**: p. 948-953.
39. Appella, E. and C.W. Anderson, *Post-translational modifications and activation of p53 genotoxic stresses*. Eur. J. Biochem., 2001. **268**: p. 2764-2772.
40. Di Lello, P., et al., *Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIID and the activation domain of p53*. Mol Cell, 2006. **22**(6): p. 731-40.
41. Sakaguchi, K., et al., *DNA damage activates p53 through a phosphorylation-acetylation cascade*. Genes & Dev., 1998. **12**: p. 2831-2841.

42. Sakaguchi, K., et al., *Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by acasein 1-like kinase*. J. Biol. Chem., 2000. **275**: p. 9278-9283.
43. Siatecka, M., L. Xue, and J.J. Bieker, *Sumoylation of EKLF promotes transcriptional repression and is involved in inhibition of megakaryopoiesis*. Molecular and Cellular Biology, 2007. **27**(24): p. 8547-8560.
44. Langlois, C., et al., *NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: Structural similarities between VP16 and p53*. Journal of the American Chemical Society, 2008. **130**(32): p. 10596-10604.
45. Di Lello, P., et al., *NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIID and characterization of its phosphoinositide and VP16 binding sites*. Biochemistry, 2005. **44**(21): p. 7678-86.
46. Di Lello, P., et al., *p53 and TFIIEalpha share a common binding site on the Tfb1/p62 subunit of TFIID*. Proc Natl Acad Sci U S A, 2008. **105**(1): p. 106-11.
47. Omichinski, J.G., et al., *NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1*. Science, 1993. **261**(5120): p. 438-46.
48. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
49. Johnson, B.A. and R.A. Blevins, *NMRView: A computer program for the visualization and analysis of NMR data*. J. Biomol. NMR, 1994. **4**: p. 603-614.
50. Cornilescu, G., F. Delaglio, and A. Bax, *Protein Backbone angle restraints from searching a database for chemical shift and sequence homology*. J. Biomol. NMR, 1999. **13**: p. 289-302.
51. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination*. Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.
52. Choy, W.Y., et al., *Direct structure refinement of high molecular weight proteins against residual dipolar couplings and carbonyl chemical shift changes upon alignment: an application to maltose binding protein*. J Biomol NMR, 2001. **21**(1): p. 31-40.
53. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR*. J Biomol NMR, 1996. **8**(4): p. 477-86.
54. Koradi, R., M. Billeter, and K. Wüthrich, *MOLMOL: a program for display and analysis of macromolecular structures*. J. Mol. Graphics, 1996. **14**: p. 51-55.
55. Kaiser C., M.S., and Mitchell A. , *Laboratory Course Manual for Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, 1994.
56. Handbook, Y.P., *Clontech Laboratories Inc*. 2001: p. 28-29.

57. Miller, J.H., *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory Press, 1992.
58. Gorokhovatsky, A.Y., et al., *Homogeneous assay for biotin based on Aequorea victoria bioluminescence resonance energy transfer system*. *Analyt. Biochem.*, 2003. **313**(1): p. 68-75.
59. Blau, J., et al., *Three functional classes of transcriptional activation domain*. *Mol Cell Biol*, 1996. **16**(5): p. 2044-55.
60. Vernimmen, D., et al., *Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression*. *EMBO J*, 2007. **26**(8): p. 2041-51.
61. Glover-Cutter, K., et al., *TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II*. *Mol Cell Biol*, 2009. **29**(20): p. 5455-64.
62. Bird, G., D.A.R. Zorio, and D.L. Bentley, *RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3' end formation*. *Molecular and Cellular Biology*, 2004. **24**(20): p. 8963-8969.
63. Viprakasit, V., et al., *Mutations in the general transcription factor TFIIH result in beta-thalassaemia in individuals with trichothiodystrophy*. *Hum Mol Genet*, 2001. **10**(24): p. 2797-802.
64. Harding, M.M., et al., *Interactions of the antitumor agent molybdocene dichloride with oligonucleotides*. *Inorganic Chemistry*, 1998. **37**(10): p. 2432-2437.
65. Avantaggiati, M.L., et al., *Recruitment of p300/CBP in p53-dependent signal pathways*. *Cell*, 1997. **89**(7): p. 1175-1184.
66. Gu, W. and R.G. Roeder, *Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain*. *Cell*, 1997. **90**(4): p. 595-606.
67. Ferreon, J.C., et al., *Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2*. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. **106**(16): p. 6591-6596.
68. Livengood, J.A., et al., *p53 transcriptional activity mediated through the SRC1-interacting domain of CBP/p300*. *J. Biol. Chem.*, 2002. **277**: p. 9054-9061.
69. Teufel, D.P., et al., *Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53*. *Proc. Natl. Acad. Sci. USA*, 2007. **104**(17): p. 7009-7014.
70. Ohkuma, Y., et al., *Analysis of the role of TFIIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIIE α* . *Mol. Cell. Biol.*, 1995. **15**: p. 4856-4866.
71. Okuda, M., et al., *Structural insight into the TFIIIE-TFIIH interaction: TFIIIE and p53 share the binding region on TFIIH*. *EMBO J*, 2008. **27**(7): p. 1161-71.

72. O'Hare, P. and G. Williams, *Structural studies of the acidic transactivation domain of the Vmw65 protein of herpes simplex virus using proton NMR*. *Biochemistry*, 1992. **31**: p. 4150-4156.
73. Lee, H., et al., *Local structural elements in the mostly unstructured transcriptional activation domain of human p53*. *J. Biol. Chem.*, 2000. **275**: p. 29426-29432.
74. Donaldson, L. and J.P. Capone, *Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1*. *J. Biol. Chem.*, 1992. **267**: p. 1411-1414.
75. Wojciak, J.M., et al., *Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains*. *EMBO J*, 2009. **28**(7): p. 948-58.
76. Ferreon, J.C., et al., *Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein*. *Proc Natl Acad Sci U S A*, 2009. **106**(32): p. 13260-5.
77. Uesugi, M., et al., *Induced ζ -helix in the VP16 activation domain upon binding to a human TAF*. *Science*, 1997. **277**: p. 1310-1313.
78. Uesugi, M. and G.L. Verdine, *The ζ -helical FXXFF motif in p53: TAF interaction and discrimination by MDM2*. *Proc. Natl. Acad. Sci. USA*, 1999. **96**(26): p. 14801-14806.
79. De Guzman, R.N., et al., *Structural basis for cooperative transcription factor binding to the CBP coactivator*. *J Mol Biol*, 2006. **355**(5): p. 1005-13.
80. De Guzman, R.N., et al., *Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligands for non-identical binding sites*. *J Biol Chem*, 2004. **279**(4): p. 3042-9.
81. Dames, S.A., et al., *Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response*. *Proc Natl Acad Sci U S A*, 2002. **99**(8): p. 5271-6.
82. Bochkareva, E., et al., *Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A*. *Proc Natl Acad Sci U S A*, 2005. **102**(43): p. 15412-7.
83. Liu, J.T., J.R. Faeder, and C.J. Camacho, *Toward a quantitative theory of intrinsically disordered proteins and their function*. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. **106**(47): p. 19819-19823.
84. Pandya, K., D. Donze, and T.M. Townes, *Novel transactivation domain in erythroid Kruppel-like Factor (EKLF)*. *Journal of Biological Chemistry*, 2001. **276**(11): p. 8239-8243.
85. Brown, R.C., et al., *Distinct domains of erythroid Kruppel-like factor modulate chromatin remodeling and transactivation at the endogenous beta-globin gene promoter*. *Molecular and Cellular Biology*, 2002. **22**(1): p. 161-170.

86. SenBanerjee, S., et al., *KLF2 is a novel transcriptional regulator of endothelial proinflammatory activation*. Journal of Experimental Medicine, 2004. **199**(10): p. 1305-1315.
87. Evans, P.M., et al., *Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation*. Journal of Biological Chemistry, 2007. **282**(47): p. 33994-34002.
88. Zhang, Z.P. and C.T. Teng, *Phosphorylation of Kruppel-like factor 5 (KLF5/IKLF) at the CBP interaction region enhances its transactivation function*. Nucleic Acids Research, 2003. **31**(8): p. 2196-2208.
89. Hussain, S., E. Lara-Pezzi, and N. Brand, *Role of the KLF15 transcription factor in an anti-hypertrophic pathway involving GSK3 beta*. Heart, 2008. **94**: p. A53-A54.
90. Black, J.C., et al., *A mechanism for coordinating chromatin modification and preinitiation complex assembly*. Mol. Cell, 2006. **23**(6): p. 809-818.

Tables

Table 4. 1. Comparison of the K_d (μM) Values for the Binding of EKLFTAD2 with Tfb1PH, p62PH and domains of CBP/p300.

		K_d (μM)
Tfb1PH	EKLFTAD2	1.0 ± 0.2
p62PH	EKLFTAD2	1.1 ± 0.1
CBP KIX	EKLFTAD2	1.2 ± 0.7
CBP IBiD	EKLFTAD2	4.2 ± 2.0
CBP TAZ1	EKLFTAD2	3.0 ± 1.0
CBP TAZ2	EKLFTAD2	0.74 ± 0.08

Table 4. 2. Structural Statistics of the Tfb1PH/EKLFTAD2 Complex^a

Restrains used for the structure calculations	
Total number of NOE distances restraints	1686
Short-range (intraresidue)	675
Medium-range ($ i-j \leq 4$)	575
Long-range	389
Intermolecular	47
Hydrogen bond	36
Number of dihedral angle restraints (ϕ , ψ)	116
Structural statistics	
Rms deviations from idealized geometry	
Bonds (Å)	0.0027±0.00008
Angles (deg)	0.4363±0.0065
Impropers (deg)	0.2871±0.0098
Rms deviations from distance restraints (Å)	0.0207±0.0007
Rms deviations from dihedral restraints (deg)	0.5408±0.0359
Ramachandran statistics (%) ^b	
Residues in most favored regions	80.0
Residues in additional allowed regions	17.3
Residues in generously allowed regions	2.6
Residues in disallowed regions	0.1
Coordinate precision ^c	
Atomic pair wise rmsd (Å)	
Tfb1PH/EKLFTAD2 complex	
Backbone atoms (C', C ^α , N)	0.70±0.15
All heavy atoms	1.35±0.19
Tfb1PH alone	
Backbone atoms (C', C ^α , N)	0.65±0.15
All heavy atoms	1.32±0.19
EKLFTAD2 alone	
Backbone atoms (C', C ^α , N)	0.68±0.22
All heavy atoms	1.35±0.41

^a The 20 conformers with the lowest energies were selected for statistical analysis. Because of the absence of medium-range, long-range, and intermolecular NOEs involving residues 51-58 and 86-90 of EKLFTAD2, these amino acids were not included in the calculations.

^b Based on PROCHECK-NMR analysis.

^c Only residues 4–63 and 86–114 of Tfb1PH and residues 70-79 of EKLFTAD2 were used for the rmsd calculations. Residues at the N terminus (1–3), at the C terminus (113–115), and in the flexible loop (64-85) of Tfb1PH, as well as residues at the N terminus (59-69) and at the C terminus (80-85) of EKLFTAD2, were not included in the calculation.

Table 4. 3 :Comparison of the K_d (μM) Values for the Binding of EKLFTAD2 to Tfb1PH and Tfb1PH mutants.

		K_d (μM)
p62PH	EKLFTAD2	1.1 ± 0.2
Tfb1PH	EKLFTAD2	1.0 ± 0.1
Tfb1PH Q49A	EKLFTAD2	2.4 ± 0.4
Tfb1PH M59A	EKLFTAD2	1.0 ± 0.2
Tfb1PH R61A	EKLFTAD2	n.b. ¹
Tfb1PH K57E	EKLFTAD2	n.b. ¹
Tfb1PH M88A	EKLFTAD2	1.2 ± 0.1

¹ n.b. no binding detected $K_d \geq 100 \mu\text{M}$.

Figures

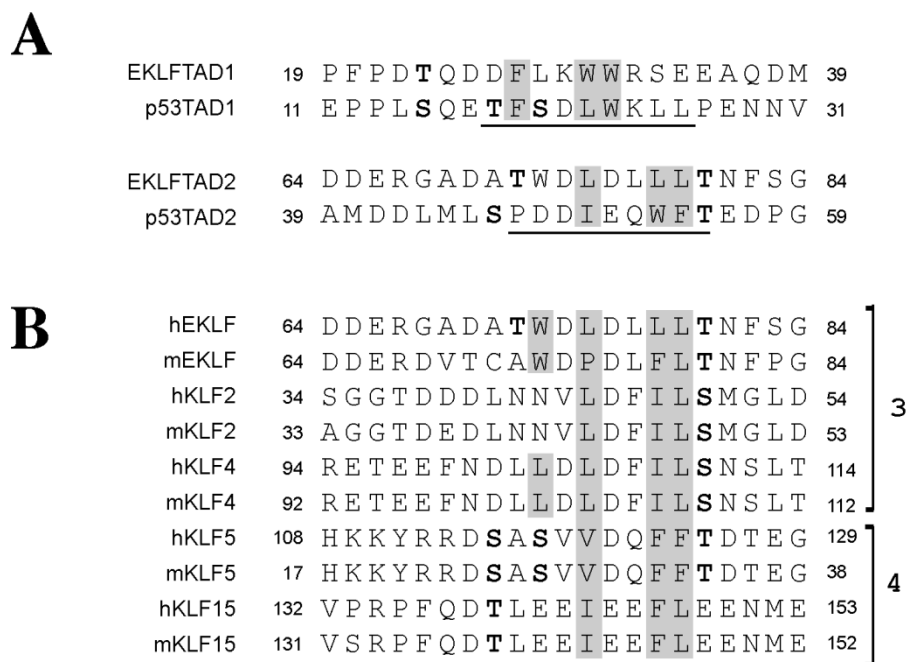


Figure 4. 1 : EKLFTAD1 and EKLFTAD2 are homologous to p53TAD1 and p53TAD2. (A) Sequence alignments of EKLFTAD1 and EKLFTAD2 from human EKLF with p53TAD1 and p53TAD2 from human p53. The residues of p53 that form helices in the p53TAD1/MDM2 complex [38] and the p53TAD2/Tfb1PH complex [40] are underlined in black. The three key hydrophobic residues at the binding interfaces of the p53TAD1/MDM2 complex and the p53TAD2/Tfb1PH are highlighted in grey. Several known or potential phosphorylation sites are highlighted in bold. (B) Sequence alignment of regions of other KLF proteins (mouse and human) that share homology with EKLFTAD2. The KLF proteins are members of either KLF group 3 or KLF group 4 based on phylogenetic analysis of all known KLF proteins [1]. Key hydrophobic residues and potential phosphorylation sites are highlighted in bold.

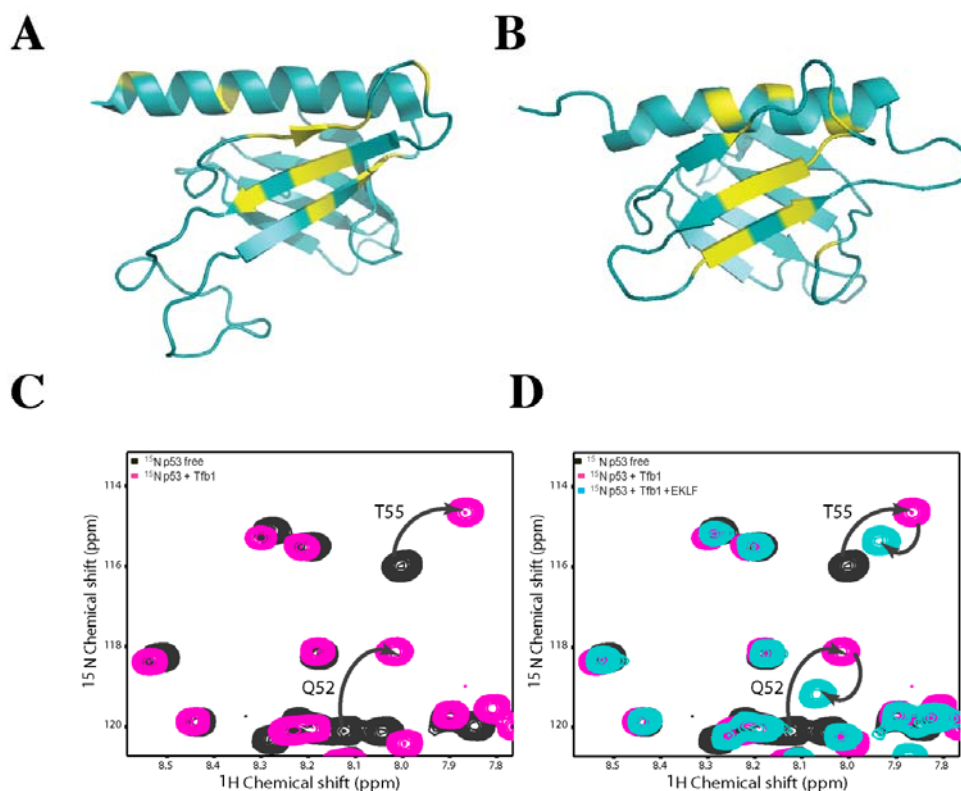


Figure 4. 2 : EKLFTAD2 and p53TAD2 share a common binding site on Tfb1PH/p62PH. Ribbon models of the 3D structures of Tfb1PH (A) and p62PH (B). The amino acids showing a significant chemical shift change ($\Delta\delta_{(\text{ppm})} > 0.15$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) upon formation of a complex with EKLFTAD2 are shown in yellow (A-B). (C) Overlay of a selected region from the two-dimensional ^1H - ^{15}N HSQC spectra for a 0.5 mM sample of ^{15}N -labeled p53TAD2 in the free form (black) and in the presence of 0.4 mM unlabeled Tfb1PH (pink). (D) Overlay of a selected region from the two-dimensional ^1H - ^{15}N HSQC spectra for a 0.5 mM sample of ^{15}N -labeled p53TAD2 in the free form (black), in the presence of 0.4 mM unlabeled Tfb1PH (pink), and after addition of 0.5 mM unlabeled EKLFTAD2 (aqua). Signals that undergo significant changes in ^1H and ^{15}N chemical

shifts upon formation of the complex with Tfb1PH (C) and that return towards their original position following the addition of EKLFTAD2 (D) are indicated by arrows.

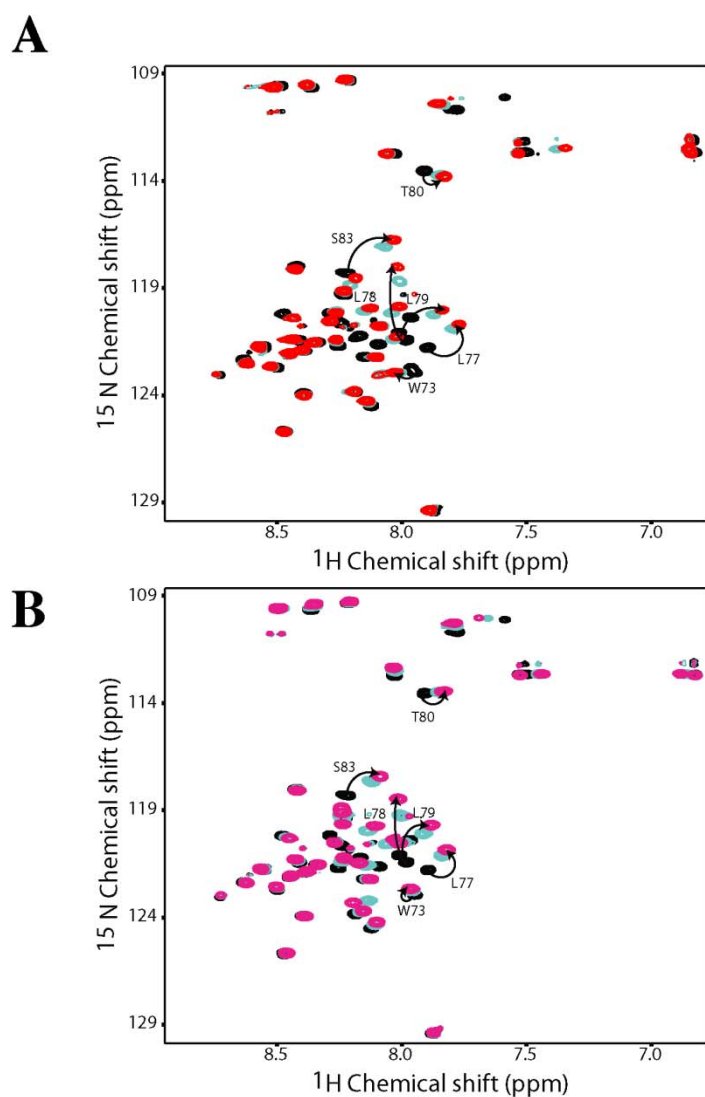


Figure 4. 3 The KIX and IBiD domains of CBP interact with EKLFTAD2. **(A)** Overlay of the two-dimensional ^1H - ^{15}N HSQC spectra of ^{15}N -labeled EKLFTAD2 in the free form (black) and in the presence of 0.8 (aqua) and 1.0 (red) equivalents of the KIX domain of CBP. **(B)** Overlay of the two-dimensional ^1H - ^{15}N HSQC spectra of the ^{15}N -labeled EKLFTAD2 in the free form (black) and in the presence of 0.8 (aqua) and 1.0 (pink) equivalents of the IBiD domain of CBP. Arrows in both

spectra highlight select signals of EKLFTAD2 that undergo significant changes in ^1H and ^{15}N chemical shifts upon formation of the complexes.

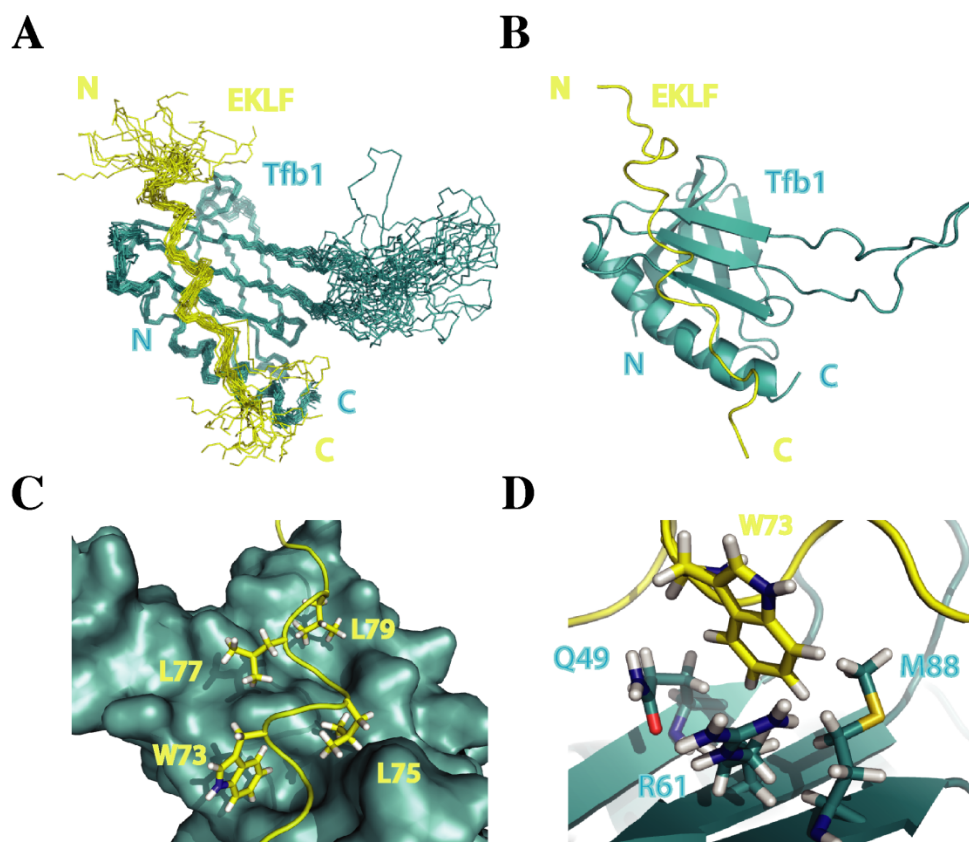


Figure 4. 4: Structure of the Tfb1PH/EKLFTAD2 Complex. (A) Overlay of the 20 lowest-energy structures of the complex between Tfb1PH (blue) and EKLFTAD2 (yellow). The structures were superimposed using the backbone atoms C', C^α, and N of residues 4–65 and 85–112 of Tfb1PH and residues 59-84 of EKLFTAD2. (B) Ribbon model of the lowest energy conformer of the complex between Tfb1PH (blue) and EKLFTAD2 (yellow). (C) 3D structure of Tfb1PH is shown as molecular surface (blue), and EKLFTAD2 is represented as a tube (yellow). The side chains of several hydrophobic residues of EKLFTAD2 (Trp73, Leu75, Leu77, and Leu79) that form the interface with Tfb1PH are shown as sticks. (D) 3D structure of Tfb1PH is shown as ribbon (blue), and EKLFTAD2 is represented as a tube (yellow). The side chain of Trp73 (EKLFTAD2) is

positioned in a pocket formed by Gln49, Arg61 and M88 of Tfb1. In this pocket, Trp73 (EKLFTAD2) forms cation- π , and amino-aromatic interactions with Arg61 and Gln49 of Tfb1PH, respectively.

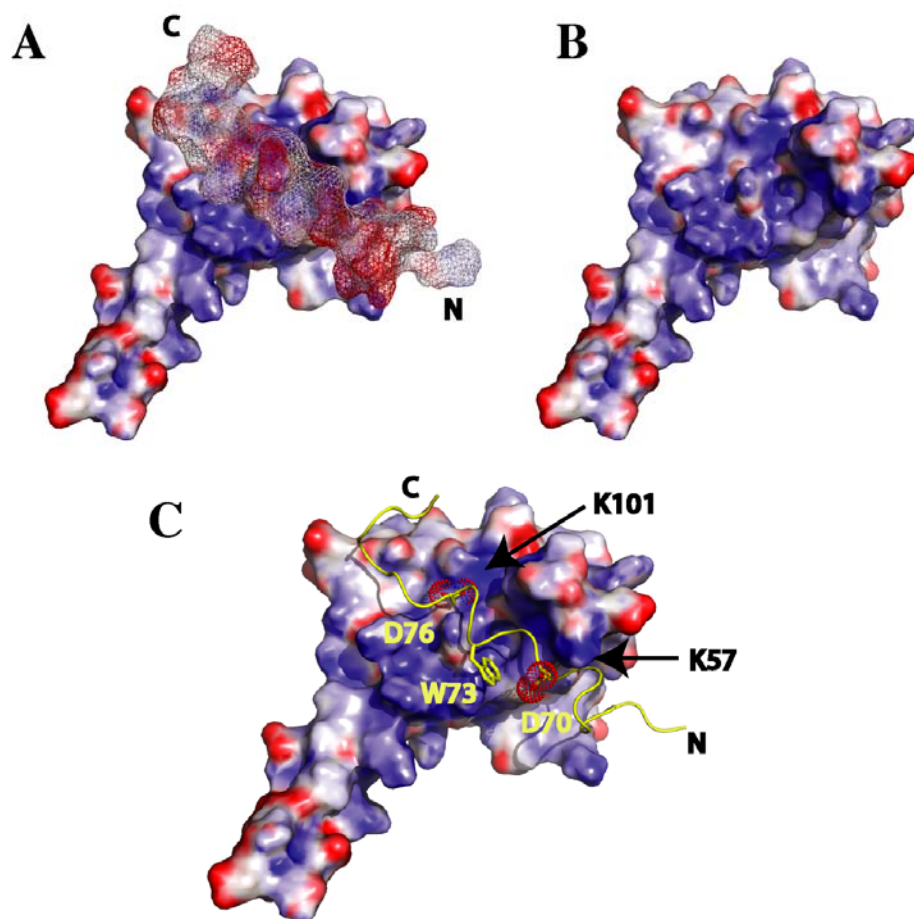


Figure 4. 5: Electrostatic interactions in the Tfb1PH/EKLFTAD2 complex. Three-dimensional structure of Tfb1PH shown as a molecular surface in which the electrostatic potential is mapped between -10 kT (red) and +10 kT (blue) either in the presence (A and C) or absence of EKLFTAD2

(B). In (A), the 3D structure of EKLFTAD2 is shown as mesh molecular surface with the same electrostatic potential mapped on the surface. In (C), the two salt bridges that appear to contribute to the stability of the Tfb1PH/EKLFTAD2 complex are highlighted. One salt bridge involves Asp70 of EKLFTAD2 and Lys57 of Tfb1PH. The second salt bridge is between Asp76 of EKLFTAD2 and Lys101 of Tfb1PH. Asp70 and Asp76 of EKLFTAD2 are highlighted with red mesh and the backbone trace of EKLFTAD2 and the side chain of Trp73 are shown in yellow.

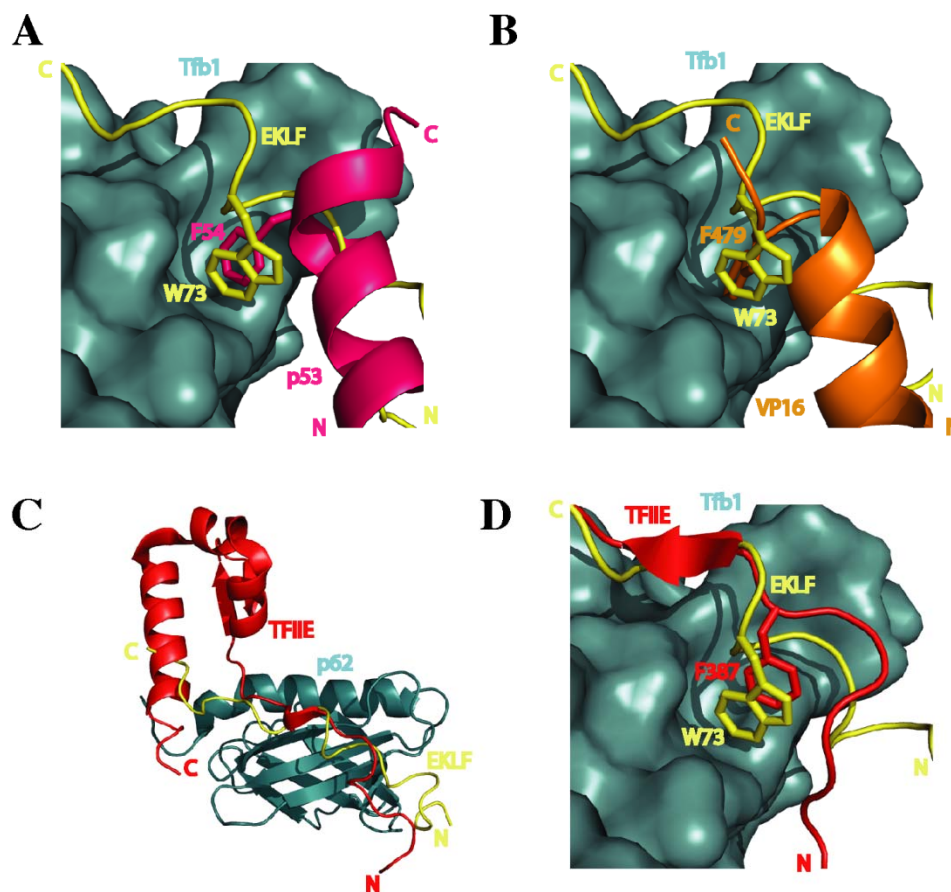


Figure 4. 6: EKLFTAD2 and TFIIIE α bind similarly to Tfb1PH/p62PH, but different then p53TAD2 and VP16C. (A) Overlay of EKLFTAD2 (yellow) and p53TAD2 (pink) shown as tubes on the surface of Tfb1PH (blue). EKLFTAD2 is in an extended conformation and p53TAD2 forms a 9-residue α -helix. In the overlay, the side chains of Trp73 of EKLFTAD2 and Phe54 of p53TAD2 are located in the same pocket on Tfb1PH. (B) Overlay of EKLFTAD2 (yellow) and VP16 (orange) represented as tubes in complex with Tfb1PH (blue). EKLFTAD2 is in an extended conformation and VP16 forms a 9-residue α -helix and the side chains of Trp73 of EKLFTAD2 and Phe479 of VP16C are located in the same binding pocket on Tfb1PH. (C) Overlay of TFIIIE α CTD and EKLFTAD2 on p62PH. (D) Overlay of TFIIIE α CTD and EKLFTAD2 on Tfb1PH.

the 3D structure of p62PH. EKLFTAD2 (yellow) and TFII α CTD (red) are represented as tubes on the surface of p62PH (blue). (D) Overlay of TFII α CTD and EKLFTAD2 on the 3D structure of Tfb1PH. EKLFTAD2 (yellow) and TFII α CTD (red) are represented as tubes on the surface of Tfb1PH (blue). And in this overlay the side chains of Trp73 of EKLFTAD2 and Phe387 of TFII α CTD are highlighted to show that they are located in the same binding pocket on Tfb1PH/p62PH.

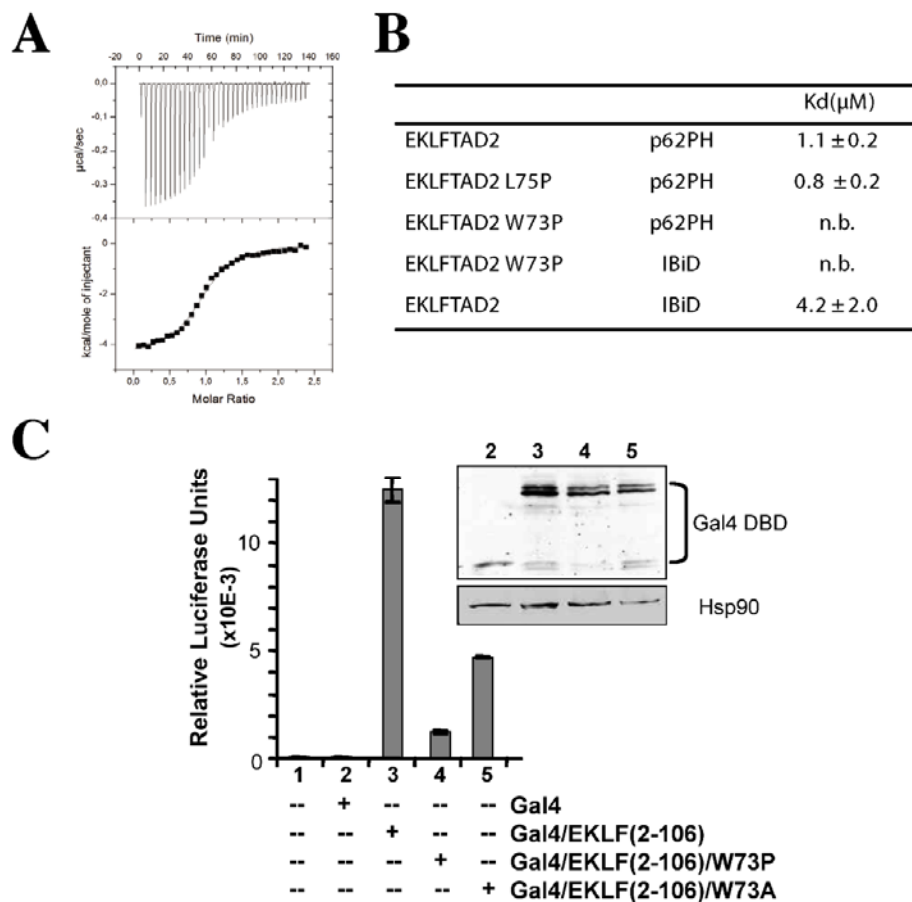
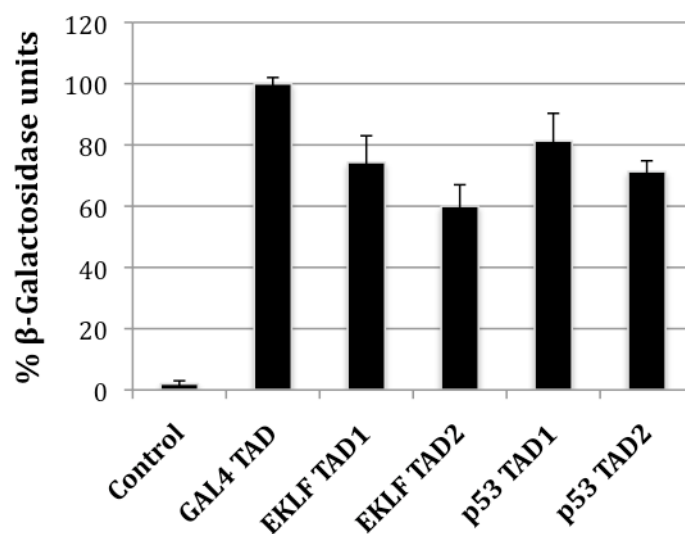


Figure 4. 7: Trp73 is crucial for the binding and in vivo activity of the minimal EKLFTAD. (A) Representative ITC thermogram obtained by successive additions of EKLFTAD2 into p62PH. **(B)** Dissociation constants (K_d) for the binding of EKLFTAD2 and Trp73 mutants of EKLFTAD2 with either p62PH or the IBiD domain of CBP/p300 determined by ITC. All injections fit the single binding site mechanism with 1:1 stoichiometry. **(C)** K562 cells were co-transfected with the 5xGAL-TK- β -globin-reporter gene together with plasmids expressing the indicated GAL4-fusion proteins

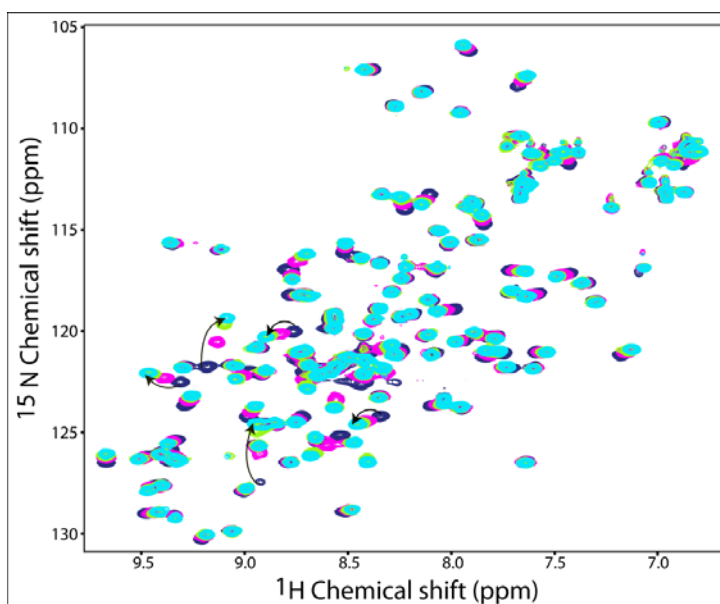
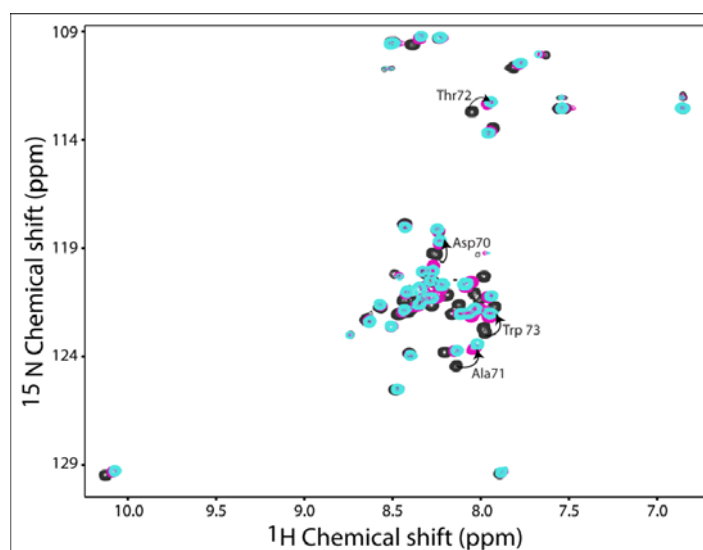
[22]. A plasmid expressing *Renilla* luciferase was included as a control for normalization of transfection efficiency. The transcriptional activity of these proteins was normalized. Error bars represent standard error about the mean of multiple independent experiments. The insert shows the level of expression of the GAL4-fusion proteins as monitored by western blot with an antibody against GAL4. Hsp90 levels are shown as a loading control.

Supplementary Figures

LexA fusion protein	% β -Galactosidase units
LexA-EKLFTAD1	74.3 \pm 8.7
LexA-EKLFTAD2	60.0 \pm 7.0
LexA-p53TAD1	81.3 \pm 9.0
LexA-p53TAD2	71.3 \pm 3.5
LexA-Gal4TAD	100 \pm 2
LexA	2 \pm 1



Supplementary Figure 4. 1. EKLFTAD1 and EKLFTAD2 function as independent activation domains in yeast. LexA-EKLF and LexA-p53 constructs were co-transformed into yeast with the reporter LexA operator-Lac-Z fusion plasmid pSH18-34. Results are presented as the mean of the percentages of the β -galactosidase units of the tested fusion proteins on the β -galactosidase units of the LexA-GAL4TAD positive control. Error bars represent standard error about the mean of three independent experiments.

A**B**

Supplementary Figure 4. 2. EKLFTAD2 binds to Tfb1PH. (A) Overlay of the two dimensional ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Tfb1PH in its free form (dark blue), in the presence of 0.5 equivalents (pink), 1 equivalent (green) and 2 equivalents (light blue) of EKLFTAD2. (B) Overlay of the 2D ^1H - ^{15}N HSQC spectra for ^{15}N -labeled EKLFTAD2 in its free form (black), in the presence of 0.5 equivalent (pink), 1.5 equivalent (light blue) of Tfb1PH. Arrows highlight several signals that

undergo significant changes in ^1H and ^{15}N chemical shifts in EKLFTAD2 upon formation of the TFb1PH/EKLFTAD2 complex.

Chapitre V

Conclusions et Perspectives

A - Conclusions

Le principal objectif de cette thèse fut d'entreprendre des études structurales et fonctionnelles d'interactions impliquées dans la régulation de la transcription au cours de la différenciation des érythrocytes avec pour but d'améliorer notre compréhension du mode d'action des domaines de transactivation.

Les travaux présentés dans le chapitre II relatent la découverte et l'étude de l'interaction entre le facteur de transcription érythropoïétique GATA-1 et la protéine suppresseur de tumeur p53. En collaboration avec le Dr. Cecelia Trainor (National Institutes of Health, Bethesda, MD), nous avons identifié et caractérisé une interaction entre les facteurs de transcription GATA-1 et p53. Nous avons démontré que, *in vitro*, le TAD de p53 interagit spécifiquement avec un domaine de GATA-1 comprenant le « linker » et le doigt de zinc en position carboxy-terminale (L+CF). Par contre, des études par RMN ont révélé que la formation d'un complexe ternaire impliquant GATA-1, l'ADN et p53 n'était pas possible, suggérant que p53 et l'ADN AGATAA sont en compétition pour lier GATA-1 L+CF. Parallèlement, les études *in vivo* ont démontré que ces deux partenaires interagissent ensemble et peuvent être co-immunoprécipités. Il semble que cette interaction a pour conséquence l'inhibition réciproque de l'activation des gènes cibles de ces protéines. À la lumière de ces résultats, nous suggérons que seule l'inhibition de p53 par GATA-1 est fonctionnellement importante et nécessaire afin d'inhiber la voie p53 et ainsi éviter l'apoptose durant le processus d'énucléation lors du développement des globules rouges.

Curieusement, GATA-1 est impliqué dans une interaction similaire avec PU.1. En effet, l'inhibition réciproque de GATA-1 et de PU.1 passe par une interaction directe entre ces deux partenaires. Le domaine de liaison à l'ADN de GATA-1 peut interagir avec le DBD du PU.1 ainsi qu'avec le TAD de PU.1 afin de diriger la différenciation soit vers l'érythropoïèse soit vers la thrombocytopoïèse. Ces travaux démontrent l'importance des TAD dans la régulation de l'activité des facteurs de transcription.

Cependant, la fonction principale des TAD demeure l'activation de la transcription *via* des interactions avec la machinerie transcriptionnelle. Les travaux présentés dans les Chapitres III et IV présentent des études structurales d'interactions impliquant des TAD acides avec la machinerie transcriptionnelle de base. L'implication de la sous-unité p62/Tfb1 (humain/levure) du facteur général de transcription TFIIH lors du recrutement de ce facteur dans le PIC, ainsi que son rôle dans l'activation de la transcription, font de cette sous-unité un acteur important de la transcription. De fait, l'interaction directe de p62/Tfb1 avec de nombreux TAD acides corrèle avec l'initiation de la transcription. Le Chapitre III de cette thèse présente la structure RMN de Tfb1 en complexe avec le TADC de la protéine virale VP16. Ces travaux démontrent clairement que l'activateur viral VP16 est très semblable au TAD2 de l'activateur p53 en complexe avec Tfb1. En effet, ces TAD partagent le même site de liaison sur Tfb1/p62 et adoptent tous deux une structure en hélice- α amphipatique de neuf résidus en complexe avec Tfb1. De façon similaire, dans les deux complexes, un résidu aromatique compris dans l'hélice- α (la Phe⁵⁴ de p53 et la Phe⁴⁷⁹ de VP16) vient se positionner dans une poche située entre les brins β 5, β 6 et β 7 de Tfb1. De plus, le positionnement de ces TAD est facilité par la présence de résidus chargés positivement à la surface de Tfb1 formant ainsi des interactions ioniques avec les charges négatives des TAD. La comparaison de ces deux structures a permis de constater que le TADC de la protéine virale VP16 et le TAD2 de la protéine humaine p53 interagissent avec un partenaire commun de façon comparable et a permis de proposer une séquence consensus pouvant interagir avec Tfb1/p62. La composition en acides aminés de cette séquence serait la suivante : D/E Φ EQ Φ T, où Φ est un résidu hydrophobe préférentiellement aromatique.

Le Chapitre IV rapporte la structure RMN du complexe formé par Tfb1 et le TAD d'EKLF. Ces résultats démontrent que les TAD d'EKLF et de p53 partagent un certain nombre de partenaires communs, tels que Tfb1/p62 et divers domaines des co-activateurs CBP/p300. Cependant, en dépit la similarité de séquence et de fonction de ces TAD et malgré le fait que le TAD2 d'EKLF partage le même site de liaison sur Tfb1 que les TAD de p53 et de VP16, EKLF utilise un mode de liaison entièrement différent de l'hélice- α amphipatique habituellement observée chez les activateurs. Effectivement, lors de son interaction avec Tfb1, EKLF TAD2

adopte une conformation majoritairement étendue allant des résidus Ala⁶⁹ à Thr⁸⁰. Par conséquent, ce TAD interagit avec Tfb1 sur une plus grande surface que les TAD de p53 et VP16, allant des feuillets β 5, β 6 et β 7 jusqu'à l'hélice H1 de Tfb1. Malgré des différences dans le mode de liaison, un résidu aromatique d'EKLF, le Trp⁷³, vient s'insérer dans la même cavité de Tfb1 que les résidus Phe⁴⁷ de p53 et Phe⁴⁷⁹ de VP16 dans les complexes Tfb1/p53 et Tfb1/VP16 respectivement. Récemment, la structure RMN de l'homologue humain de Tfb1, p62, en complexe avec TFIIIE α a été résolue [40]. Cette structure a révélé que TFIIIE α et les TAD de p53, VP16 et EKLF partagent le même site de liaison sur Tfb1/p62. Étonnamment, tout comme EKLF en complexe avec Tfb1, TFIIIE α adopte une conformation étendue en interaction avec p62. De plus, on observe que, de façon récurrente, les partenaires de Tfb1/p62 insèrent un acide aminé aromatique dans une poche située entre les brins β 5 et β 6 de Tfb1/p62. Dans le cas de TFIIIE α , c'est la Phe³⁸⁷ qui vient se positionner dans cette cavité.

À la lumière de ces résultats, il apparaît donc que les facteurs de transcription, les activateurs, de même que les co-activateurs ont la capacité d'interagir différemment avec leurs multiples partenaires. En effet, on compte au minimum deux modes d'interactions différents juste pour le domaine PH de p62 /Tfb1, le premier, où le ligand adopte une conformation tridimensionnelle en hélice α , et le second, où ce dernier s'allonge à la surface Tfb1/p62. Malgré cette différence majeure, il demeure certaines similitudes dans leur mode d'interaction, telle que la présence d'un résidu hydrophobe à l'intérieur de la poche formée des brins β 5 et β 6 de p62 et de Tfb1 qui semble crucial. Le site de liaison situé sur la face β 5, β 6 et β 7 de p62 et de Tfb1 semble donc posséder une certaine flexibilité lui permettant d'interagir avec une grande variété de facteurs de transcription, d'activateurs et de co-activateurs afin de réguler la transcription ponctuelle et sélective de gènes impliqués dans des voies biochimiques très diversifiées. Cette plasticité des protéines intervenant dans la transcription, brièvement démontrée lors de mes travaux, permet une fine régulation de la transcription impliquée dans des processus aussi diversifiés que la maturation des cellules sanguines et la prévention de la cancérogénèse.

B - Perspectives

1 - GATA-1 et ses partenaires

Ayant identifié les domaines minima requis pour la formation du complexe GATA-1 L+CF/p53 TAD2 (Chapitre III), il s'avèrerait intéressant de réaliser une étude approfondie du mécanisme de liaison entre ces deux partenaires. Dans un premier temps, des titrages RMN permettraient de définir les résidus engagés dans l'interaction et de cartographier la zone d'interaction sur chacun des partenaires. L'étude de l'effet de mutants ponctuels sur l'affinité de cette liaison pourrait permettre d'évaluer la contribution relative de résidus présents à l'interface de l'interaction. L'identification de mutants de GATA_1 ou de p53 perturbant l'interaction *in vitro* pourrait être corrélée avec des études *in vivo*. En effet, il serait possible de vérifier, dans un système modèle chez la souris, si les mutants ainsi ciblés affectent l'activité transcriptionnelle de ces facteurs de transcription. Par ailleurs, il serait intéressant de mener des études similaires sur l'interaction entre GATA-1 et le TAD acide de PU.1. Le mécanisme d'interaction entre GATA-1 et PU.1 est dépendant de la lignée sanguine dans laquelle elle a lieu, et s'avère cruciale pour diriger la différenciation des cellules hématopoïétiques vers la lignée érythroïde ou vers la lignée myéloïde [192]. Dans les cellules myéloïdes, le DBD de GATA-1 interagit avec le TAD de PU1 [141], tandis que dans les cellules érythroïdes le DBD de GATA-1 semble interagir avec le domaine ETS de PU.1 [138, 139]. Ceci suggère que le mécanisme d'interaction entre GATA-1 et PU.1 diffère dans les deux lignées sanguines. La caractérisation des interactions impliquant le DBD de GATA-1 avec les différents domaines de PU.1 et de p53 fournirait des informations importantes concernant les mécanismes par lesquels les interactions avec le DBD de GATA-1 régulent spécifiquement l'expression de gènes au cours du développement hématopoïétique ainsi que sur les mécanismes de reconnaissance de partenaires multiples par les TAD acides.

Par ailleurs, une autre interaction joue un rôle primordial dans l'érythropoïèse : l'interaction entre EKLF et GATA-1. Il a été démontré que cette interaction implique le DBD de GATA-1 et le DBD d'EKLF [193]. Cependant, aux vues des similitudes entre les TAD de p53 et

d'EKLF il serait intéressant de vérifier si le TAD d'EKLF peut interagir avec le DBD de GATA-1.

2 - Effet de modifications post-traductionnelles

Différents types de modifications post-traductionnelles, tel que l'acétylation, la phosphorylation, l'ubiquitination, ou encore la sumoylation, ont pour rôle de réguler l'activité des protéines. De par leur nature, les TAD acides, comporte peu ou pas de résidus pouvant être acétylés. La sumoylation et l'ubiquitination des TAD, sont elles, des modifications jouant un rôle important pour l'activation et la dégradation des facteurs de transcription. En effet, le dégron (séquence cible de l'ubiquitine) se retrouve souvent dans les domaines de transactivation. La phosphorylation des TAD acides, quant à elle, est un moyen couramment retrouvé pour réguler l'activité des facteurs de transcription et contrôler leur affinité pour leurs partenaires[194]. Il serait intéressant d'étudier l'effet de l'état de phosphorylation d'EKLF sur son affinité pour ses partenaires. En effet, la régulation des affinités relatives du TAD de p53 pour ses cibles, dépendamment de son état de phosphorylation, illustre ce phénomène (**Figure 1. 12** en page 37). L'activité de cette protéine est régulée *via* des interactions impliquant ses TAD avec l'ubiquitine ligase MDM2 et les co-activateurs CBP/p300 [90]. La liaison entre p53 TAD1 et MDM2, à l'origine de l'inhibition de p53, est abolie par la phosphorylation de la Thr¹⁸ du TAD1 de p53. Cela a pour conséquence l'activation de p53 et une augmentation de l'affinité du TAD1 de p53 pour le domaine KIX de CBP. De plus, deux sites de phosphorylation sont situés dans le TAD2 de p53 : Ser⁴⁶ [195] et Thr⁵⁵ [196]. Il a été démontré que la phosphorylation de ces résidus augmentent l'affinité de p53 pour p62 d'un facteur 30 [38], suggérant qu'une cascade de phosphorylations impliquant la Ser⁴⁶ et la Thr⁵⁵, joue un rôle crucial dans la régulation des gènes cibles de p53 nécessitant le TAD2.

De précédentes études ont mis en évidence que l'état de phosphorylation de plusieurs résidus compris dans le TAD d'EKLF pourrait être crucial pour l'activité de cette protéine [153] et que la phosphorylation de la Thr²³ du TAD1 d'EKLF par la casein kinase II stimule la transcription [174].

L'ensemble de ces observations laissent à penser que, de la même manière que p53, l'état de phosphorylation du TAD d'EKLF pourrait moduler son affinité pour ses partenaires et ainsi jouer un rôle dans son activité. Mener des études sur l'effet de la phosphorylation des TAD d'EKLF sur leur affinité avec leurs partenaires permettrait de mieux comprendre les mécanismes de régulation de ces gènes cibles. En outre, basé sur la structure du complexe EKLF TAD2/Tfb1, nous avons pu identifier deux sites potentiels de phosphorylation sur EKLF TAD2 : Thr⁷² et Thr⁸⁰. La phosphorylation pourrait entraîner respectivement la formation d'une interaction ionique avec les Lys⁵⁴ et Lys¹⁰⁴ de p62. Il serait intéressant de vérifier si, *in vivo*, EKLF est modifié à ces positions, puis de voir si la phosphorylation des TAD d'EKLF module son affinité pour Tfb1/p62. Étant donné le rôle que joue la phosphorylation du TAD de p53 sur son interaction avec CBP, il serait aussi important de regarder l'effet de l'état de phosphorylation des TAD d'EKLF sur leur interaction avec les divers domaines de CBP. L'utilisation de peptides phosphorylés dans des études de liaison *in vitro* avec divers partenaires d'EKLF pourrait en conséquence nous éclairer quand au rôle potentiel de telles modifications.

3 - Mécanismes de liaison des TAD

Depuis plusieurs années, plusieurs travaux ont été menés afin de comprendre le mode d'action des domaines de transactivation et, de fait, de plus en plus de données structurales sont à notre disposition. Malgré cela, les mécanismes par lesquels ces domaines sont capables d'interagir avec de multiples partenaires restent méconnus. Les premières structures de TAD ayant été résolues relataient la formation d'hélice- α en interaction avec leurs partenaires. Quelques exemples sont les interactions p53 TAD1/MDM2 [60], p53 TAD2/Tfb1 [38] ou encore le domaine KIX de CBP en interaction avec CREB [78]. Depuis quelques années, les co-activateurs CBP/p300 ont fait l'objet de nombreuses études structurales. Un nombre croissant de structures impliquant différents domaines de CBP/p300 en complexes avec divers domaines de transactivation sont à notre disposition. Récemment, la structure du domaine de transactivation de STAT1 en complexe avec le domaine TAZ2 de CBP a mis en évidence un nouveau mécanisme d'interaction d'un TAD pour sa cible. En effet, malgré le fait que STAT1 partage le même site de

liaison sur TAZ2 avec plusieurs autres TAD (tels CITED2 et HIF-1 α), il n'adopte pas de structure secondaire bien définie en complexe avec ce partenaire et sa conformation est majoritairement non structurée. La structure du TAD2 d'EKLF en complexe avec le domaine PH de Tfb1 est le deuxième exemple connu de TAD interagissant avec sa cible sous forme étendue.

Nous savons que les cibles d'activateurs, Tfb1/p62 et CBPp300, peuvent interagir avec plusieurs TAD de diverses façons suggérant une adaptabilité de ces protéines face à la diversité de leurs partenaires. Mais il reste encore à déterminer si un TAD peut adopter différentes structures en interaction avec différentes cibles. Un exemple de TAD adoptant une structure similaire en complexe avec deux partenaires est le TAD1 de p53. La structure nouvellement résolue du TAD1 de p53 en complexe avec le domaine TAZ2 de CBP/p300 a révélé que cet activateur met en place un mécanisme d'interaction analogue avec ces deux partenaires : MDM2 [60] et TAZ2 [85]. En effet, la formation d'une hélice- α dans le TAD1 de p53 est induite par sa liaison à MDM2 (hélice allant des résidus 18 à 25) et à TAZ2 (hélice allant des résidus 16 à 25). Il serait donc intéressant de poursuivre des études structurales des complexes dans lesquels le TAD2 d'EKLF est impliqué, afin de déterminer si ce TAD interagit avec différents partenaires de la même manière qu'il interagit avec Tfb1/p62. De plus, des études structurales de l'interaction Tfb1 en interaction avec le TAD1 d'EKLF ainsi qu'avec le TADN de VP16 seraient des plus intéressantes puisque ces TAD activent la transcription et qu'ils partagent avec VP16 TADC, p53 TAD2 et EKLF TAD2, le même site de liaison sur Tfb1/p62.

4 - La transcription, une cible thérapeutique

Actuellement, de nombreux travaux visent à développer des inhibiteurs d'interactions protéine/protéine clefs dans le développement et la progression de cancers. Ces inhibiteurs se présentent habituellement sous la forme de petites molécules organiques ou de peptides, et tentent de mimer l'interaction ciblée afin de l'inhiber. Un exemple est le développement d'inhibiteurs de l'interaction p53 TAD1/MDM2 [60] ayant pour but de rétablir l'activité de p53. La surexpression de MDM2 dans diverses tumeurs étant responsable de l'inactivation de p53, l'inhibition de

l'interaction entre ces partenaires aurait pour effet de réactiver l'activité de p53, menant à un effet anti tumoral.

Les interactions de TAD viraux et cellulaires avec la sous-unité p62/Tfb1 du facteur général de transcription TFIID, ou avec les histones acétyltransférases CBP/p300, corrélient avec l'activation de la transcription. Ainsi, ces complexes représentent des cibles potentielles d'inhibiteurs pour la lutte contre le développement de cancers ou contre les infections virales. En effet, la perturbation de l'interaction entre le TAD viral de VP16 avec le facteur général de transcription TFIID entrainerait l'inhibition de l'activation des gènes cibles viraux, arrêtant ainsi l'infection virale. Par ailleurs, plusieurs protéines virales contiennent des TAD acides s'associant à CBP/p300 tels qu'E1A ou HIV-1 Tat [197, 198]. Le mode de fonctionnement de ces protéines virales passe par une séquestration de CBP/p300 afin de concurrencer des protéines cellulaires telles que p53 pour la liaison à CBP/p300. Ces mécanismes seraient donc responsables de l'activation de la transcription de gènes viraux et de la reprogrammation de l'expression des gènes cellulaires dans la cellule infectée, fournissant ainsi un environnement optimal pour la réplication virale. L'inhibition de ces interactions inactiverait l'expression des gènes viraux et endiguerait l'infection virale.

Dans le cas de l'interaction impliquant le facteur général de transcription TFIID et le TAD de p53, on pourrait imaginer pouvoir réactiver les gènes cibles de p53 dans des cellules cancéreuses en mimant l'interaction p62/p53. D'autre part, dans le contexte d'une pathologie telle que la β -thalassémie, une molécule parodiant l'interaction EKLF/p62 pourrait rétablir l'expression de la β -globine.

Par ailleurs, depuis plusieurs années divers travaux ont révélé des rôles contradictoires de CBP/p300 dans le cancer. En effet, ces co-activateurs sont la cible d'oncoprotéines virales [199]. De plus, des translocations chromosomiques ont été rapportées entre la protéine MOZ (**M**onocytic leukemia **Z**inc finger) et CBP chez des patients atteint de leucémie myéloïde [200], ainsi qu'entre p300 et le gène de MLL (lineage leukemia protein) chez des patients atteint d'AML [201]. De surcroît, des mutations faux-sens de p300 ont été identifiées dans des cellules épithéliales cancéreuses [202, 203]. Cependant, d'autres travaux ont démontré que CBP/p300

peuvent avoir un rôle opposé. En effet, l'inhibition de p300 dans des mélanocytes empêche la croissance tumorale [204]. Ces résultats indiquent que le rôle de CBP/p300 dans le développement de cancers est versatile et contexte-dépendant. Cependant, ces observations soulèvent l'idée que des petites molécules inhibitrices de ces co-activateurs pourraient être exploitées à des fins thérapeutiques dans le but de traiter certaines formes de cancers. Pour cela, les études structurales des complexes formés par ces protéines et leurs partenaires sont des atouts pour la conception d'inhibiteurs.

C - Conclusions générales

Via des interactions protéine/protéine impliquant leurs domaines de transactivation, les activateurs transcriptionnels régulent l'expression de gènes en temps et lieu approprié. Les TAD, non structurés sous forme libre, offrent l'avantage d'être accessible à des modifications post-traductionnelles et ont la capacité d'interagir avec de multiples partenaires. L'impact de nos travaux s'avère être capital tant du point de vue de la compréhension détaillée des mécanismes de fonctionnement des TAD, que dans l'optique du développement de nouveaux agents thérapeutiques. En effet, nos études ont mis en évidence la diversité des mécanismes de reconnaissance des TAD cellulaire et viraux par leurs divers partenaires, et illustrent le fait que les facteurs généraux de transcriptions possèdent une flexibilité de reconnaissance, puisque ceux-ci sont capables de lier divers TAD de différentes manières. Les travaux réalisés au cours de cette thèse apportent une contribution notable à la compréhension des mécanismes d'action des domaines de transactivation acide non structurés.

Bibliographie

1. Orkin, S.H., *Development of the hematopoietic system*. Curr Opin Genet Dev, 1996. **6**(5): p. 597-602.
2. Orkin, S.H. and L.I. Zon, *Hematopoiesis: an evolving paradigm for stem cell biology*. Cell, 2008. **132**(4): p. 631-44.
3. Kim, S.I. and E.H. Bresnick, *Transcriptional control of erythropoiesis: emerging mechanisms and principles*. Oncogene, 2007. **26**(47): p. 6777-6794.
4. Grzesiek, S. and A. Bax, *Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein*. J. Magn. Reson., 1992. **96**(2): p. 432-40.
5. Grzesiek, S., et al., *Proton, carbon-13, and nitrogen-15 NMR backbone assignments and secondary structure of human interferon-gamma*. Biochemistry, 1992. **31**(35): p. 8180-90.
6. Guardiola, A.R. and T.P. Yao, *Molecular cloning and characterization of a novel histone deacetylase HDAC10*. Journal of Biological Chemistry, 2002. **277**(5): p. 3350-3356.
7. Orkin, S.H., *Diversification of haematopoietic stem cells to specific lineages*. Nature Rev. Genet., 2000. **1**: p. 57-64.
8. Shivdasani, R.A. and S.H. Orkin, *The transcriptional control of hematopoiesis*. Blood, 1996. **87**(10): p. 4025-39.
9. Cantor, A.B. and S.H. Orkin, *Transcriptional regulation of erythropoiesis: an affair involving multiple partners*. Oncogene, 2002. **21**(21): p. 3368-76.
10. Davison, B.L., et al., *Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences*. Nature, 1983. **301**(5902): p. 680-6.
11. Burley, S.K. and R.G. Roeder, *Biochemistry and structural biology of transcription factor IID (TFIID)*. Annu Rev Biochem, 1996. **65**: p. 769-99.
12. Conaway, R.C. and J.W. Conaway, *General initiation factors for RNA polymerase II*. Annu Rev Biochem, 1993. **62**: p. 161-90.
13. Robert, F., et al., *Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF*. Mol Cell, 1998. **2**(3): p. 341-51.
14. Orphanides, G., T. Lagrange, and D. Reinberg, *The general transcription factors of RNA polymerase II*. Genes & Dev., 1996. **10**: p. 2657-2673.
15. Kobor, M.S. and J. Greenblatt, *Regulation of transcription elongation by phosphorylation*. Biochimica et Biophysica Acta, Gene Structure and Expression, 2002. **1577**(2): p. 261-275.

16. Lu, H., et al., *Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II*. *Nature*, 1992. **358**: p. 641-645.
17. Tirode, F., et al., *Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymic subunits, XPB, XPD, and cdk7*. *Molecular Cell*, 1999. **3**(1): p. 87-95.
18. Price, D.H., *P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II*. *Mol Cell Biol*, 2000. **20**(8): p. 2629-34.
19. Ohkuma, Y., et al., *Analysis of the role of TFIIH in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIH α* . *Mol. Cell. Biol.*, 1995. **15**: p. 4856-4866.
20. Ohkuma, Y. and R.G. Roeder, *Regulation of TFIIH ATPase and kinase activities by TFIIH during active initiation complex formation*. *Nature*, 1994. **368**: p. 160-163.
21. Maxon, M.E., J.A. Goodrich, and R. Tjian, *Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance*. *Genes & Dev.*, 1994. **8**: p. 515-524.
22. Fisher, R.P., et al., *Alternative mechanisms of CAK assembly require an assembly factor or an activating kinase*. *Cell*, 1995. **83**: p. 47-57.
23. Fisher, R.P. and D.O. Morgan, *A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase*. *Cell*, 1994. **78**: p. 713-24.
24. Fisher, R.P. and D.O. Morgan, *CAK in TFIIH: crucial connection or confounding coincidence?* *Biochimica Biophysica Acta*, 1996. **1288**: p. O7-O10.
25. Rochette-Egly, C., *Nuclear receptors: integration of multiple signalling pathways through phosphorylation*. *Cellular Signalling*, 2003. **15**: p. 355-366.
26. Beck, B.D., D.S. Hah, and S.H. Lee, *XPB and XPD between transcription and DNA repair*. *Adv Exp Med Biol*, 2008. **637**: p. 39-46.
27. Coin, F., V. Oksenyshyn, and J.M. Egly, *Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair*. *Mol Cell*, 2007. **26**(2): p. 245-56.
28. Feaver, W.J., et al., *CTD Kinase associated with yeast RNA polymerase II initiation factor b*. *Cell*, 1991. **67**: p. 1223-1230.
29. Glover-Cutter, K., et al., *TFIIH-associated Cdk7 kinase functions in phosphorylation of CTD Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II*. *Mol. Cell. Biol.*, 2009: p. MCB.00637-09.
30. Viprakasit, V., et al., *Mutations in a general transcription factor, TFIIH result in a novel unlinked form of beta thalassaemia in individuals with trichothiodystrophy*. *Blood*, 2001. **98**(11): p. 1837.

31. Viprakasit, V., et al., *Mutations in the general transcription factor TFIIH result in beta-thalassaemia in individuals with trichothiodystrophy*. Human Molecular Genetics, 2001. **10**(24): p. 2797-2802.
32. Bushnell, D.A., C. Bamdad, and R.D. Kornberg, *A minimal set of RNA polymerase II transcription protein interactions*. J Biol Chem, 1996. **271**(33): p. 20170-4.
33. Pai, M.T., et al., *Solution structure of the Ubp-M BUZ domain, a highly specific protein module that recognizes the C-terminal tail of free ubiquitin*. Journal of Molecular Biology, 2007. **370**(2): p. 290-302.
34. Pearson, A. and J. Greenblatt, *Modular organization of the E2F1 activation domain and its interaction with general transcription factors TBP and TFIIH*. Oncogene, 1997. **15**: p. 2643-2658.
35. Xiao, H., et al., *Binding of basal transcription factor TFIIH to the acidic activation domain of VP16 and p53*. Mol. Cell. Biol., 1994. **14**: p. 7013-7024.
36. Lemmon, M.A., K.M. Ferguson, and C.S. Abrams, *Pleckstrin homology domains and the cytoskeleton*. FEBS Lett., 2002. **513**: p. 71-76.
37. Di Lello, P., et al., *NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites*. Biochemistry, 2005. **44**(21): p. 7678-86.
38. Di Lello, P., et al., *Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53*. Mol Cell, 2006. **22**(6): p. 731-40.
39. Gervais, V., et al., *TFIIH contains a PH domain involved in DNA nucleotide excision repair*. Nat Struct Mol Biol, 2004. **11**(7): p. 616-22.
40. Okuda, M., et al., *Structural insight into the TFIIIE-TFIIH interaction: TFIIIE and p53 share the binding region on TFIIH*. EMBO J, 2008. **27**(7): p. 1161-71.
41. Di Lello, P., et al., *p53 and TFIIIEalpha share a common binding site on the Tfb1/p62 subunit of TFIIH*. Proc Natl Acad Sci U S A, 2008. **105**(1): p. 106-11.
42. Nedialkov, Y.A. and S.J. Triezenberg, *Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region*. Arch. Biochem. and Biophys., 2004. **425**(1): p. 77-86.
43. Tansey, W.P. and W. Herr, *Selective use of TBP and TFIIB revealed by a TATA-TBP-TFIIB array with altered specificity*. Science, 1997. **275**(5301): p. 829-831.
44. Tansey, W.P., et al., *Multiple regions of TBP participate in the response to transcriptional activators in vivo*. Genes & Development, 1994. **8**(22): p. 2756-69.
45. Tong, X., et al., *The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein*. J. Virol., 1995. **69**(1): p. 585-8.

46. Uhlmann, T., et al., *The VP16 activation domain establishes an active Mediator lacking CDK8 in vivo*. J. Biol. Chem., 2007. **282**(4): p. 2163-2173.
47. Yang, F., et al., *An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis*. Nature, 2006. **442**(7103): p. 700-704.
48. Giebler, H.A., I. Lemasson, and J.K. Nyborg, *p53 recruitment of CREB binding protein mediated through phosphorylated CREB: a novel pathway of tumor suppressor regulation*. Molecular and Cellular Biology, 2000. **20**(13): p. 4849-4858.
49. Zhang, W. and J.J. Bieker, *Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases*. Proc Natl Acad Sci U S A, 1998. **95**(17): p. 9855-60.
50. Lee, D., et al., *SWI/SNF complex interacts with tumor suppressor p53 and is necessary for the activation of p53-mediated transcription*. J. Biol. Chem., 2002. **277**(25): p. 22330-22337.
51. Gutierrez, J.L., et al., *Activation domains drive nucleosome eviction by SWI/SNF*. EMBO J., 2007. **26**(3): p. 730-740.
52. Neeley, K.E., et al., *Transcription activator interactions with multiple SWI/SNF subunits*. Mol. Cell. Biol., 2002. **22**: p. 1615-1625.
53. Ptashne, M. and A.A.F. Gann, *Activators and targets*. Nature, 1990. **346**: p. 329-331.
54. Ptashne, M. and A.A.F. Gann, *Transcriptional activation by recruitment*. Nature, 1997. **386**: p. 569-577.
55. Schmitz, M.L., et al., *Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation*. J. Biol. Chem., 1994. **269**: p. 25613-25620.
56. Vandel, L. and T. Kouzarides, *Residues phosphorylated by TFIIH are required for E2F-1 degradation during S-phase*. EMBO J., 1999. **18**: p. 4280-4291.
57. Fink, A.L., *Natively unfolded proteins*. Curr Opin Struct Biol, 2005. **15**(1): p. 35-41.
58. Iakoucheva, L.M., et al., *Intrinsic disorder in cell-signaling and cancer-associated proteins*. J Mol Biol, 2002. **323**(3): p. 573-84.
59. Liu, J., J.R. Faeder, and C.J. Camacho, *Toward a quantitative theory of intrinsically disordered proteins and their function*. Proc Natl Acad Sci U S A, 2009. **106**(47): p. 19819-23.
60. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain*. Science, 1996. **274**: p. 948-953.
61. Bochkareva, E., et al., *Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A*. Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15412-7.

62. Jonker, H.R.A., et al., *Structural properties of the promiscuous VP16 activation domain*. Biochemistry, 2005. **44**(3): p. 827-839.
63. Uesugi, M., et al., *Induced ζ -helix in the VP16 activation domain upon binding to a human TAF*. Science, 1997. **277**: p. 1310-1313.
64. Whitehouse, I., et al., *Nucleosome mobilization catalysed by the yeast SWI/SNF complex*. Nature, 1999. **400**(6746): p. 784-7.
65. Rachez, C. and L.P. Freedman, *Mediator complexes and transcription*. Curr Opin Cell Biol, 2001. **13**(3): p. 274-80.
66. Casamassimi, A. and C. Napoli, *Mediator complexes and eukaryotic transcription regulation: an overview*. Biochimie, 2007. **89**(12): p. 1439-46.
67. Hampsey, M. and D. Reinberg, *RNA polymerase II as a control panel for multiple coactivator complexes*. Curr Opin Genet Dev, 1999. **9**(2): p. 132-9.
68. Rickert, P., et al., *Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II*. Oncogene, 1996. **12**(12): p. 2631-40.
69. Eckner, R., et al., *Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor*. Genes Dev, 1994. **8**(8): p. 869-84.
70. Ogryzko, V.V., et al., *The transcriptional coactivators p300 and CBP are histone acetyltransferases*. Cell, 1996. **87**(5): p. 953-9.
71. Lundblad, J.R., et al., *Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP*. Nature, 1995. **374**(6517): p. 85-8.
72. Felzien, L.K., et al., *Specificity of cyclin E-Cdk2, TFIIB, and E1A interactions with a common domain of the p300 coactivator*. Mol Cell Biol, 1999. **19**(6): p. 4241-6.
73. Dallas, P.B., P. Yaciuk, and E. Moran, *Characterization of monoclonal antibodies raised against p300: both p300 and CBP are present in intracellular TBP complexes*. J Virol, 1997. **71**(2): p. 1726-31.
74. Cho, H., et al., *A human RNA polymerase II complex containing factors that modify chromatin structure*. Mol Cell Biol, 1998. **18**(9): p. 5355-63.
75. Blobel, G.A., *CREB-binding protein and p300: molecular integrators of hematopoietic transcription*. Blood, 2000. **95**(3): p. 745-755.
76. Lin, C.H., et al., *A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies*. Mol Cell, 2001. **8**(3): p. 581-90.
77. Zor, T., et al., *Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb*. J Mol Biol, 2004. **337**(3): p. 521-34.

78. Radhakrishnan, I., et al., *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions*. Cell, 1997. **91**: p. 741-752.
79. De Guzman, R.N., et al., *Structural basis for cooperative transcription factor binding to the CBP coactivator*. J Mol Biol, 2006. **355**(5): p. 1005-13.
80. Lee, C.W., et al., *Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP*. Biochemistry, 2009. **48**(10): p. 2115-24.
81. Wojciak, J.M., et al., *Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains*. EMBO J, 2009. **28**(7): p. 948-58.
82. Waters, L., et al., *Structural diversity in p160/CREB-binding protein coactivator complexes*. J Biol Chem, 2006. **281**(21): p. 14787-95.
83. Demarest, S.J., et al., *Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators*. Nature, 2002. **415**(6871): p. 549-53.
84. De Guzman, R.N., et al., *Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligands for non-identical binding sites*. J Biol Chem, 2004. **279**(4): p. 3042-9.
85. Feng, H., et al., *Structural basis for p300 Taz2-p53 TAD1 binding and modulation by phosphorylation*. Structure, 2009. **17**(2): p. 202-10.
86. Dames, S.A., et al., *Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response*. Proc. Natl. Acad. Sci. U.S.A., 2002. **99**: p. 5271-5276.
87. Ferreon, J.C., et al., *Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein*. Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13260-5.
88. Demarest, S.J., et al., *Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators*. Nature, 2002. **415**: p. 549-553.
89. Grossman, S.R., *p300/CBP/p53 interaction and regulation of the p53 response*. European J. Biochemistry, 2001. **268**(10): p. 2773-2778.
90. Ferreon, J.C., et al., *Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(16): p. 6591-6596.
91. Ito, A., et al., *p300/CBP-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by MDM2*. Embo Journal, 2001. **20**(6): p. 1331-1340.
92. Blobel, G.A., et al., *CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation*. Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 2061-2066.

93. Hung, H.-L., et al., *CREB-binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites*. Mol. Cell. Biol., 1999. **19**: p. 3496-3505.
94. HuiJie, J.J., et al., *CREB-binding proteins (CBP) as a transcriptional coactivator of GATA-2*. Science in China Series C-Life Sciences, 2008. **51**(3): p. 191-198.
95. Hong, W., et al., *Inhibition of CBP acetyltransferase activity by the Ets family oncoprotein PU.1*. Blood, 2001. **98**(11): p. 3448.
96. Cheng, X., et al., *The transcriptional integrator CREB-binding protein mediates positive cross talk between nuclear hormone receptors and the hematopoietic bZip protein p45/NF-E2*. Mol Cell Biol, 1997. **17**(3): p. 1407-16.
97. Orkin, S.H., *GATA-binding transcription factors in hematopoietic cells*. Blood, 1992. **80**: p. 575-581.
98. Trainor, C.D., et al., *Structure and evolution of a human erythroid transcription factor*. Nature, 1990. **343**(6253): p. 92-96.
99. Evans, T. and G. Felsenfeld, *The erythroid-specific transcription factor eryf1: A new finger protein*. Cell, 1989. **58**: p. 877-885.
100. Omichinski, J.G., et al., *A small single-'finger' peptide from the erythroid transcription factor GATA-1 binds specifically to DNA as a zinc or an iron complex*. Proc. Natl. Acad. Sci. USA, 1993. **93**: p. 1676-1680.
101. Omichinski, J.G., et al., *NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1*. Science, 1993. **261**(5120): p. 438-46.
102. Crossley, M., et al., *Regulation of the erythroid Kruppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1*. J. Biol. Chem., 1994. **269**: p. 15440-4.
103. McInerney, J.M., et al., *Formation of a beta-globin LCR core structure in nuclear chromatin requires binding sites for multiple factors including GATA-1, NF-E2, EKLF and Sp-1*. Blood, 2000. **96**(11): p. 2405.
104. Merika, M. and S.H. Orkin, *Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins SP1 and EKLF*. Molecular and Cellular Biology, 1995. **15**(5): p. 2437-2447.
105. Rekhtman, N., et al., *Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells*. Genes & Dev., 1999. **13**: p. 1398-1411.
106. Shimizu, R., et al., *GATA-1 self-association controls erythroid development in vivo*. Journal of Biological Chemistry, 2007. **282**(21): p. 15862-15871.
107. Trainor, C.D., et al., *A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction*. Mol Cell Biol, 1996. **16**(5): p. 2238-47.

108. Weiss, M.J. and S.H. Orkin, *Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis*. Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 9623-7.
109. De Maria, R., et al., *Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1*. Nature, 1999. **401**(6752): p. 489-493.
110. Pevny, L., et al., *Development of hematopoietic cells lacking transcription factor GATA-1*, in *Development*. 1995. p. 163-72.
111. Grass, J.A., et al., *GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(15): p. 8811-8816.
112. Nerlov, C., et al., *GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription*. Blood, 2000. **95**: p. 2543-2551.
113. Zhang, P., et al., *Negative cross-talk between the hematopoietic regulators: GATA proteins repress PU.1*. Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 8705-8710.
114. Li, Q.L., et al., *Binary transgenic mouse model for studying the trans control of globin gene switching: Evidence that GATA-1 is an in vivo repressor of human epsilon gene expression*. Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(6): p. 2444-2448.
115. Fujiwara, Y., et al., *Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(22): p. 12355-12358.
116. Yamaguchi, Y., et al., *Forced GATA-1 expression in the murine myeloid cell line M1: induction of c-Mpl expression and megakaryocytic/erythroid differentiation*. Blood, 1998. **91**: p. 450-457.
117. Lowry, J.A. and J.P. Mackay, *GATA-1: One protein, many partners*. International Journal of Biochemistry & Cell Biology, 2006. **38**(1): p. 6-11.
118. Welch, J.J., et al., *Global regulation of erythroid gene expression by transcription factor GATA-1*. Blood, 2004. **104**(10): p. 3136-3147.
119. Anguita, E., et al., *Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2*. EMBO J., 2004. **23**: p. 2841-2852.
120. Ackley, C.J. and C.H. Lowrey. *GATA-1, NF-E2 and EKLF/Sp1 binding elements cooperate to specify patterns of histone hyper-acetylation at beta-globin LCR HS core elements*. 2004.
121. Goodwin, A.J., et al., *In vivo formation of a human beta-globin locus control region core element requires binding sites for multiple factors including GATA-1, NF-E2, erythroid*

- Kruppel-like factor, and Sp1*. Journal of Biological Chemistry, 2001. **276**(29): p. 26883-26892.
122. Gallagher, P.G., et al., *GATA-1 and Oct-1 are required for expression of the human alpha-hemoglobin-stabilizing protein gene*. Journal of Biological Chemistry, 2005. **280**(47): p. 39016-39023.
 123. Boulanger, L., et al., *Erythroid expression of the human alpha-spectrin gene promoter is mediated by GATA-1- and NF-E2-binding proteins*. J Biol Chem, 2002. **277**(44): p. 41563-70.
 124. Suzuki, N., et al. *Transgenic expression of erythropoietin receptor (EPOR) under the GATA-1 gene regulatory domain rescued EPOR knockout mouse from lethality*. 2003.
 125. Shirihai, O.S., et al., *ABC-me: A novel GATA-1-induced mitochondrial erythroid transporter involved in heme biosynthesis*. Blood, 1999. **94**(10): p. 2855.
 126. Shirihai, O.S., et al., *ABC-me: a novel mitochondrial transporter induced by GATA-1 during erythroid differentiation*. Embo Journal, 2000. **19**(11): p. 2492-2502.
 127. Zhou, S.P., et al., *Biochemical fates of alpha hemoglobin bound to alpha hemoglobin-stabilizing protein AHSP*. Journal of Biological Chemistry, 2006. **281**(43): p. 32611-32618.
 128. Crispino, J.D., et al., *Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex*. Mol. Cell., 1999. **3**: p. 219-228.
 129. Wang, X., et al., *Cooperativity between GATA-1, FOG-1 and an Ets family transcription factor is required for expression of a megakaryocyte-specific gene in vivo*. Blood, 2001. **98**(11): p. 3446.
 130. Kowalski, K., et al., *The solution structure of the N-terminal zinc finger of GATA-1 reveals a specific binding face for the transcriptional co-factor FOG*. J. Biomol. NMR, 1999. **13**: p. 249-262.
 131. Chang, A.N., et al., *The essential role of FOG-1 in early megakaryopoiesis requires an interaction with either GATA-1 or GATA-2*. Blood, 2001. **98**(11): p. 3301.
 132. Cantor, A.B., S.G. Katz, and S.H. Orkin, *Distinct domains of the GATA-1 cofactor FOG-1 differentially influence erythroid versus megakaryocytic maturation*. Mol. Cell. Biol., 2002. **22**: p. 4268-4279.
 133. Muntean, A.G. and J.D. Crispino, *Differential requirements for the activation domain and FOG-interaction surface of GATA-1 in megakaryocyte gene expression and development*. Blood, 2005. **106**(4): p. 1223-1231.
 134. Vakoc, C.R., et al., *Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1*. Molecular Cell, 2005. **17**(3): p. 453-462.
 135. Wang, X., et al., *Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors*. EMBO J., 2002. **21**: p. 5225-5234.

136. Gregory, R.C., et al., *Functional interaction of GATA1 with erythroid Kruppel-like factor and SPI1 at defined erythroid promoters*. Blood, 1996. **87**(5): p. 1793-1801.
137. Liew, C.W., et al., *Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1*. Journal of Biological Chemistry, 2006. **281**(38): p. 28296-28306.
138. Rekhtman, N., et al., *PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation*. Mol. Cell. Biol., 2003. **23**: p. 7460-7474.
139. Stopka, T., et al., *PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure*. Embo Journal, 2005. **24**(21): p. 3712-3723.
140. Yamada, T., et al., *Reduction of DNA binding activity of the GATA-1 transcription factor in the apoptotic process induced by overexpression of PU.1 in murine erythroleukemia cells*. Experimental Cell Res., 1998. **245**: p. 186-194.
141. Zhang, P., et al., *PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding*. Blood, 2000. **96**: p. 2641-2648.
142. Takahashi, T., et al., *Inhibitory interaction of c-Myb and GATA-1 via transcriptional co-activator CBP*. Oncogene, 2000. **19**: p. 134-140.
143. Lamonica, J.M., C.R. Vakoc, and G.A. Blobel, *Acetylation of GATA-1 is required for chromatin occupancy*. Blood, 2006. **108**(12): p. 3736-3738.
144. Levine, A.J., *p53, the cellular gatekeeper for growth and division*. Cell, 1997. **88**(3): p. 323-31.
145. Gregory, T., et al., *GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression*. Blood, 1999. **94**(1): p. 87-96.
146. Motoyama, N., et al., *bcl-x prevents apoptotic cell death of both primitive and definitive erythrocytes at the end of maturation*. J Exp Med, 1999. **189**(11): p. 1691-8.
147. Aispuru, G.R., et al., *Erythroid expansion and survival in response to acute anemia stress: the role of EPO receptor, GATA-1, Bcl-xL and caspase-3*. Cell Biol Int, 2008. **32**(8): p. 966-78.
148. Ribeil, J.A., et al., *Hsp70 regulates erythropoiesis by preventing caspase-3-mediated cleavage of GATA-1*. Nature, 2007. **445**(7123): p. 102-5.
149. Maetens, M., et al., *Distinct roles of Mdm2 and Mdm4 in red cell production*. Blood, 2007. **109**(6): p. 2630-2633.
150. Schechter, A.N., *Hemoglobin research and the origins of molecular medicine*. Blood, 2008. **112**(10): p. 3927-38.
151. Miller, I.J. and J.J. Bieker, *A novel, erythroid cell-specific murine transcription factor that binds the CACCC element and is related to the Kruppel family of nuclear proteins*. Molecular and Cellular Biology, 1993. **13**(5): p. 2776-2786.

152. Philipsen, S. and G. Suske, *A tale of three fingers: the family of mammalian Sp/XKLF transcription factors*. Nucleic Acids Research, 1999. **27**(15): p. 2991-3000.
153. Chen, X. and J.J. Bieker, *Erythroid Krueppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions*, in *EMBO J*. 1996. p. 5888-5896.
154. Pandya, K.R. and T.M. Townes, *Multiple transactivation domains of Erythroid Kruppel-like Factor (EKLF)*. Blood, 2000. **96**(11): p. 1008.
155. Anderson, K.P., S.C. Crable, and J.B. Lingrel, *Regulation of the EKLF gene requires an E-box binding motif and multiple GATA-1 sites*. Blood, 1997. **90**(10): p. 1357-1357.
156. Bieker, J.J. and C.M. Southwood, *The erythroid kruppel-like factor transactivation domain is a critical component for cell-specific inducibility of a beta-globin promoter*. Molecular and Cellular Biology, 1995. **15**(2): p. 852-860.
157. Pandya, K., D. Donze, and T.M. Townes, *Novel transactivation domain in erythroid Kruppel-like Factor (EKLF)*. Journal of Biological Chemistry, 2001. **276**(11): p. 8239-8243.
158. Bieker, J.J., *Probing the onset and regulation of erythroid cell-specific gene expression*. Mount Sinai Journal of Medicine, 2005. **72**(5): p. 333-338.
159. van den Wijngaard, A., et al., *EKLF levels and globin gene switching: Importance of gene order and distance to the LCR*. Blood Cells Molecules and Diseases, 2000. **26**(5): p. 56.
160. Perkins, A.C., A.H. Sharpe, and S.H. Orkin, *Lethal beta-thalassemia in mice lacking the erythroid CACCC-transcription factor EKLF*. Nature, 1995. **375**(6529): p. 318-322.
161. Shyu, Y.C., et al., *Subcellular transport of EKLF and switch-on of murine adult beta(maj) globin gene transcription*. Molecular and Cellular Biology, 2007. **27**(6): p. 2309-2323.
162. Perkins, A.C., K.M.L. Gaensler, and S.H. Orkin, *The transcription factor, EKLF, is required for completion of the developmental switch from human gamma- to beta-globin gene expression*. Blood, 1995. **86**(10): p. 977-977.
163. Sengupta, T., et al., *Acetylation of EKLF is essential for epigenetic modification and transcriptional activation of the beta-globin locus*. Molecular and Cellular Biology, 2008. **28**(20): p. 6160-6170.
164. Zhang, W., et al., *Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex*. Mol Cell Biol, 2001. **21**(7): p. 2413-22.
165. Armstrong, J.A., J.J. Bieker, and B.M. Emerson, *A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro*. Cell, 1998. **95**(1): p. 93-104.

166. Brown, R.C., et al., *Interaction of hBRG1 with the erythroid transcription factor EKLf is essential, but not sufficient for maximal activation of the beta-globin promoter*. Blood, 2000. **96**(11): p. 1002.
167. Kadam, S. and B.M. Emerson, *Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes*. Mol. Cell, 2003. **11**: p. 377-389.
168. Kim, S.I., et al., *BRG1 requirement for long-range interaction of a locus control region with a downstream promoter*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(7): p. 2259-2264.
169. Bottardi, S., et al., *Lineage-specific activators affect beta-globin locus chromatin in multipotent hematopoietic progenitors*. Embo Journal, 2006. **25**(15): p. 3586-3595.
170. Bouilloux, F., et al., *EKLf restricts megakaryocytic differentiation at the benefit of erythrocytic differentiation*. Blood, 2008. **112**(3): p. 576-584.
171. Starck, J., et al., *Functional cross-antagonism between transcription factors FLI-1 and EKLf*. Molecular and Cellular Biology, 2003. **23**(4): p. 1390-1402.
172. Sengupta, T., et al., *Distinct modes of gene regulation by a cell-specific transcriptional activator*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4213-8.
173. Chen, X., M. Reitman, and J.J. Bieker, *Chromatin structure and transcriptional control elements of the erythroid Kruppel-like factor (EKLf) gene*. J. Biol. Chem., 1998. **273**: p. 25031-25040.
174. Ouyang, L., X. Chen, and J.J. Bieker, *Regulation of erythroid Kruppel-like factor (EKLf) transcriptional activity by phosphorylation of a protein kinase casein kinase II site within its interaction domain*. J Biol Chem, 1998. **273**(36): p. 23019-25.
175. Quadrini, K.J. and J.J. Bieker, *EKLf/KLF1 is ubiquitinated in vivo and its stability is regulated by activation domain sequences through the 26S proteasome*. Febs Letters, 2006. **580**(9): p. 2285-2293.
176. Siatecka, M., L. Xue, and J.J. Bieker, *Sumoylation of EKLf promotes transcriptional repression and is involved in inhibition of megakaryopoiesis*. Molecular and Cellular Biology, 2007. **27**(24): p. 8547-8560.
177. Freson, K., et al., *Platelet characteristics in patients with X-linked macrothrombocytopenia because of a novel GATA1 mutation*. Blood, 2001. **98**(1): p. 85-92.
178. Mehaffey, M.G., et al., *X-linked thrombocytopenia caused by a novel mutation of GATA-1*. Blood, 2001. **98**(9): p. 2681-8.
179. Freson, K., et al., *Different substitutions at residue D218 of the X-linked transcription factor GATA1 lead to altered clinical severity of macrothrombocytopenia and anemia and are associated with variable skewed X inactivation*. Hum Mol Genet, 2002. **11**(2): p. 147-52.

180. Wechsler, J., et al., *Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome*. Nat Genet, 2002. **32**(1): p. 148-52.
181. Yu, C., et al., *X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction*. Blood, 2002. **100**(6): p. 2040-5.
182. Zhang, S.J., J.Y. Shi, and J.Y. Li, *GATA-2 L359 V mutation is exclusively associated with CML progression but not other hematological malignancies and GATA-2 P250A is a novel single nucleotide polymorphism*. Leuk Res, 2009. **33**(8): p. 1141-3.
183. Liew, C.K., et al., *Zinc fingers as protein recognition motifs: structural basis for the GATA-1/friend of GATA interaction*. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 583-8.
184. Newton, A., J. Mackay, and M. Crossley, *The N-terminal zinc finger of the erythroid transcription factor GATA-1 binds GATC motifs in DNA*. J. Biol. Chem., 2001. **276**: p. 35794-35801.
185. Mundschau, G., et al., *Mutagenesis of GATA1 is an initiating event in Down syndrome leukemogenesis*. Blood, 2003. **101**(11): p. 4298-300.
186. Gurbuxani, S., P. Vyas, and J.D. Crispino, *Recent insights into the mechanisms of myeloid leukemogenesis in Down syndrome*. Blood, 2004. **103**(2): p. 399-406.
187. Taub, J.W., et al., *Prenatal origin of GATA1 mutations may be an initiating step in the development of megakaryocytic leukemia in Down syndrome*. Blood, 2004. **104**(5): p. 1588-9.
188. Steidl, U., et al., *A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia*. J Clin Invest, 2007. **117**(9): p. 2611-20.
189. Moi, P., et al., *A novel silent beta-thalassemia mutation in the distal CACCC box affects the binding and responsiveness to EKLF*. Br J Haematol, 2004. **126**(6): p. 881-4.
190. Gonzalez-Redondo, J.M., et al., *A C----T substitution at nt--101 in a conserved DNA sequence of the promotor region of the beta-globin gene is associated with "silent" beta-thalassemia*. Blood, 1989. **73**(6): p. 1705-11.
191. Pierce, M.M., C.S. Raman, and B.T. Nall, *Isothermal titration calorimetry of protein-protein interactions*. Methods, 1999. **19**(2): p. 213-21.
192. Nerlov, C. and T. Graf, *PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors*. Genes & Dev., 1998. **12**: p. 2403-2412.
193. Merika, M. and S.H. Orkin, *Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF*. Mol Cell Biol, 1995. **15**(5): p. 2437-47.
194. Hunter, T. and M. Karin, *The regulation of transcription by phosphorylation*. Cell, 1992. **70**(3): p. 375-87.

195. Bode, A.M. and Z. Dong, *Post-translational modification of p53 in tumorigenesis*. Nature Rev. Cancer, 2004. **4**(10): p. 793-805.
196. Li, H.-H., et al., *Phosphorylation on Thr-55 by TAF1 mediates degradation of p53: a role for TAF1 in cell G1 progression*. Molecular Cell, 2004. **13**(6): p. 867-878.
197. Vendel, A.C. and K.J. Lumb, *Molecular recognition of the human coactivator CBP by the HIV-1 transcriptional activator Tat*. Biochemistry, 2003. **42**(4): p. 910-6.
198. Vendel, A.C. and K.J. Lumb, *NMR mapping of the HIV-1 Tat interaction surface of the KIX domain of the human coactivator CBP*. Biochemistry, 2004. **43**(4): p. 904-8.
199. Arany, Z., et al., *A family of transcriptional adaptor proteins targeted by the E1A oncoprotein*. Nature, 1995. **374**(6517): p. 81-4.
200. Kitabayashi, I., et al., *Fusion of MOZ and p300 histone acetyltransferases in acute monocytic leukemia with a t(8;22)(p11;q13) chromosome translocation*. Leukemia, 2001. **15**(1): p. 89-94.
201. Ida, K., et al., *Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13)*. Blood, 1997. **90**(12): p. 4699-704.
202. Muraoka, M., et al., *p300 gene alterations in colorectal and gastric carcinomas*. Oncogene, 1996. **12**(7): p. 1565-9.
203. Gayther, S.A., et al., *Mutations truncating the EP300 acetylase in human cancers*. Nat Genet, 2000. **24**(3): p. 300-3.
204. Bandyopadhyay, D., et al., *Down-regulation of p300/CBP histone acetyltransferase activates a senescence checkpoint in human melanocytes*. Cancer Res, 2002. **62**(21): p. 6231-9.