INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Protein-Nucleic Acid Interactions of Wilms' Tumor and TFIIIA Zinc Finger Proteins.

by

TATYANA HAMILTON
B. Sc., Novosibirsk State University, Russia, 1990

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY
in the Department of Biochemistry and Microbiology

We accept this thesis as conforming
to the required standard

Dr. Paul J. Romaniuk, Supervisor (Dept. of Biochemistry & Microbiology)

Dr. Robert W. Olafson, Departmental Member
(Dept. of Biochemistry & Microbiology)

Dr. Francis E. Nano, Departmental Member
(Dept. of Biochemistry & Microbiology)

Dr. Michael J. Ashwood-Smith,
Outside Member (Dept. of Biology)

Dr. Terry Pearson, Departmental Member
(Dept. of Biochemistry & Microbiology)

Dr. David Setzer, External Examiner
Case Western Reserve University)

Supervisor: Dr. Paul J. Romaniuk

Abstract

This Ph.D. work represents the study of nucleic acid interactions of two zinc finger proteins: mammalian Wilms' tumor suppressor (WT1) and *Xenopus* transcription factor IIIA proteins (TFIIIA).

WT1 is a putative transcriptional regulatory protein which is inactivated in a subtype of Wilms' tumours. Using selection and amplification binding assay (SAAB) we determined that the highest affinity binding sites for WT1[-KTS] consist of a 12 base pair sequence GCG-TGG-GCG-**(T/G)(G/A/T)(T/G)**. These sequences have a four-fold higher affinity for the protein than the nonselected sequence GCG-TGG-GCG-**CCC**, as measured by a quantitative nitrocellulose filter binding assay.

The effects of Denys-Drash syndrome (DDS) point mutations on the DNA binding activity of WT1 were determined. SAAB assay revealed that none of the DDS mutant proteins give rise to a new sequence specificity. One mutation, R394W abolishes specific binding of the protein. The remaining mutations result in reduced DNA-binding activity, ranging from 1.4 to 14-fold, which suggests that even small changes in DNA-binding activity may precipitate the clinical phenotype of Denys-Drash syndrome.

Comparative analysis of the DNA binding characteristics of Wilms' tumour and Early growth response proteins was conducted. The stoichiometry of the DNA-protein complexes, their stability to dissociation, and the effects of pH, temperature and salt concentration on the equilibrium binding of these proteins to their cognate DNA

sequences have been determined. Under the conditions of 0.1 M salt, pH 7.5, and 22 ° C WT1-ZF has an apparent dissociation constant (K_d) of 1.14± 0.2 x 10^{-09} M, and EGR-1 protein has a K_d of 3.55 ± 0.4 x 10^{-09} M. In addition, we tested relative contribution of each base pair in the consensus binding site to the high affinity binding by point mutational analysis, and identified important differences that exist in the binding modes of the two proteins.

Transcription factor IIIA controls the expression of the 5S ribosomal RNA genes during development of *Xenopus laevis.*, and specifically interacts with both 5S DNA and 5S rRNA molecules. The present study assesses contributions of the central zinc fingers four through seven to specific DNA and RNA binding activities of the protein. The results demonstrate that each zinc finger in the zf 4-7 region contributes to both the high affinity DNA and RNA interactions: the largest effect on TFIIIA-DNA binding (10-fold) was produced when zinc finger 5 of TFIIIA was replaced with the donor sequences of either p43 or WT1. However, while all the zinc fingers 4-7 contribute to the high affinity 5S rRNA binding, substitution of an α-helical portion of zinc finger 6 with the equivalent sequences from WT1 abolished RNA-binding activity of TFIIIA, suggesting that zinc finger 6 plays a particularly important role in binding to RNA.

Examiners:

Dr. Paul J. Romaniuk, Supervisor (Dept. of Biochemistry & Microbiology)

Dr. Robert W. Olafson, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Francis E. Nano, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Terry W. Pearson, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Michael J. Ashwood-Smith, Outside Member (Dept. of Biology)

Dr. David Setzer, External Examiner (Case Western Reserve University)

Table of Contents

List of Tables

List of Figures

List of Abbreviations

bp: base pair

BSA: bovine serum albumin

BWS: Beckwith-Wiedemann syndrome

CAT: ChloramphenicolAcetylTransferase

cDNA: complementary deoxyribonucleic acid

cpm: counts per minute

DDS: Denys-Drash syndrome

deoxynucleotide triphosphates:

dATP, deoxyadenosine triphosphate

dCTP, deoxycytidine triphosphate

dGTP, deoxyguanine triphosphate

dTTP, deoxythymidine triphosphate

DEPC: diethy pyrocarbonate

DNA: deoxyribonucleic acid

DTT: dithiotreitol

E.coli: *Escherichia coli*

EDTA: ethylenediamine-tetraacetic acid

EGR1: Early Growth Response 1

F.U.P.: Forward Universal Primer

HEPES: *N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulphonic acid

ICR: internal control region

IE: intermediate element

IGFII: Insulin-like Growth Factor II

IPTG: isopropyl-β-D-thiogalactopyranoside

KTS: lysine, threonine, serine

LB: Luria-Benton broth

LOH: Loss Of Heterozygosity

mRNA: messenger ribonucleic acid

MW: molecular weight

NMR: nuclear magnetic resonance

nt: nucleotide

Nucleotide triphosphates:

GTP, guanine triphosphate

CTP, cytidine triphosphate

ATP, adenosine triphosphate

UTP, uridine triphosphate

nucleotide bases:

G, guanine

C, cytosine

A, adenine

T, thymidine

N, either A, A, G, or T

PAGE: polyacrylamide gel electrophoresis

PDGF-A: Platelet-Derived Growth Factor A

PEG: polyethylene glycol

PMSF: phenylmethylsulfonyl fluoride

PPO: 2,5-diphenyloxazole

R.U.P.: Reverse Universal Primer

RB: retinoblastoma

RNA: ribonucleic acid

RNase: ribonuclease

RNasin: ribonuclease inhibitor

RNP: ribonucleoprotein

rRNA: ribosomal nucleic acid

S: Svedberg unit

SAAB: selected amplification and binding

SAAB: Selection And Amplification Binding assay

SDS: sodium dodecyl sulphate

TBE: Tris base, borate, EDTA

TFIIIA: transcription factor IIIA

TFIIIB: transcription factor IIIB

TFIIIC: transcription factor IIIC

Tris-HCl: tris-(hydroxymethyl) aminomethane hydrochloride

tRNA: transfer ribonucleic acid

TTK: TramTracK protein

WAGR: Wilms' tumour, Aniridia, uroGenital malformations, mental Retardation

WT1: Wilms' Tumour 1

Xbo: *Xenopus borealis* oocyte

Xbs: *Xenopus borealis* somatic

Xlo: *Xenopus laevis* oocyte

Xls: *Xenopus laevis* somatic

Xlt: *Xenopus laevis* trace-oocyte

Acknowledgments

I would like to thank my supervisor, Dr. Paul J. Romaniuk for giving me the wonderful opportunity to join his lab, and for his continuous support, encouragement and guidance over the years.

Valuable assistance was provided by all the members of my supervisory committee: Dr. Robert W. Olafson, Dr. Francis E. Nano, Dr. Terry W. Pearson and Dr. Michael J. Ashwood-Smith, for which I am grateful. I would also like to mention Dr. Al T. Matheson, Dr. Ed E. Ishiguro and Dr. Juan Ausio for their assistance.

My sincere thanks to all the members of Dr. Romaniuk's lab - Kathy Barilla, Frank Borel, John Ferris, Steve Hendy, Maya Iskandar, Colleen Nelson, Nik Veldhoen, Judy Wise, Qimin You, and Wei-Qing Zang - for their help and friendship. My thanks also to the other graduate students at the department who didn't hesitate to help: Gerry Bearon, Siobhan Cowley, Ben Forward, LeeAnn Howe, Armando Jardim, Kizzy Mdluli, Dmitrii Rodionov, Caroline Stebeck.

I am grateful to Albert Labossiere and Scott Scholz, who were always there to help with all manner of technical assistance from fixing a computer to setting up a photo room.

I would very much like to acknowledge the help of the office: Rosanne Poulson, Maree Roome, and Claire Tugwell.

I would also like to thank Marion and Arthur Fontaine for connecting me to Dr. Paul Romaniuk, for their infinite kindness, and for being like second parents to me.

Finally, I would like to thank Arthur Hamilton, my husband and best friend, for his love, patience and support. Without him my life would never be complete.

Dedication

This work is dedicated to my mother, Lina Kletskova.

CHAPTER 1.0 EARLY GROWTH RESPONSE PROTEIN 1 (EGR-1): A PROTOTYPICAL MEMBER OF A C_2H_2 ZINC FINGER FAMILY OF TRANSCRIPTION FACTORS

1.1 Introduction

1.1.1 Overview of the EGR-1 family of transcription factors

The proliferation and differentiation of eukaryotic cells is influenced by a multitude of stimuli, including growth factors, adhesion molecules and other extracellular ligands. These complex long-term cellular responses are mediated by changes in gene expression, and are coupled to the initial signal transduction events occurring at the level of the plasma membrane. The immediate-early genes are the earliest downstream nuclear targets for these events. The activation of these genes is generally very rapid, transient, and independent of *de novo* protein synthesis. A subclass of these genes encodes transcription factors, which form the first step in initiation of genetic programs that would lead to an appropriate cellular response.

The best characterized members of this group of immediate-early genes include *c-jun*, *c-fos*, and *Egr-1*. Each of these genes, in turn, represents a prototype for a family of closely related proteins. Numerous studies have demonstrated that EGR-1 induction is universal. Various stimuli that induce EGR-1 expression, as well as diverse cell types in which EGR-1 expression has been described, are summarized in Figure 1.1. Extracellular stimuli that induce EGR-1 can be grouped into the following categories: (1) mitogens; (2) developmental or differentiation cues; (3) tissue or radiation injury; (4) signals that cause neuronal excitation. In practically every cell type examined, EGR-1 expression is rapidly induced by such mitogens as hormones, growth factors, and the tumor promoter TPA (phorbol) (Sukhatme et al., 1987; Sukhatme et al., 1988; Lau and Nathans, 1987; Lemaire et al., 1988). Various

Response	Stimulus	Cell type
Mitogenic	serum	fibroblasts
	PDGF, EGF, FGF	fibroblasts
	insulin	hepatocytes
	phytohemagglutinin	peripheral blood lymphocytes
	anti-μ	B lymphocytes
	adenosine diphosphate	kidney epithelial cells
	PDGF, vasopressin	kidney mesangial cells
	serum	skeletal muscle Sol8 cells
	bFGF, PDGF BB	skeletal muscle Sol8 cells
	partial hepatectomy	liver
	GM-CSF, LPS	peritoneal macrophages
	endothelin	astrocytes
	angiotensin II	vascular smooth muscle cells
Hypertrophic	endothelin, angiotensin II	myocyte
Differentiative	NGF	pheochromocytoma PC12 (neural)
	retinoic acid, DMSO	embryonal carcinoma P19
	retinoic acid	embryonal calvarial cells, RCT-1 (osteoblast)
	TPA, DMSO	myeloid leukemia HL60 and U-937
Tissue/ radiation injury	ischemia	kidney
	ionizing radiation	293, SQ-20B
Neuronal excitation	potassium ions	PC12 (depolarization)
	metrazole	seizures in vivo
	NMDA	hippocampus
	visual stimuli	visual cortex
	electroconvulsive shock therapy, dopamine receptor activation, opiate withdrawal	CNS
	peripheral nervous system	sciatic nerve transection
Other	urea	MDCK, LLC-PK$_1$ renal epithelial cells

Figure 1.1 Biological processes mediated by EGR-1 in response to diverse stimuli (Gashler & Sukhatme, 1995).

differentiative stimuli such as retinoic acid, DMSO (dimethyl sulfoxide), and phorbol give rise to EGR-1 expression, which can be correlated with differentiative processes of the kidney, spleen, brain, and most other tissues (Sukhatme et al., 1988; Lemaire et al., 1988; Christy et al., 1988). In the context of tissue injury such as ischemia or ionizing radiation, EGR-1 has been shown to be strikingly induced (Bonventre et al., 1991; Hallahan et al., 1991). It is likely that EGR-1 is not induced by the injury itself, but rather acts in response to post-injury events, mediating subsequent processes of cellular proliferation and differentiation. Numerous lines of evidence exist indicating an important role of EGR-1 in neuronal signaling. For example, EGR-1 levels increase dramatically in the brain following seizure activity (Sukhatme et al., 1988). A dramatic increase in EGR-1 levels is observed following electroconvulsive shock therapy, dopamine receptor activation, and opiate withdrawal (Bhat et al., 1992). Finally, the expression of EGR-1 in developing and adult brain emphasizes the importance of EGR-1 in neurophysiological processes (Watson and Milbrandt, 1990). Thus, EGR-1 mediates responses of enormous complexity. It acts in different cellular contexts and is able to respond to a multitude of extracellular signals.

All members of the EGR family share highly similar, C_2H_2 zinc finger DNA-binding motifs. At present, the closest members of the EGR family of transcriptional regulators are: EGR-2/Krox20 (Chavrier et al., 1988; Joseph et al., 1988), EGR-3 (Patwardhan et al., 1991), and EGR-4/NGFI-C/pAT133 (Patwardhan et al., 1991; Crosby et al., 1991; Muller et al., 1991). All of the above proteins have zinc finger domains that are virtually identical to that of EGR-1 (Figure 1.2). The EGR-1 zinc-finger domain is over 95% identical to that of EGR-2 (Joseph et al., 1988) and 91% identical to that of EGR-3 at the amino acid level (Patwardhan et al., 1991). The residues important for specific

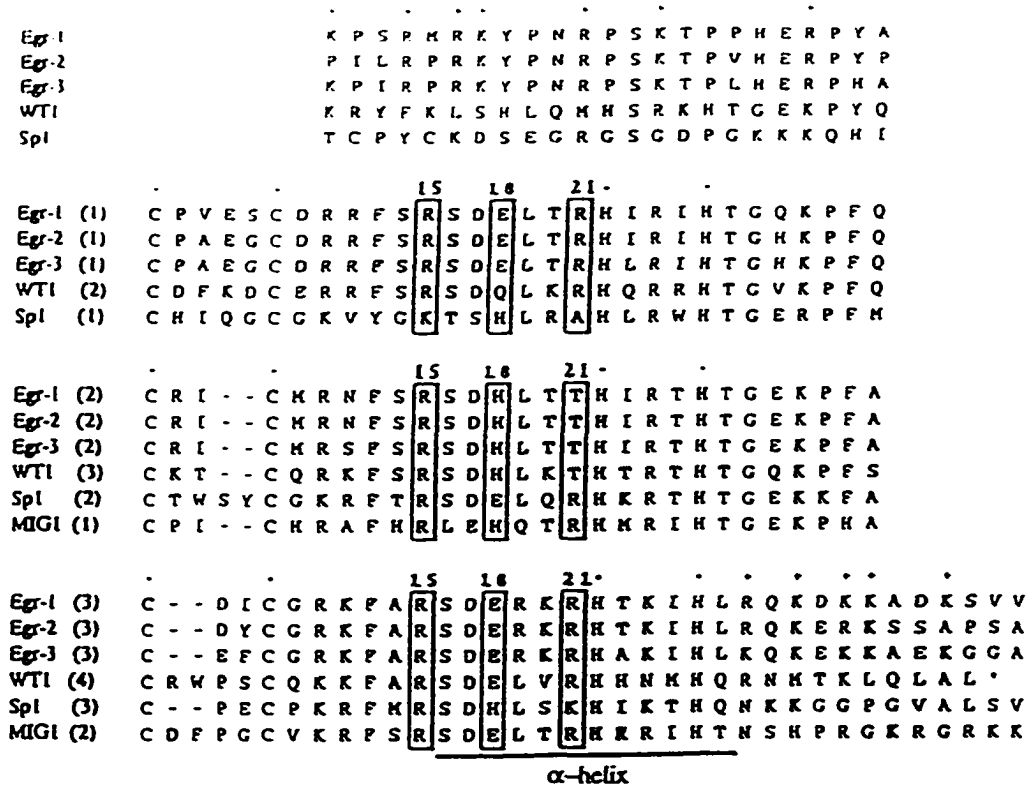


Figure 1.2 Comparison of DNA-binding domains of the EGR-1 family of proteins. Zinc fingers and adjacent sequences of the EGR-1 family of proteins are aligned. Conserved cysteine and histidine residues are marked (•); the α-helical region is underlined; conserved residues important for determining binding specificity are enclosed; basic residues are denoted (+) (Gashler & Sukhatme, 1995).

DNA recognition are conserved, and most of the changes are conservative substitutions (Figure 1.2). The homology between these proteins extends to short stretches of adjacent basic sequences but sharply drops outside this region. This suggests that these proteins may recognize the same DNA target sequences through their zinc fingers, but that interactions with other transcriptional regulatory proteins are distinct for each family member. This multigene family offers a great system to study the relationship between signal transduction and gene expression in both normal and transformed cells.

The Wilms' tumor gene product has four zinc fingers, three of which possess related but less homologous zinc finger DNA-binding domains (approximately 65% identity to the EGR-1 zinc finger domain) (Gessler et al., 1990; Call et al., 1990). The mammalian ubiquitous transcriptional activator Sp1 has three related zinc fingers (Kadonaga et al., 1987). Sp1 finger 2 is most similar to EGR-1 fingers 1 and 3 (Kadonaga et al., 1987). Another distant member of the EGR family of proteins is a yeast protein MIG1 involved in responses to glucose repression (Nehlin and Ronne, 1990). It contains two zinc fingers that are most similar to the second and third fingers of EGR-1 protein, with 60% identical residues (Nehlin and Ronne, 1990). Outside the zinc fingers, MIG-1 has no obvious similarity to other proteins. This absence of sequence conservation outside the zinc finger motifs is a common finding among C_2H_2 zinc finger proteins. Therefore, a comparison of the family members must rely on an alignment of finger sequences. Studies of the immediate-early proteins and downstream promoter elements they target will continue to enhance our knowledge of protein-DNA interactions and general mechanisms of transcriptional activation and repression.

1.1.2 Gene targets of EGR-1 regulation

Consistent with its role in cellular proliferation and differentiation, EGR-1 binds and regulates genes involved in these functions. Some discussion follows on the best documented instances that involve regulation by EGR-1.

Platelet-derived growth factor A (PDGF-A) is a potent mitogen which is present at elevated levels in response to growth factors or cytokines (Wang and Deuel, 1992). An EGR-1 binding site has been defined in the promoter region of the PDGF-A gene (Wang and Deuel, 1992). The DNA fragment bound by EGR-1 has the following sequence: GAG-GAG-GAG-GAGGA. Although this site deviates from the consensus EGR-1 binding sequence, which is GCG-G/TGG-GGG, it competes equally well in the gel-shift competition experiments. Similar motifs have been identified in promoters of other growth-related genes, such as the insulin receptor, the tumor growth factor β, the epidermal growth factor receptor, c-myc and c-Ki-ras (Wang and Deuel, 1992).

Another gene whose expression is regulated by EGR-1 is thymidine kinase (tk), an important player in DNA biosynthesis. Using monoclonal anti - EGR-1 antibodies, it was shown that EGR-1 was one of the components of the tk promoter complex obtained from serum-stimulated nuclear extract (Molnar et al., 1994). Further transient transfection experiments demonstrated that EGR-1 activates a reporter driven by a tk promoter element (Molnar et al., 1994).

EGR-1 has been shown to be required for differentiation of myeloblasts along the macrophage lineage (Nguen et al., 1993). The use of EGR-1 antisense oligomers in the cell culture medium resulted in blocked macrophage differentiation in normal myeloblasts as well as myeloid

leukemia cells. Thus, the authors clearly demonstrated that expression of EGR-1 is essential for macrophage differentiation.

The myosin heavy chain α gene (α-MHC) was shown to be expressed in concordance with EGR-1 in stimulated cultures of cardiac myocytes (Gupta et al., 1991). Transient transfection assays using a CAT reporter containing α-myosin heavy chain promoter showed 10-fold induction of the promoter activity by an EGR-1 expression vector (Gupta et al., 1991). The region of the rat α-MHC promoter that is responsive to EGR-1 has been defined, and shown to contain a potential EGR-1 binding site (Gupta et al., 1991). The α-MHC induction appears to be tissue-specific since it was observed in the myogenic Sol8 cell line, but not in NIH3T3 fibroblasts (Gupta et al., 1991).

EGR-1 is expressed at high level in the rat adrenal gland and in PC12 cells as demonstrated by Elbert et al., (1994). A proposed function for EGR-1 in adrenergic differentiation may be via regulation of phenylethanolamine *N*-methyltransferase (PNMT) (PNMT is an adrenal enzyme that converts norepinephrine to epinephrine) (Elbert et al., 1994). The analysis of the PNMT promoter revealed that it contains two potential EGR-1 binding sites, one of them differing only by one nucleotide from the EGR-1 consensus DNA binding sequence. Transient transfections showed that EGR-1 can stimulate a PNMT reporter by fourfold (Elbert et al., 1994).

Despite the abundance of data on EGR-1 induction by mitogenic signals, additional EGR-1 gene targets await elucidation. Finally, *in vitro* studies of the EGR-1 involvement in cellular proliferation and differentiation need to be correlated with the *in vivo* role of EGR-1.

1.1.3 Identification of EGR-1 cDNA, and characterization of its protein product

Using differential screening techniques, several groups set out to identify novel genes which have a low level of expression in nondividing cells but which are rapidly up-regulated in cells stimulated by mitogen. The following criteria were used to isolate such novel genes: (1) transcripts should be rapidly and transiently induced by serum stimulation of quiescent cells; (2) these genes should be induced without intervening protein synthesis, i.e. the induction should not be affected by inhibitors of protein synthesis, such as cycloheximide; (3) expression should be induced by a wide spectrum of mitogens, such as growth factors, hormones, and other ligands; (4) expression should be induced in a broad array of cell types; and (5) the genes should be highly conserved in evolution.

The novel immediate-early gene has been cloned by a number of research groups. Sukhatme et al. (1987, 1988) identified a gene designated *Egr-1* . The authors used a differential screening technique to screen a library from BALB/c 3T3 cells stimulated with serum in the presence of cycloheximide. Clones which preferentially hybridized to cDNA from serum and cycloheximide-treated fibroblasts were identified by comparing them to cDNA from quiescent cells. A 3.4-kb transcript appeared upon mitogenic stimulation of a variety of cell types, and was designated *Egr-1*. Using a similar differential screening strategy, Milbrandt (1987) independently isolated a transcript named *NGFI-A*, which was activated in rat pheochromocytoma PC12 cells by nerve growth factor, and is a rat analog of the mouse *Egr-1*. The same gene has been independently cloned by other groups, using a similar approach: *zif268* was cloned from serum-stimulated 3T3 fibroblasts from BALB/c mice (Lau and Nathans, 1987); *tis8* was identified as a phorbol-inducible gene in 3T3 cells (Lim et al., 1987); the

chicken homolog, *cef5*, was cloned as a *v-src* - inducible gene from chicken embryo fibroblasts (Simmons et al., 1989); gene 225 was identified as a T-cell-activated transcript (Wright et al., 1990), and *Krox24* was isolated from serum-stimulated 3T3 cells by hybridization to a highly conserved domain of the *Drosophila* factor Kruppel (Lemaire et al., 1988).

Sequence analysis of EGR-1 revealed that the protein contains three tandemly repeated zinc finger motifs of C_2H_2 type that govern the DNA-binding function of this protein (Sukhatme et al., 1988; Lau and Nathans, 1987; Lemaire et al., 1988; Christy et al., 1988) (Figure 1.3A). The coding sequence of the protein is highly conserved across species, indicating the importance of the *Egr-1* gene product. Complementary DNAs from human (Suggs et al., 1990), rat (Milbrandt, 1987), mouse (Sukhatme et al., 1988; Lemaire et al., 1988; Christy et al., 1988), chicken (Simmons et al., 1989), and zebrafish (Drummond et al., 1994) are highly similar. The product of the *Egr-1* gene is a protein of 80-82 kDa (Cao et al., 1990). Cell fractionation studies and immunocytochemistry demonstrated that EGR-1 is localized to the nucleus, which is consistent with its putative DNA-binding function (Cao et al., 1990; Day et al., 1990). Further examination of the amino acid sequence of the EGR-1 protein showed it to contain basic residues clustered in the zinc fingers and the adjacent sequence (Figure 1.3B). The amino-terminal amino acids are rich in proline and serine-threonine residues, organized in stretches of tree to five consecutive amino acids. There is one series of seven consecutive serine-threonine residues, which is followed by seven glycines (Figure 1.3B). The carboxyl terminus of the EGR-1 protein is also rich in proline and serine/threonine residues. The regions of the protein containing proline amino acids are predicted to lack α-helical secondary structure, whereas the high content of serine and threonine residues indicates that

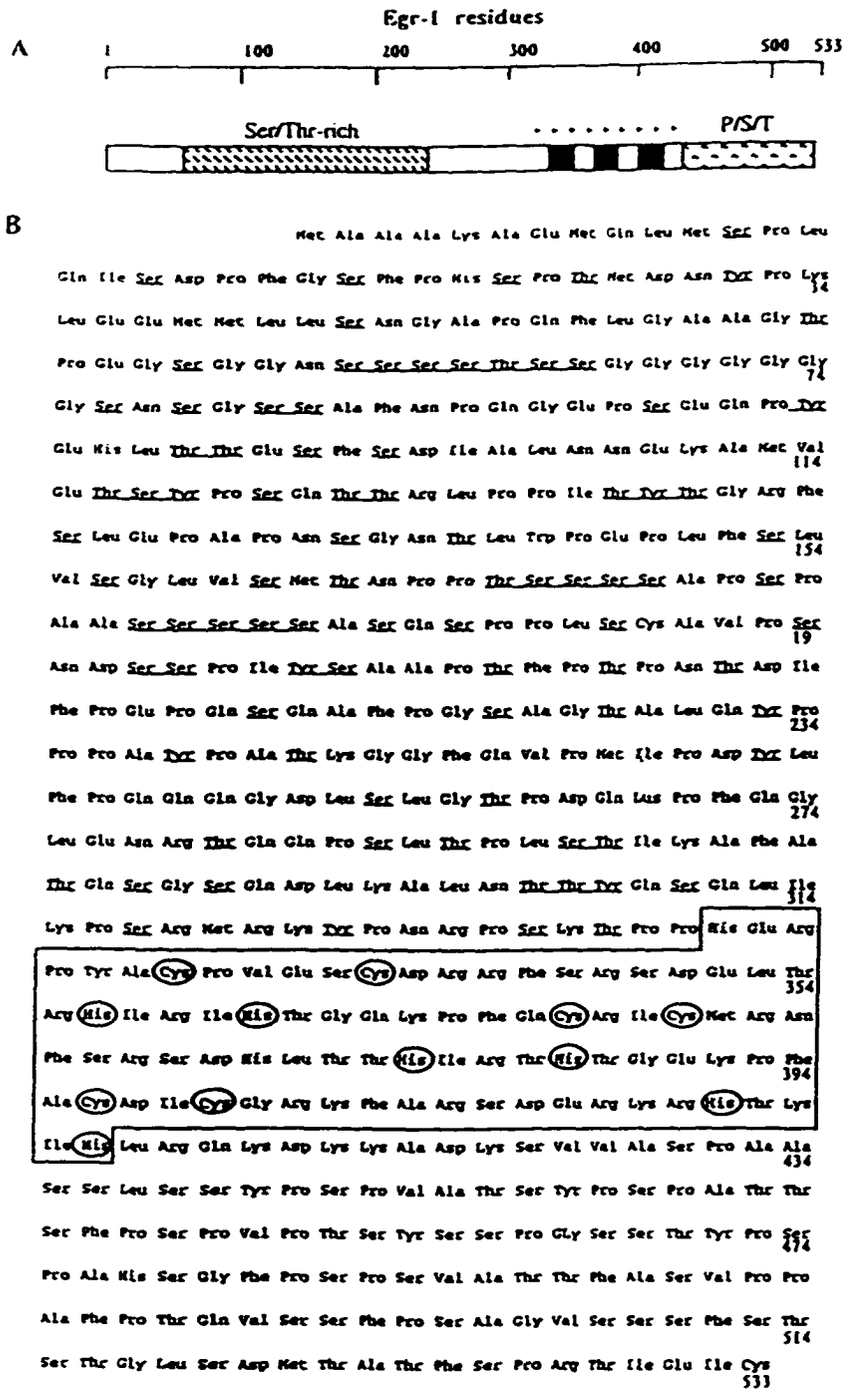


Figure 1.3 Schematic structure and amino acid sequence of EGR-1 protein.

Figure 1.3 (**A**) A diagram of the functional domains of EGR-1 protein: the serine-threonine N-terminal domain is shown in the hatched box on the left; black bars represent zinc finger motifs; the C-terminal proline-serine-threonine domain is designated by a hatched box on the right. (**B**) Coding sequence of murine EGR-1. The three zinc finger motifs are enclosed in a box; zinc-coordinating cysteine and histidine residues are circled; serine, threonine and tyrosine residues in the N-terminal portion of the sequence are underlined (Gashler & Sukhatme, 1995).

EGR-1 may be subject to phosphorylation. Indeed, alkaline phosphatase was shown to convert slow migrating form of EGR-1 to the faster migrating form as seen on SDS-PAGE (Cao et al., 1990; Day et al., 1990).

Further structure-function analysis of EGR-1 protein defined its activation, repression, DNA binding and nuclear localization domains (Gashler et al., 1993), summarized in a diagram in Figure 1.4. Thus, the organization of EGR-1 is modular in nature, which is characteristic of many classical transcription factors. A potent activational domain was mapped to the serine- and threonine-rich amino terminal domain, using deletion analysis of EGR-1 (Gashler et al., 1993). These EGR-1 activation sequences were shown to be independent domains by placing them in the context of other proteins. When residues 3-281, or 3-138, or 138-281 were fused to the DNA-binding domain of the yeast factor GAL4, they activated transcription 100-fold as GAL4 fusions (Gashler et al., 1993). These findings were confirmed by studying NGFI-A protein, the rat homolog of EGR-1 (Russo et al., 1993). A weak transactivation domain was mapped by several laboratories to the C-terminus of EGR-1: amino acids 420-533 (Russo et al., 1993; Gashler et al., 1993).

A repression domain of EGR-1 was identified when a small internal deletion 5' to the zinc-finger domain (amino acids 284-330) resulted in almost 5-fold increased transactivation in HeLa cells (Gashler et al., 1993). This was an unexpected finding. Further experiments have shown that this repression domain is also modular in nature, since it could be fused to the DNA-binding domain of GAL4, and represses transcription 7- to 10-fold (Gashler et al., 1993). This serine/threonine-rich domain is highly conserved in evolution (Drummond et al., 1994), and is distinct from repression domains of other transcriptional regulators, such as the alanine- and glycine-rich repression

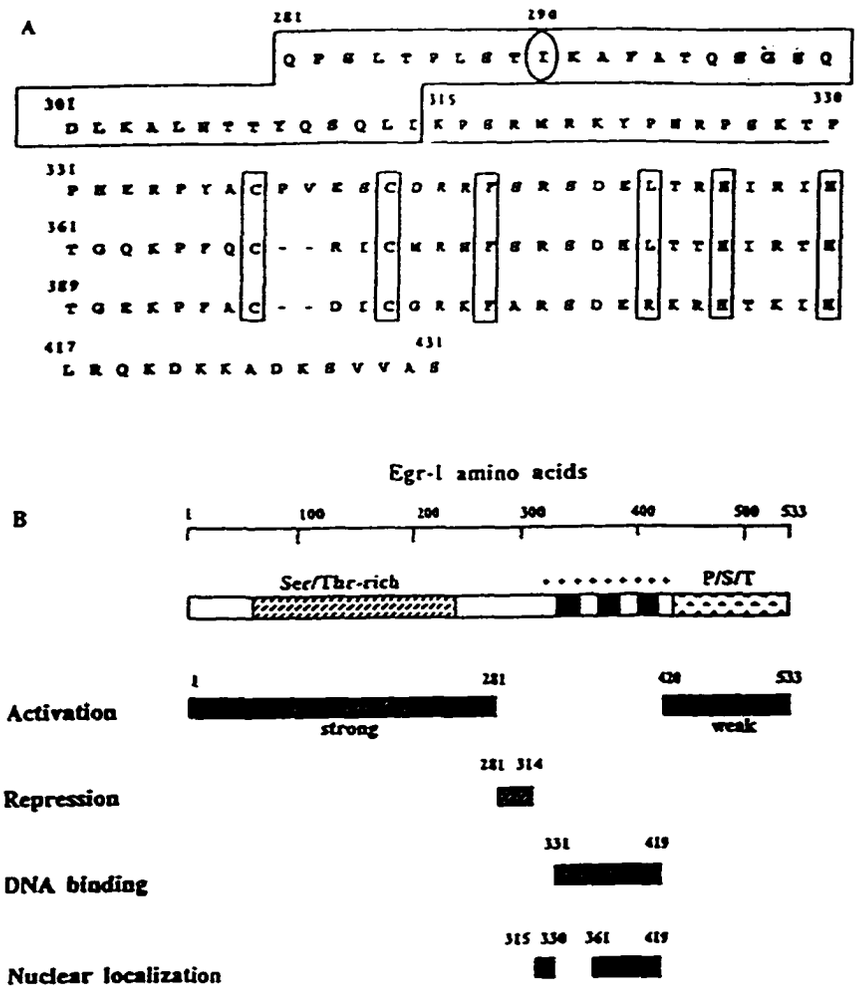


Figure 1.4 Summary of EGR-1 functional domains.

(A) Amino acid sequence of EGR-1 repression domain and zinc fingers. (B) Schematic map of EGR-1 domains (Gashler & Sukhatme, 1995).

domain in Kruppel (Licht et al., 1990), hydrophobic- and proline-rich Even-skipped repressor (Han and Manley, 1993), or proline- and glycine-rich repressor of WT1 (Madden et al., 1993). The fact that 7 out of 24 amino acid residues of the EGR-1 repression domain are serine or threonine suggests that the repression function may be regulated by phosphorylation.

Consistent with its transcriptional regulatory function, EGR-1 was shown to localize to the nucleus (Cao et al., 1990; Day et al., 1990). Generally, nuclear localization signals are short stretches of 8-10 amino acids rich in basic and proline residues (Silver et al., 1984). In EGR-1, the only regions rich in basic residues are the three zinc finger region and the adjacent stretch of basic amino acids, suggesting that the nuclear localization signal resides in those sequences. Using a series of deletion mutants of EGR-1 along with subcellular fractionation and Western blot analysis, amino acids 315 to 429 were shown to be important for proper targeting to the nucleus (Day et al., 1990). However, the zinc finger region alone (amino acids 331 to 419) was not sufficient for nuclear localization of the protein. The basic sequence 5' to the zinc finger domain (amino acids 315 to 330) was also required for nuclear targeting (Day et al., 1990). Thus, EGR-1 has a bipartate nuclear localization signal, the largest portion of which coincides with the DNA-binding domain.

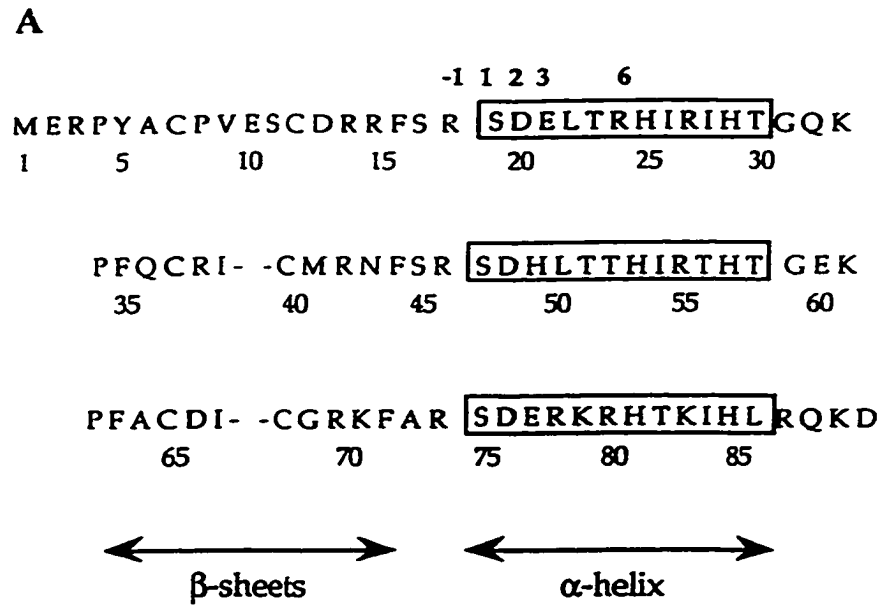
1.1.4 DNA-binding function of EGR-1

The DNA sequence to which EGR-1 binds was initially identified based on the assumption that EGR-1 regulates its own expression and might, therefore, bind to the 5' upstream flanking sequence of the *Egr-1* gene (Christy and Nathans, 1989). Using bacterially expressed EGR-1 protein, specific binding to one promoter fragment was observed in gel mobility shift assays (Christy and Nathans, 1989). This binding site was located within 650 base

pairs 5' to the transcription start site of *Egr-1.* DNase-I footprinting experiments identified the sites of contact in the Egr-1 promoter binding sequence, as well as a number of sites 5' to other genes (Christy and Nathans, 1989). These sequences, taken together, were used to establish a consensus high-affinity binding site sequence GCG-G/TGG-GCG. Further methylation interference experiments indicated that EGR-1 makes extensive contacts along this sequence (Christy and Nathans, 1989).

EGR-1 is the first C_2H_2 zinc finger protein for which a high resolution structure was obtained. The structure of the complex consisting of the cloned three zinc finger peptide and the 12-bp DNA containing the specific 9-bp sequence has been solved at 2.1-Å resolution (Pavletich and Pabo, 1991) (Figure 1.5). The structural information obtained in this study served as a topological blueprint for other members of the zinc finger family. This EGR-1 - DNA complex was recently refined at 1.6 Å (Elrod-Erickson et al., 1996). The new structure confirms all the basic features of the 2.1 Å model, and reveals additional critical details of the complex.

The crystal structures show that the three fingers wrap around the double helix, describing a C shape (Figure 1.6). The overall alignment of the fingers on the DNA is antiparallel (finger 1 binds near the 3' end, and finger 3 binds near the 5' end of the 5'-GCGTGGGCG-3' consensus binding site). [The mode of TFIIIA - DNA interaction is also antiparallel, with most of the contacts involving the guanine-rich strand of the DNA (Smith et al., 1984; Vrana et al., 1988)]. The α-helices of the zinc fingers are tipped at ca.45° relative to the plane of the base pairs; therefore, they are only approximately aligned with the major groove. The N-terminal end of each α-helix makes specific contacts with the base pairs, each helix interacting primarily with the 3-bp subset of the 9-bp consensus binding site.



B

1 2 3 4 5 6 7 8 9 10 11

A G C G T G G G C G T

C G C A C C C G C A T

Figure 1.5 Sequence of the EGR-1 zinc finger peptide and of the DNA binding site used in cocrystallization. **(A)** Sequences of the three zinc fingers are aligned by conserved residues and secondary structure elements. α-helices are enclosed in boxes, and β-sheets are indicated by arrows. The conserved cysteine and histidine residues are highlighted in bold. **(B)** Sequence of the duplex oligonucleotide used in cocrystallization (Pavletich & Pabo, 1991).

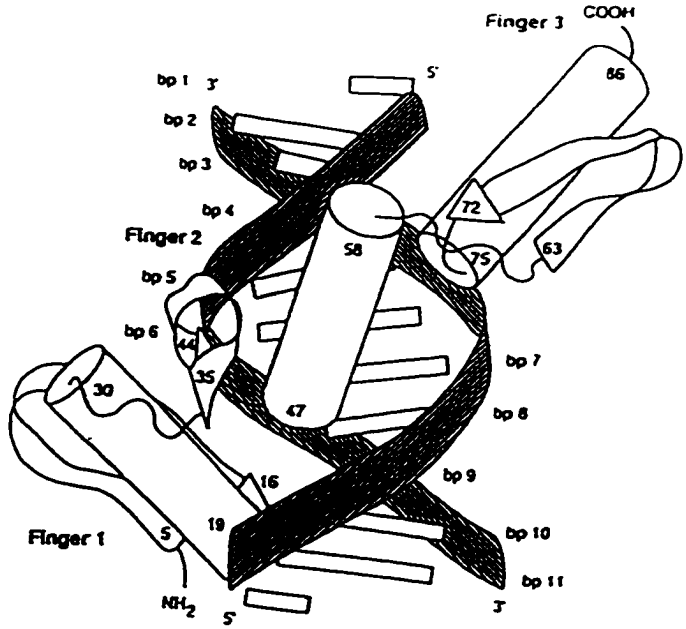


Figure 1.6 The overall arrangement of the three zinc fingers of EGR-1 in the major groove of DNA (Pavletich & Pabo, 1991).

The N-terminal strand of each β-sheet makes no contact with DNA, while the more C-terminal strand contacts the sugar-phosphate backbone along one DNA strand. The amino acid side chains directly involved in contacts with the DNA bases lie at α-helical positions -1, 2, 3, and 6 (numbered relative to the first residue in each α-helix). Base specific contacts identified in the original crystal structure are primarily hydrogen bonds with the G-rich strand of the DNA (Figure 1.7). The higher resolution structure provides a more detailed view of the EGR-1-DNA interface with more direct and water-mediated contacts. The summary of direct base and phosphate contacts is shown in Figure 1.8. Water-mediated and van der Waals contacts to bases and phosphates are summarized in Figure 1.9.

Many features of this complex were correctly predicted by sequence analyses and mutational studies (Nardelli et al., 1991). Site-directed mutagenesis experiments with DNA-binding domains of Krox-20 and Sp1 proteins substituted the nonhomologous amino acids in the recognition α-helices and resulted in interconversion of the specificity of the mutant finger (Nardelli et al., 1991). Thus, even prior to the solution of the EGR-1 crystal structure, some of the base-contacting positions had already been identified.

Fingers 1 and 3 have identical residues at positions -1, 2, 3 and 6 of the α-helix: R, D, E, R, whereas finger 2 has R, D, H, and T at the corresponding positions. Fingers 1 and 3 make two primary contacts to the guanines of the GCG subsite using the guanidinium group of the arginines to hydrogen bond with the N7 and O6 of the guanines in each subsite (Figure 1.10A). The interaction of arginine with guanine was predicted to play an important role in sequence-specific recognition based on its unique stereochemical complementarity (Seeman et al., 1976). Indeed, these contacts are the most prominent ones in the EGR-1 - DNA complex. These arginines also interact

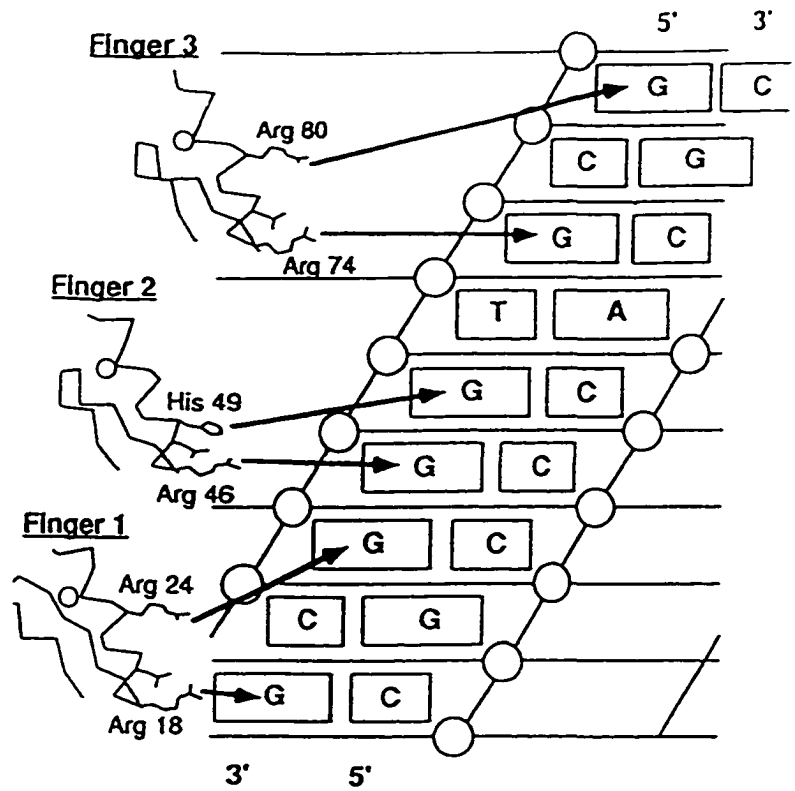


Figure 1.7 Sketch summarizing the principal amino acid-base contacts as seen in the original EGR-1 X-ray structure (Pavletich & Pabo, 1991).

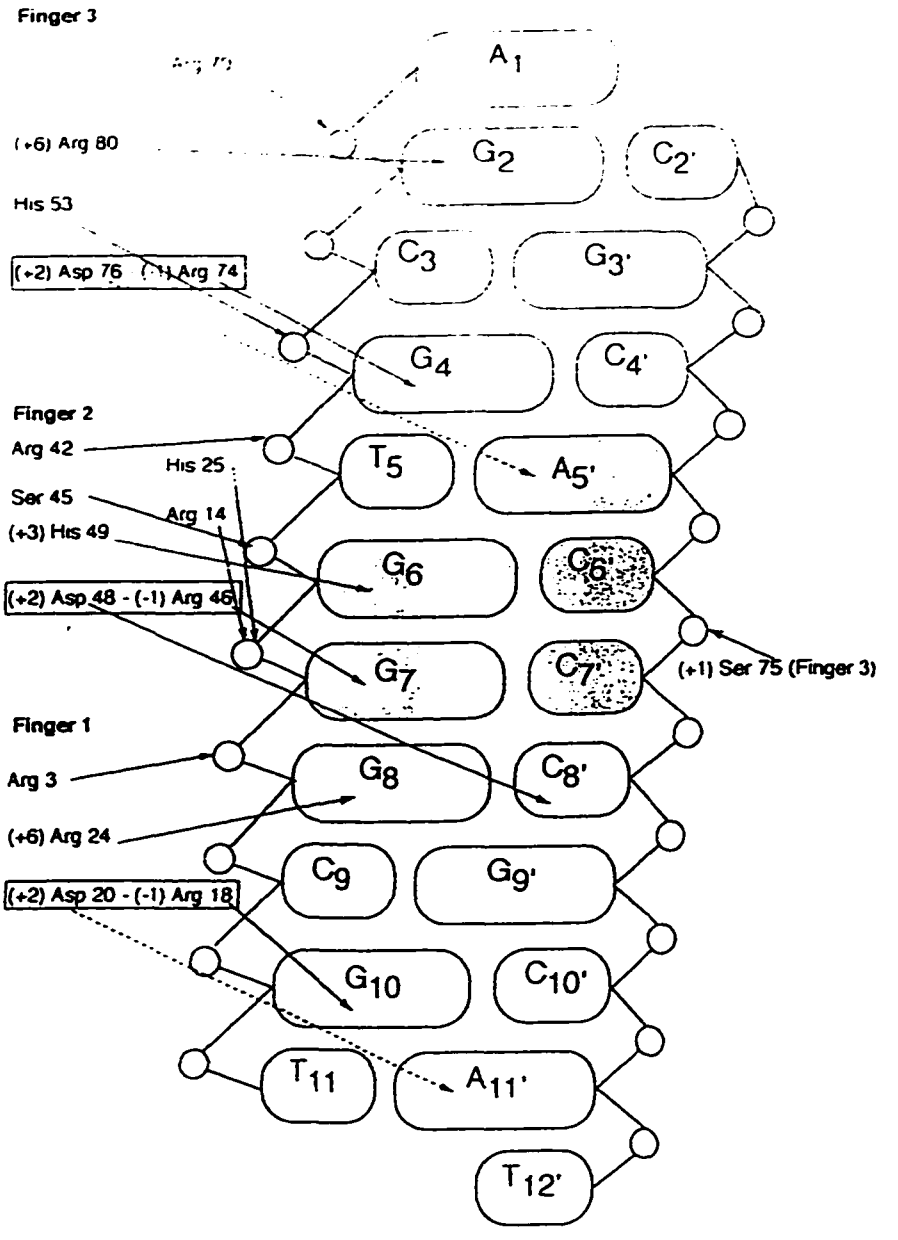


Figure 1.8 Summary of direct base and phosphate contacts in EGR-1-DNA complex. The numbers in parenthesis in front of amino acids are their positions within α-helices. Boxes indicate coupled residue pairs. Arrows indicate hydrogen bonds; dotted arrows represent bonds with marginal geometry (derived from Elrod-Erickson et al., 1996).

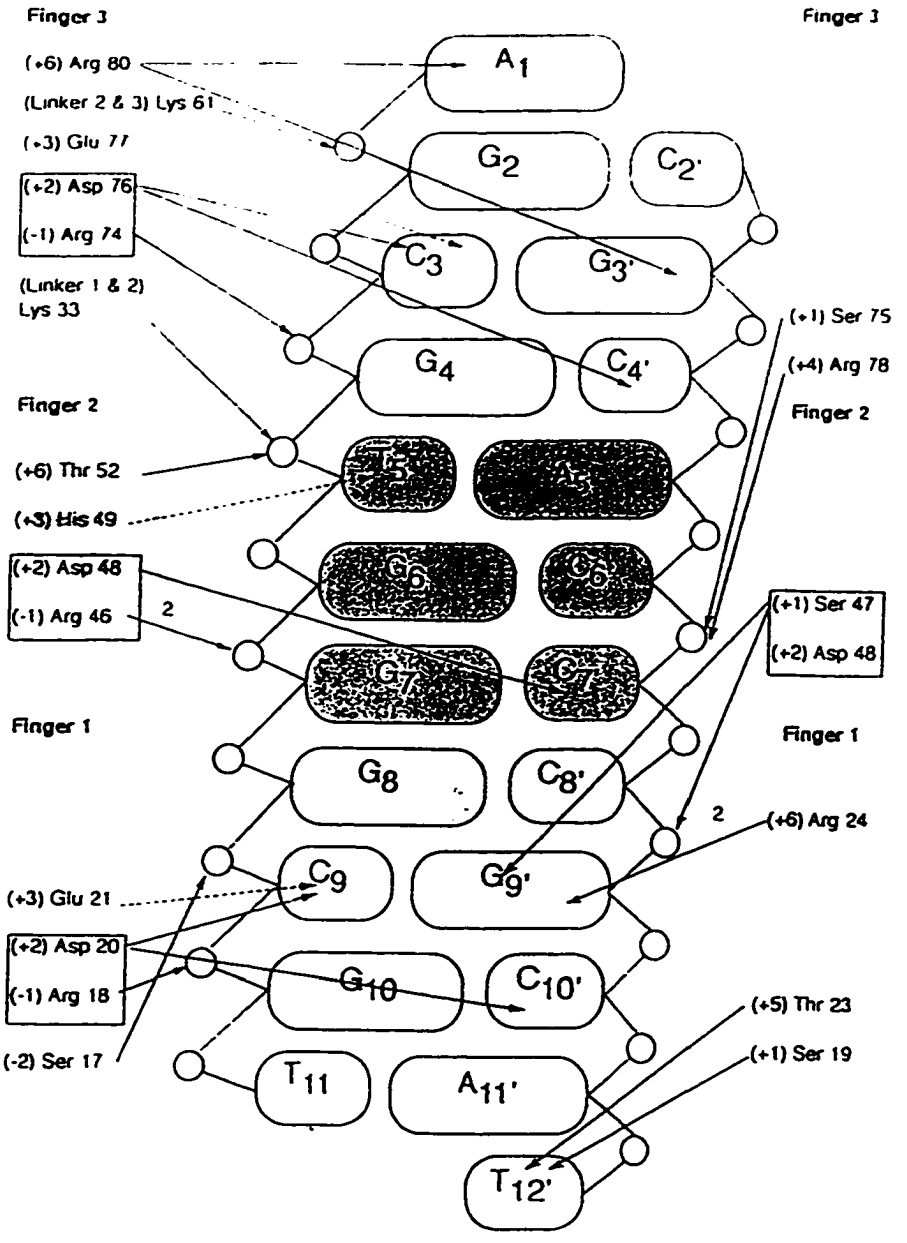


Figure 1.9 Summary of water-mediated and van der Waals contacts to bases and phosphates in EGR-1 - DNA complex. The numbers in parenthesis in front of amino acids are their positions within α–helices. Boxes indicate coupled residue pairs. Arrows indicate water-mediated contacts. Small numbers shown over the arrows indicate the number of water-mediated bonds. Dotted arrows represent van der Waals contacts.

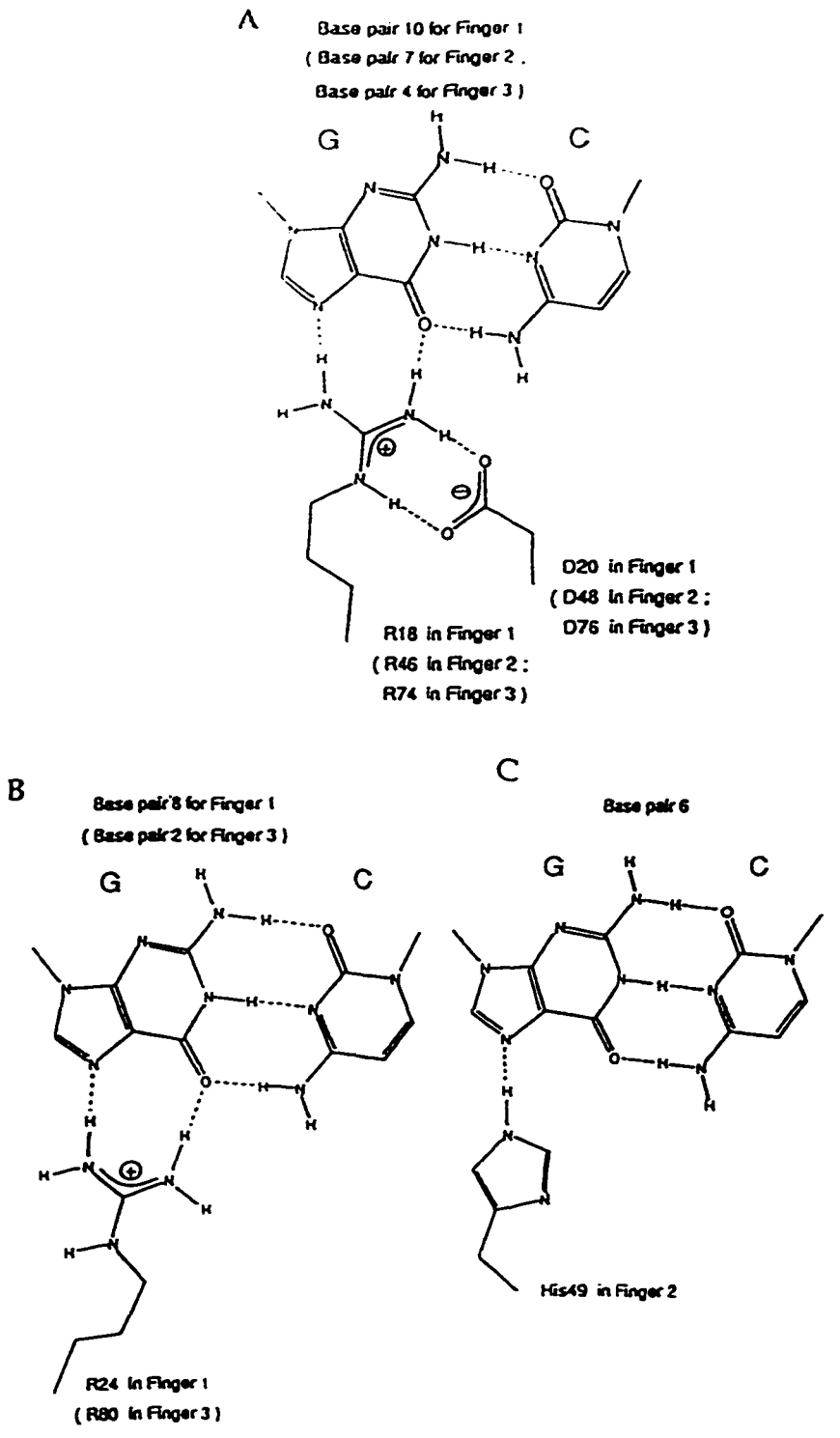


Figure 1.10 Drawings of amino acid-base pair contacts of the EGR-1-DNA complex: **(A)** Drawing of the Asp-Arg-guanine interaction that occurs in all three zinc fingers; (B) Drawing of the Arg-guanine interaction that is present in fingers 1 and 3; (C) Drawing of the His-guanine interaction seen in finger 2 (Pavletich & Pabo, 1991).

with an aspartic acid, which is the second residue within each α-helix. The carboxylate oxygens of the aspartic acid make hydrogen bond-salt bridges with the Nε and Nη of the guanidinium group of the arginines that are already bound to the first G of each subsite, thus stabilizing the interaction of the long arginine side chain with the base (Figure 1.10B). Finger two uses an arginine and a histidine to contact the guanines of the middle base pair subsite TGG. A histidine-guanine interaction that occurs in finger 2 is shown in Figure 1.10C.

The new 1.6 Å structure helps explain the roles of the aspartic acid residues at position 2 of the recognition helices. The crystal structure shows that the Arg -1/Asp2 residue pairs make water-mediated contacts with the cytosine which is base paired to the guanine contacted by Arg -1 (as well as with the phosphate on the 5' side of this guanine). These contacts are seen in all three zinc fingers. In fingers one and three, these aspartates also make water-mediated contacts with the neighboring base 5' of the critical guanine. For instance, Asp +2 in finger 1 makes a water-mediated contact with cytosine 9, and an analogous contact is made by Asp +2 of finger 3 to the N4 of cytosine 3 (Figure 1.9). The crystal structure shows that Asp +2 of finger 2 clearly contacts cytosine 8', and tentatively suggests that Asp +2 residues from fingers 1 and 3 forms weak interactions with a base positioned outside the canonical triplet on the secondary, C-rich strand of the DNA (Figure 1.8). In the original X-ray structure one of the oxygens of the aspartic acid was shown to be within hydrogen bonding distance to a neighboring base on the parallel, C-rich strand of the DNA. However, since the H-bonding geometry was not ideal, these contacts were presumed unimportant for recognition (Pavletich and Pabo, 1991). More favorable geometries were observed in the refined structure (Elrod-Erickson et al., 1996). For example, when Arg -1 of finger 3 contacts G_4, this interaction is stabilized by Asp +2 of the same finger, which in turn

contacts $A_{5'}$ (the complementary base to T_5 on the opposite strand) (Figure 1.8). Such contact was clearly demonstrated for finger 2 (Asp48 interacting with $C_{8'}$). The corresponding interactions for fingers 1 and 3 had less favorable stereochemistry.

There are several pieces of evidence suggesting that this contact contributes to recognition in other zinc finger proteins. In the cocrystal structure of the *Drosophila* regulatory protein Tramtrack complexed with its DNA the homologous aspartic acid in the second zinc finger was shown to make direct hydrogen bonds with the secondary DNA strand (Fairall et al., 1993). Replacement of the aspartic acid by alanine in the second zinc finger of the *S. cerevisiae* ADR1 protein resulted in significant reduction of protein binding to its cognate site (Thukral et al., 1991; Thukral et al., 1992).

Several biochemical studies also propose the interaction of Asp +2 with the parallel strand of the DNA. Zinc finger phage display selection studies provided original support for this interaction. The technique of phage display makes use of the expression of zinc finger regions that carry partially randomized recognition helices as part of a phage coat protein. By then allowing the phage carrying zinc finger peptides on its surface to equilibrate with a target DNA sequence, it is possible to isolate and amplify the specific zinc finger proteins that recognize desired DNA target sites. In the study by Choo and Klug (1994a) the sequence of the middle finger of EGR-1 was randomized to test new binding specificities. Strikingly, in almost all of the selected fingers in which arginine interacts with the 3' guanine, aspartic acid was selected at position +2. When position -1 was not arginine, aspartic acid at position +2 was practically never selected. This suggests that the aspartic acid residue +2 plays an important role in contributing to the specificity of the protein-DNA complex. This result also emphasizes the importance of

arginine in conferring DNA-binding ability. In our study of Denys-Drash WT1 protein mutants the substitution of arginine for a tryptophan was detrimental to DNA-binding activity (Borel et al., 1996). In a complementary study by Choo and Klug (1994b), a bias toward G or T at the 3rd position of the recognition subsites was demonstrated, and it was concluded that the amino acid at position +2 was able to modulate the specificity of the amino acid at other positions. Thus, it was proposed that Asp +2 makes a cross-strand contact. Swirnoff and Milbrandt (1995) carried out binding site selections to derive the high-affinity consensus binding site for EGR-1 family of proteins. Their findings support the importance of aspartic acid for sequence recognition by EGR-1 protein family members. The authors noted a very strong selection for thymine at position 5 and less frequently for guanine, while adenine or cytosine were never selected. Both A and C, the complementary bases to T or G at position 5 have hydrogen bond donor groups (N6 and N4) that can interact with the carboxylate oxygen of Asp +2. These results support the idea that the hydrogen bond potential on the secondary DNA strand contributes to recognition. Another implication of these findings is that a single zinc finger can actually specify not 3 bp, as originally proposed, but 4 bp DNA sites, and these binding sites overlap by one base pair. Recently published work by Isalan, et al. (1997) demonstrates that EGR-1-like zinc fingers can, in fact, specify overlapping, 4-bp subsites. In this study, the third zinc finger of EGR-1 protein was modified to remove the finger's potential for DNA interactions. Normally, the third zinc finger interacts with the first triplet of the EGR-1 consensus sequence. However, as a result of finger three alterations, changes were observed in the 5' position of the middle triplet of the DNA binding site. Thus, deleting the contact from finger 3 affects the specificity for the 5' base of the binding site of finger 2.

Further selection and gel mobility shift experiments pointed to the same conclusion (Isalan et al., 1997). Thus, the potential of zinc fingers to specify overlapping, 4-bp subsites was demonstrated, and the role of aspartic acid at helical position 2 was emphasized. The specificity of each recognition subsite is, therefore, mediated by two adjacent fingers that are synergistic in their mode of binding. This has led the authors to redefine the binding subsites for each zinc finger, and the new schematic diagram of recognition is shown in Figure 1.11.

The roles of the third α-helical residues are also refined in the 1.6 Å structure (Elrod-Erickson et al., 1996). The role of His49 (the third residue in finger 2 α-helix) was previously discussed in the original EGR-1 crystal structure (Pavletich and Pabo, 1991). This residue was previously shown to make a direct hydrogen bond to N7 of guanine 6. In the new structure, this histidine may be contacting O6 of the guanine instead. In addition, His49 also makes a series of van der Waals interactions with the methyl group of thymine 5. The significance of the histidine-thymine interaction was demonstrated by Swirnoff and Milbrandt (1995) in their binding site selection experiments. Thus, histidine contributes to the recognition of two bases on the G-rich DNA strand.

Glutamic acid residues which occur at α-helical position three of fingers 1 and 3 were suggested to contribute to specificity because biochemical site selection and binding studies all show a preference for cytosine at the center of triplets recognized by fingers one and three (Christy and Nathans, 1989; Jamieson et al., 1994; Swirnoff & Milbrandt, 1995). As revealed in the refined crystal structure, the carboxylate groups of Glu21 and Glu77 do not make any base-specific contacts. However, the side chains of these glutamates do hydrogen bond to the backbone amides of the arginines immediately

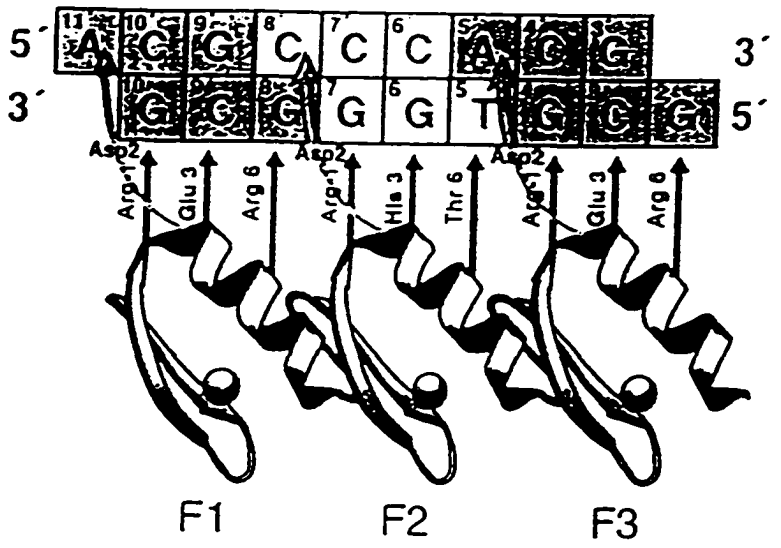


Figure 1.11 Schematic diagram of EGR-1 zinc fingers interacting with the proposed overlapping, 4-bp DNA subsites (Isalan et al., 1997).

preceding the α helix (Arg18 in finger 1, and Arg74 in finger 3). Glu21 can also hydrogen bond to the backbone amide of Ser17, while Glu77 makes a corresponding contact to Ala73. In addition, these glutamic acid residues make van der Waals contacts to the edge of the corresponding cytosines. These contributions, albeit modest, may play a role in favoring cytosines over other bases. Another good observation was made by Nardelli et al. (1992) and Swirnoff & Milbrandt (1995): the reason for cytosines and not guanines in the middle of triplet 1 and 3 might be an electrostatic repulsion between the carboxylate oxygens of the glutamates and the O6 and N7 groups of guanines. Cytosines do not have either of these groups, and, therefore, are better tolerated at these positions.

A number of water-mediated contacts were detected at the protein-DNA interface that were not seen in the original X-ray structure. They are summarized in Figure 1.9. A set of contacts with the sugar-phosphate backbone of the DNA is also detected in the X-ray structure, and most of these are made to the G-rich strand of the consensus site (Figure 1.8). Each finger uses an arginine in the second β-strand and the first zinc-coordinating histidine to contact phosphodiester oxygens. These contacts include one unusual arrangement in which the imidazole ring of the first histidine, involved in zinc coordination through Nε, makes contacts to the phosphodiester oxygens through Nγ (Figure 1.12). Since zinc coordinating histidines are conserved in all zinc fingers, it is possible that they all use these types of contacts. The 1.6 Å structure revealed additional, numerous water-mediated phosphate contacts. They are summarized in Figure 1.9. Serine residues that occur at position 1 of the α helices all provide phosphate contacts: Ser75 (in finger 3) makes a hydrogen bond to phosphate 7', while Ser47 (in finger 2) makes a water-mediated contact with phosphate 9'.

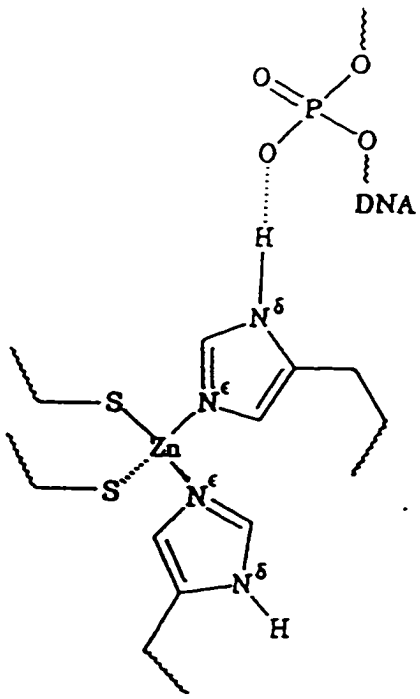


Figure 1.12 Schematic representation of hydrogen bonding between Zn-coordinated histidine and DNA backbone (O'Halloran, 1993).

Similarly, Ser19 (in finger 1) makes a water-mediated contact with O5' of thymine in position 12'. Other residues also contribute to the network of water-mediated phosphate contacts, namely Ser17 makes a water-mediated contact to the 5' phosphate of base 9; Arg78 (the fourth α-helical residue in finger 3) makes a water-mediated contact with phosphate 7'; and Thr23 (residue 5 in the helix of finger 1) contacts the O5' of base 12'. In addition, the conserved lysines that occur in the linker region between the zinc fingers also make several water-mediated contacts: Lys33 (located between fingers 1 and 2) makes a pair of water-mediated contacts to the 5' phosphate of base 5; Lys61 (located between fingers 2 and 3) makes an analogous water-mediated contact with the 5' phosphate of base 2. An analogous phosphate contact involving lysine residue was described for DNA-binding peptide of TFIIIA (Choo and Klug, 1993). Mutation of this residue in TFIIIA reduces its affinity for DNA by about sevenfold.

The important role of the linker sequences was pointed out in the study by Wilson et al. (1992). The authors used a reversion-based analysis of DNA-binding domains: they constructed a chimeric protein containing the DNA-binding domain of EGR-1 inserted between the LexA DNA-binding domain and the GAL4 transcriptional activating domain (LAG). This dimeric protein when expressed in yeast resulted in retardation of their growth, while a dimeric protein containing only the LexA and GAL4 domains acted as an activator. Some yeast cells harbouring LAG eventually reverted to wild-type growth by inactivation of LAG. Those revertants were analysed for mutations in LAG, which all mapped to the EGR-1 zinc finger DNA-binding domain. The majority of the mutations were found to comprise the DNA-contacting amino acids identified in the crystal structure, while several of the mutations lay in the linker regions of the zinc fingers. The mutation of the

two lysines in the linker regions was disruptive, and the authors proposed that the removal of the positive charge imparted by the lysines may destabilize the tertiary structure of the protein. Secondly, it was noted that the two linker regions were not equally sensitive to mutation: the second linker was mutated seventeen times whereas the first was altered only three times, suggesting that the linkers may not play identical roles in orienting the zinc fingers.

Finally, the role of DNA conformation was addressed in the recent X-ray study (Elrod-Erickson et al., 1996). It is well known that DNA conformation either before or subsequent to protein binding is an important determinant of sequence-specific recognition (Steitz, 1993). Thus, stretches of A/T residues in the binding sites of the zinc finger proteins Tramtrack and MIG1 are required for recognition (Fairall et al., 1993; Lundin et al., 1994). While there was no bending detected in the crystal structures of EGR-1-DNA complex, circular dichroism studies show a striking difference occurring in the EGR-1 binding site depending on the presence or absence of the protein (Elrod-Erickson et al., 1996). This suggests that EGR-1 DNA changes conformation upon complex formation. The DNA also adopts a novel conformation, both wide and deep, that has been described as $B_{\text{enlarged groove}}$ -DNA. The binding sites in Tramtrack and GLI have been shown to adopt the same conformation (Fairall et al., 1993; Nekludova & Pabo, 1994). The significance of this $B_{\text{enlarged groove}}$-DNA conformation was not clear until the modeling studies showed that the observed interfinger contacts and the linker length would not allow binding to canonical B-DNA (Elrod-Erickson et al., 1996).

The original crystal structure of EGR-1 complex (Pavletich & Pabo, 1991) gave an initial view of zinc finger-DNA interactions. The new, refined

EGR-1 - DNA structure (Elrod-Erickson et al., 1996) revealed considerable additional complexity, and changed the widely held perception that EGR-1 uses a simple recognition scheme in its interaction with DNA.

1.1.5 Structures of other zinc finger-DNA complexes

The elucidation of high resolution structures of less closely related zinc finger proteins - the human oncogene GLI (Pavletich & Pabo, 1993) and the *Drosophila* regulatory protein Tramtrack (Fairall et al., 1993) provide a broader perspective on the issue of specificity of zinc finger-DNA interactions. The human oncogene protein GLI contains five zinc fingers, and has been cocrystallized with the sequence of DNA derived from *in vitro* genomic selection studies (since the physiologically relevant GLI-binding sites are not yet known) (Figure 1.13). The comparison of GLI-DNA complex with the original EGR-1-DNA structure had revealed more differences than similarities. However, with the elucidation of the refined, 1.6Å structure of the EGR-1-DNA complex, more similarities between the two protein-DNA complexes became apparent. The overall structure of the GLI-DNA complex shows that fingers 2 to 5 are arranged similar to the EGR-1 fingers, wrapping around the DNA with the N-terminal portion of their α-helices fitting into the DNA major groove. The difference between the two complexes is that, surprisingly, finger 1 of GLI does not make any contacts with the DNA at all (Figure 1.14). Fingers 2 and 3 bind a section of the DNA largely non-specifically, contacting the DNA backbone, with finger 2 providing the only base contact in this region. On the other hand, fingers 4 and 5 do bind a conserved 9-bp region of the site in a sequence-specific manner. These fingers are most important for recognition. In EGR-1, all fingers make generally similar contributions to sequence-specific recognition. Alignment of the zinc

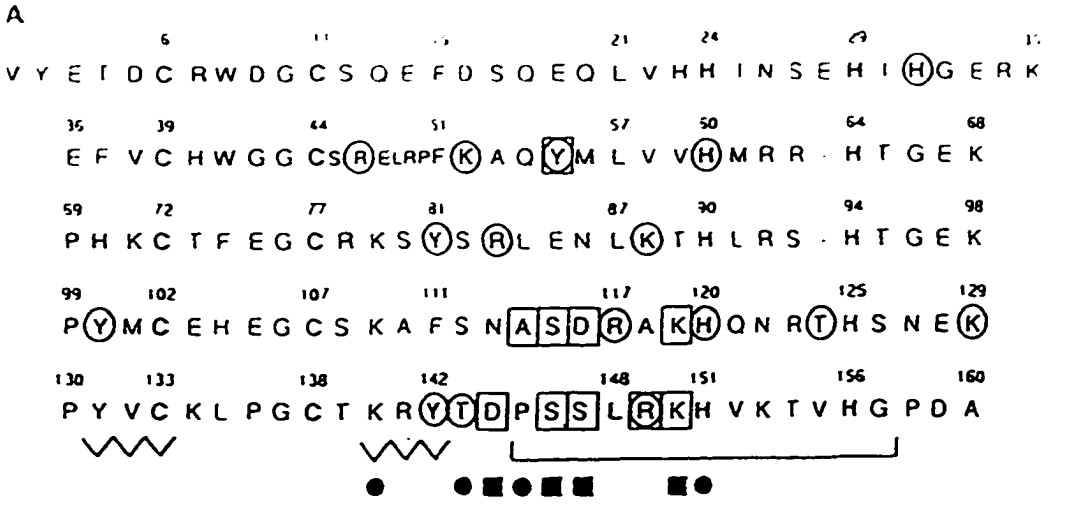


Figure 1.13 Sequences of the GLI zinc finger domain and the DNA-binding site used for cocrystallization. **(A)** The five zinc fingers of GLI are aligned to show the conserved amino acid residues and secondary structure elements. The position of α-helices is underlined, and β-sheets are indicated by zig-zag lines. Open boxes highlight residues that make base contacts in the crystal structure, and open circles indicate residues that make phosphate contacts. Symbols below the GLI sequence indicate the corresponding positions of side chain-base contacts (filled boxes) and side chain-phosphate contacts (filled circles) that were observed in EGR-1 complex. **(B)** DNA duplex used in cocrystallization (Pavletich and Pabo, 1993).

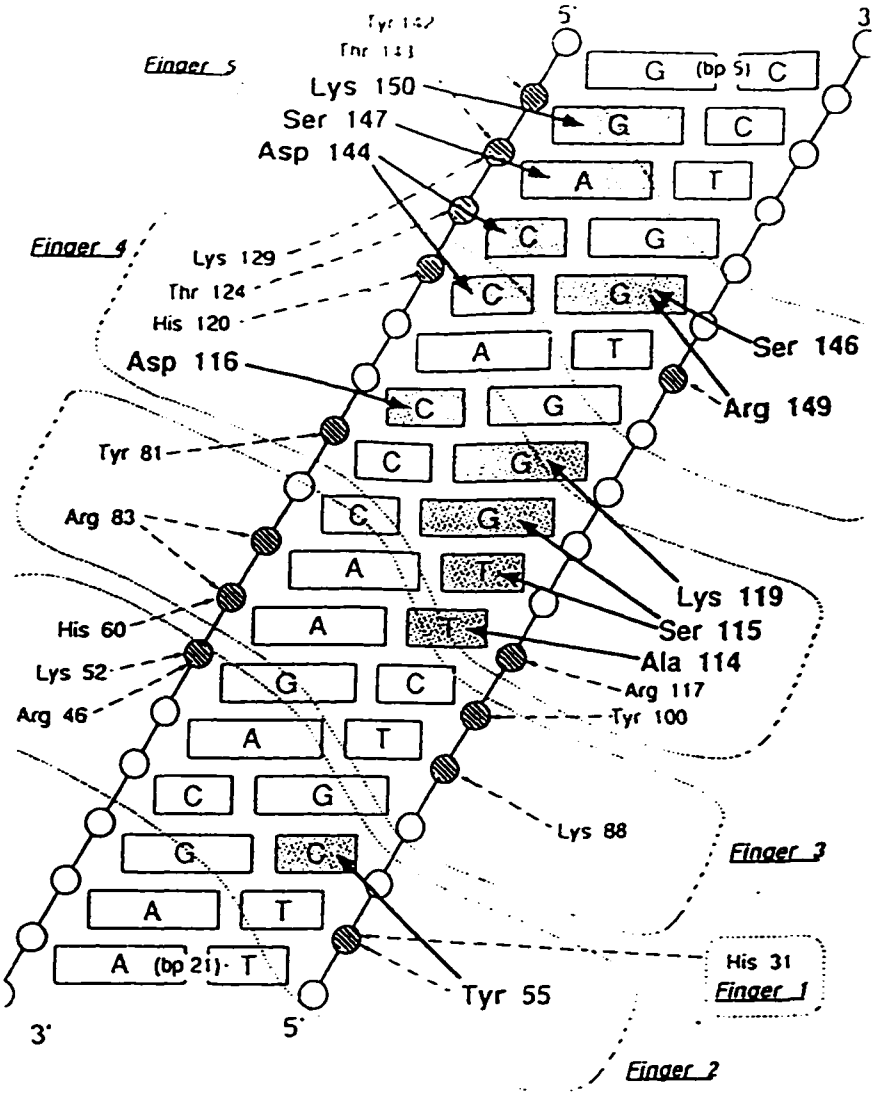


Figure 1.14 Scetch summarizing base and phosphate contacts made by the GLI peptide. Solid arrows indicate base contacts and dotted arrows indicate phosphate contacts (Pavletich & Pabo, 1991).

finger sequences of GLI and EGR-1 showed a clear correlation between the positions of the base contacting residues in the two proteins. However, the mode of docking zinc fingers against DNA shows differences, which are reflected in the different pattern of amino acid-base interactions at the protein-DNA interface. Despite the fact that GLI and EGR-1 belong to different subfamilies, with very little amino acid similarity in the zinc finger region, and different DNA sequences to which they bind, the two proteins have the following features of the protein-DNA complex in common: (1) each protein contacts primarily four bases in each subsite, and these subsites can overlap; (2) four amino acids in the α helix are usually involved in these contacts; (3) both proteins contact bases on both DNA strands, except that the majority of EGR-1 contacts involve primary, G-rich strand of the DNA, while GLI contacts are more evenly distributed between the two strands; (4) in both complexes the DNA assumes a conformation that is intermediate between A- and B-form; (5) the DNA-contacting residues of EGR-1 fall at positions -1, 2, 3, and 6 of the α-helix, GLI also uses the DNA-contacting residues corresponding to positions -1, 2, 3, and 6. In GLI, however, additional residues of the α-helix make specific contacts, including positions 1 and 5. While GLI fingers make some use of the primary contacts identified in EGR-1, several of these interactions are unique in terms of the subsite position of the base contacted, the strand contacted, and the nature of the sidechain-base interaction.

The structure of a two zinc finger protein Tramtrack complexed with its recognition DNA sequence shows many similarities with the EGR-1-DNA complex. The contacts between Tramtrack and its DNA are summarized in Figure 1.15. A comparison of base-specific contacts formed by EGR-1 and Tramtrack proteins is diagrammed in Figure 1.16. The following features of

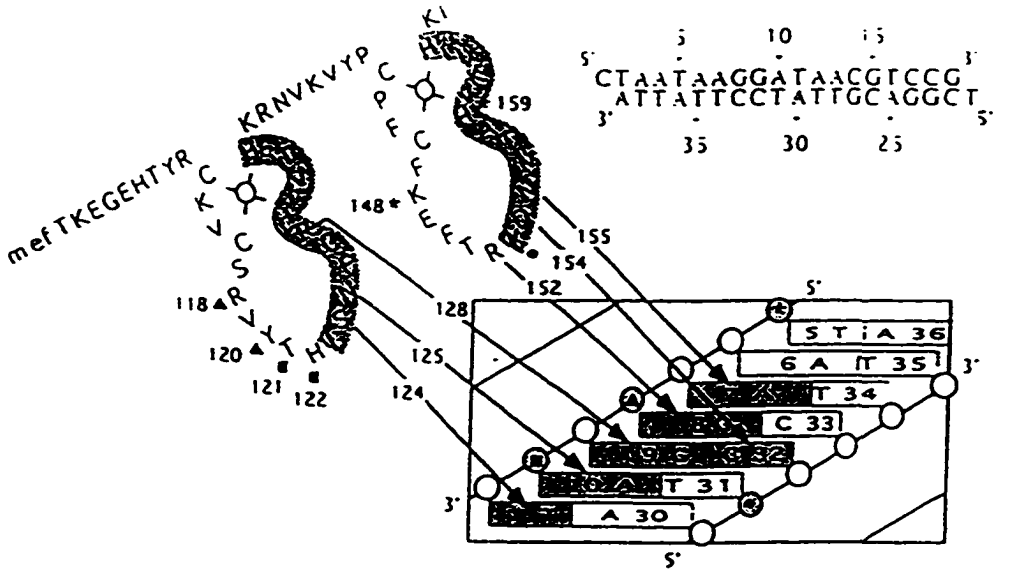


Figure 1.15 The summary of Tramtrack-DNA contacts (Fairall et al., 1993). The sequence of the DNA used for crystallization is shown in the top right of the figure. The details of the protein-DNA contacts are discussed in the text.

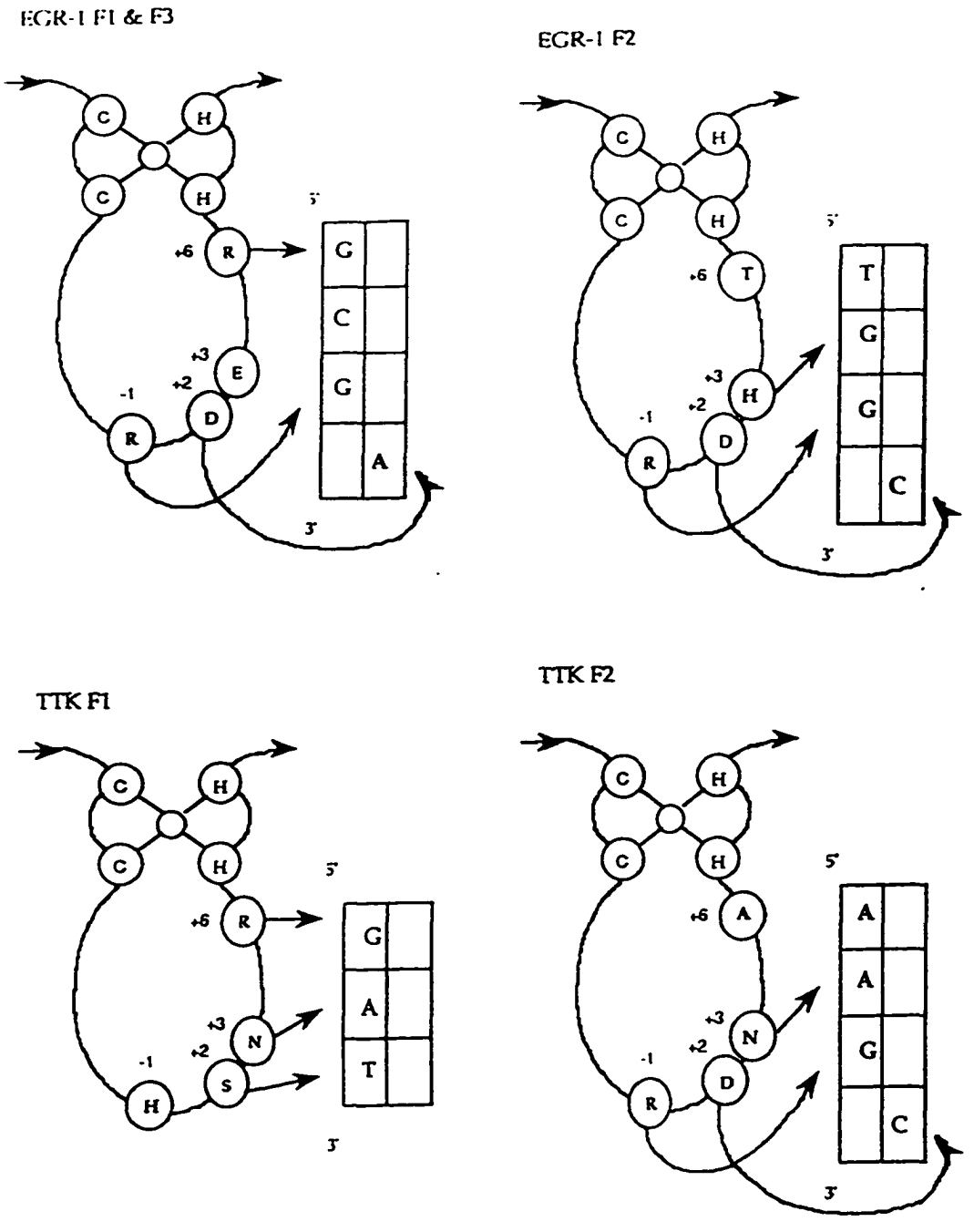


Figure 1.16 Schematic summary of the principal protein-DNA contacts observed in the cocrystal structures of the EGR-1 and Tramtrack. Residues are numbered relative to their position in the α helix. Arrows represent hydrogen bonds observed in the respective crystal structure (Choo & Klug, 1997).

the Tramtrack-DNA structure are similar to those of EGR-1: (1) both fingers of Tramtrack use the N-terminus of their α-helices to contact specific bases in the DNA major groove; (2) the majority of contacts are made to one DNA strand; (3) the binding sites for the adjoining fingers overlap; (4) α–helical positions -1, 2, 3, and 6 are used to make base-specific contacts, which coincide with the same residue positions in EGR-1 complex; (5) finger 2 makes very similar base and phosphate contacts to those of EGR-1 fingers. Finger 2 in both proteins uses an arginine at position -1 for contacting guanine. The use of position 2 in finger 2 observed in the Tramtrack cocrystal structure is equivalent to the contact seen in finger 2 of EGR-1. In both proteins, an aspartic acid residue at this position accepts two hydrogen bonds from an arginine and makes an optimal hydrogen bond to a cytosine on the opposite DNA strand. The crystal structural data of the Tramtrack complex also illustrated novel features. Thus, finger 1 of Tramtrack has an additional β strand, which doesn't make contacts to DNA, but is probably required for the overall stability of finger 1. The pattern of amino acid/base contacts in finger 1 differs from those seen in finger 2 and in all three fingers of EGR-1: the histidine at position -1 does not make a base-specific contact; instead, it contacts a phosphate; a short serine sidechain at position 2 contacts a thymine (this is facilitated by the DNA bending), whereas in zinc fingers 1 and 3 of EGR-1analogous contacts involve aspartates hydrogen bonding to adenine bases on the opposite strand of the DNA. Curiously, about 50 % of zinc finger sequences have a serine residue conserved at this position (Fairall et al., 1993). Overall, although differing in some of the exact contacts, there is a remarkable concordance between the two structures.

Recently, the structure of a designed zinc finger protein bound to consensus DNA sequence has been solved (Kim & Berg, 1996). The overall

mode of binding resembles that of EGR-1, with most of the amino acid-base contacts occurring as predicted. Finally, the cocrystal structure of the four zinc finger protein YY1 (Ying-Yang 1) complexed with a 20-bp adeno-associated virus P5 initiator has been reported (Houbaviy et al., 1996). YY1 is a human GLI-related protein capable of supporting specific initiation of mRNA production (Houbaviy et al., 1996). A diagram summarizing YY1-DNA interactions is shown in Figure 1.17. All four fingers of YY1 bind in the major groove, with fingers 2 to 4 making multiple contacts with DNA. Finger 1, in contrast, contacts only one base. YY1 uses similar α-helical positions within the fingers as those used by EGR-1, GLI, and Tramtrack for making base-specific contacts: -1, 2, 3, and 6 (Figure 1.18 A). A schematic comparison of the DNA contacts made by EGR-1, GLI, and YY1 is shown in Figure 1.18 B. An unusual feature of the YY1-DNA complex is the abrupt strand switching, which is proposed to play an essential role in establishing the polarity of transcription from the initiator element. The distribution of DNA contacts is almost even between the two strands of DNA (template and non-template) until the strand switch occurs in the middle of the binding site, at A10 - T31 and T11 - A30 base pairs (Figure 1.17). From A10 to A17, the remaining interactions are made to the template strand with only one contact involving the nontemplate strand (Figure 1.17). The mechanism of strand switching was observed in other zinc finger - DNA complexes, but not to the same extent as seen in YY1 protein - DNA complex. In conclusion, this crystal structure reveals EGR-1 - like binding by fingers 3 and 4, while finger 1 mainly makes phosphate contacts, but only one contact to a DNA base, and finger 2 uses an unusual lysine - guanine contact from position 3 (Figure 1.17). Overall, the YY1 illustrates new versatility of the zinc finger DNA binding motif.

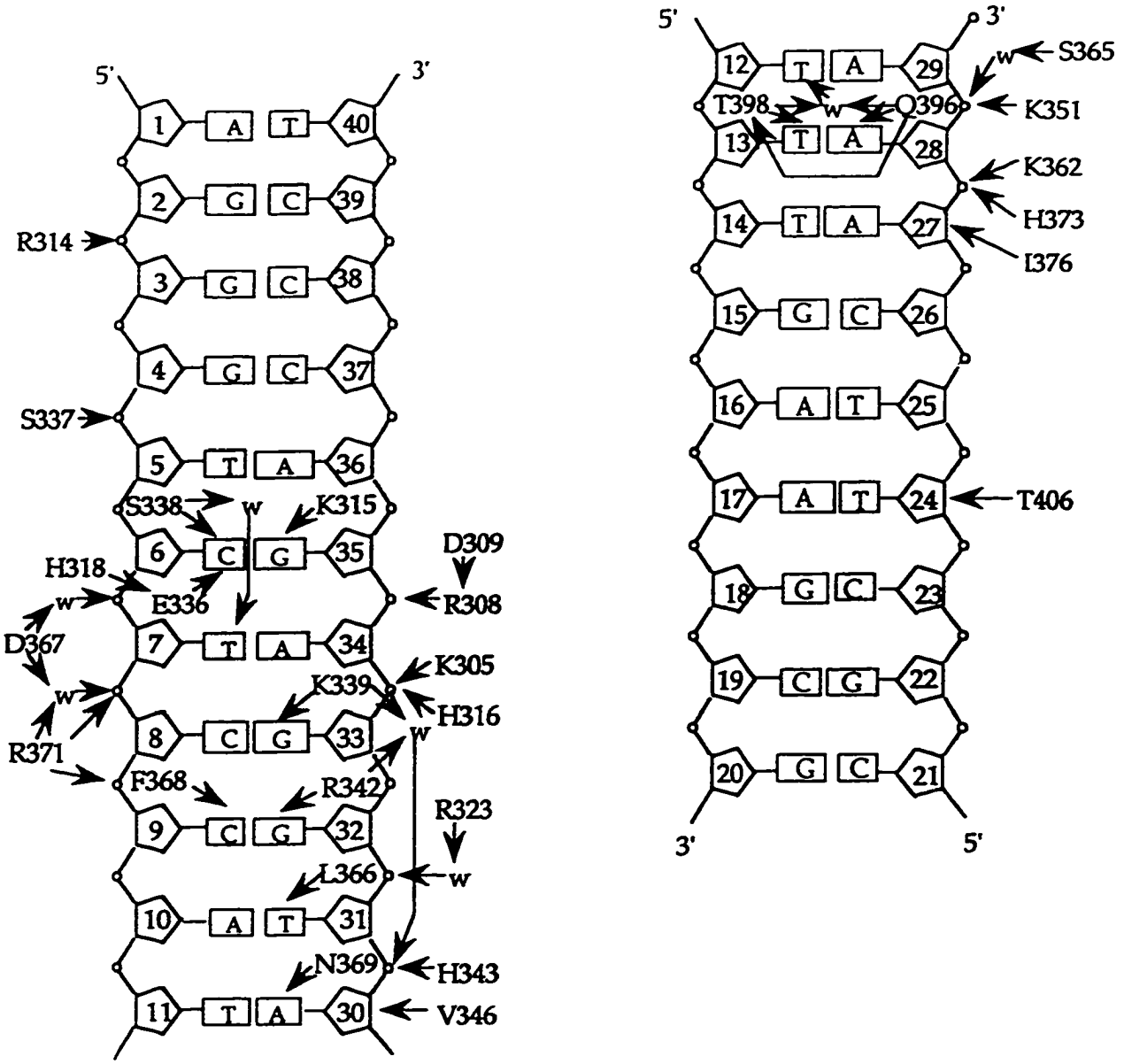


Figure 1.17 Schematic representation of the YY1-DNA interactions. The template strand is labeled 'T'. Protein-DNA contacts are denoted with arrows (Houbaviy et al., 1996). The details of the complex are discussed in the text.

Figure 1.18 Comparison of YY1 with other zinc finger structures.

Figure 1.18 Comparison of YY1 with other zinc finger structures. (**A**) Amino acid sequence of YY1. The zinc fingers are aligned by their zinc-coordinating residues. The YY1 contacts are indicated as follows: •, backbone contact; +, base contact, =, touches both base and backbone. These contacts are also presented with those seen in the structures of EGR-1 (Z1-Z3), GLI (G2-G5), and Tramtrack (T1, T2). (**B**) Schematic comparisons of the DNA contacts made by EGR-1, GLI, and YY1. DNA bases and backbone are represented by rectangles and circles, respectively. Hatched rectangles and solid circles denote sites of protein contact (Houbaviy et al., 1996).

1.1.6 Studies toward a zinc finger recognition code

The recognition principals deduced from the original EGR-1 crystal structure prompted researchers to apply the recognition rules of EGR-1 to other C_2H_2 proteins, such as Sp1 transcription factor, the mammalian Wilms' tumor protein (Haber and Backler, 1992), *Xenopus* transcription factor IIIA, the yeast MIG1 repressor (Nehlin and Ronne, 1990) and others. Comparison of the sequences of more than 1000 zinc finger motifs has shown that amino acids in positions involved in direct base pair contacts in EGR-1 complex are particularly variable in the other zinc fingers (Jacobs, 1992). Observations such as these seem to indicate that EGR-1 and at least its close relatives recognize DNA in a similar manner. This raised an exciting possibility of predicting DNA binding sites for proteins with unknown specificity, or creating new proteins designed to recognize specific DNA sequences. This, in turn, could open new opportunities for both the study of gene regulation and *de novo* design of proteins with preprogrammed specificities for therapeutic use.

Indeed, such ideas have been developed in mutational studies involving the contact residues, some of which resulted in proteins with changed DNA binding specificities (Nardelli et al., 1991; Desjarlais and Berg, 1992). Desjarlais and Berg designed several polypeptides of the Sp1 DNA-binding domain that were made up of fingers of known specificity, and every polypeptide displayed essentially the expected sequence preference. The observed "recognition rules" were analyzed and compiled into databases (Nardelli et al., 1991; Desjarlais and Berg, 1992). Subsequent studies showed that application of some of these rules successfully enables zinc fingers to bind to predetermined DNA-binding sites, and that those rules can be transferred from one protein to another (Desjarlais & Berg, 1993). Hoffman et al. (1993)

introduced amino acid changes at one of the three critical positions of the yeast transcription factor ADR1, which reduced DNA binding affinity while not disrupting the structure of the finger. This suggests that the mutated amino acids were those involved in direct contacts with DNA.

The work of Kriwaki et al. (1992) and Kuwahara et al. (1993) presented evidence that fingers 2 and 3 of the transcription factor Sp1 bind with higher affinity than does finger 1. This finding is also consistent with the use of EGR-1 contact positions, where fingers 2 and 3 provide more hydrogen bonds than finger 1. Furthermore, in one of the earlier studies derivatives of the Sp1 protein were generated, and bound to a predicted site with an affinity of 2 nM (Nardelli et al., 1991).

A different approach to designing a protein with the predetermined binding specificity involves a selection rather than a rational design strategy. This technique is known as phage display, and is now being extensively used by different laboratories (reviewed in Choo & Klug, 1995) (a brief description of the technique can be found in Chapter 1). This has proved to be a powerful selection method used in many studies of DNA-binding proteins. Some of the examples of using phage display to study zinc fingers include the studies by Rebar and Pabo (1994), Choo and Klug (1994a, 1994b), and Jamieson et al. (1994; 1996).

Rebar and Pabo (1994) used a phage display library expressing EGR-1-phage coat protein fusions which contained randomized critical amino acids in the first finger. Rounds of affinity selections using oligonucleotides randomized in the region recognized by the first finger yielded variants of EGR-1 carrying different amino acids in zinc finger 1, that bound specifically to the new DNA sites.

Choo & Klug (1994a) in a very comprehensive study constructed a phage display library containing over 2 million variants of EGR-1 in which the amino acid sequence of finger 2 was randomized at eight positions. This library was then screened against different DNA triplets representing DNA binding subsites. As a result, 16 successful triplets were isolated and matched to a subpopulation of fingers able to bind them. A strong bias was observed towards the base contacting positions revealed in the crystal structure of EGR-1 (Choo & Klug, 1994a). Other studies used similar strategies, and demonstrated that the phage display system is a powerful tool for selecting zinc fingers with novel DNA binding specificities, which could be used in research or even therapy.

Another interesting study reports *in vivo* repression of transcription by a C_2H_2 zinc finger protein designed to specifically recognize an oncogenic sequence (Choo et al., 1994c). A protein containing three zinc fingers was designed using a rationale based on phage display selection as well as the available structural data. A nuclear localization signal was attached to the protein, which directed it to the nucleus where it bound its chromosomal DNA target sequence, and caused a specific inhibition of transcription *in vivo*.

Several research groups have made attempts to describe a DNA recognition code for transcription factors. The group led by Masashi Suzuki is one of the most determined ones, trying to reconcile the available structural data to deduce general rules for protein-DNA recognition (Suzuki, 1993; Suzuki et al. 1995). A recent publication by Choo and Klug (1997) discusses the framework of the zinc finger-DNA recognition code. However, at the present time there isn't sufficient data to draw any far-reaching conclusions. The remark by Harrison and Aggarwal that "the variety of protein-base pair

contacts is more evident than their uniformity" (Aggarwal et al. 1988) still holds today.

The new high resolution structures of zinc finger proteins (discussed in the previous sections of this thesis) reveal very complex interactions at protein-DNA interfaces, and emphasize the difficulties inherent in trying to develop simple rules of protein-DNA recognition. More structural data is necessary for elucidation of a more comprehensive set of rules which could be used to predict DNA binding sites from the amino acid sequence of a protein.

CHAPTER 2.0 THE WT1 WILMS' TUMOR GENE PRODUCT: A TUMOR SUPPRESSOR INVOLVED IN REGULATION OF KIDNEY DEVELOPMENT.

2.1 Introduction

2.1.1 The concept of tumor suppressor genes.

The concept of tumor suppressor genes which play a role in negatively regulating cell growth was introduced in the late 1960s (Harris et al., 1969). Until then, the concept that DNA and RNA tumor viruses introduced new oncogenes that acted in a genetically dominant fashion was the only accepted mechanism leading to neoplastic transformation. The isolation and identification of tumor suppressor genes took many years of exploring using various cloning and somatic cell hybridization techniques. These extensive studies have yielded a number of human tumor suppressor genes including retinoblastoma (RB) (reviewed by Witman, 1993; Hamel, 1993), p53 (reviewed by Zambetti and Levine, 1993), neuroblastoma (Brodeur, 1990), and Wilms' tumor gene (Haber and Housman, 1992). It is without doubt that the discovery of many new tumor suppressor genes is forthcoming, and that the understanding of these genes and their products will eventually lead to the development of new clinical strategies and preventative therapies.

2.1.2 Biology of Wilms' tumor and associated syndromes.

Wilms' tumor is a pediatric kidney malignancy that arises through aberrant differentiation of the kidney stem cells. It represents the most common solid tumor of childhood, affecting approximately 1 in 10, 000 children, and accounting for 85% of all pediatric nephroblastomas (Stanbridge, 1990; Haber and Housman, 1992). The average age of children affected by Wilms' tumor is 3.5 years old, but many of them die before the age of 2. In rare cases, this tumor may develop at an older age, typically in the

mid-20s. The treatment of this malignancy includes surgery and chemotherapy, and, in the case of advanced disease, radiation therapy. These treatments usually result in good long-term survival (5 years follow-up): about 80% of Wilms tumor patients are successfully cured (Lemerle et al., 1983; D'Angio et al., 1989).

Wilms' tumor can occur in either familial or sporadic forms, with almost 99% of the tumors arising sporadically. Most of the tumors develop unilaterally, and only 5-10% of children are affected with the bilateral form of the disease, which is usually inherited. Histologically, Wilms tumors originate from residual islands of immature kidney, and consist of blastemal cells as well as epithelial cells and muscle cells of the supporting stroma. In the majority of the cases Wilms' tumor is not associated with any other systemic disease. However, it can be part of the following syndromes: the WAGR syndrome, which accounts for 1-2% of cases of Wilms' tumor, and includes Aniridia (malformation or absence of the iris), uroGenital malformations, and mental Retardation (Baird et al., 1992a); the Denys-Drash syndrome (2% of all Wilms' tumors), which also includes various anomalies of the genitourinary tract and nephropathy (Coppes et al., 1992b; Baird et al., 1992b); Beckwith-Wiedemann syndrome (BWS) which is characterized by growth abnormalities and a predisposition to Wilms' tumor (Koufos et al., 1989; Henry et al., 1989). All these syndromes have in common the more frequent occurrence of bilateral tumors which appear at an earlier age than in sporadic cases not associated with syndromes. Based on the similarities of tumor presentation between the Wilms' tumor and retinoblastoma (RB) (namely, an association between familial inheritance and bilateral onset of tumors), a "two-hit" hypothesis was proposed by Knudson et al. (Knudson, 1971; Knudson and Strong, 1972). According to this hypothesis, two

independent events leading to inactivation of both alleles of a tumor suppressor gene are required for the tumor to develop. Whereas some Wilms' tumors (the familial cases in particular) may conform to the two-hit hypothesis, the genetics of the majority of Wilms' tumors are likely to be much more complex. There are several lines of evidence to suggest a more complex development of Wilms' tumor. First, the incidence of familial Wilms' tumor cases is only 1-2%, which is very low compared to the 35-40% for retinoblastoma (Brodeur, 1990; Pelletier et al., 1991c). Second, retinoblastoma results from inactivation of a single tumor susceptibility gene located at chromosome 13q14 (Hamel et al., 1993). In the case of Wilms' tumor, three separate genetic loci have been linked to predisposition to the disease.

The association of Wilms' tumor with WAGR syndrome has been of particular importance in identifying a Wilms' tumor susceptibility gene. WAGR patients carry cytogenetically visible chromosomal deletions involving band p13 of the short arm of one of the two chromosomes 11 (Riccardi et al., 1978; Francke et al., 1979; Hastie, 1994). This observation initially generated interest in the chromosomal region 11p13. Additionally, about 20% of sporadic Wilms' tumors show loss of heterozygosity (LOH) for 11p13 markers, as indicated by analysis of restriction fragment - length polymorphisms (RFLPs) (Lewis et al., 1988; Ton et al., 1991). The WAGR-associated deletion was later resolved into two separate loci: an aniridia gene (AN2/Pax-6: small eye in chromosome 2 in mouse), and a putative Wilms' tumor suppressor gene WT1 (Gessler et al. 1990). Further physiological evidence for the involvement of chromosome 11 in Wilms' tumors was provided by the observation that transfer of a normal chromosome 11 (but

not chromosomes 13 or X) into cultured Wilms' tumor cells can suppress their tumorigenic phenotype (Weissman et al., 1987).

A second locus was mapped to chromosomal region 11p15 which has been implicated in Beckwith-Wiedemann syndrome and Wilms' tumor (Coppes et al., 1992; Henry et al., 1989; Koufos et al., 1989). Loss of heterozygosity for chromosome 11p15 markers occurs in about 40% of Wilms' tumors (Coppes et al., 1992). This locus contains genes for IGF-II (insulin-like growth factor) and H19 mRNA. Changes in imprinting of these genes have been linked to tumorigenesis (Zhang and Tycko, 1992; Ogawa et al., 1993; Rainier et al., 1993).

A third locus, albeit not involved in hereditary predisposition to Wilms' tumor, is frequently lost during tumor progression, and is thus considered important in Wilms' tumor development (Coppes et al., 1992; Maw et al., 1992). Tumor LOH for markers at chromosome 16q is found in almost 20% of Wilms' tumor, and is always accompanied by LOH at 11p13 and 11p15 (Coppes et al., 1992).

2.1.3 Identification and characterization of the WT1 gene and its protein product.

Three independent groups have reported cloning of the WT1 gene. Genomic probe WiT-13 was isolated by Call et al. from a cosmid library of the short arm of chromosome 11 (Call et al., 1990), which was deleted in sporadic Wilms' tumor. This deletion was initially identified by loss of DNA markers at 11p13 (Lewis et al., 1988). The WiT-13 DNA fragment was then used to probe several cDNA libraries, and a cDNA clone designated WT33 was isolated. By conducting Northern blot analysis, this cDNA clone was further shown to hybridize to 3 kb-long mRNA species in baboon kidney and spleen,

while no detectable hybridization was observed in RNA derived from muscle, liver, or brain (Call et al., 1990). Similarly, the WIT-2 gene was isolated by Bonetta et al. (1990), using yeast artificial chromosomes, and shown to be identical to the WT33 sequence. Gessler et al. (1990) used an alternative strategy of chromosome jumping to isolate the WT1 gene. First, a genomic probe was obtained using rare-cutting enzymes, which produced sequences indicating the 5'-end of a gene. This probe mapped within chromosomal region 11p13 which is deleted in WAGR patients. Subsequently, using this probe, a cDNA clone termed LK15 was isolated from a human fetal kidney cDNA library.

Analysis of the WT1 gene sequence revealed that it spans approximately 50 kbp of DNA, and contains 10 exons. The WT1 mRNA is about 3.5 kilobases long, and encodes a protein with a molecular weight of 47 to 49 kDa, depending on the presence or absence of two alternatively spliced exons (Haber et al., 1991). The WT1 mRNA is expressed in a limited number of cell types, primarily kidney cells and some hematopoetic cells (Gessler et al., 1990; Call et al., 1990). The predicted protein sequence has functional domains considered to be hallmarks of transcriptional regulatory proteins (Figure 2.1). The amino terminus of WT1 is rich in proline and glutamine residues (Lewis et al., 1988), which has been found in transactivation domains of a number of transcription factors, such as EGR-1, SP1, CTF/NF1, and others (reviewed in Pabo and Sauer, 1992; Mitchell and Tijian, 1989). The carboxyl terminus of WT1 contains four zinc fingers of C_2H_2 type, which represents a putative DNA binding motif characteristic of a family of sequence-specific DNA-binding proteins (Pabo and Sauer, 1992). Fingers 2-4 of WT1 possess a significant similarity to the three zinc fingers of the immediate-early growth response genes, EGR.

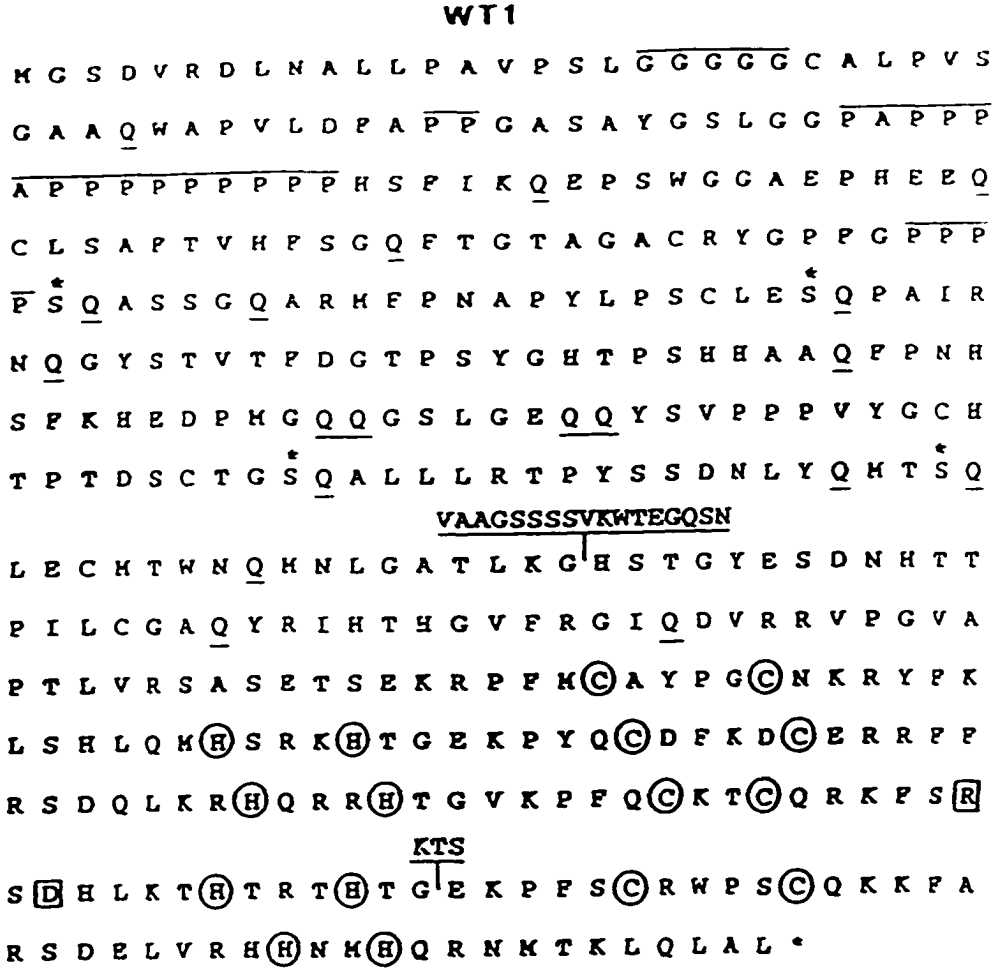


Figure 2.1 The amino acid sequence of WT1 protein: the amino terminus is rich in proline, glutamine and glycine residues, commonly occuring in clusters; the DNA-binding zinc finger region is in the carbooxyl terminus; the zinc-coordinating cysteines and histidines are circled; serine residues that are potential targets for the DNA-activated protein kinase are marked by an asterisk; the 17 amino acid, and the three amino acid insertions resulting from alternative splicing are indicated above; the arginine and aspartic acid residues of finger three found mutated in Denys-Drash patients are boxed (Rauscher, 1993).

Four distinct isoforms of WT1 can be generated by differential splicing of the primary transcript, reflecting the presence or absence of two splice insertions. The 17 amino acid insertion is located in the amino terminus, and contains a number of contiguous serine residues. The three amino acid insertion (lysine-threonine-serine, or KTS) occurs between fingers 3 and 4 of the highly conserved DNA binding domain (Figure 2.2). The most abundant WT1 mRNA species is the one containing both insertions, and the least common is the one without both insertions. The ratio of the four alternative splice variants is conserved in human and mouse, and does not seem to be developmentally regulated (Haber et al., 1991). A mechanism of alternative splicing to generate different binding specificities has been documented for a number of other zinc finger proteins. The chorion transcription factor CP2 and the Tramtrack proteins, both from *Drosophila*, use differential splicing to alter DNA-binding specificities (Gogos et al., 1992; Hsu et al., 1992; Read et al., 1992). Another protein, proto-oncogene *Evi-1* also undergoes alternative splicing in its zinc finger region (Bordereaux et al., 1990). However, it is not yet known what effect this has on DNA binding. Further heterogeneity of WT1 isoforms may be generated by a non-AUG translational initiation event, as recently reported by Bruening and Pelletier (1996).

The presence of serine residues, especially in the 17 amino acid alternative splice insertion, suggests potential targets for phosphorylation. However, despite the abundance of serine residues, the WT1 protein appears to be poorly phophorylated in COS-1 cells (Morris et al., 1991). More recent information suggests that the +KTS splice variant of WT1 may be a better candidate for phosphorylation (Anant et al., 1994). Recently, Ye et al. (1996) demonstrated that both +KTS and -KTS isoforms of WT1 are subject to phosphorylation by both protein kinase A (PKA) and protein kinase C. These

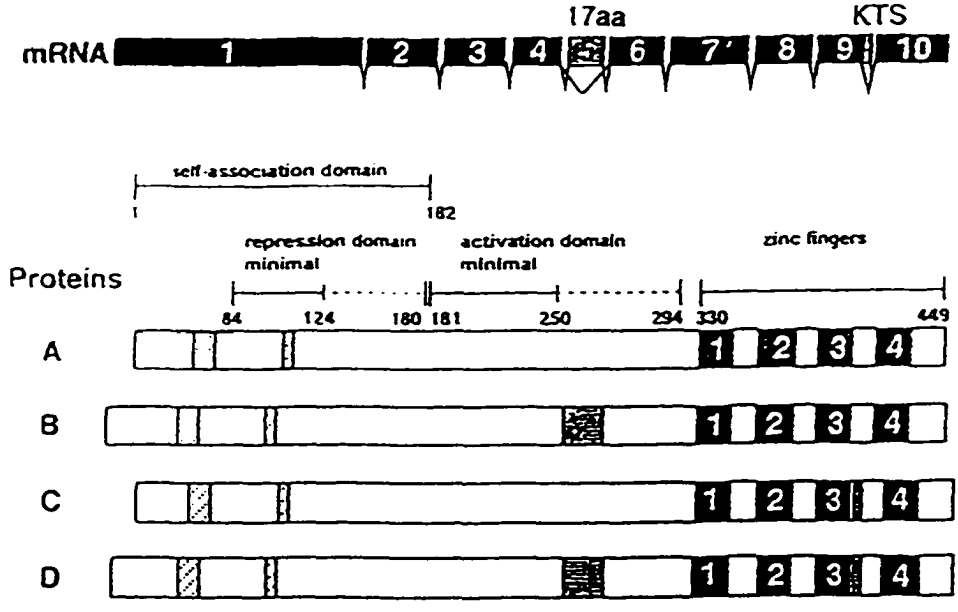


Figure 2.2 Schematic diagram of the structure of the WT1 mRNA and proteins. Stippled boxes represent nucleotides or amino acids that make up the alternative splice insertions. Hatched boxes represent proline-rich regions of the protein. The exons of the mRNA and the zinc fingers of the protein are numbered (Reddy & Licht, 1996).

phosphorylation events inhibited the DNA-binding ability of both proteins, as measured in gel mobility shift assays. Moreover, activation of PKA resulted in the reversal of WT1 suppression of a reporter construct, as well as in alteration in nuclear translocation of the proteins (a change from nuclear to cytoplasmic retention). Thus, phosphorylation of WT1 can have significant and pleiotropic effects on the function of the protein.

WT1 protein also contains potential sites for N-linked glycosylation in the 17 amino acid insertion, as well as a number of potential recognition motifs for a DNA-activated protein kinase (Lees-Miller, et al., 1991). The functional significance of these features needs to be investigated further.

Consistent with its proposed function as a transcriptional regulator, WT1 has been shown to localize to the nucleus (Pelletier et al., 1991; Morris, et al., 1991). Using immunostaining and confocal microscopy, it was recently observed that WT1[+KTS] isoform co-localized predominantly with the splicing factors, while the WT1[-KTS] variant primarily associated with transcription factors (Larsson et al., 1995). It remains to be seen whether WT1[+KTS] or other isoforms of the protein play a role in mRNA splicing.

With respect to the subnuclear localization of WT1 isoforms, Englert et al. (1995) have recently demonstrated that WT1 truncation mutants have distinct subnuclear compartmentalization properties. Using immunofluorescence studies it was demonstrated that WT1[-KTS] had a diffuse pattern of expression throughout the entire nucleus, whereas WT1[+KTS] produced a speckled appearance in subnuclear structures. Similarly, truncation mutants of WT1 exhibited a speckled nuclear distribution. These speckled structures are similar to the interchromatin granules which contain spliceosomal components, and are thought to play a role in spliceosomal assembly. Moreover, the dominant-negative WT1

mutants were shown to physically associate with the wild-type WT1[-KTS] *in vivo*, and thus alter the nuclear localization of the wild-type protein, changing the distribution of WT1[-KTS] form from a diffuse to a speckled pattern and its subsequent functional inactivation. It is of interest to point out that many of the truncation mutations isolated occur in a heterozygous state, with the wild WT1 allele still present (Haber et al., 1990; Pelletier et al., 1991). Pelletier et al. (1991) proposed that these isoforms, deficient for DNA binding, can act as dominant negative mutants. Thus, dominant mutations of WT1 may lead to tumorigenesis without requiring the loss of the wild type allele (Haber et al., 1990). The mechanism by which mutant WT1 proteins can sequester wild-type WT1 is analogous to that proposed for dominant-negative p53 mutants (Michalivitz et al., 1990; Martinez et al., 1991). The difference between the two is that mutant p53 may associate with the wild-type protein in the cytoplasm, while WT1 changes its sublocalization in the nucleus. The molecular mechanism for the dominant-negative effect observed for WT1 truncation mutants was investigated by several groups, and found to be due to protein self-association (Moffett et al., 1995; Reddy et al., 1995). WT1 protein was found to self-associate *in vitro* and in the yeast two-hybrid assay (Reddy et al., 1995). It was found, in fact, that some WT1 truncation mutants can oligomerize more efficiently than wild-type WT1 (Moffett et al., 1995), thus antagonizing the function of the wild-type allele. In addition, the domain of WT1 necessary for self-association was mapped to the N-terminal 182 amino acids (Reddy et al., 1995).

2.1.4 DNA-binding activity of Wilms' tumor protein.

Once it became clear that the WT1 gene encoded a transcription factor containing four zinc fingers of the C_2H_2 class, the search began to identify the

DNA sequence the WT1 protein recognizes. The first DNA-binding sequence was identified by Rauscher et al. (1990). They used a strategy of selecting a DNA-binding sequence out of a pool of randomized synthetic oligonucleotides. The authors used a bacterially expressed recombinant protein encoding the four zinc fingers of the WT1[-KTS] isoform. The resulting selected sequence was very GC-rich, and similar (although not identical) to the EGR-1 consensus binding site 5'-GCGGGGGCG-3' (Cao et al., 1990). Further analysis of the DNA-binding properties of WT1 and EGR-1 proteins, involving competition and binding site mutation experiments have confirmed that EGR-1 and WT1 recognize a common DNA core element (Rauscher et al., 1990) (Figure 2.3 row A). This result was perhaps not so surprising since both proteins share a moderate degree of amino acid sequence similarity (about 60%) in the zinc finger region, with 6 out of 6 residues proposed to be important for DNA recognition conserved between the two proteins (Pavletich and Pabo, 1991). Subsequently, using transient transfection assays, WT1 protein was shown to regulate transcription from promoters containing an EGR-1 consensus binding site (Madden et al., 1991). For quite some time the EGR-1 consensus site was presumed to be a canonical WT1[-KTS] binding sequence. However, the results of other studies aimed at identification of WT1 DNA binding sites have revealed extended and additional DNA binding sequences. These results are summarized in Figure 2.3.

A novel, TC-rich binding site for WT1[-KTS] was identified from DNAse footprinting analysis of the PDGF A-chain promoter by Wang et al. (1993a) (Figure 2.3 row C). This second WT1 binding motif functioned equally well in co-transfection experiments as the EGR-1 consensus sequence (Wang et al., 1993). The TCC repeat sequences were found within a number of

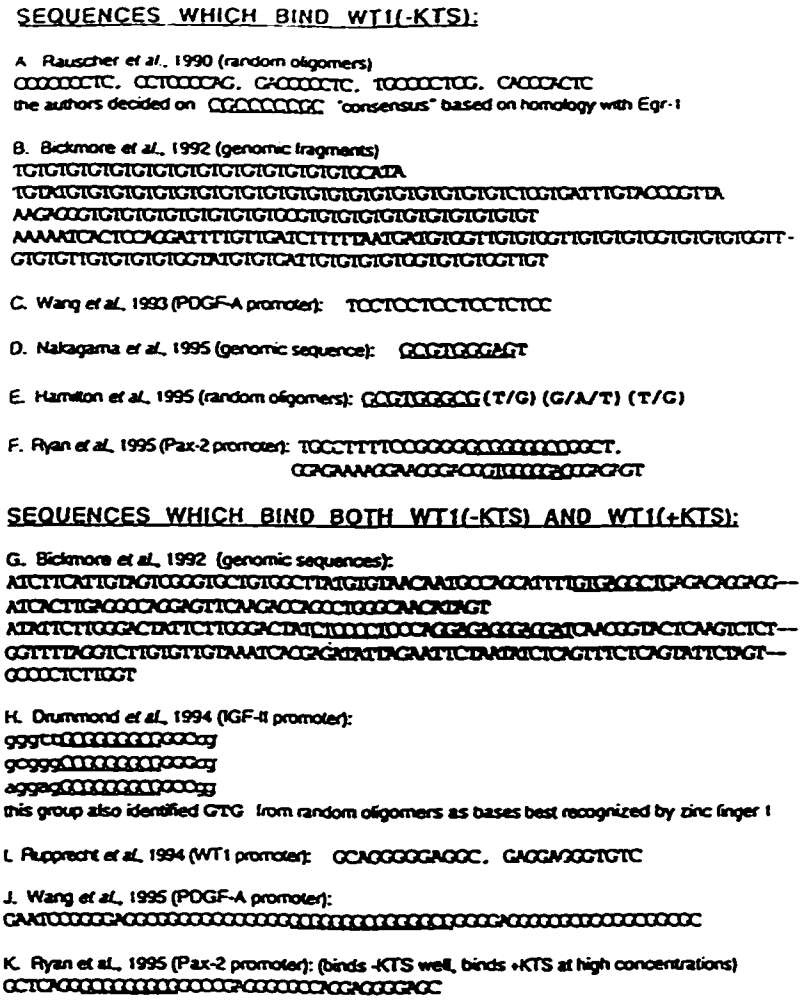


Figure 2.3 DNA sequences which bind WT1 proteins (Reddy & Licht, 1996).

promoters of other growth-related genes, such as Ki-ras, epithelial growth factor receptor, insulin receptor, c-myc, and tumor growth factor β3 (Wang et al., 1993). Thus, as evident from the transient-transfection assays, WT1 can function on a variety of promoter sequences. However, the relevance of the *in vitro* determined sequences to promoter regulation in the chromosomal context has not been determined.

The most abundant splice variant, WT1[+KTS], containing the tripeptide lysine, threonine and serine in the H-C linker region between the third and forth WT1 zinc fingers, does not bind to the EGR-1 core sequence appreciably (Rauscher et al., 1990). Since WT1[+KTS] represents the most abundant isoform of the WT1 protein, subsequent attempts were aimed at identification of a binding site for the WT1[+KTS].

Bickmore et al. (1992) used a whole-genome PCR approach to isolate genomic DNA sequences that bind to either -KTS or +KTS splice variants of WT1. The two WT1 isoforms were shown to have different DNA binding specificities. Sequences isolated with WT1[-KTS] form had runs of GT nucleotides (Figure 2.3 row B). In contrast, WT1[+KTS] did not select any GT-rich sequences. Instead, two extended sequences (+P2 and +P5) of about 100 bp were identified, each of which containing an EGR-1-like consensus binding site (Figure 2.3 row G). Interestingly, DNAse I footprinting or methylation interference analysis failed to identify the precise nucleotide sequences bound by the WT1[+KTS] zinc fingers. In addition, the analysis performed in this study was only qualitative, without quantitative assessment of the binding affinities.

Rupprecht et al. (1994) had carried out footprinting analysis of the WT1 promoter, and found multiple WT1 binding sites, on which WT1 acted as a transcriptional repressor. Two of the WT1 binding sites could bind both -KTS

and +KTS isoforms, and were not related to any other, previously identified WT1 binding sequences (Figure 2.3 row I). However, methylation interference analysis again failed to reveal a distinct pattern of protection due to the WT1[+KTS] binding.

Drummond et al. (1994) analyzed binding of -KTS and +KTS isoforms of WT1 to the fetal IGF-II P3 promoter, which is an active human kidney promoter (Drummond et al., 1992). Footprinting experiments identified a 12-bp sequence which bound both splice variants. This site represents an extended sequence containing features common to the EGR-1 consensus site. A study of PDGF-A promoter also identified a GC-rich sequence capable of binding both -KTS and +KTS variants (Wang et al., 1995) (Figure 2.3 row J). This GC-rich sequence represents an extension of the EGR-1 binding motif. However, the precise location of binding of WT1[+KTS] to this sequence is not clear. Recently, EGR-1-like sequences were also found to be contained in the Pax-2 promoter (Ryan et al., 1995) (Figure 2.3 rows F and K). Three sequences were identified by footprinting assays; two of them bound WT1[-KTS], and the third site could also bind WT1[+KTS] , although only with a low affinity.

Thus, it remains unclear what constitutes a WT1[+KTS] binding site. Therefore, at this point, it wouldn't be prudent to assign a specific DNA-binding function to the WT1[+KTS]. Since it was demonstrated that the two protein isoforms have distinct nuclear localizations, namely, that WT1[-KTS] is found associated predominantly with the transcription factors, whereas the WT1[+KTS] was found with the splicing factors, it is reasonable to suggest that the latter isoform may have a cellular function other than binding DNA.

Since the three zinc fingers of EGR-1 and fingers 2-4 of WT1 have structural similarity as well as recognize a common DNA sequence, it was

assumed that WT1 zinc fingers 2-4 will interact with the EGR-1 binding site in a fundamentally similar manner. Namely, that the structural basis of DNA recognition for WT1 will be the same as found in the EGR-1 X-ray crystal structure, which is discussed in detail in Chapter 1 (Figure 2.4). However, WT1 protein has an additional, first finger which suggests that WT1 may recognize extended DNA sequences.

To date, several groups, including ours, have undertaken studies to identify higher affinity binding sites for the WT[-KTS] protein (Figure 2.3). Drummond et al. (1994) performed footprinting analysis of the IGF-II (insulin-like growth factor) promoter, and identified an extended, 12 bp sequence containing three additional base pairs immediately downstream of the EGR-1 consensus sequence. The authors had shown that both WT1[-KTS] and WT1[+KTS] isoforms could recognize the same 12-bp DNA sequence, albeit with relatively low binding affinities (1.1 X 10^{-05} to 1.4 X 10^{-07} M for WT1[-KTS] and 2 - 7 X 10^{-06} M for WT1 [+KTS]) (Drummond et al., 1994). Additionally, the authors also conducted binding site selection experiments aimed at identification of the WT1 finger 1 recognition subsite. They used a peptide containing only zinc fingers 1-3 of WT1, and found an additional, three-base pair sequence 5'-GTG-3' immediately downstream of the EGR-1 consensus sequence.

Recently, a whole-genome PCR was used to identify physiological high-affinity sites for the WT1[-KTS] (Nakagama et al., 1995). The identified 10-bp sequence contained an EGR-1 - like site (except for an adenine at the eighth position) and an additional thymine flanking the 3' end of the sequence: GCG TGG GAGT (compare to the EGR-1 consensus sequence GCG GGG GCG) (Figure 2.3 row D). This sequence was shown to bind WT1[-KTS] with a 20 to 30-fold higher affinity than the EGR-1 consensus sequence, which suggests

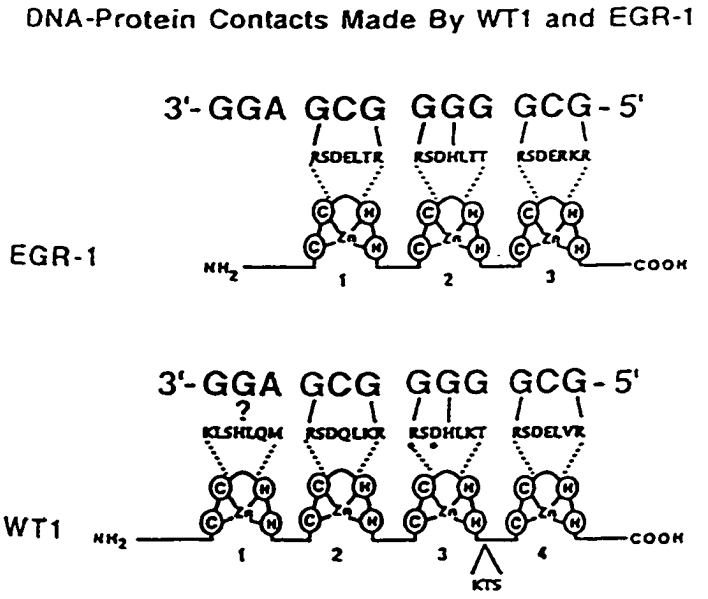


Figure 2.4 Protein-DNA contacts made by EGR-1 protein, and proposed DNA contacts for WT1 zinc fingers. The important recognition helix amino acids are shown. The asterisks on finger 3 of the WT1 protein are residues frequently mutated in Denys-Drash patients (Rauscher, 1993).

that WT1 and EGR-1 protein share related, but not identical recognition sequences.

In parallel with these studies, we conducted a binding site selection and amplification assay in an attempt to identify higher-affinity binding sites for WT1[-KTS] (Hamilton et al., 1995). We determined that the highest affinity binding sites for WT1[-KTS] consist of a 12-bp sequence, including an EGR-1 consensus and a 3-bp subsite recognized by zinc finger 1 (Figure 2.3 row E). These experiments are described in detail in Chapter 3. Our results have recently been once again confirmed in the study by Hewitt et al. (1996). It was shown that zinc finger 1 recognizes an extended DNA binding motif, where guanines constitute the dinucleotide downstream of the 9-bp EGR-1 - like sequence.

On the basis of the above studies, as well as the EGR-1-DNA crystal structural data, it is possible to propose contacts that are formed in the WT1-DNA complex (Figure 2.5). However, more mutagenesis and structural data are necessary to unambiguously assign specific contacts to individual amino acids and DNA base pairs.

In conclusion, despite extensive studies, there is little agreement as to the DNA-binding specificity of WT1[+KTS] protein. The reported dissociation constants (K_ds) for the WT1[+KTS] - DNA complexes are typically several orders of magnitude higher that those for the WT1[-KTS]. The K_ds for the WT1[-KTS] also vary from study to study, perhaps reflecting different sources of proteins, as well as different binding assays and conditions used by different researchers. Furthermore, many groups do not perform quantitative analysis of affinity constants, thus making it difficult to make direct comparisons between different studies.

Recent studies described a specific RNA binding activity for WT1

	ZF4	ZF3	ZF2	ZF1		Protein, Source	K_d	
	R E R	T H R	R Q R	M H K				
5'-	G C G	T G G	G A G	T ? ?	-3'	WT1(-KTS) full-length *in vitro* translated	5×10^{-10}	Nakagama et al., 1995
5'-	G C G	T G G	G C G	G A G T G T T	-3'	ZF(-KTS) bacterial	8.4×10^{-8}	Hamilton et al., 1995
5'-	G C G	G G G	G C G	G G C	-3'	ZF(-KTS) ZF(+KTS) bacterial	1.4×10^{-7} 2.0×10^{-6}	Drummond et al., 1994
5'-	G C G	G G G	G C G	G T G	-3'	ZF 1,2,3 bacterial	N/D	
	G C G	G G G T	G A G C	G A G T G T T C		consensus		

Figure 2.5 Proposed WT1 zinc finger-DNA contacts (Reddy & Licht, 1996).

proteins (Caricasole et al., 1996; Ye et al., 1996). These findings further complicate the question of the biological function of WT1 proteins and, at the same time, introduce a new exciting possibility that one of the important roles of WT1 proteins may be to bind to RNA. It was first shown that the normal nuclear localization of WT1 proteins is not sensitive to deoxyribonuclease, but rather to ribonuclease (Caricasole et al., 1996). Next, the authors report that WT1, but not EGR-1 proteins can bind to specific IGF-II (insulin-like growth factor II) RNA sequences, employing zinc fingers for this sequence-specific interaction. Mutational analysis revealed that zinc finger 1 of WT1, which has no counterpart in EGR-1 protein, is more important for this RNA binding activity. Therefore, a posttranscriptional regulatory function for WT1 was proposed. A study by Ye et al. (1996) demonstrated that WT1 can also bind an antisense RNA, designated WIT-1, which is complementary to WT1 mRNA. Thus, WT1 may be another protein (besides TFIIIA) to regulate gene expression by binding to both DNA and RNA. Identification of other RNA binding targets for WT1 awaits further investigation.

2.1.5 Transcriptional regulatory functions of WT1

Since identification of the WT1 DNA binding site as an EGR1-consensus sequence, studies into transcriptional functions of WT1 began, using transient transfection assays in which WT1 bound to EGR-1 sites. Madden et al. (1991) first demonstrated that WT1 expressed from a vector under control of a cytomegalovirus promoter functioned as a repressor of transcription when bound to a CAT reporter construct containing three contiguous EGR-1 binding sites. It was also shown that DNA binding was necessary, but not sufficient for transcriptional repression. Using a set of

truncated mutant proteins, the repression domain was mapped to the proline- and glutamine-rich N-terminus of WT1. In addition, when the WT1 repression domain was fused to the EGR-1 zinc finger domain, this converted EGR-1 from a transcriptional activator to a repressor (Madden et al., 1991). Thus, a discrete, modular character of the transcriptional effector domain of WT1 was demonstrated, and a transcriptional repression function was first proposed for WT1.

Subsequent studies further defined the transcriptional effector domain (Madden et al., 1993; Wang et al., 1993b). Madden et al. (1993) had demonstrated that WT1 contains a functionally transferable repressor domain by fusing amino acids 84-180 of WT1 to the GAL4 1-147 DNA-binding domain. Wang et al. (1993b) confirmed that amino acid region 84-189 was sufficient to confer repression, while amino acids 180-294 encoded a transcriptional activation domain that activated transcription from the PDGF-A promoter sites (Figure 2.2). Thus, WT1 protein was shown to contain active repression and activation domains. Other transcription factors that have modular transcriptional effector domains are the *Drosophila* proteins Kruppel, even-skipped, and engrailed (reviewed in Levine & Manley, 1989).

Curiously, Madden et al. (1993) found that the minimal repressor domain of WT1 (amino acids 84-180) functioned in NIH 3T3 mouse fibroblast cells, but not in human embryonic kidney 293 cells. To achieve repression in 293 cells, a domain consisting of amino acids 1-298 was required. Therefore, in some circumstances the transcriptional activation domain 180-294 is needed to help the repression function of the 84-180 repressor domain. This observation also suggests that the transcriptional repression function of WT1 is context-dependent, and probably involves interactions with cell-type specific protein co-factors. A point mutation at position 201 of the repressor

domain had been described (Park et al., 1993). This mutation results in a substitution of aspartic acid for glycine, and converts the function of WT1 protein from a transcriptional repressor to an activator.

To date, a variety of natural promoters regulated by WT1 have been identified and analyzed. They include the promoters of the following genes: insulin-like growth factor II (Drummond et al., 1992), insulin-like growth factor I receptor (Werner et al., 1993; Werner et al., 1994), platelet-derived growth factor-A chain (Wang et al., 1992; Gashler et al., 1992; Wang et al., 1993b; Wang et al., 1995), transforming growth factor-β1 (Dey et al., 1994), colony-stimulating factor-1 (Harrington et al., 1993), retinoic acid receptor-α (Goodyer et al., 1995), PAX-2 (Ryan et al., 1995), as well as the WT1 promoter (Rupprecht et al., 1994). All of these promoters contain WT1 binding sites, and are repressed by WT1 when assayed in transient transfection experiments.

The finding that IGF-II and PDGF-A are overexpressed in Wilms' tumors (Reeve et al., 1985; Scott et al., 1985; Fraizer et al., 1987) suggested that deregulated expression of growth factor-encoding genes such as IGF-II at a critical time could contribute to Wilms' tumor development. Further studies of PDGF-A promoter by Wang et al. (1993b) had shown that WT1 can activate or suppress transcription, and that separate domains of the protein were required for either transcriptional activation or suppression. The authors also found that WT1 binding sites both upstream and downstream of the transcriptional start site are required for the suppression function of WT1, whereas activation only requires either an upstream or a downstream site (Wang et al., 1993b).

Promoters such as IGF-II and IGF-I-R (Drummond et al., 1992; Werner et al., 1994) have also been shown to require both upstream and downstream

binding sites for repression by WT1. However, other promoters can be repressed from either upstream or downstream sites, complicating the mechanism of transcriptional regulation by WT1 even further (Madden et al., 1991; Rackley et al., 1993; Madden et al., 1993; Rupprecht et al., 1994; Hewitt et al., 1996). Furthermore, Hewitt and coworkers (1996) propose that multiple WT1 binding motifs are required for transcriptional repression by WT1, and that the multiple binding sites can be upstream of the transcriptional start site.

Further studies of the WT1 repressor activity had defined the roles of alternative splice variants of WT1. The study by Rupprecht et al. (1994) examined the repression abilities of all four WT1 splice isoforms, which were all capable of repressing the WT1 promoter. However, the splice variant containing both the 17-amino acid and KTS insertions (the most abundant splice variant) was the strongest repressor, capable of 58.8-fold repression of WT1 promoter, whereas the isoform carrying the KTS insertion only could confer a weak, 2.2-fold repression of the WT1 promoter construct. In contrast, the 17-amino acid insertion had no added repressor function on the WT1 protein lacking the KTS insertion. It appears that the 17-amino acid insertion converts the WT1 into a powerful repressor. This finding was confirmed in the study of PDGF-A promoter by Wang et al. (1995). Furthermore, it was shown that the 17 amino acid insertion could confer repressor activity when fused to either the zinc finger domain of WT1 or GAL-4. This repression ability was abolished when four consecutive serine residues were deleted from the 17 amino acid peptide. This indicates that the serine residues are important for the repressor activity of the isoforms carrying the 17 amino acid insertion. To confer the transcriptional repression function, the splice variants with the 17 amino acid peptide were bound upstream, downstream,

or on both sides relative to the transcriptional start site. In contrast, the splice variants without the 17 amino acid insertion could only repress transcription when they were bound to both 5' and 3' sites; binding on either side alone resulted in transcriptional activation.

The mechanism of transcriptional repression by WT1 is not clear. If WT1 is bound upstream of the transcriptional start site, perhaps it may interfere with recruitment and/or stability of the initiation complex. One might envision that binding of WT1 protein to the 3' sites could block elongation of newly synthesized RNA molecules. However, it is not known whether WT1 represses transcription at the level of elongation or initiation. On the other hand, the requirement of sites both 3' and 5' relative to the transcriptional start site could mean that functional interaction between WT1 molecules bound on both sites of the transcriptional start is necessary for repression. A similar mechanism has been proposed for the *Drosophila* even-skipped protein, acting on the *Ultrabithorax* promoter, where repression proceeds via a looping mechanism (TenHarmsel et al., 1993).

Recently, WT1 protein was found to physically associate with another tumor-suppressor protein, p53 (Maherwaran et al., 1993). This was shown by coimmunoprecipitation and chemical cross-linking experiments. Transfection assays using EGR-1 promoter and WT1 expression vector, but no p53, demonstrated that WT1 acted as a transcriptional activator. However, cotransfection of a p53 expression vector along with the WT1, resulted in repression of the same promoter. Thus, the association of WT1 with other proteins such as p53 can modulate its regulatory activity. The first two zinc fingers of WT1 were recently shown to be the site of interaction with p53 (Maheswaran et al., 1995).

It should be kept in mind, however, that p53 is known to have pleiotropic effects on transcription, activating or repressing transcription from a variety of promoters in either binding site-dependent or independent fashion (Zambetti and Levine, 1993). The fact that transgenic mice with a disrupted p53 gene undergo normal development, albeit being prone to tumor formation (Donehower et al., 1992; Jacks et al., 1994), while mice lacking the intact WT1 gene (Kreidberg et al., 1993) show prenatal lethality and other developmental abnormalities suggests that not all developmental functions of WT1 require the presence of p53. It remains to be seen what protein factors in different cells determine whether WT1 will activate or repress transcription.

Recently, several groups have provided evidence that WT1 protein may associate with itself (Wang et al., 1995; Reddy et al., 1995; Moffett et al., 1995). Wang et al. first suggested that mutant WT1 proteins may antagonize the activity of the wild-type protein. Reddy et al. had mapped the self-association domain of WT1 to the first 182 amino acids of the protein. When this domain was expressed, or when two other proteins deficient for DNA-binding were expressed, they resulted in inhibition of transcriptional activity of WT1 (Reddy et al., 1995). Using an *in vitro* biochemical assay and the *in vivo* yeast two-hybrid assay WT1 was found to self-associate. It was thus proposed that WT1 mutants act in a dominant-negative fashion, contributing to the tumorigenic development. A similar study by Moffett et al. (1995) confirmed the above findings, demonstrating both *in vitro* and *in vivo* that WT1 can self-associate. The authors used a Denys-Drash syndrome allele to demonstrate that it behaves in a dominant-negative fashion, and therefore proposed a mechanism via which WT1 mutations may act in deregulating cellular proliferation and differentiation.

Despite the accumulated wealth of studies on the mechanisms of transcriptional regulation by WT1 proteins, there are still many unresolved problems. A number of issues need to be addressed in future studies, including the identification of cellular genes that are actual targets for WT1 regulation in both normal and cancerous cells. It is not clear what protein co-factors modulate WT1 activity in particular cell types. It will be necessary to identify more promoters that are controlled by WT1 proteins in order to understand the *in vivo* gene targets. This is complicated by the heterogeneity of the WT1 binding sites identified *in vitro* . Moreover, promoters reported to be regulated by WT1 proteins in transient cotransfection assays may not be relevant when analyzed *in vivo.* Cellular promoters that require WT1 binding sites for tissue-specific expression in cell culture or in transgenic mice need to be identified and analyzed. Finally, transcriptional regulation by WT1 proteins should be studied in the context of native chromatin structure to gain an insight into gene regulatory events as they occur within the nucleus.

The question remains as to the transcriptional effect of WT1: is it at the level of transcriptional initiation, elongation, or both? Does WT1 play a role in posttranscriptional regulation, including splicing of nascent RNA molecules? What proteins of the basal transcriptional machinery does WT1 interact with? How is the expression of WT1 itself differentially regulated to account for the tissue-specific and development-specific expression of the WT1 gene? All these questions need to be answered to provide an understanding of the role of WT1 in both normal development and neoplastic transformation.

CHAPTER 3.0 IDENTIFICATION OF HIGH AFFINITY BINDING SITES FOR THE WILMS' TUMOUR SUPPRESSOR PROTEIN WT1

3.1 Introduction

Despite the fact that both WT1[-KTS] and EGR-1 can bind to the same DNA sequence, it is not clear what the preferred target sequence for WT1[-KTS] is, and how the affinity and specificity of WT1[-KTS] for such sites would compare with competing factors like EGR-1. Recent crystallographic studies of a complex formed between the zinc finger domain of the EGR-1 protein and a consensus DNA recognition sequence (Pavletich & Pabo, 1991; Elrod-Erickson et al., 1996) provided a structural basis for the observations that EGR-1 and WT1[-KTS] recognize the same DNA sequence. These studies identified those critical amino acids which contact specific base pairs in the consensus EGR-1 binding site GCG-TGG-GCG. These contact amino acids are completely conserved in the WT1[-KTS] protein, thus implying a similar role in the binding of the last three zinc fingers of WT1[-KTS] to the nine base pair EGR-1 recognition sequence (Figure 3.1A) .

The crystallographic data indicate that each of the three fingers of EGR-1 interact with a specific 4 bp subsite within the 9 base pair consensus binding site (Figure 3.1B). The yeast MIG1 repressor protein has two zinc fingers corresponding in sequence to fingers 2 and 3 of EGR-1. This protein recognizes a 6 base pair G rich sequence (Nehlin & Ronne, 1990). This trend predicts that WT1[-KTS] will recognize a specific 12 base pair sequence, the first 9 base pairs of which will correspond to the EGR-1 consensus sequence. The effects on WT1-ZFP binding of mutated sequences within and flanking an EGR-1 consensus site suggested that this zinc finger peptide derived from WT1[-KTS] recognizes a longer sequence (Rauscher et al., 1990; Morris et al., 1991).

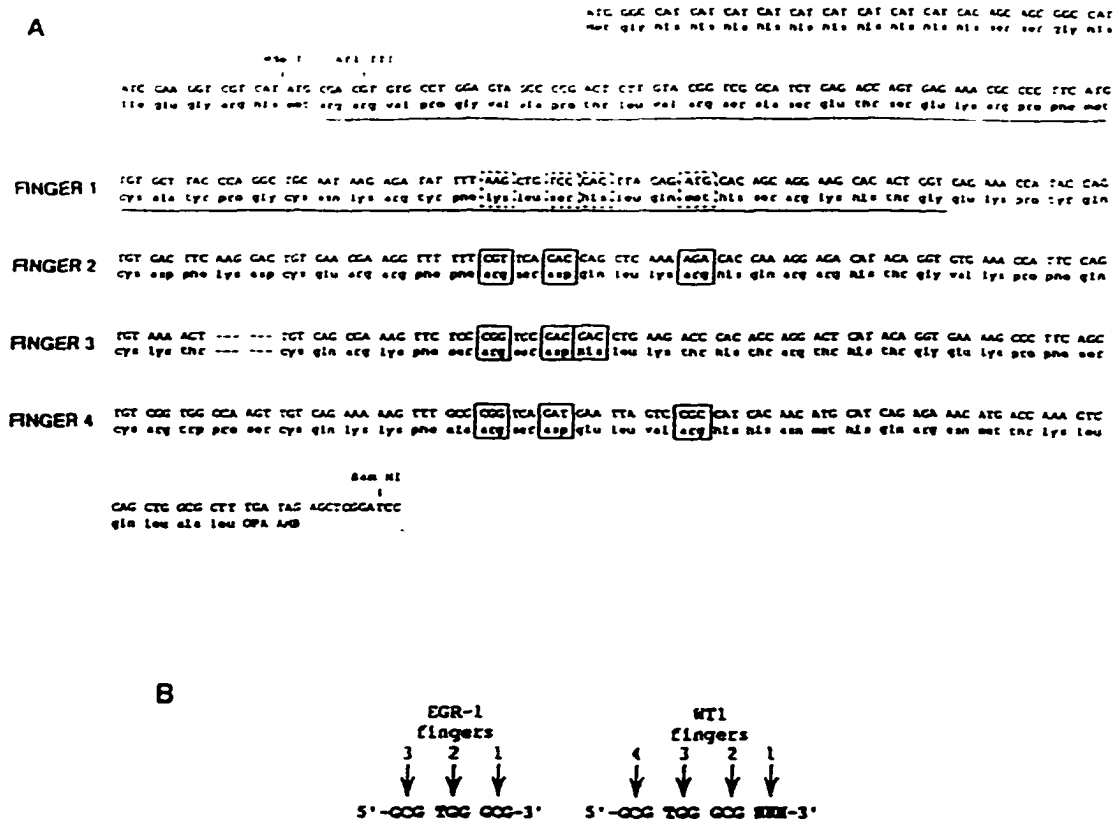


Figure 3.1 DNA sequence of the peptide encoding insert in plasmid pET-WTZFP, with the amino acid sequence of the peptide. The region of sequence underlined is deleted in plasmid pET-WT1ΔF1, the peptide product lacking the first zinc finger of the WT1 protein. Amino acids in fingers 2-4 that are boxed correspond to the amino acids of EGR-1 that are involved in specific DNA recognition, and that in WT1 are proposed to play an analogous role. Amno acids in finger 1 that are enclosed with dashed boxes occupy positions in the alpha helix that could be involved in DNA interactions. (B) Diagrammatic representation oof the zinc finger:3 bp DNA subsite interactions in the EGR-1-DNA complex, and the proposed interaction of the four WT1 zinc fingers with DNA.

To determine the role of the first zinc finger of WT1[-KTS] in DNA binding, I used an *in vitro* binding site selection assay (SAAB) (Blackwell & Weintraub, 1990). From the crystal structure data available for the EGR-1:DNA complex, it was reasonable to hypothesize that the first finger of WT1[-KTS] should contact the 3-5 base pairs immediately downstream of the EGR-1 consensus site (Figure 3.1B). Therefore, an oligonucleotide with flanking primer sites and including the target sequence -GCGTGGGCGNNNNN- was synthesized and used as a template in the SAAB assay. After four cycles of selection/amplification with either WT1-ZFP, a recombinant peptide containing the four zinc finger motifs of WT1, or WT1ΔF1-ZFP, a recombinant peptide containing fingers 2-4 of WT1, high affinity binding sites were cloned and sequenced. The selected high affinity binding sites for WT1-ZFP had the sequence: GCG-TGG-GCG-**(T/G)(A/T/G)(T/G)NN,** whereas the WT1ΔF1-ZFP did not demonstrate any sequence preference at the corresponding DNA binding subsite. These results indicate that the first zinc finger of the WT1[-KTS] protein is indeed required for the recognition of additional specific DNA sequences contiguous with an EGR-1 binding site. A quantitative binding assay was used to demonstrate that WT1-ZFP has approximately a four fold higher affinity for the selected sites versus a nonselected sequence. The ability of a consensus high affinity binding site for WT1[-KTS] to act as a regulatory sequence in mammalian cells was confirmed using a reporter gene assay. DNA sequences in the fetal promoter of the insulin-like growth factor II gene that confer WT1 responsiveness in a transient transfection assay bind to the WT1-ZFP with affinities that vary according to the number of consensus bases each sequence possesses in the finger 1 subsite. All the experiments in this study were performed by myself, with the exception of the oligonucleotide synthesis, which was accomplished

by Dr. Paul J. Romaniuk, and the transient transfection assays, which were carried out by Kathleen C. Barilla.

3.2 Materials and Methods

3.2.1 Bacterial strains and DNA vectors

Plasmids pET-WT1ZFP and pET-WT1ΔF1 contain the coding sequences for the four zinc finger domain of WT1 and the last three zinc fingers of the domain, respectively. The zinc finger cassettes were cloned into the *Nde*I/*Bam*HI restriction sites of the T7 expression vector pET-16b (Studier et al., 1990). The selected DNA sequences were cloned into the *Bam*HI and *Eco*RI restriction sites of pUC19. All DNA vectors were maintained in *E.coli* strain JM109. Protein expression was carried out in *E.coli* strain BL21(DE3) (F-*omp*T $r_b^- m_b^-$).

The mammalian expression vector pCB6+ and its derivative pWT1, which contains a complete cDNA encoding the WT1 tumour suppressor protein lacking either the 17 amino acid insertion in the transcriptional regulatory domain or the three amino acid insertion in the zinc finger domain (Wang et al., 1993), were generously provided by Dr. T.F. Deuel. Plasmid ΔMTV-CAT, which encodes a cDNA for chloramphenicol acetyltransferase (CAT) under control of a mouse mammary tumour virus promoter (Thompson & Evans, 1989) was a kind gift from Dr. R. Evans.

3.2.2 Other Materials

Bacterial cell culture: yeast extract, tryptone and agar were purchased from BDH Inc.

Chemicals: commonly used chemicals were purchased from Sigma Chemical Inc., Fisher Scientific Company or BDH Inc.. Urea was purchased from ICN

Pharmaceutical Inc.. Toluene was purchased from VWR Scientific Inc.. 2, 5-diphenyl oxazole (PPO) was purchased from Sigma Chemical Inc.

Enzymes: Restriction enzymes and T4 DNA ligase were purchased from either New England Biolabs or Pharmacia. Taq DNA polymerase used in PCR reactions was purchased from Promega.

Nucleotides: NTPs and dNTPs were purchased from Pharmacia. Oligonucleotides were synthesized by Dr. Romaniuk using an Applied Biosystems Inc. 391 DNA synthesizer.

Isotopes: [α-^{32}P] dATP (specific activity 3000Ci/mmol) was purchased from New England Nuclear (Du Pont).

Other: Agarose and low melting temperature agarose were purchased from FMC Bioproducts. Gelatin was purchased from Sigma. Poly (dI-dC) was purchased from Pharmacia. XAR-5 film was purchased from Kodak. Oligonucleotide purification cartridges were purchased from Applied Biosystems Inc.. Qiagen DNA purification columns with pre-packed anion-exchange resin were purchased from Qiagen Inc.. Mermaid kit used for recovering DNA from agarose gels was purchased from Bio 101 Inc.. His-Bind resin and columns for protein purification were purchased from Novagen. Nitrocellulose filters were purchased from Millipore.

3.2.3 Construction and Purification of Recombinant WT1-ZFP and WT1ΔF1-ZFP Proteins.

The synthetic genes encoding the four zinc finger domain (WT1-ZFP) and the last three zinc fingers of the domain (WT1ΔF1-ZFP) of WT1[-KTS] were created by PCR, using cDNA prepared from human fetal kidney tissue (Clontech) as a template. For the construction of WT1-ZFP, appropriate primers were synthesized to allow for the amplification of a cassette

incorporating the WT1[-KTS] coding sequence from the unique AflIII site to the termination codon (Figure 3.1A). PCR products were initially cloned by blunt-end ligation into Sma I site of pUC18 (Figure 3.2). The desired clones were identified by DNA sequencing using an ABI 373A automated DNA sequencer, and designated pUC-WT1ZFP. A PCR cassette encoding only the last three zinc fingers of WT1[-KTS] (WT1ΔF1-ZFP) was generated using the pUC-WT1ZFP construct as a template, the original downstream PCR primer and an upstream primer with the sequence GCGCGCGGCCGAATTCATATGGAGAAA CCATACCAGTGT (Figure 3.1A). This PCR product was cloned into *Sma*I site of pUC18 and the correct clone identified by DNA sequencing using an ABI 373A automated DNA sequencer. Once the sequences were verified, the zinc finger cassettes were cloned into the NdeI/BamHI restriction sites of the T7 expression vector pET-16b (Studier et al., 1990) (Figure 3.3) to yield the plasmids pET-WT1ZFP and pET-WT1ΔF1. The resulting synthetic genes encoded zinc finger peptides with a histidine tag and factor Xa cleavage site fused to the amino terminus (Figure 3.1A).

Recombinant proteins were expressed by inducing log phase *E.coli* (strain BL21(DE3)) containing the pET-16b constructs with isopropyl-β-D-thiogalactopyranoside (1 mM). The cells were harvested 3 hours later by centrifugation at 8630 x g in a Beckman JA-14 rotor. The cell pellets were washed with buffer A (10 mM TrisHCl, pH 7.5 at 4 °C, 5 mM $MgCl_2$, 250 mM NaCl, 10 mM PMSF, 10% glycerol, 5 mM DTT, 5 mM imidazole), then resuspended in the same buffer and lysed by ultrasonication (8 X 15 second pulses on ice with 1.5 minute intervals between pulses). Cell debris was pelleted by centrifugation in 15 ml corex tubes at 16,000 rpm for 15 minutes in a Beckman JA-21 rotor, and the supernatant was discarded. The proteins were extracted from the inclusion bodies by incubation overnight at 4 °C in 10 ml

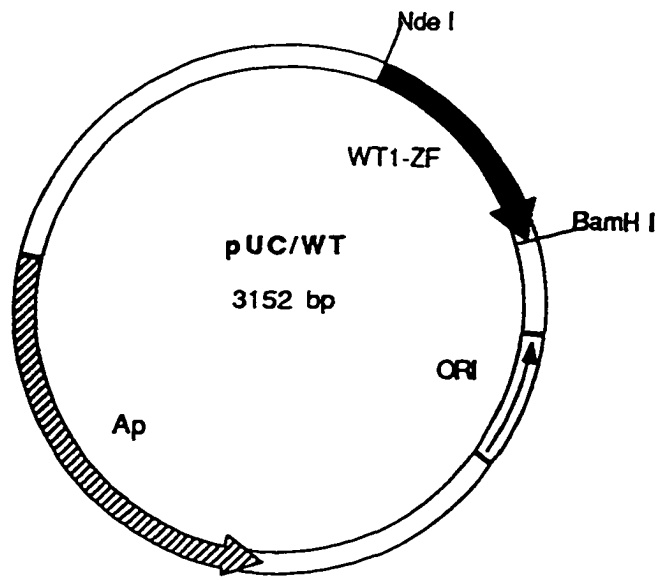


Figure 3.2 Schematic representation of the recombinant WT1-ZF plasmid, pUC18/WT1-ZF. Restriction enzyme sites are indicated.

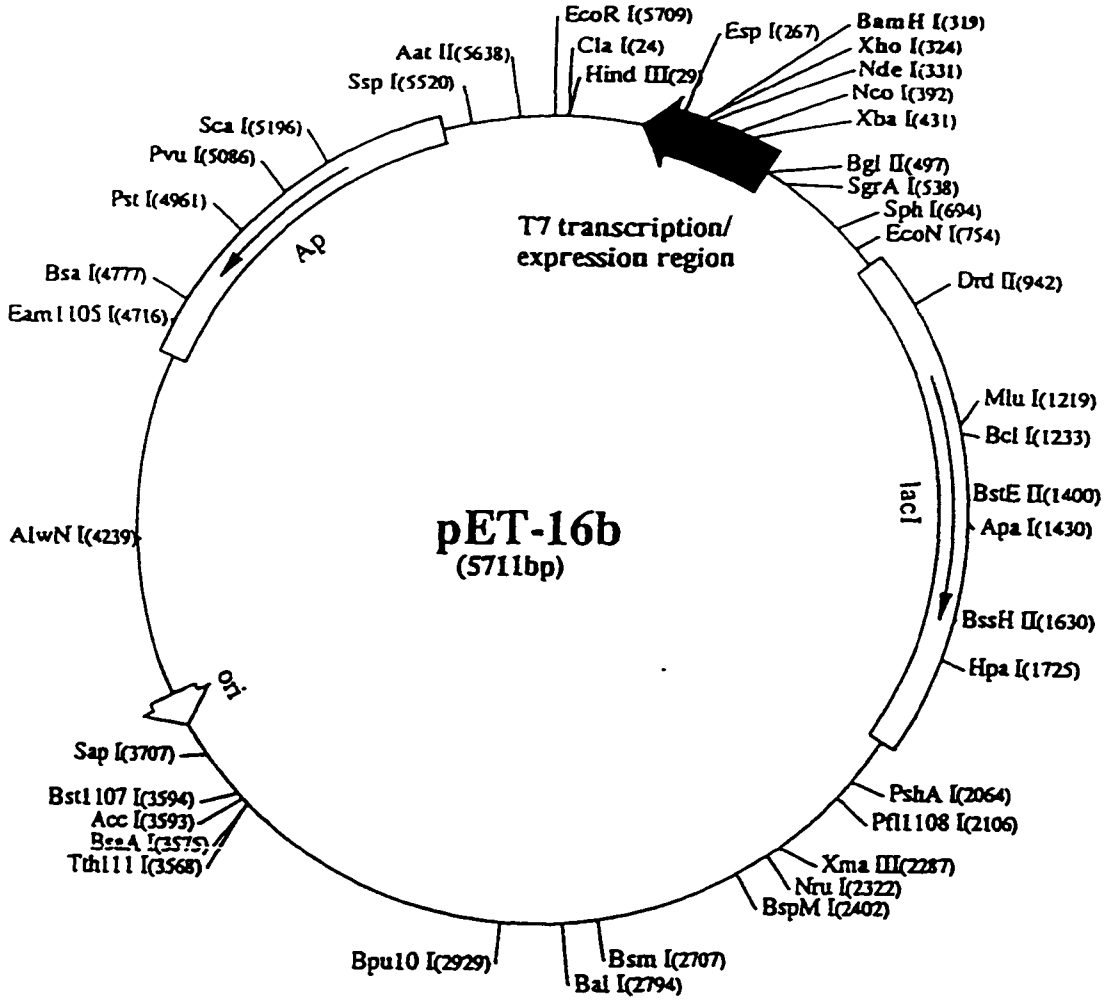


Figure 3.3 Plasmid map of pET-16b expression vector (Novagen technical bulletin). Protein coding sequences were cloned into *Bam*HI/*Nde*I sites.

of buffer B (buffer A containing 5M urea). Recombinant zinc finger peptides were purified to homogeneity using a 1 ml nickel-chelate affinity column. After washing the column with 2 ml of buffer B containing 50 mM imidazole, the recombinant protein was eluted with 1 ml of buffer B containing 150 mM imidazole followed by 1 ml of buffer B with 250 mM imidazole. Protein concentrations were determined by the Bradford method, using BSA as the standard (Bradford, 1976). Protein yield averaged 10 mg per litre of bacterial culture. On SDS-PAGE (15%), the purified WT1-ZFP migrated with an apparent molecular mass of 22 kilodaltons (kDa), and the WT1ΔF1-ZFP with an apparent molecular mass of 17 kDa (Figure 3.4).

3.2.4 Selection Amplification and Binding (SAAB) Assay.

The SAAB assay was used according to the published procedure (Blackwell & Weintraub, 1990) with several minor modifications (Figure 3.5). A 69 bp oligonucleotide was synthesized that included the target sequence -GCGTGGGCGNNNNN- flanked by BamH1 and EcoR1 restriction sites and by binding sites for M13 universal forward and reverse sequencing primers (Figure 3.6).

Primer extension was used to convert the template oligonucleotide into a mixture of 1024 double stranded DNA sequences. The labeling reaction for the initial round of selection had a final volume of 40 µl, and consisted of: 10 pmol of SAAB template, 80 pmol R.U.P. (Reverse Universal Primer), NTB buffer (50mM Tris-HCl, pH 7.2, 10 mM $MgCl_2$, 1 mM DTT, 50 mg/ml BSA). First, the primer was annealed by incubating the mixture at 65°C for five minutes, then slowly cooling to ambient temperature. The reaction was started by adding 8 nmol each of dCTP, dGTP, dTTP, 20 µCi [α-^{32}P]-dATP and 10 units of Klenow fragment of *E.coli* DNA polymerase I. The reaction was

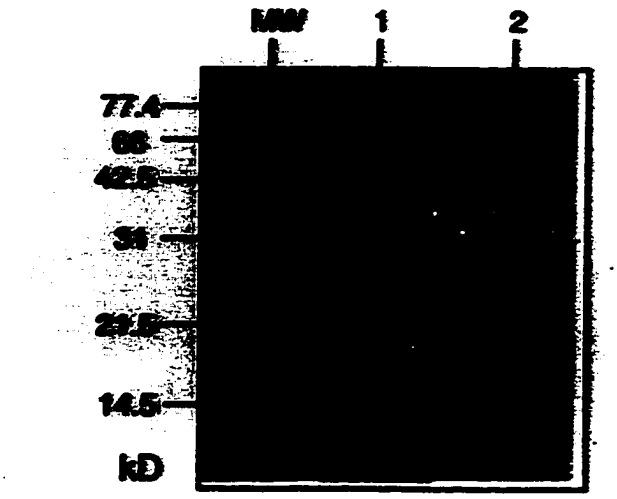


Figure 3.4 Coomassie blue-stained 15% SDS-polyacrylamide gel showing purifiied WT1-ZFP and WT1ΔF1-ZFP proteins. Lane designated MW represents protein molecular weight marker; Lane 1: WT1-ZFP; Lane 2: WT1ΔF1-ZFP.

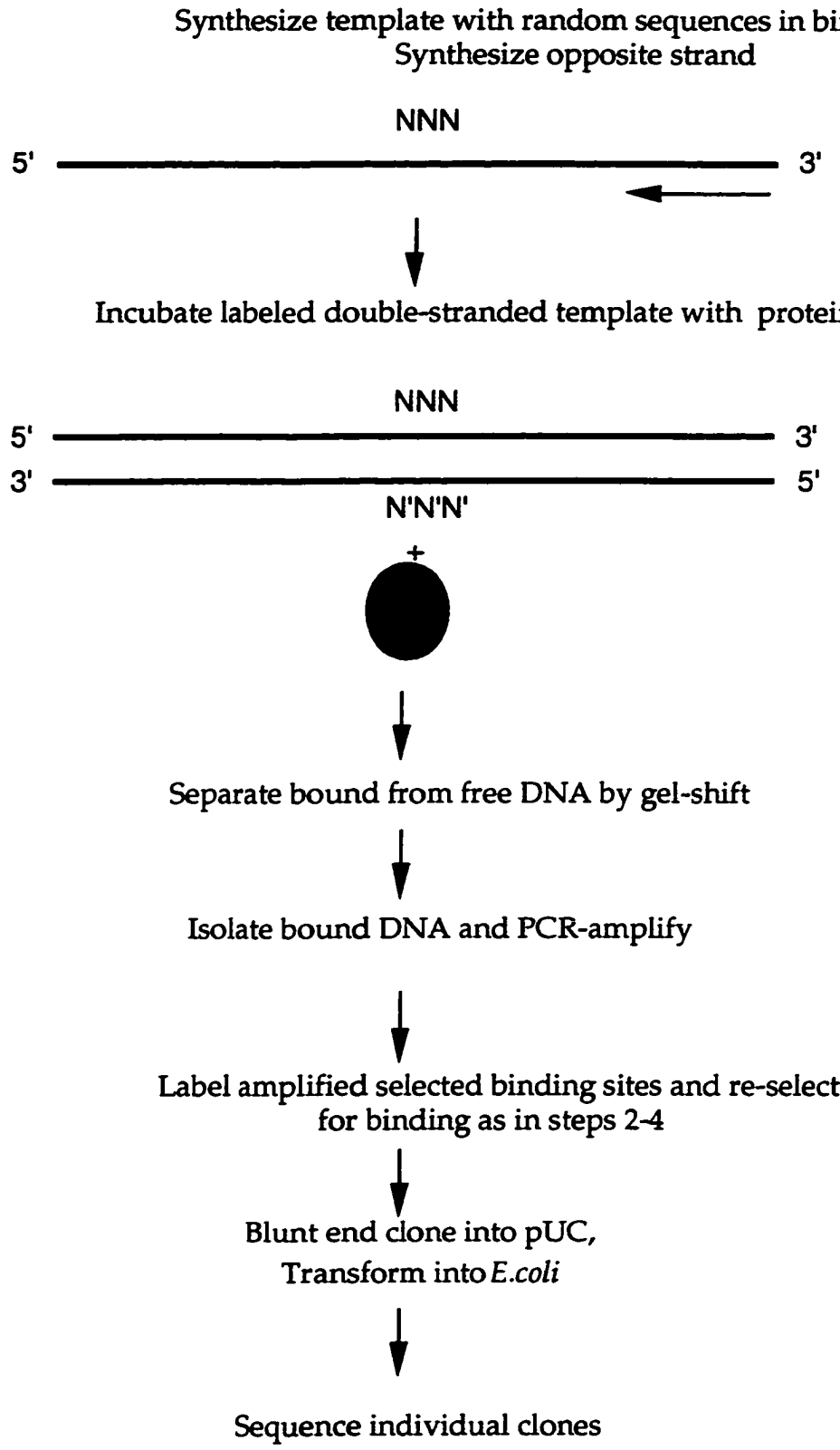


Figure 3.5 Protocol for the SAAB (Selection and Amplification of Binding site assay).

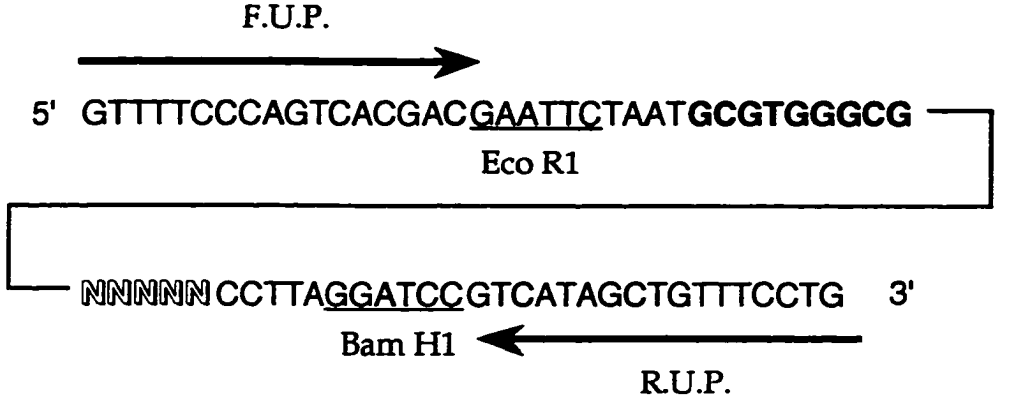


Figure 3.6 The oligonucleotide containing a randomized target sequence for SAAB analysis of WT1[-KTS] finger 1 subsite. Arrows represent the forward (R.U.P.) and reverse (F.U.P.) universal primers. Endonuclease restriction sites are underlined, the EGR-1 consensus sequence is shown in bold, and the randomized sequence is outlined.

allowed to proceed for 45 minutes, then the enzyme was inactivated by incubation at 65°C for five minutes. Following second strand synthesis, the DNA was labeled using [α-^{32}P]-dATP, in the presence of M13 universal primers and the Klenow fragment of DNA polymerase I. The 40 µl labeling reaction consisted of the following: NTB buffer, 40 pmol each of F.U.P. (Forward Universal Primer) and R.U.P., 10pmol DNA template, 40 µCi of [α-^{32}P]-dATP, 8nmol each of dCTP, dTTP, dGTP, and 10 units Klenow fragment of DNA polymerase I. The reaction was incubated for 30 minutes, and stopped with the addition of 8 nmol dATP. The labeled DNA product was purified on a 9% non-denaturing polyacrylamide gel (29:1 crosslinking ratio of acrylamide to bis; 16 x 16 cm x 0.75 mm), and visualized by autoradiography. Gel slices containing labeled DNA were excised, DNA was eluted overnight in 250 µl elution buffer (0.6 M NH_4Ac, 1.0 mM EDTA, 0.1% SDS) (Sambrook et al., 1989). The eluted DNA was recovered by ethanol precipitation, and resuspended in deionized water.

Labeled DNA (10-20,000 cpm, 0.005 pM) was incubated with either WT1-ZFP or WT1ΔF1-ZFP for 20 minutes at room temperature in a buffer containing 20 mM HEPES, pH 7.5, 70 mM NH_4Cl, 7 mM $MgCl_2$, 0.1% NP-40, 10 µM $ZnCl_2$, 2.5 mM DTT, 100 mg/ml BSA, 30 µg/ml poly dI-dC, 6% glycerol. Bromophenol blue and xylene cyanol were added and the reactions were loaded immediately onto a non-denaturing 6% polyacrylamide gel in a buffer containing 89 mM Tris, 89 mM borate, 2 mM EDTA. Following electrophoresis at 200 V for 3 h at 4 °C, the gels were subjected to autoradiography (Figure 3.7). The band corresponding to the DNA-protein complex was cut out the gel, and the bound DNA in the gel slice was eluted by incubation overnight in 250 µl of 0.6 M ammonium acetate, 1 mM EDTA, 0.1% SDS at 37 °C. The DNA was recovered by ethanol precipitation, and

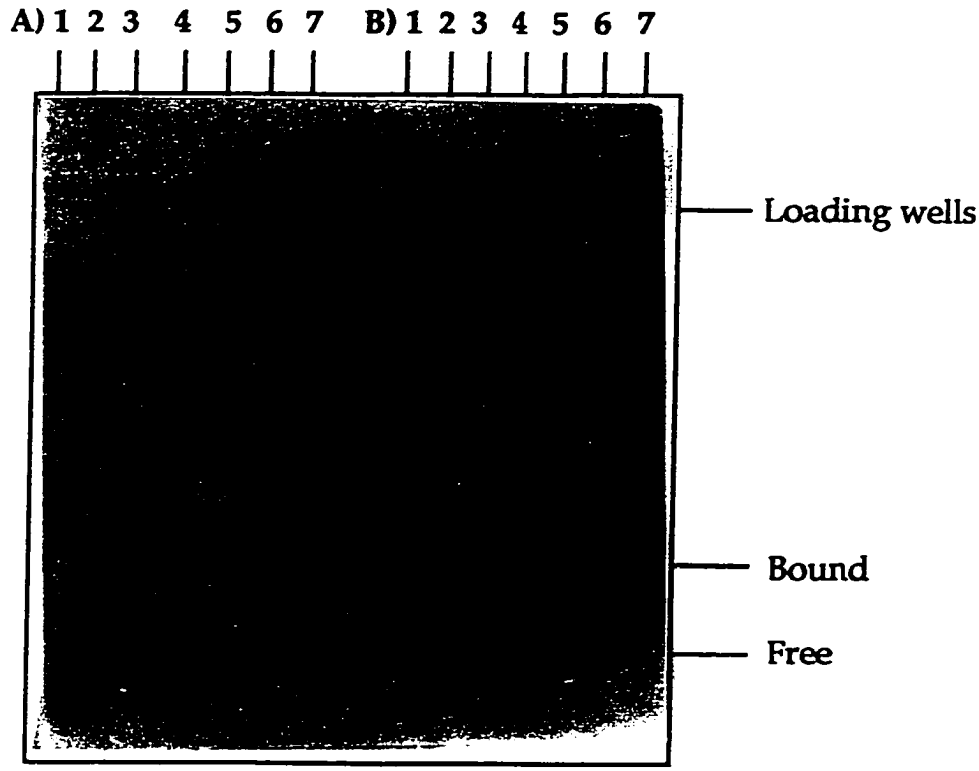


Figure 3.7 An autoradiogramm of round 3 of SAAB. (A) WT1-ZF+DNA template; (B) WT1ΔF1-ZFP+DNA template. Lanes 1-7 represent protein concentrations 0; 1.56; 3.1; 6.25; 12.5; 25 and 50 nM, respectively.

resuspended in 40 µl of deionized water. Labeled DNA was then prepared for the next selection round using the eluted DNA as a template in a PCR reaction (Blackwell & Weintraub, 1990). A standard PCR reaction had a volume of 100 µl, and contained the following: 10 mM Tris-HCl, pH 8.3 at 20°C, 1.5 mM $MgCl_2$, 25 mM KCl, 50 µg/ml gelatin, 10 µl selected template DNA, 20 nmol each of dTTP, dCTP, dGTP, and dATP, 50 pmol each F.U.P. and R.U.P. primers, and 2.5 units Taq DNA polymerase. The reaction was overlaid with an equal volume of mineral oil, and carried out using a Techne PCR thermocycler. The reaction was carried out through 25 rounds of thermal cycling, each cycle consisting of 1.5 minutes at denaturation temperature of 94°C, 1.5 minutes at an annealing temperature of 55°C, and 1.5 minutes at an extension temperature of 72°C. The amplified DNA product was purified by phenol:$CHCl_3$ extraction, and ethanol precipitated with the addition of glycogen as a carrier. The purified DNA was resuspended in 20 µl of deionized water.

In the first round of selection, the protein concentrations were 25-100 nM; in the next three rounds the concentrations were reduced to 6, 3, and 0.75 nM respectively. DNA isolated from the last round was amplified and the product digested with BamH1 and EcoR1 for cloning into pUC19. After transformation of *E. coli* JM109, thirty six to fifty clones were randomly selected, and plasmid DNA prepared. The base sequences of the EcoR I/BamH I inserts in these constructs were determined using an ABI 373A automated DNA sequencer.

3.2.5 Construction of plasmids containing sequence elements of the insulin-like growth factor II fetal promoter, and the non-selected sequence with the CCC subsite for finger 1 binding

The three WT1 recognition sequences identified in the insulin-like growth factor II (IGF-II)gene (Drummond et al., 1992), as well as a non-selected sequence, GCG TGG GCG CCC were created *de novo* by Dr. Romaniuk using an Applied Biosystems Inc. 391 DNA synthesizer. The top strands of oligonucleotides containing promoter elements had the following sequences:
IGF-A: GAT CCG GTT GCG CGG GGG CGA CG
IGF-B: GAT CCG GTT GCG GGG GCG GGC CG
IGF-E: GAT CCA GGA GGC GGG GGC GGC CGG AAG G,
and incorporated a *Bam*HI site for cloning into pUC19.

The top strand of the oligonucleotide with the non-selected sequence had the sequence: AAT TCT AAT GCG TGG GCG CCC ACC CTT AG, and incorporated *Eco*RI and *Bam*HI restriction sites for cloning into pUC19 plasmid. The base sequences of the cloned inserts in these constructs were determined using an ABI 373A automated DNA sequencer.

3.2.6 Purification of oligonucleotides

Synthesized oligonucleotides were deprotected by incubation in 1.0 ml of concentrated ammonium hydroxide for two hours at ambient temperature. The supernatant was separated from the resin by centrifugation for 30 seconds in a microcentrifuge, and incubated at 50°C overnight. The next day, the solution was cooled prior to purification, and half of it was purified as per protocol provided by Applied Biosystems Inc. The resulting elutant was evaporated in a RH40-11 Speed Vac concentrator, and the pellet resuspended in 100 µl deionized water. Oligonucleotide concentrations were determined by absorbance at 260 nm measured using DU-65 Spectrophotometer (Beckman) (taking into account the OD^{260} of a 10 M oligonucleotide solution).

3.2.7 End-labeling of DNA

A series of DNA sequences containing WT1 binding sites were released from plasmid pUC19 using *Eco*RI and *Hind*III and end-labeled with [α-^{32}P]dATP and the Klenow fragment of *E.coli* DNA polymerase I (Sambrook et al., 1989). The fill-in labeling reaction had a final volume of 40 µl, and included the following: 50mM Tris-HCl, pH 7.2, 10mM $MgSO_4$, 0.1 mM DTT, 50 µg/ml BSA, 0.1 mM each dATP, dTTP, dCTP and dGTP, 25 µCi [α-32P]dATP, and 2.5 units of Klenow fragment of *E.coli* polymerase I. The labeling reaction was incubated at ambient temperature for 30 minutes; inactivated by incubation at 65° C for five minutes, then loaded onto a 9% non-denaturing polyacrylamide gel (29:1 crosslinking ratio of acrylamide to bis; 16 x 16 cm x 0.75 mm). Labeled DNA product was visualized by autoradiography, the gel band excised, and the DNA was eluted by incubation overnight in 250 µl of 0.6 M ammonium acetate, 1 mM EDTA, 0.1% SDS. The DNA was recovered by ethanol precipitation, and resuspended in 40 µl of deionized water.

3.2.8 Nitrocellulose Filter Binding Assay

The binding affinities of a series of DNA sequences for the WT1[-KTS] zinc finger peptides were quantified using a nitrocellulose filter binding assay developed to study zinc finger protein-DNA interactions (Romaniuk, 1985; Romaniuk, 1990). The TMK binding reaction buffer contained 20 mM Tris-HCl, 5mM $MgCl_2$, 100 mM KCl, 100 µg/ml BSA, 1mM DTT, 10 µM $ZnCl_2$, and was adjusted to pH 7.5 20°C. The proteins were serially diluted in TMK buffer to give final concentrations ranging from 60 nM to 0.2 nM. Prior to incubation with DNA, the proteins were allowed to equilibrate in the binding buffer for 15 minutes. The binding reactions were initiated by the addition of

approximately 10,000 cpm (approximately 0.005 pmol) of labeled DNA along with the nonspecific competitor poly (dI-dC) (Pharmacia) at 30 ng per binding reaction. After incubation for an appropriate amount of time (minimum 30 minutes), a 180 µl aliquot was removed, and vacuum filtered through pre-soaked 0.45 µm nitrocellulose filters (Millipore). The filters were dried and counted in toluene-based scintillant using a liquid scintillation counter (LKB). Non-specific retention of labeled DNA on the filters was typically in the range of 10-15 % of input, and was taken into account to normalize measurements of complex formation. Experimental data (at least 3 independent determinations per data point) were fitted using a non-linear least squares function: m2 x m1 x m0/(1 + m1 x m0); where m0 = varied ligand concentration; m1 = 10^9, initially; and m2 = 1, initially (Kaleidagraph version 3.0 Synergy Software, Reading, PA) designed for an Apple Macintosh computer. The dissociation constant (K_d) for a simple bimolecular equilibrium (single component pseudo first-order reaction) was determined as a concentration of protein required to give 50 % saturation.

3.2.9 Transient Transfection Assays

The CAT reporter plasmids were constructed by cloning oligonucleotides containing either one or three repeats of the indicated WT1 binding motif into the unique HindIII site upstream of the transcription start of the CAT gene in ΔMTV-CAT. HepG2 cells (ATCC HB8065) were cultured in minimal essential media with Earle's salts, L-glutamine and non-essential amino acids (Gibco BRL) with 1 mM sodium pyruvate and 10% fetal bovine serum. Cotransfection of HepG2 with 8 µg of either pWT1 or pCB6+, 2 µg of CAT reporter plasmid and 3 µg of pSVβ-gal (Promega) in 60 mm plates utilized a calcium phosphate protocol (Chen & Okayama, 1987). Cell extracts

were prepared in 1X Reporter Lysis Buffer (Promega) 48 h after transfection. β-galactosidase activity was measured (Sambrook et al., 1989) and was then used to normalize the cell extracts for transfection efficiency before the extracts were assayed for CAT activity. CAT activity was measured by a non-chromatographic technique utilizing solvent partitioning of the labeled substrate ([^{14}C]-acetyl coenzymeA) from the product ([^{14}C]-acetylated chloramphenicol) (Sleigh, 1986).

2.3 Results

2.3.1 Isolation of DNA binding subsite for finger 1 of WT1-ZF by SAAB.

The selection/amplification assay (Blackwell & Weintraub, 1990) was used to identify the preferred target sequences for the WT1-ZFP peptide, which lacks the KTS insertion between fingers 3 and 4 introduced by alternative splicing. Four cycles of SAAB analysis were carried out to select high affinity binding sites from a low-degeneracy pool containing the target sequence -GCGTGGGCGNNNNN- flanked by PCR primer sites. This oligonucleotide was synthesized assuming, based upon the crystal structure data available for the EGR-1:DNA complex (Pavletich & Pabo, 1991), that fingers 2-4 of WT1 would interact with the EGR consensus sequence GCGTGGGCG. By extension of the analogy to the EGR-1 protein, the first finger of WT1 should contact the 3-4 bases immediately downstream of the EGR consensus site, which is the region randomized in the template oligonucleotide. The WT1ΔF1-ZFP protein which lacks the first zinc finger was used in a parallel selection procedure to serve as a negative control. After the final round of selection and amplification, high affinity binding sites from the WT1-ZFP selection, as well as sequences from the WT1ΔF1-ZFP selection, were cloned into pUC-19 and 36 or more clones of each were sequenced. The

```
                      1  4  7  10  14
                      |  |  |  |   |
SAAB Template:   ..GCGTGGGCGNNNNN..

Sequences of bases 10-14 of fifty clones
            GAGCG          TAACA
                           TAACC
            GANCT
                           TACAC
            GCACC          TACAT
                           TACCC
            GCTCT
                           TAGCA
            GGGCA          TAGGG
            GGGCG          TAGTA
            GGGGT          TAGTG
            GGGTG
            GGGTG          TATCT
                           TATGC
            GGTAG          TATGT
            GGTAT          TATNN
            GGTGA
            GGTGC          TCCAA
            GGTGG          TCCCA
            GGTTG
                           TGACC
            GTAAA          TGACC
            GTACC
                           TGCTA
            GTTGA
                           TGGAC
            TGTAC          TGGGA
            TGTAC          TGGGG
            TGTCA          TGGGG
            TGTCA
            TGTCT          TTTAC
                           TTTCC
            TTGAC
            TTGGA
            TTGGC
```

Figure 3.8 Sequences of bases 10-14 arising from the random SAAB template after four rounds of selection with WT1-ZFP. Fifty clones were isolated after ligation into plasmid pUC18 and sequenced by standard methods. The sequences have been sorted according to their identity at bases 10-12. Nucleotides shown as "N" could not be determined unambiguously, and these were not used in calculating the distribution frequencies shown in Table 3.1.

Table 3.1 Frequencies of Nucleotides Selected by WT1-ZFP and WT1ΔF1-ZFP

	WT1-ZFP[a]				
Position:	10	11	12	13	14
A	0.0 (0.0)	30.0 (13.9)	14.0 (11.1)	24.0 (9.8)	30.0 (43.0)
C	0.0 (0.0)	8.0 (0.0)	12.0 (13.9)	38.0 (16.1)	34.0 (21.4)
G	36.0 (55.6)	46.0 (55.6)	34.0 (38.9)	26.0 (41.9)	22.0 (17.9)
T	64.0 (44.4)	16.0 (30.5)	38.0 (36.1)	12.0 (32.2)	14.0 (17.9)
	WT1ΔF1-ZFP[b]				
Position:	10	11	12	13	14
A	15.2	24.2	6.5	19.3	32.3
C	27.3	33.3	22.5	25.9	19.4
G	27.3	24.2	38.7	38.7	25.8
T	30.3	18.2	32.2	16.1	22.6

[a]Compiled from the sequences of 50 clones chosen randomly after the final selection round. Numbers in brackets compiled from the sequences of 36 clones obtained from the final selection round of a second, independent SAAB experiment.

[b]Compiled from the sequences of 36 clones chosen randomly after the final selection round.

sequences of 50 clones chosen from the final pool of DNAs selected by WT1-ZFP are shown in Figure 3.8. The percentage frequencies of the four bases at each of the five randomized positions are shown in Table 3.1 for both WT1-ZFP and WT1ΔF1-ZFP. A chi-square test of the base distribution at each position indicated that only the distributions at positions 10, 11 and 12 in the WT1-ZFP selection experiments were non-random ($\alpha \leq 0.025$), while all other distributions can be considered random ($\alpha \geq 0.05$). The high affinity binding site for WT1-ZFP had the sequence: GCG-TGG-GCG-**(T/G)(A/G)(T/G)NN**. In contrast, WT1ΔF1-ZFP, that had a deletion of the first zinc finger, showed no preference at all for DNA sequences derived from the five degenerate positions. The reproduciblity of the SAAB results was tested by repeating the selection of sequences with the WT1-ZFP. The results obtained were essentially the same (Table 3.1), with a slight difference in the preference demonstrated for position 11 (G/T rather than A/G). It is interesting to note that only a three base pair sequence was selected for the WT1-ZFP finger 1 binding subsite, while 4 bp recognition subsites are observed for EGR-1 protein.

3.3.2 Quantitative Binding of WT1-ZFP to Various DNA Sequences

A nitrocellulose filter binding assay (Romaniuk, 1990) was used to determine the affinities of the WT1-ZFP for several high affinity binding sites selected in the SAAB assay. A non-selected sequence, GCGTGGGCG**CCC**, was also tested for binding to WT1-ZFP. Figure 3.9A shows typical binding curves measured for WT1 recognition sequences containing TGT and CCC finger 1 subsites, and for an unrelated sequence, a thyroid hormone receptor response element (TRE). The values of the apparent association constants (K_a) for the binding of WT1-ZFP to these DNA sequences can be derived from the data

assuming a simple bimolecular interaction. The results of several independent experiments indicate that the WT1 element with the consensus TGT finger 1 subsite binds WT1-ZFP with an apparent K_a of 8.40±1.21 X 10^8 M^{-1}, while the nonconsensus CCC element has an apparent K_a of 2.10±0.41 X 10^8 M^{-1}. In comparison, the TRE sequence binds WT1-ZFP with approximately two orders of magnitude less affinity (K_a estimated to be ca. 3 X 10^6 M^{-1}). K_a values for a number of the selected high affinity binding sites were measured using the same assay, and these sequences have affinities for WT1-ZFP that are equal to that of the TGT sequence (Table 3.2). In comparison, the WT1 element containing the CCC finger 1 subsite significantly reduced the binding of the WT1-ZFP, decreasing the affinity four-fold compared to the selected, high affinity binding sites. The affinities of the TGT element (K_a = 1.67±0.05 X 10^8 M^{-1}) and the CCC element (K_a = 0.92±0.09 X 10^8 M^{-1}) for the mutant peptide WT1DF1-ZFP, which lacks the first zinc finger of WT1, are almost equivalent (Figure 3.9B). It is apparent that the difference in affinities of selected vs. non-selected sequences for WT1-ZFP is almost completely attributable to the interaction of finger 1 with the base 10-12 subsite defined by the SAAB experiments (Table 3.2). These data give a clear indication of the contribution made by finger 1 of the WT1-ZFP to the overall free energy of DNA binding. These results also indicate that the SAAB assay easily discriminates between elements that differ in their binding affinity for a specific protein by as little as four-fold.

WT1 binding sites have been identified in a number of eukaryotic promoters (Drummond et al., 1990; Gashler et al., 1992; Wang et al., 1992; Harrington et al., 1993; Rupprecht et al., 1994; Dey et al., 1994). To determine the biological significance of the WT1 recognition sequences selected in the

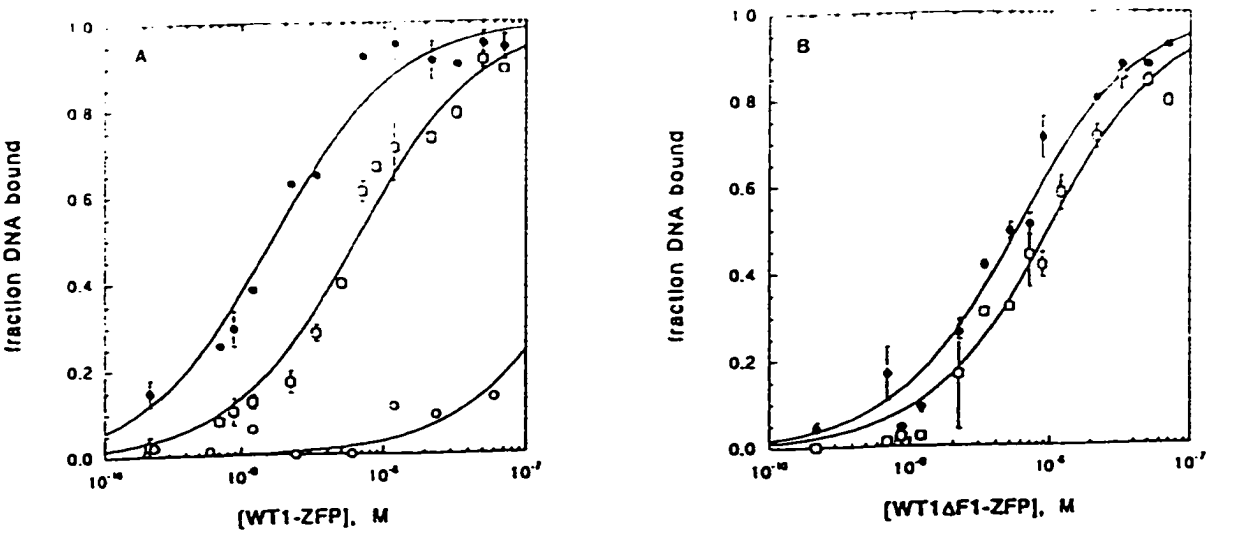


Figure 3.9 **(A)** Results of an nitrocellulose filter binding assay measuring the equilibrium binding of WT1-ZFP to the WT1 element with a TGT subsite for finger 1 (closed circles), the WT1 element with a CCC subsite for finger 1 (open squares) or a thyroid hormone receptor element (open circles). Curves represent the best fit of the data in each case to a simple bimolecular equilibrium. **(B)** Results of an nitrocellulose filter binding assay measuring the equilibrium binding of WT1ΔF1-ZFP to the WT1 element with a TGT subsite for finger 1 (closed circles) and the WT1 element with a CCC subsite for finger 1 (open squares).

Table 3.2 Relative Affinities of WT1 Binding Sites for WT1-ZFP and WT1ΔF1-ZFP

DNA Sequence[a]	Relative Affinity for WT1-ZFP[b]	Relative Affinity for WT1ΔF1-ZFP[b]
TGT	1.00	1.00
CCC	0.22±0.04	0.56±0.07
GGG	0.98±0.26	n.d.
TAG	0.98±0.37	n.d.
GGT	0.81±0.30	n.d.

[a]Sequence of bases 10-12 (finger 1 subsite) are shown

[b]Determined as the ratio of the apparent K_a measured for each binding site to that measured for the TGT binding site in parallel. Mean ± standard deviation reported for two or more independent determinations, n.d. = not determined.

SAAB assay, we compared the binding affinity of WT1-ZFP for three sites identified in the promoter of the insulin-like growth factor II gene (Drummond et al., 1992), with the affinity of the protein for the high affinity TGT target sequence identified by SAAB analysis. As the results in Figure 3.10 show, the three WT1 sites isolated from the *IGF-II* promoter act as high affinity target sequences for the WT1-ZFP, with affinities roughly equal to that of the idealized TGT element. In relative terms, the three elements have affinities for WT1-ZFP in the order (from highest to lowest) of IGF-B≥IGF-E>IGF-A (relative affinites being 1.00, 0.77±0.22 and 0.28±0.08 respectively). This correlates well with the SAAB results: IGF-B has bases 10 and 11 that conform to the SAAB consensus, IGF-E has only 1 consensus base but it is in the critical 10 position, and IGF-A has only 1 consensus base but it is in the less critical 11 position. In summary, our data demonstrated that WT1[-KTS] binds to a 12 bp DNA sequence with high affinity, and that the relative affinities of WT1 binding sites in genomic DNA can be understood based upon this consensus binding sequence.

3.3.3 An *In vitro* Selected WT1 Binding Site Acts as a Strong Transcriptional Regulator.

It has been demonstrated that WT1 can activate, or repress, transcription from a promoter depending upon the number and location of WT1 binding sites (Madden et al., 1991; Wang et al., 1993; Werner et al., 1993). To demonstrate that a high affinity WT1 binding site selected *in vitro* can act as a transcriptional regulator *in vivo*, we tested a series of reporter constructs in which the chloramphenicol acetyltransferase (CAT) gene was fused to the mouse mammary tumor virus LTR promoter containing a variable number of the TGT sites cloned upstream of the transcription start site. This reporter

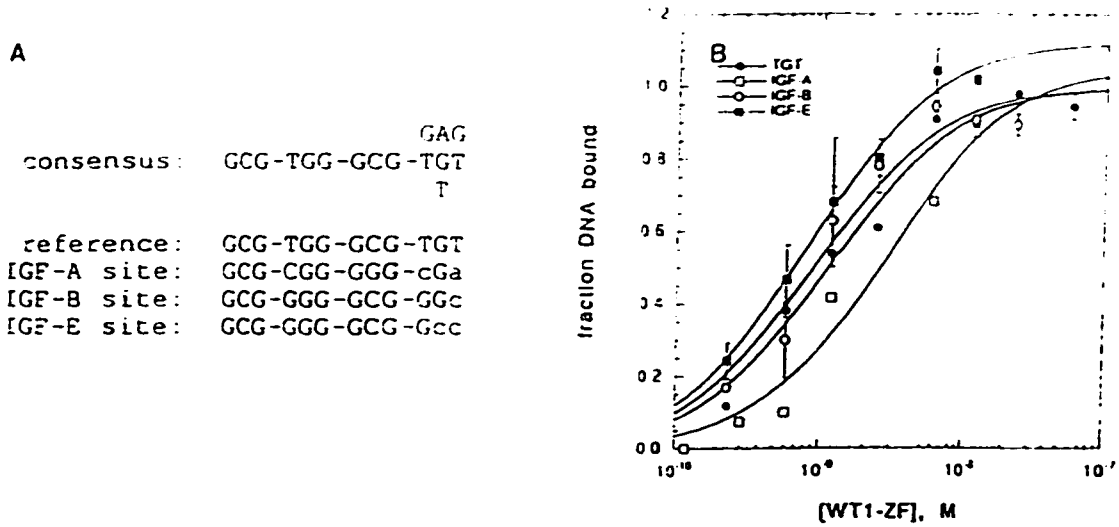


Figure 3.10 (**A**) Sequences of WT1 elements identified in the fetal promoter of the *IGF-II* gene. (**B**) Results of an nitrocellulose filter binding assay measuring the equilibrium binding of WT1-ZFP to the WT1 element with a TGT subsite for finger 1 (closed circles), the IGF-A element (open squares), the IGF-B element (closed squares) or the IGF-E element (open circles). Curves represent the best fit of the data in each case to a simple bimolecular equilibrium.

gene construct has been used to demonstrate the gene regulatory function of thyroid hormone reponsive elements cloned into the unique Hind III site (Thompson & Evans, 1989). Similar tests of the regulatory activity of other WT1 binding sites have utilized an analogous tk-CAT reporter system (Madden et al., 1991). The MTV-CAT reporter constructs were co-transfected into HepG2 cells with a eukaryotic expression vector containing the human *WT1* cDNA, or the expression vector without an insert. After 48 hours, cell extracts were prepared and after normalization for transfection efficiency, were assayed for CAT activity. As the results in Figure 3.11 show, insertion of a single TGT site in the MTV promoter gives rise to a significant stimulation of transcription by WT1 compared to basal expression. Insertion of three tandem TGT sites gives a much larger responsiveness to WT1. This increased responsiveness is primarily the result of a significant decrease in basal promoter activity with the insertion of three tandem sites, compared to the basal promoter activity with the insertion of one site.

A number of the putative WT1 binding sites identified in several promoters do not conform to the consensus site we have identified by the SAAB assay. To confirm that such sites also act as transcriptional regulators responsive to WT1 in this system, we cloned either one or three copies of the lower affinity CCC site into the MTV LTR promoter. As the results in Figure 3.11 show, a single CCC site has approximately the same effect on promoter activity as a single TGT site has. In contrast, three tandem CCC sites provide approximately twice the responsiveness to WT1 than that observed for three tandem TGT sites. In this case, the increased responsiveness observed for three vs. one CCC site is directly the result of a decrease in the basal activity of the promoter: the level of CAT activity expressed as the result of WT1

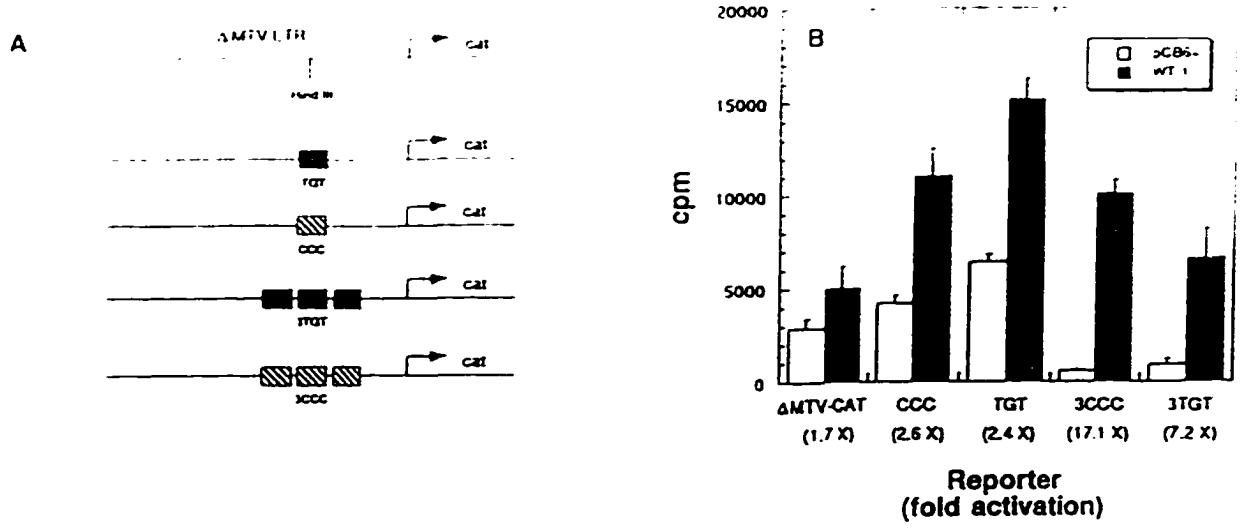


Figure 3.11 (**A**) CAT reporter gene constructs used in transient transfection assays to measure WT1 responsiveness. Boxes represent the number of WT1 elements with the indicated finger 1 subsite sequence inserted into the unique Hind III site of the DMTV-CAT plasmid. (**B**) Assay of chloramphenicol acetyltransferase activity in cell extracts prepared from transient transfections. Each transfection of HepG2 cells included the indicated reporter plasmid, and either the eukaryotic expression plasmid pCB6+ which does not express WT1 (open bars), or pWT1 (closed bars), a pCB6+ derivative that does express WT1. Cpm values are reported as the mean ± standard deviation from 5 or more independent experiments. Fold activation is reported as the ratio of cpm obtained in transfections with pWT1 to the cpm obtained in transfections with pCB6+.

activation is the same whether the MTV promoter contains one or three CCC sites.

3.4 Discussion

The DNA binding domain of WT1[-KTS] consists of four zinc fingers, the last three of which have significant homology to the three zinc fingers of EGR-1. Molecular details of the interaction of the zinc fingers of EGR-1 with a consensus binding site have been resolved by X-ray crystallography (Pavletich & Pabo, 1991; Elrod-Erickson et al., 1996), and are discussed in detail in Chapter 1 of this thesis. The three homologous zinc fingers found in WT1[-KTS] retain the specific amino acids of EGR-1 that are involved in DNA binding. The first, unique zinc finger of WT1[-KTS] might be expected by analogy to interact with a subsite contiguous to the EGR-1 consensus site, implying that WT1[-KTS] recognizes a 12 (or longer) base pair sequence in target promoters.

In addition, the structures of two other zinc finger:DNA complexes have been solved by X-ray crystallography: the SWI:DNA complex (Fairall et al., 1993) and the GLI:DNA complex (Pavletich & Pabo, 1993). The results of these investigations demonstrated that zinc finger interactions with double stranded DNA are not restricted to one specific strand, nor to contiguous three base pair subsites. Interactions can span both strands of the DNA, subsites can be larger than 3 base pairs and can be overlapping, and some zinc fingers might not interact with the DNA at all.

Given this variability in the ways by which zinc fingers can interact with DNA, we designed an experiment to determine the full binding site for WT1[-KTS] that would favour the most likely hypothesis (mechanism of interaction identical to EGR-1) without preventing the identification of other

modes of binding. The template for the selection of high affinity sites was designed to have a random sequence of five base pairs immediately downstream of the nine base pair EGR-1 consensus site. Based upon the antiparallel binding of EGR-1 to its nine base pair consensus site, the first finger of WT1[-KTS] would be expected to interact with a contiguous subsite immediately downstream. To verify that any sequences selected from this template by the complete WT1[-KTS] zinc finger domain were the result of the specific interaction of finger 1 with the DNA, the selection assay was carried out in parallel with a zinc finger peptide containing only fingers 2 to 4 of WT1[-KTS].

The results of this experiment indicate that the first zinc finger of WT1[-KTS], like the three fingers homologous to the EGR-1 protein, binds to a three base pair subsite within a 12 base pair consensus binding site for WT1[-KTS]. This subsite has the consensus sequence (T or G),(G or A or T), (G or T). If the first zinc finger of WT1[-KTS] interacts specifically with the same strand of DNA as the other zinc fingers, it would appear from the selected consensus sequence that this finger interacts preferentially with keto substituents in the major groove of the DNA (Figure 3.12). The absolute discrimination for G or T in the first position (base 10) suggests that this nucleotide may form the strongest interaction with an amino acid side chain within the α-helix of the first zinc finger. Selection at the other two positions was somewhat less stringent (and varied from experiment to experiment for base 11), and so it is not clear exactly what role these two positions might play in protein binding: they could provide specific DNA-protein contacts, or they might provide a DNA sequence context that optimizes the interaction of finger 1 with the DNA.

While this work was in progress, a report was published describing the

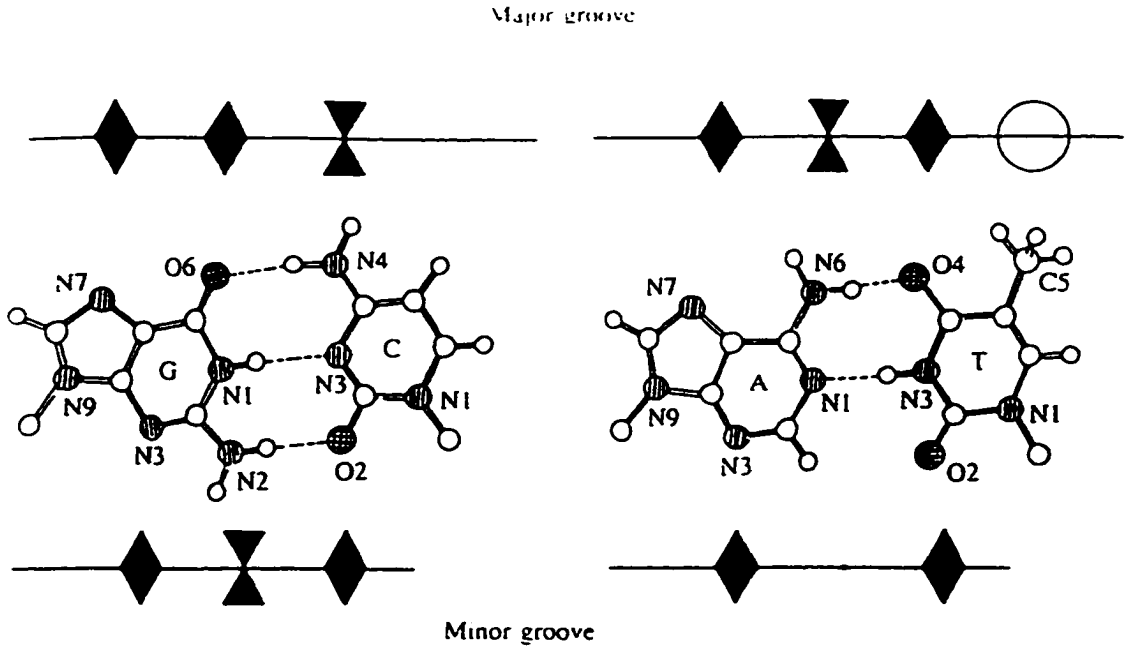


Figure 3.12 The hydrogen-bond donors and acceptors presented by Watson-Crick base pairs to the major groove and the minor groove (Steitz, 1990). Two inverted triangles represent hydrogen bond donors; diamond-shaped symbols represent hydrogen bond acceptors; a circle - a methyl group.

results of a SAAB experiment conducted with a recombinant peptide encompassing fingers 1-3 of WT1[-KTS] (WT1ΔF4-ZFP) (Drummond et al., 1994). Apparently selection experiments conducted by this group with the full four finger domain did not result in the selection of any specific sequences, most probably because of the high protein concentrations (≥5 μM) used in the assay. However, when the experiment was repeated with the WT1ΔF4-ZFP protein, a final selection round at 10 μM peptide yielded a strong selection of a three base pair subsite with a GT(C or T) consensus sequence. My data shows a somewhat less stringent sequence was selected at approximately 10,000-fold lower concentrations of a full four finger peptide. Nevertheless, the sequence selected with the three finger WT1ΔF4-ZFP peptide is a member of the consensus sequence we identified with the full zinc finger domain of WT1[-KTS].

While the SAAB experiment is useful for identifying high affinity binding sites for a DNA binding protein, it provides little insight into the relative affinities that consensus sequence sites have for the protein in comparison to nonconsensus sequences. We used a nitrocellulose filter binding assay to quantify the affinities that WT1-ZFP and WT1ΔF1-ZFP had for a number of the selected sequences, and the non-selected sequence GCG-TGG-GCG-CCC. Four of the selected sequences chosen at random had equal affinities for WT1-ZFP, while the non-selected CCC element had a four-fold weaker affinity. In comparison, the consensus TGT sequence and the non-selected CCC element had roughly equal affinities for the WT1ΔF1-ZFP. Therefore the target sites identified by the SAAB experiment for WT1-ZFP have 4 fold higher affinities for the protein than a non-selected site as the result of specific interaction of the first zinc finger with DNA.

In addition to providing information on the relative affinities of selected vs. non-selected sequences for WT1-ZFP, the nitrocellulose filter binding assay also provides a measure of the apparent association constant for the equilibrium binding of this peptide to DNA. A previous study reported the equilibrium constants measured by a gel mobility shift assay for the binding of a WT1 zinc finger peptide to a 12 base pair WT1 selected sequence to be 7.1 X 10^6 M^{-1} vs. a value of 4.5 X 10^6 M^{-1} for a non-selected sequence including the 9 base pair EGR-1 consensus site (Drummond et al., 1994). The zinc finger peptide WT1-ZFP that we have constructed binds to a selected DNA sequence with a K_a value measured by a nitrocellulose filter binding assay of 8.40±1.21 X 10^8 M^{-1} vs. 2.10±0.41 X 10^8 M^{-1} for a non-selected sequence. In the gel mobility shift assay that was used as part of the SAAB experiment, we observed strong binding of 0.75 nM WT1-ZFP to selected DNA sequences in agreement with the affinities measured by the filter binding assay. The values for equilibrium constants that we measure are about two orders of magnitude higher than those reported earlier, and consistent with the reported equilibrium binding constant of 1.7 X 10^8 M^{-1} for the interaction of the zinc finger domain of EGR- 1with its consensus DNA sequence (Pavletich & Pabo, 1991). Thus the zinc finger domain of WT1 binds with high affinity to a twelve base pair consensus sequence.

Putative WT1 binding sites have been identified in a number of native promoters, and it has been demonstrated by transient transfection assays that these promoters can be regulated by WT1[-KTS] (Drummond et al., 1990; Gashler et al., 1992; Wang et al., 1992; Harrington et al., 1993; Rupprecht et al., 1994; Dey et al., 1994). However, many of these putative WT1 recognition sites do not conform completely to the finger 1 subsite that we have identified by the *in vitro* SAAB assay. To determine whether these naturally occuring

sites bind WT1 with high affinity, we measured the affinities of three WT1 sites from the fetal promoter of the IGF-II gene for WT1-ZFP. These three sites had affinities for WT1-ZFP that were similar to the affinity of the consensus TGT selected site. However, the three IGF sites were not equivalent in their affinities for WT1-ZFP, but varied in affinity in relationship to their divergence from the finger 1 subsite consensus sequence. These results link the *in vitro* SAAB experiment to biologically relevant sites in genomic DNA, and again demonstrate the contribution of the first zinc finger of WT1 to the interaction of the protein with DNA.

WT1 acts as a transcriptional regulator *in vivo*, and natural selection will have designed WT1 recognition sites based both on DNA binding affinity and this regulatory function in a way that the *in vitro* SAAB assay is unable to. To probe this relationship, we compared the regulatory action of the selected, high affinity TGT site with the regulatory action of the non-selected, lower affinity CCC site in the MTV-CAT reporter gene system. In a single copy, both sites confer a similar positive regulatory response of the MTV LTR promoter in this reporter to WT1. This result is consistent with reports that the WT1 protein acts as a transcriptional activator on promoters that contain binding sites located exclusively upstream or downstream of the transcription start site (Wang et al., 1993). However, when the sites are present in three copies, it becomes apparent that the lower affinity CCC site confers greater responsiveness to WT1 than does the high affinity TGT site. This result is intriguing because it suggests that the strength of the interaction of finger 1 with DNA may in some way modulate the ability of the protein to regulate transcription. In addition, the multiple copies of either WT1 site significantly dampen the basal level of transcription from the MTV promoter, another

effect that may be active in natural selection. Further investigation will be required to completely understand these phenomena.

In summary, our data have demonstrated that WT1[-KTS] binds to a 12 base pair DNA sequence with high affinity, and that the relative affinities of WT1 binding sites in genomic DNA can be understood based upon this consensus binding sequence. Investigation of the specific nature of amino acid-DNA interactions of finger 1 of WT1 will enhance further our understanding of the structure and function of this tumor suppressor protein.

CHAPTER 4.0 WT1 MUTATIONS AND THE DENYS-DRASH SYNDROME

4.1 Introduction

Denys-Drash syndrome (DDS) is a rare human disease. Its symptoms include Wilms' tumor, genital anomalies and nephropathy (Drash et al., 1970). Heterozygous WT1 point mutations have been identified in over 95 % of DDS patients, pointing to a crucial role WT1 plays in the genesis of the Denys-Drash syndrome (Pelletier et al., 1991). The WT1 mutations observed in DDS patients can be classified into four categories: (1) missense mutations of DNA-binding amino acids - the most common group of mutations; (2) substitution of zinc binding amino acids; (3) nonsense mutations leading to a frameshift and removal of entire zinc fingers, and (4) early termination. These mutations are summarized in Figure 4.1 A.

The most common mutations affect the key amino acids proposed by analogy to EGR-1 to contact DNA. Arginine 394 appears to be a hot spot of the missense mutations reported in 19 of 38 cases (50 %) of DDS patients (Pelletier et al., 1991; Coppes et al., 1992; Bruening et al., 1992; Baird et al., 1992; Sakai et al., 1993; Akasaka et al., 1993; Baird & Cowell, 1993; Nordenskjold et al., 1994). This mutation is located in exon 9, which encodes zinc finger 3 of the WT1 protein (Figure 4.1B). Other mutations identified in exon 9 involve aspartate 396 substitutions (reported in 6 individuals) (Bruening et al., 1992; Baird et a., 1992b; Nordenckjold et al., 1994; Pelletier et al., 1991; Little et al., 1993), as well as frame shift (at position 386; Ogawa et al., 1993), and truncation mutations (at position 390; Little et al., 1992; Baird et al., 1992a; Quek et al., 1993; Varanasi et al., 1994). In addition, a mutation interfering with splicing at exon 9 was isolated (Huff et al., 1991). Mutations mapping to exon 8 (zinc finger 2) involve arginine 366 (Henry et al., 1989; Bruening et al., 1992), histidine 373 (Little et al., 1993; Sharma et al., 1994), and histidine 377 (Pelletier et al., 1991).

A

Mutation	No cases
340 C -> Y	1
348 S -> Y	1
355 C -> Y	1
360 C -> Y	1
360 C -> G	1
361 frameshift	1
362 R -> stop	5
365 S -> F	1 (rat)
366 R -> H	2
366 R -> C	1
373 H -> Q	1
373 H -> Y	1 (rat)
377 H -> Y	1
377 H -> R	1
384 frameshift	1
386 frameshift	1
390 R -> stop	3
394 R -> W	18
394 R -> P	1
396 D -> N	5
396 D -> G	1
432 frameshift	2

B

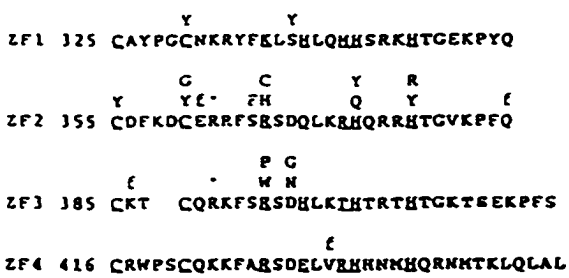


Figure 4.1 (**A**) WT1 mutations associated with Denys-Drash syndrome. (**B**) Point mutations and frameshifts in the zinc finger region of WT1. *, mutation to stop codon; f = frameshift after the indicated amino acid (Reddy et al., 1996).

Cysteine 330 to tyrosine transition was observed in zinc finger 1 (Huff et al., 1991).

Thus, most mutations described are single amino acid substitutions clustered in the zinc finger region of WT1 protein. They were therefore postulated to interfere with DNA binding of WT1 by either disrupting the DNA-binding activity, or lowering the affinity of the WT1 for DNA. The DDS phenotype elicited by the loss of DNA binding can also be explained by altered WT1 dosage. Given that the levels of WT1 protein are very tightly regulated during development, the disruption of the dosage of functional WT1 protein may precipitate the observed DDS symptoms.

It is intriguing that the clinical manifestations of DDS syndrome in different patients do not clearly correlate with any specific type of WT1 mutation detected. Individuals with identical mutations in WT1 protein exhibit the entire spectrum of clinical symptoms. The simplest explanation may be that other genes influence the phenotypic expression of various WT1 mutations. Finally, there is a case reported of an individual who carries a WT1 mutation, and yet is phenotypically normal (Coppes et al., 1992). This individual's son has DDS syndrome, and carries the same mutation as the father: exon 9 arginine 394 to tryptophan missense mutation, the most common mutation found in DDS patients (Coppes et al., 1992). This observation suggests that even though most DDS patients have constitutional WT1 mutations, not all of them result in the disease. A possible explanation for this observation is that perhaps the mutant WT1 allele is not expressed during the crucial stages in development. The genetic background and other epigenetic factors of an individual may also influence the effects of a reduction in WT1 gene dosage. In addition, WT1 was suggested to play a more important role in male genitourinary development, since genetic

females with WT1 mutations associated with DDS have a significantly lower frequency of urogenital abnormalities than males (Coppes et al., 1993).

The majority of mutations in DDS patients have been found in a heterozygous state, suggesting that the DNA-binding - deficient isoforms may act in a dominant-negative fashion. This hypothesis was first proposed by Pelletier et al. (1991). Later, Bradeesy et al. (1994) demonstrated that a protein containing amino acids 1 - 222 represents a minimal domain required for the dominant negative effect. Reddy et al. (1995) confirmed that protein containing amino acids 1 - 182 interferes with transactivation by wild-type WT1, acting in a dominant-negative fashion. Thus, dominant negative WT1 alleles could inhibit the WT1 transcriptional activation function, eventually leading to progressive renal failure. Another mechanism by which mutant WT1 isoforms may alter appropriate gene expression is through conferring new DNA-binding specificity, thus resulting in a change of gene networks regulated by WT1 protein.

To test this hypothesis, we conducted selection and amplification binding assays (SAAB) for the Denys-Drash mutant proteins. Five different mutations were introduced into the zinc finger domain of WT1 (Figure 4.2). Two missense mutations in the second zinc finger substitute arginine 366 with a histidine or a cysteine. This substitution was proposed to disrupt an arginine-guanine interaction, based on analogy with the EGR-1 - DNA complex (Pavletich and Pabo, 1991). The remaining three mutations are in finger 3 of WT1, and substitute arginine 394 with a tryptophan, and aspartic acid 396 with either an asparagine or a glycine residue. DNA templates were randomized for finger 2 and 3 binding subsites. DNA binding affinities of the Denys-Drash mutant proteins for the selected sequences were measured using a nitrocellulose filter binding assay. This work is a collaborative study with

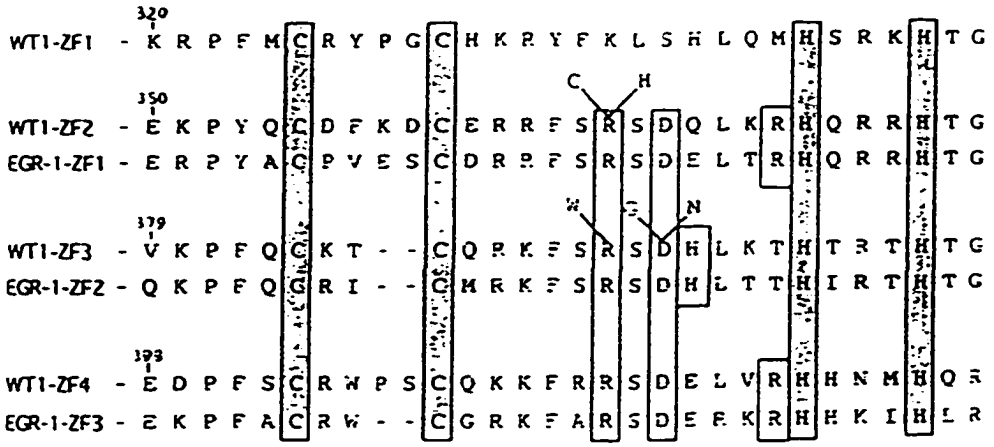


Figure 4.2 Comparison of the sequences of WT1 and EGR-1 zinc fingers. Residues involved in DNA binding and their counterparts in WT1 are enclosed in clear boxes. Shaded boxes indicate residues that coordinate to zinc. Finger 2 and 3 Denys-Drash missense mutations are also indicated.

Frank Borel, Kathleen Barilla and May Iskandar from Dr. Romaniuk's research group. I performed site-directed mutagenesis and expression and purification of the WT1-ZF Denys-Drash protein mutants, as well as the initial DNA-binding assays with the mutant proteins.

4.2 Materials and Methods

4.2.1 Construction and expression of Denys-Drash mutant proteins

The WT1-ZFP Denys-Drash mutants (R366H, R366C, R394W, D396G, D396N) were constructed by site-directed mutagenesis (Nelson & Long, 1989). The plasmid pUC WT1ZFP which contains the zinc finger domain of WT1 cloned into pUC19 (construction described in Chapter 3), was used as a template and the following mutagenic primers were used. The base alteration that introduces each mutation is underlined:

R366C: AGG TTT TCT TGT TCA GAG CAG

R366H: AGG TTT TCT CAT TCA GAG CAG

R394W: AGG TCC TCC TGG TCC GAC CAC

D396G: TCC CGG TCC GGC CAC CTG AAG

D396N: TCC CGG TCC AAC CAC CTG AAG

DNA sequencing using an ABI 373A automated DNA sequencer was used to verify that only the desired mutation was introduced into the WT1-ZFP cDNA. Mutant cDNAs were cloned into the *Nde* I and *Bam*H I sites of the T7 expression vector pET16b (Studier & Moffatt, 1986, Studier et al., 1990). The resulting constructs encode zinc finger peptides with a histidine tag and a factor Xa cleavage site fused to the amino terminus and were used to transform *Escherichia coli* BL21 (DE3) pLysS. Expression and purification of the zinc finger proteins were carried out as described previously (Chapter 3). Protein concentrations were determined by the Bradford method, using BSA

as the standard (Bradford, 1976). SDS-polyacrylamide gel analysis was used to determine protein purity (Figure 4.3). Protein yield averaged 10 mg per litre of bacterial culture. Fractional activity of each protein preparation (75%-100%) was determined by a saturation DNA binding assay.

4.2.2 Selection amplification and binding (SAAB) assay

The SAAB assay was carried out according to a published method (Blackwell & Weintraub, 1990), with minor modifications (Figure 3.5). The 67 base pair template oligonucleotides contained a central region including the WT1 consensus sequence randomized in either the finger 2 or finger 3 recognition sites (5' TAAT GCG TGG **NNN** TGT CCTAA 3' and 5' TAAT GCG **NNN** GCG TGT CCTAA 3' respectively) flanked by *Eco* RI (5') and *Bam*H I (3') recognition sequences and the M13 forward universal primer (F.U.P.) sequence (5') and a sequence complementary to the M13 reverse universal primer (R.U.P., 3') (Figure 4.4).

The template was labeled for the initial round of selection by primer extension using 10 pmol of the oligonucleotide annealed to 80 pmol R.U.P. in 1X Klenow buffer (New England Biolabs), and extended in 0.2 mM dGTP, 0.2 mM dCTP, 0.2 mM dTTP, 20 μCi [α-^{32}P]-dATP and 9 units Klenow (New England Biolabs) with incubation at room temperature for 45 minutes. The resulting double-stranded DNA containing a mixture of 64 possible sequences in the randomized region was purified on a 12% non-denaturing polyacrylamide gel, and eluted overnight at room temperature in 250 μl of 0.6 M ammonium acetate, 1 mM EDTA, 0.1% SDS.

Labeled DNA (200,000 cpm) was incubated with the proteins as described previously (Chapter 3), except the incubation mixture used Tris rather than HEPES to buffer the solution. The reactions were loaded onto a 5% non-

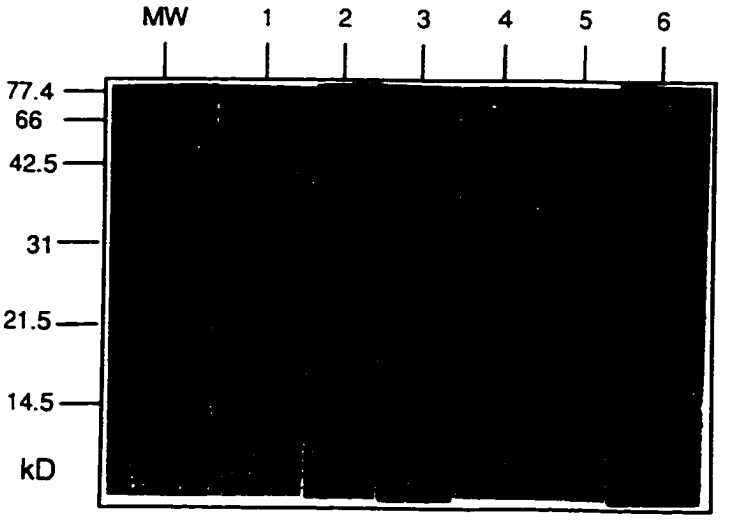
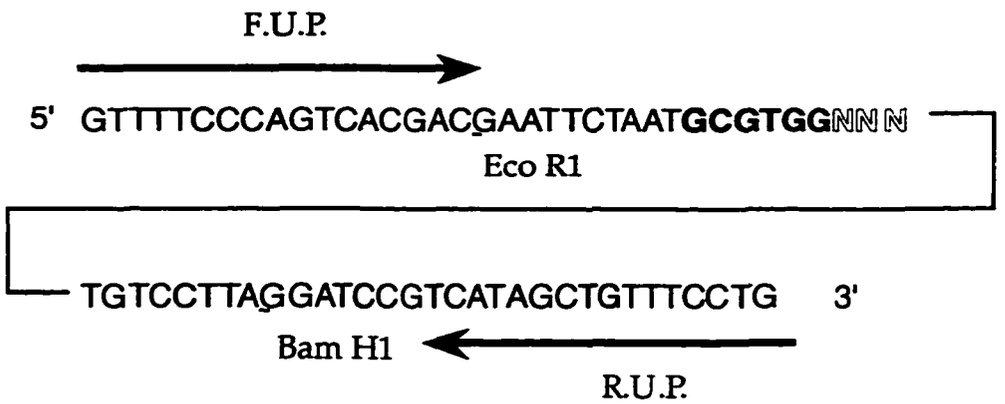
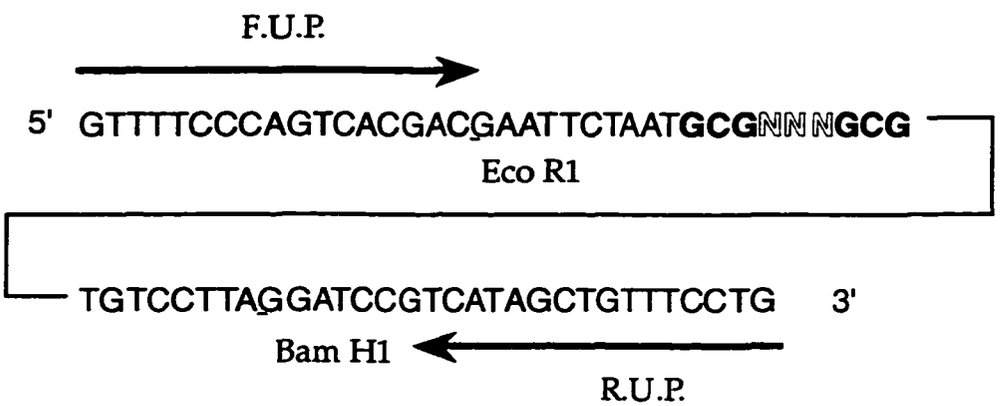


Figure 4.3 Coomassie blue-stained 15% SDS-polyacrylamide gel showing purifiied WT1-ZFP Denys-Drash mutants. Lane designated MW represents protein molecular weight marker; Lane 1: wild-type WT1-ZFP; Lane 2: R366H; Lane 3: R366C; Lane 4: R394W; Lane 5: D396G; Lane 6: D396N.



(A) SAAB template for WT1 finger 2



(B) SAAB template for WT1 finger 3

Figure 4.4 The SAAB template oligonucleotides containing randomized sequences for the WT1[-KTS] finger 2 (A) or finger 3 (B) recognition.

denaturing polyacrylamide gel which had been pre-electrophoresed for 20 minutes at 250 V at 4°C. Migration on the gel was monitored by loading one lane with 5 µl of 0.1% bromophenol blue, 0.1% xylene cyanol, 25% glycerol. The gel was run in 0.3 X TBE at 250 V at 4°C until the bromophenol blue had migrated 12 cm. The gel was then subjected to autoradiography for 2 to 4 hours at 4°C, and the band corresponding to the lowest protein concentration at which a protein-DNA complex was visible was excised and the DNA eluted as described above. The eluted DNA was amplified using F.U.P. and R.U.P. as primers, radioactively labeled, and a new round of selection was begun using lower protein concentrations.

Four rounds of selection were sufficient to isolate the highest affinity binding sites for each protein from the pool of random DNA sequences. The enrichment of high affinity sites for the WT1 proteins was followed by the proportion of DNA bound to the proteins after each round of selection. The lowest protein concentrations at which a protein-DNA complex could be observed in round 4 were: WT1-ZFP, 0.5 nM; R366H, 1.25 nM; R366C, 7.5 nM; R394W, 100 nM; D396N, 12.5 nM; D396G, 12.5 nM. The final PCR products of round 4 were digested with *Eco* RI/*Bam* HI and cloned into pUC19. The original randomized oligonucleotides were converted to a double-stranded form with cold dNTPs, and also subject to PCR amplification, restriction endonuclease digestion and ligation into pUC19. A sample of each ligation was used to transform *E. coli* strain JM109. Plasmid DNA was isolated from at least 50 colonies of each transformation and sequenced using either the Sequenase kit (Amersham) or an ABI 373A automated DNA sequencer and dye primer chemistry (Applied Biosystems).

4.2.3 Nitrocellulose Filter Binding Assay

The binding affinities of a series of DNA sequences for the WT1-ZFP and the Denys-Drash mutants were quantified using a nitrocellulose filter binding assay developed to study zinc finger protein-DNA interactions (Romaniuk, 1990). The details of the assay are described in Chapter 3.

4.3 Results

4.3.1 Selection of DNA binding sites for Denys-Drash mutant proteins.

To determine whether the Denys-Drash mutant proteins could confer new DNA-binding specificities, four rounds of selection, amplification, and binding assay (SAAB) were carried out for each mutant. A wild-type WT1-ZF peptide served as a negative control. The SAAB templates were: 5'-GCG TGG **NNN** TGT-3' for finger 2 mutants, and 5'-GCG **NNN** GCG TGT-3' for finger 3 mutants. PCR products of the final rounds of selections were cloned, and about 50 randomly picked clones were sequenced for each SAAB experiment. The results of the finger 2 SAAB experiments are shown in Table 4.1, with the frequencies of the selected nucleotides summarized in Table 4.2. Apparently, the mutation of the recognition arginine at position -1 of the finger 2 α-helix did not result in a new specificity for either mutant, as all the proteins selected the wild-type guanine with an almost 100% frequency. Interestingly, adenine was selected with a high frequency at position 8 of the consensus binding site by all finger 2 mutants. This is a definite deviation from the wild-type EGR-1 consensus site, and, perhaps, reflects the real differences that exist in the binding specificities of the two proteins. Curiously, a preference for a thymine was revealed at position 9 for the R366H mutant, whereas the wild-type guanine was predominantly selected by both the wild-type WT1-ZF and the R366C proteins.

Table 4.1 Finger 2 subsite sequences selected from a randomized template by wild type WT1-ZFP and the finger 2 point mutants

Subsite	Protein		
Sequence	WT1-ZFP	R366H	R366C
CAC	–	1	–
GAA	–	1	–
GAC	–	2	1
GAG	38	13	19
GAT	3	23	3
GCA	–	2	7
GCC	1	–	1
GCG	2	1	2
GCT	–	–	4
GGA	–	–	2
GGC	–	2	–
GGG	3	–	1
GGT	–	–	1
GTA	–	–	2
GTG	1	3	5
GTT	1	1	2
TAA	–	–	1
TAG	–	1	–
TCC	1	–	–

protein	base 7				base 8			
	none	WT1	R366H	R366C	none	WT1	R366H	R366C
A	29	0	0	0	25	82	82	47
C	19	0	2	0	25	8	6	28
G	21	98	96	98	19	6	4	8
T	31	2	2	2	17	4	8	18

	base 9			
	none	WT1	R366H	R366C
A	17	0	6	24
C	17	4	10	4
G	29	88	36	53
T	37	8	48	20

Table 4.2 Frequencies of each nucleotide selected at the finger 2 subsite positions by WT1-ZFP and the finger 2 point mutants

protein:	base 4					base 5				
	none	WT1	D396G	D396N	R394W	none	WT1	D396G	D396N	R394W
A	19	0	0	0	0	20	6	0	0	22
C	22	0	0	2	0	27	0	0	0	0
G	30	34	46	10	15	27	92	100	100	76
T	29	66	54	88	85	26	2	0	0	2

	base 6				
	none	WT1	D396G	D396N	R394W
A	7	0	0	0	0
C	18	0	0	0	0
G	35	100	100	100	44
T	40	0	0	0	56

Table 4.3 Frequencies of each nucleotide selected at the finger 3 subsite positions by WT1-ZFP and the finger 3 point mutants

The data for the finger 3 mutants are shown in Table 4.3, with the selection frequencies indicated in Table 4.4. Even though there were three finger 3 mutants tested vs. two in finger 2, the number of subsites selected for finger 3 mutants is almost half that selected for finger 2. This indicates a higher stringency selection, perhaps pointing to a more important role played by finger 3 in DNA recognition. Curiously, the aspartic acid substitution mutants (D396G and D396N) along with the wild-type WT1-ZF selected the wild-type site (T/G)GG. Apparently, the loss of the buttressing interaction between this aspartic acid residue and an arginine does not result in alteration of binding selectivity: a guanine base was selected with a 100% frequency. However, a substitution of an arginine 394 for a tryptophan resulted in a change of specificity: a thymine was preferentially selected over a guanine at position 6 of the binding site.

4.3.2 Measuring the binding affinities of the mutant peptides for the selected DNA sequences

The binding affinities of the mutant proteins for the selected sequences were measured using a nitrocellulose filter binding assay (Romaniuk, 1990), and compared to those of the wild-type protein. The representative binding curves are shown in Figure 4.5. The values of the dissociation constants are shown in Tables 4.5 and 4.6. The affinities correlate well with the frequencies of the selected nucleotides at the given positions. Thus, an adenine at position 8 was preferentially selected by the wild type and the finger 2 mutant proteins. Indeed, the affinity of the wild-type protein for the GAG triplet is two-fold higher than for the GCG sequence derived from the EGR-1 consensus site. A thymine (position 4) in the finger 3 recognition triplet was also preferentially selected over a guanine, which is reflected in the two-fold

Table 4.4 Finger 3 subsite sequences selected from a randomized template by wild type WT1-ZFP and the finger 3 point mutants

Subsite	Proteins			
Sequence	WT1-ZFP	D396G	D396N	R394W
CGG	–	–	1	–
GAG	–	–	–	1
GGG	17	23	5	6
GGT	–	–	–	1
GTG	1	–	–	–
TAG	3	–	–	5
TAT	–	–	–	6
TGG	32	27	45	12
TGT	–	–	–	22
TTT	–	–	–	1

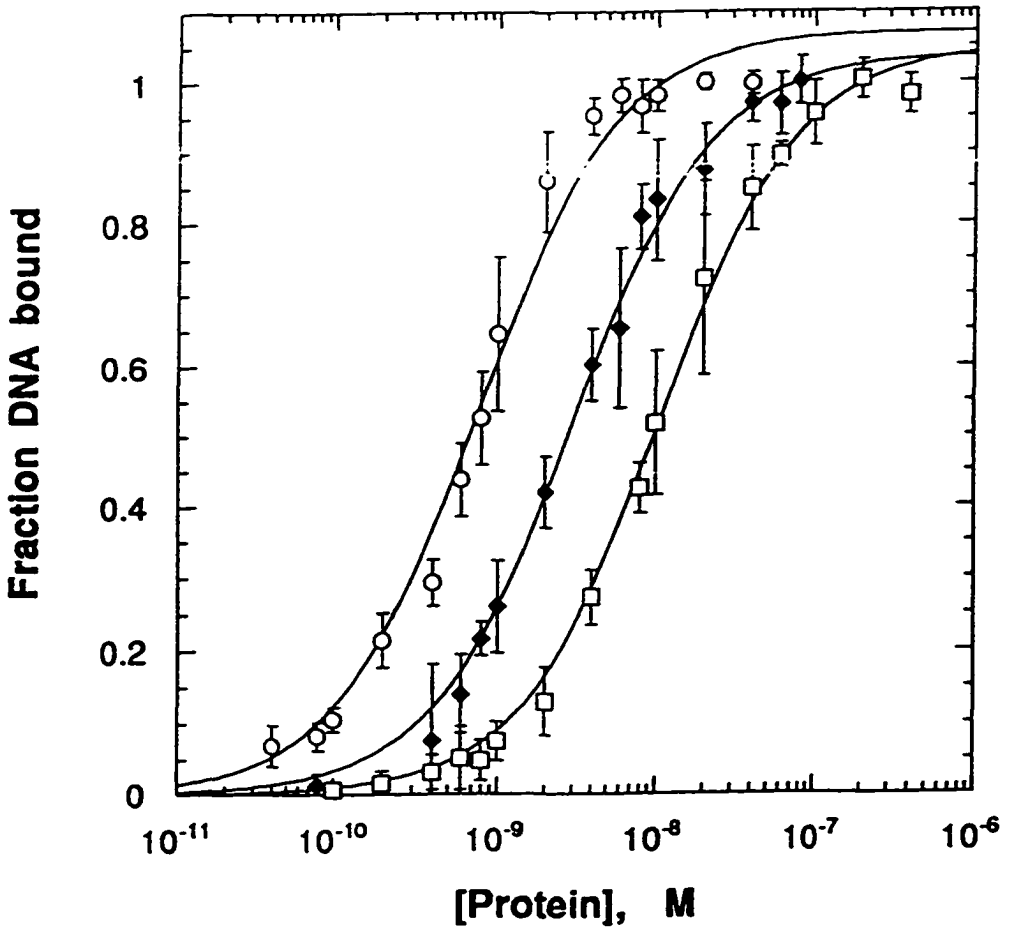


Figure 4.5 The equilibrium binding of the WT1 element GCG TGG GAG TGT to WT1-ZFP (open circles), R366H (closed diamonds) and R366C (open squares). Data points represent the mean of six or more determinations and error bars indicate the standard deviations. Curves represent the best fit of the data to a simple bimolecular equilibrium.

WT1-ZFP protein	GCG-TGG-GAG-TGT	
	K_d (nM)[a]	relative affinity[b]
wild type	0.77 ± 0.07	1
R366H	2.99 ± 0.19	0.26
R366C	10.8 ± 0.57	0.07

WT1-ZFP protein	GCG-TGG-GAT-TGT		
	K_d (nM)[a]	relative affinity[b]	selectivity[c]
wild type	1.93 ± 0.08	1	2.50
R366H	2.62 ± 0.17	0.71	0.88
R366C	14.1 ± 0.91	0.14	1.30

Table 4.5 Binding affinities of wild type and finger 2 mutant WT1-ZFP for selected DNA sequences

[c]Ratio of K_d(GAG subsite)/K_d(GAT subsite)

WT1-ZFP protein	GCG-TGG-GCG-TGT	
	K_d (nM)[a]	relative affinity[b]
wild type	1.54 ± 0.12	1
D396G	4.11 ± 0.29	0.37
D396N	12.4 ± 0.80	0.12

WT1-ZFP protein	GCG-GGG-GCG-TGT		
	K_d (nM)[a]	relative affinity[b]	selectivity[c]
wild type	3.56 ± 0.34	1	2.33
D396G	5.51 ± 0.29	0.63	1.33
D396N	26.2 ± 2.44	0.14	2.13

Table 4.6 Binding affinities of wild type and finger 3 mutant WT1-ZFP for selected DNA sequences

[a]Dissociation constants are expressed as the mean with standard deviations of a minimum of six independent determinations

[b]Ratio of K_d(mutant WT1-ZFP)/K_d(wild type WT1-ZFP)

[c]Ratio of K_d(TGG subsite)/K_d(GGG subsite)

higher affinity for the TGG vs. GGG triplet. In contrast, a thymine at position 6 of the finger 3 recognition triplet was never selected by the wild type protein or the aspartic acid mutants (D396G and D396N). Accordingly, these proteins showed no detectable binding affinity for the finger 3 subsite TGT. In addition, the R394W mutant showed no specific DNA-binding affinity for any of the selected sequences. The fact that this protein selected some sequences in the SAAB experiment probably reflects nonspecific binding due to the high protein concentrations (100 nM or more) used in the assay. This is the only protein mutant that lacks the specific DNA-binding activity. The other four Denys-Drash mutants showed varying degrees of reduced DNA-binding affinities which range from 1.4 - to 14 - fold. It is of interest to note the differences in severity of each individual amino acid mutation. Thus, substituting an arginine in zinc finger 2 for a cysteine produced a greater reduction in binding affinity: 14-fold for the GAG subsite, and 7.2-fold for the GAT subsite. Replacement of an arginine with histidine, on the other hand, resulted in a 3.9-fold and a 1.4-fold reduction in affinity for the GAG and GAT subsites, respectively.

4.4 Discussion

Almost all (95%) of the Denys-Drash patients have constitutional WT1 mutations (Coppes et al., 1993). These mutations involve key amino acids within the zinc finger domain of WT1, affecting the DNA-binding function of the protein. This study was designed to determine the effects of these mutations on the DNA-binding ability of WT1, as well as to address a question of whether these mutations confer new DNA-binding specificities.

The binding site selection and amplification experiments with the 5 zinc finger mutant proteins did not result in a selection of a new, high-

affinity DNA sequence for any of the mutant proteins. In fact, the protein with the most common mutation found in DDS patients, R394W, shows no detectable DNA binding. The remaining four protein mutants displayed altered DNA binding selectivities. Thus, both finger 2 mutants R366H and R366C had reduced sequence selectivity, with the largest effect observed for the R366H mutant (Table 4.5). However, this mutant had a four-fold lower affinity for the selected GAG relative to the wild-type protein, whereas the R366C mutant resulted in a 14-fold reduction in binding affinity. Presumably, histidine substitution at this position is less disruptive for the WT1 DNA-binding function than the cysteine substitution due to the capacity of the histidine to interact with guanines. This histidine could also form a number of van der Waals interactions with the thymine in position 9 of the triplet, thus explaining why both thymine and guanine were selected at this position. A uniform selection of an adenine residue at position 8 of the finger 2 binding subsite by the wild-type and the mutant proteins, as well as the higher affinity of the wild type protein for the adenine-containing sequence indicates that this adenine base is an important determinant in WT1 sequence specificity. The results of the human genomic DNA selection studies also identified adenine at position 8 (Nakagama et al., 1995). The authors proposed an interaction between a glutamine (Q369) of the WT1 and an adenine. In addition, the results of the mutagenesis study described in Chapter 5 of this thesis confirm the importance of this adenine for high-affinity binding. Strikingly, when the guanine at position 9 was replaced by a thymine, the affinity of the wild-type WT1-ZF protein was reduced only by a factor of 2.5. Position 9 was identified as one of the key determinants for recognition in the EGR-1 X-ray study (Pavletich & Pabo, 1991). Mutagenesis studies with EGR-1 and WT1-ZF show that substitution of guanine at

position 9 with a cytosine completely abolished binding of both proteins (Chapter 5 of this thesis). It remains to be seen why thymine is so well tolerated at this position.

In a recently refined crystal structure of EGR-1 - DNA complex, aspartic acid residues found at the second positions of the α-helices were shown to play an important role in recognition by not only stabilizing the arginine-guanine interaction, but by establishing direct hydrogen bond contacts to the secondary strand of the DNA (Elrod-Erickson et al., 1996). Thus, it is not surprising that mutations of this amino acid will have deleterious effects on the protein function. Indeed, both glycine and asparagine substitutions of aspartate in finger 3 result in a three-fold and 8-fold reduction in binding affinity, respectively. In terms of selectivity, only a glycine mutant had an appreciably reduced sequence selectivity. Interestingly, all the proteins used in the finger 3 selection experiments preferred a thymine over a guanine at position 4 of the binding sequence, while adenine or a cytosine were never selected. This agrees well with the results of the refined EGR-1 crystal structure (Elrod-Erickson et al., 1996), which showed that both thymine 4 and its complementary base on the opposite strand of the DNA (adenine 4') are involved in contacts with the protein: thymine 4 is used by a finger 2 histidine to establish a number of van der Waals contacts, whereas adenine 4' is contacted by an aspartic acid residue of the first finger of EGR-1. Finally, replacement of a guanine at position 6 abolished WT1-ZF binding. This is consistent with the role this base plays in the EGR-1 - DNA recognition, as identified in the X-ray structure (Pavletich & Pabo, 1991), as well as with a number of mutagenesis studies which observed the loss of binding upon substitution of the guanine at this position (Rauscher et al., 1990; Nakagama et al., 1995, Hamilton et al., 1997). However, while guanines at positions 6

and 9 play an equally important role in the binding of EGR-1, there must be a functional nonequivalence between these positions for the WT1 binding as thymine is tolerated at position 9, but not at position 6.

In summary, the results of this study point to several interesting conclusions. The SAAB experiments did not yield new DNA sequence specificities, thus suggesting that the Denys-Drash syndrome likely does not arise due to a change in the gene networks regulated by the WT1 protein. The most common DDS mutation, R394W, abolished specific binding of the protein. Other mutations resulted in reduced binding activities, ranging from 1.4 to 14-fold. This suggests that even small changes in the DNA-binding activity of WT1 protein can contribute to the development of the Denys-Drash syndrome, suggesting that the cellular levels of WT1 must be very tightly regulated. It is still not clear whether the reduced DNA-binding affinity of the DDS mutants is directly attributable to the phenotype seen in DDS patients, or whether the dysfunctional WT1 protein may interfere with the function of other cellular proteins. This study also provides insights into the roles of some of the key amino acids to the WT1 recognition.

CHAPTER 5.0. COMPARATIVE ANALYSIS OF THE DNA BINDING CHARACTERISTICS OF WILMS' TUMOUR AND EARLY GROWTH RESPONSE PROTEINS

5.1 Introduction

The demonstration that the WT1 protein is able to bind to the EGR-1-like sequences suggests that there may be a regulatory link between these two proteins: the WT1 protein may act as an antagonist of EGR-1 transcription factors, or may be a tissue-specific factor that is involved in maintaining a particular differentiated phenotype. The balance in the levels of EGR-1 and WT1 proteins may be critical: inactivation of WT1 could result in the onset of neoplasia.

The X-ray crystallographic studies of the EGR-1 zinc finger region bound to a 9 base pair fragment of DNA containing an EGR-1 consensus site have been serving as a topological blueprint to understand the mechanism of DNA binding by WT1 (Pavletich & Pabo, 1991; Elrod-Erickson, 1996). The details of the high-resolution analysis are discussed in detail in Chapter 1. All of the residues of the EGR-1 zinc fingers critical for specific binding to DNA are conserved in fingers 2-4 of WT1. However, despite the high degree of similarity shared between the WT1 and EGR-1 DNA-binding domains, there are a number of structural differences that suggest that the molecular details of DNA binding are not identical between the two proteins. These structural differences include approximately 50% dissimilarity between the WT1 and EGR-1, as well as an additional zinc finger present in WT1 protein.

To provide a clear quantitative picture of the interaction of WT1 and EGR-1 proteins with their DNA binding sites, we constructed peptides encompassing the zinc finger regions of these proteins (WT1-ZF and EGR1-ZF), and used a nitrocellulose filter binding assay to measure various

parameters of a bimolecular equilibrium. The stoichiometry of the DNA-protein complexes, their stability to dissociation, and the effects of pH, temperature and salt concentration on the equilibrium binding of these proteins to their cognate DNA sequences have been determined. In addition, the relative contribution of each base pair in the consensus binding site to the high affinity binding was determined by point mutational analysis. This study is a collaborative effort of several members of Dr. Romaniuk's lab: Cathy Juricic and Kathy Barilla made the DNA constructs incorporating single mutations encompassing the WT1 and EGR1 consensus binding site; Franck Borel assayed EGR1-ZF binding to the mutant DNA sequences. The remaining experiments I performed myself.

5.2 Materials and Methods

5.2.1 Construction and purification of recombinant WT1-ZF and EGR1-ZF proteins

The construction of a synthetic gene encoding the zinc finger region of EGR-1-ZF protein was conducted by analogy with that of WT1-ZF (please, refer to Chapter 3). The purification of the recombinant proteins was performed as described previously (Chapter 3). Typical protein yields averaged 10 mg per litre of bacterial culture. On SDS-PAGE (15%), the purified EGR1-ZF migrated with an apparent molecular weight of 17,000 kDa (Figure 5.1). The purified protein products were aliquoted and stored at -80'C. Protein aliquots were used only once when thawed.

5.2.2 Construction of mutant WT1-ZF and EGR1-ZF DNA binding sequences

Oligonucleotides incorporating point mutations at each one of the 12 positions of the WT1-ZF and EGR1-ZF consensus binding sequence were

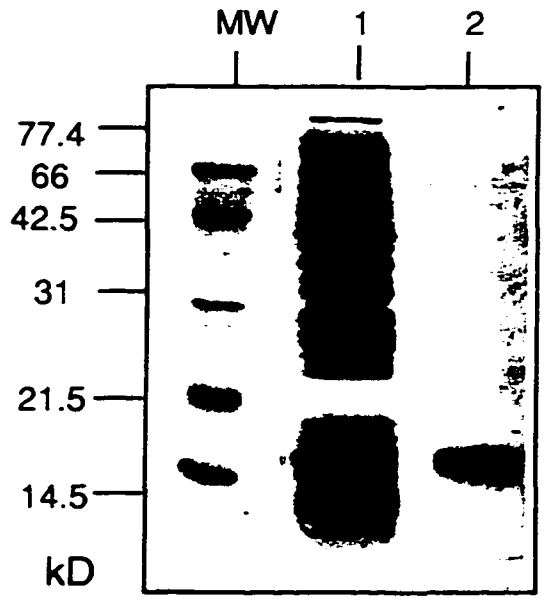


Figure 5.1 Coomassie blue-stained 15% SDS-polyacrylamide gel showing EGR1-ZF protein purified by affinity chromatography. Lane designated MW represents protein molecular weight marker; lane 1: total column pass-through; lane 2: purified EGR1-ZF.

synthesized by Dr. Paul J. Romaniuk using an Applied Biosystems Inc. 391 DNA synthesizer. The oligonucleotide with a mutation at position 1 (1C) had the following sequence: GATCCA **C**CG **TGG GCG TGT** AG, where a highlighted sequence represents the consensus binding site, and an underlined position denotes a mutated base. Similarly, other mutations were introduced: 2(T, A, G), 3C, 4(A, C, G), 5C, 6C, 7C, 8(A, G, T), 9C, 10C, 11C, 12C (Table 5.1). The oligonucleotides were cloned into *Bam*HI site of pUC19, and their sequences verified by automated DNA sequencing using an ABI 373A automated DNA sequencer and fluorescent dye primers.

5.2.3 End-labeling of DNA

Individual mutant DNA sequences were released from plasmid pUC19 using *Eco*RI and *Hind*III restriction enzymes, and end-labeled with [α-^{32}P]dATP and the Klenow fragment of *E.coli* DNA polymerase I (Sambrook et al., 1989). The details of the labeling protocol are provided in Chapter 3.

5.2.4 Nitrocellulose Filter Binding Assay

The binding affinities of a series of DNA sequences for the WT1-ZF and EGR1-ZF peptides were quantified using a nitrocellulose filter binding assay developed to study zinc finger protein-DNA interactions (Romaniuk, 1990). The details of the assay are described in Chapter 3.

5.3 Results

5.3.1 Equilibrium binding constants

We have made a detailed investigation of the mechanisms of DNA binding by both WT1-ZF and EGR1-ZF proteins, using methods that have been applied successfully to the study of the DNA and RNA binding

Table 5.1 Sequences of mutant oligonucleotides harbouring individual base pair substitutions in the EGR1-ZF and WT1-ZF consensus DNA binding site

G1C	GATCCA**C**CG TGG GCG TGT AG
C2A	GATCCA G**A**G TGG GCG TGT AG
C2G	GATCCA G**G**G TGG GCG TGT AG
C2T	GATCCA G**T**G TGG GCG TGT AG
G3C	GATCCA GC**C** TGG GCG TGT AG
T4A	GATCCA GCG **A**GG GCG TGT AG
T4C	GATCCA GCG **C**GG GCG TGT AG
T4G	GATCCA GCG **G**GG GCG TGT AG
G5C	GATCCA GCG T**C**G GCG TGT AG
G6C	GATCCA GCG TG**C** GCG TGT AG
G7C	GATCCA GCG TGG **C**CG TGT AG
C8A	GATCCA GCG TGG G**A**G TGT AG
C8G	GATCCA GCG TGG G**G**G TGT AG
C8T	GATCCA GCG TGG G**T**G TGT AG
G9C	GATCCA GCG TGG GC**C** TGT AG
T10C	GATCCA GCG TGG GCG **C**GT AG
G11C	GATCCA GCG TGG GCG T**C**T AG
T12C	GATCCA GCG TGG GCG TG**C** AG

properties of the zinc finger protein *Xenopus* transcription factor TFIIIA (Romaniuk, P.J. 1985; Romaniuk, P.J. 1990). Purified recombinant zinc finger peptides of WT1-ZF and EGR1-ZF were used to determine the parameters of the equilibrium binding of the proteins to their preferred DNA binding sequences. DNA binding was measured by titrating radiolabeled DNA with increasing amounts of purified protein and measuring the fraction of DNA bound by gel shift and filter binding assays (Romaniuk, 1990). Each data point shown in Figure 5.2 represents the mean of at least 3 independent determinations. The apparent dissociation constant (K_d) measured for WT1-ZF - DNA complex is $1.14\pm 0.2 \times 10^{-09}$ M under conditions of 0.1 M KCl, pH 7.5 and incubation at 22°C. The respective value for the EGR1-ZF - DNA binding is $3.55 \pm 0.4 \times 10^{-09}$ M under the same conditions.

To determine the nature of the equilibrium, the dissociation constants and the fraction of active protein, we performed a Scatchard analysis (Scatchard, 1949) of the interaction between the DNA consensus site and both WT1-ZF and EGR1-ZF proteins. In this analysis the protein concentrations were held constant (2 nM for WT1-ZF, and 5 nM for EGR1-ZF), and the DNA concentrations were varied from 0.31 to 30 nM. The results of the Scatchard analysis are shown in Figure 5.3. The lines indicate the theoretical curves calculated for the formation of a simple bimolecular equilibrium. The dissociation constants measured from the slopes of the lines are 9.7×10^{-10} M for WT1-ZF - DNA complex, and 2.05×10^{-09} M for EGR1-ZF - DNA complex. Both values are in close agreement with those measured using the standard filter binding assay. The results of the Scatchard analysis also indicate that the protein preparations are 100% active in the DNA binding assays.

To determine the time required for the two systems to reach a true equilibrium, we conducted a time-dependence assay, in which the

(A)

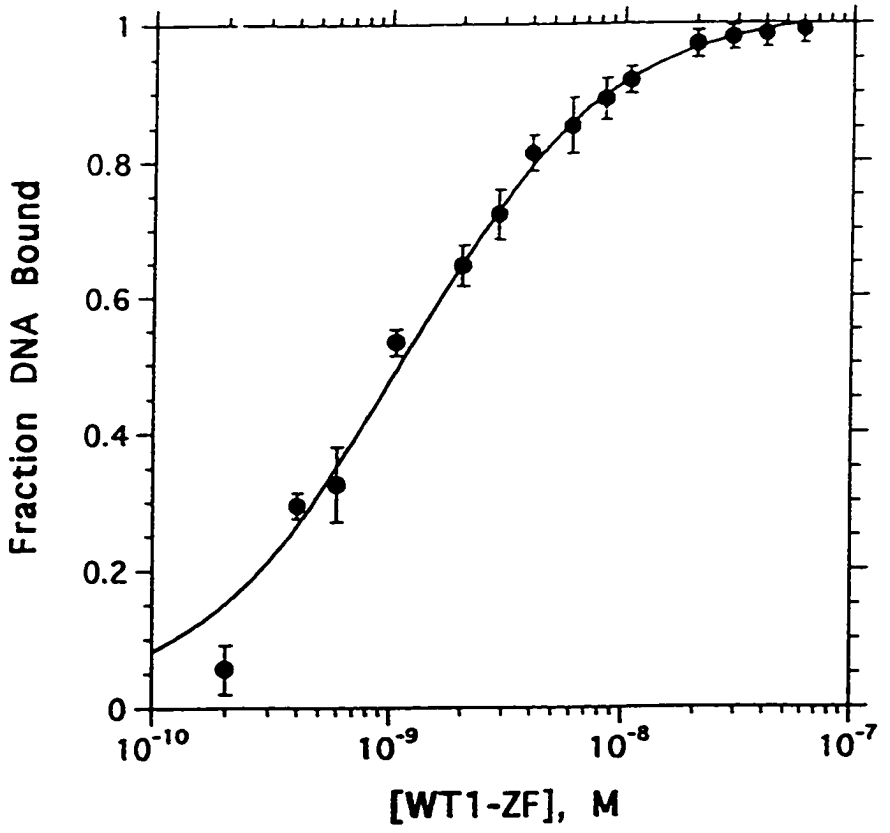


Figure 5.2 Equilibrium binding curves of WT1-ZF (A) and EGR1-ZF (B) proteins to their target DNA sequence. The affinities of WT1-ZF and EGR1-ZF proteins for their consensus DNA sequence were measured using a nitrocellulose filter binding assay. Each line represents the best fit for the formation of a simple bimolecular complex with Kd values of 1.14×10^{-09} for WT1-ZF (R = 0.994), and 3.55×10^{-09} for EGR1-ZF (R = 0.997). Each data point is the mean of three or more independent determinations, with the associated standard deviations indicated with the error bars.

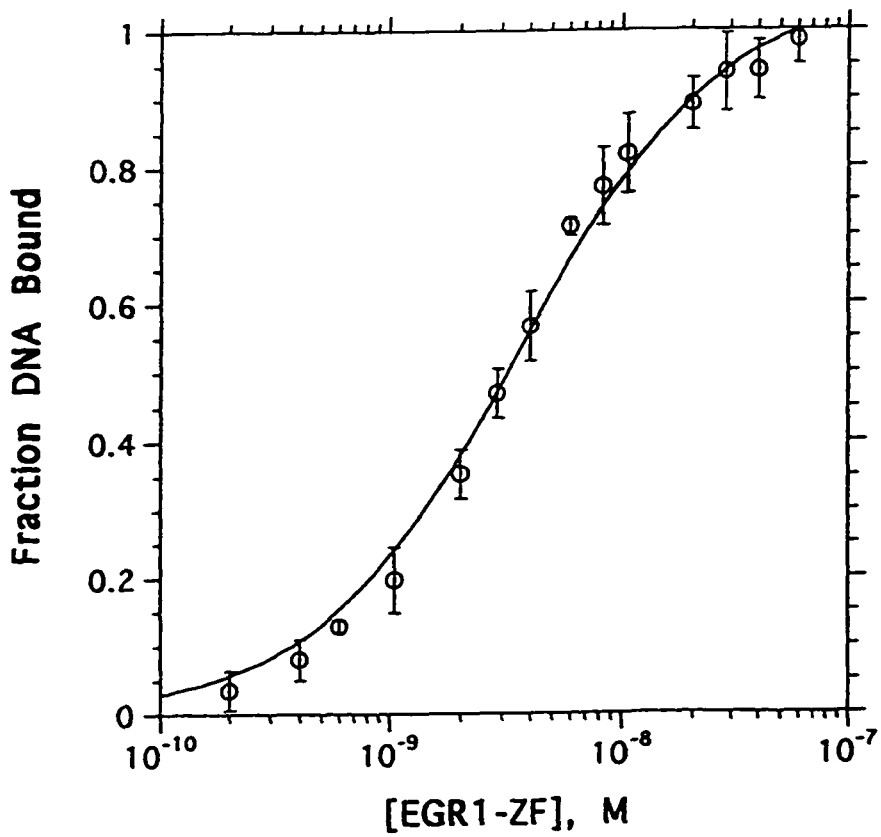


Figure 5.2 (B) Equilibrium binding curve of EGR1-ZF to its target DNA sequence.

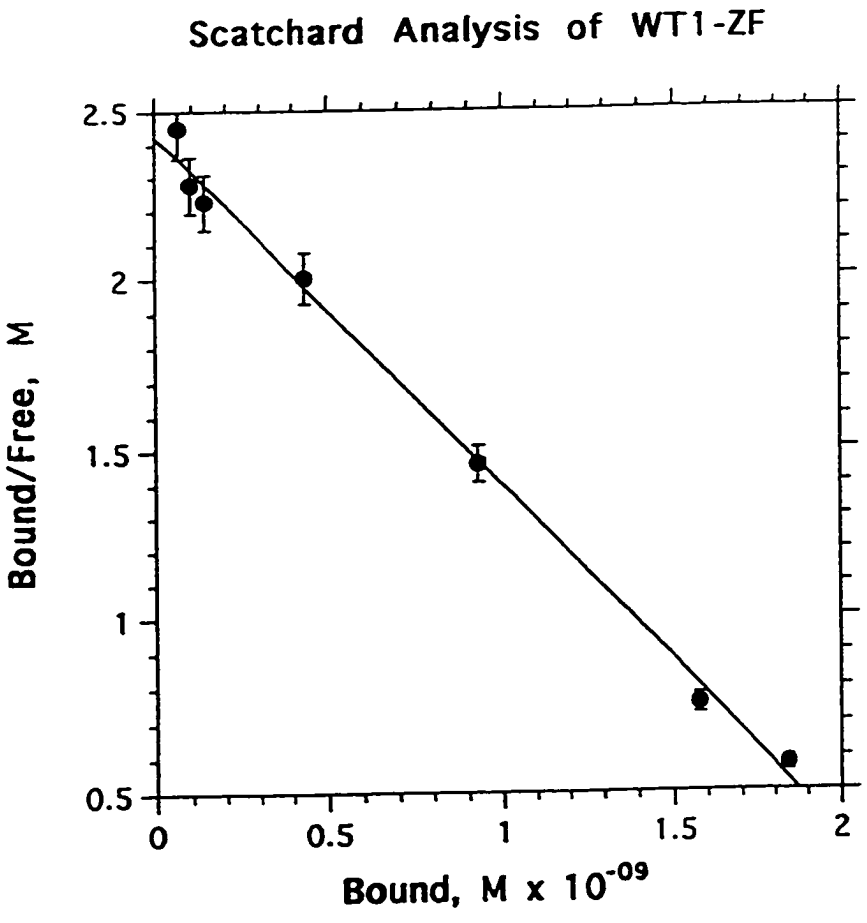


Figure 5.3 (A) Scatchard analysis of WT1-ZF. Scatchard analysis of the interaction between WT1-ZF and its consensus DNA binding site. The line represents the best fit for the formation of a simple bimolecular complex with a K_d value of 0.97×10^{-09} M (R=0.997). Each data point is the mean of two independent determinations, with the associated standard deviations indicated with the error bars.

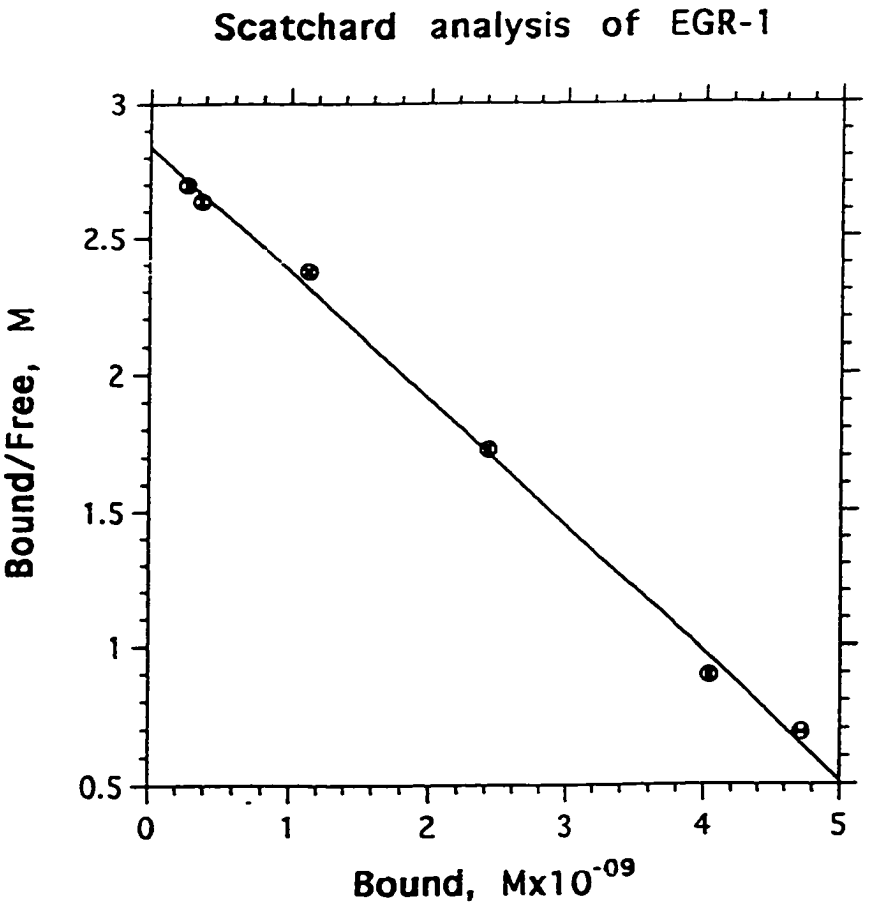


Figure 5.3 (B) Scatchard analysis of EGR1-ZF. Scatchard analysis of the equilibrium binding of EGR1-ZF to its DNA sequence. The line represents the theoretical curve calculated for a simple bimolecular equilibrium with a Kd value of 2.05 x 10^{-09} M (R=0.999). Each data point is the mean of two independent determinations, with the associated standard deviations indicated with the error bars.

dependence of the apparent K_a values were measured as a function of incubation times. Figure 5.4 shows that in a reaction mixture containing a target DNA and a competitor DNA (poly dI-dC) both proteins reach the full equilibrium after 60 minutes of incubation. The difference in the shapes of the time-dependence curves may reflect distinct modes of binding employed by each protein. The differences in the binding modes may include different kinetic rates of dissociation (not determined in this study), as well as other unidentified components of the equilibrium, such as influences of conformational changes occurring upon complex formation, sequential versus simultaneous binding by individual zinc fingers, and others. Because the time-dependence experiments were conducted last, incubation times used in the experiments described in this paper, were 45 min (with the exception of the temperature-dependence assay which was allowed to proceed for 60 minutes). Therefore, the apparent association constants were measured and used in the comparative analysis.

5.3.2 Monovalent salt dependence of the K_a for DNA binding

The affinity constants were measured at potassium concentrations that varied from 0 to 400 mM. The results given in Figure 5.5 show that the binding interaction has an optimum between 100 mM and 250 mM for the WT1-ZF protein, with affinity sharply declining at higher salt concentrations. The optimum binding for the EGR1-ZF protein is between 0 and 100 mM potassium salt with a decrease in binding affinity observed at higher salt concentrations. Both proteins exhibit dramatic dependencies of their binding constants on the monovalent salt concentration. Analysis of the salt dependence of the protein - DNA binding provides information about the number of cationic groups on the protein that interact with DNA phosphates

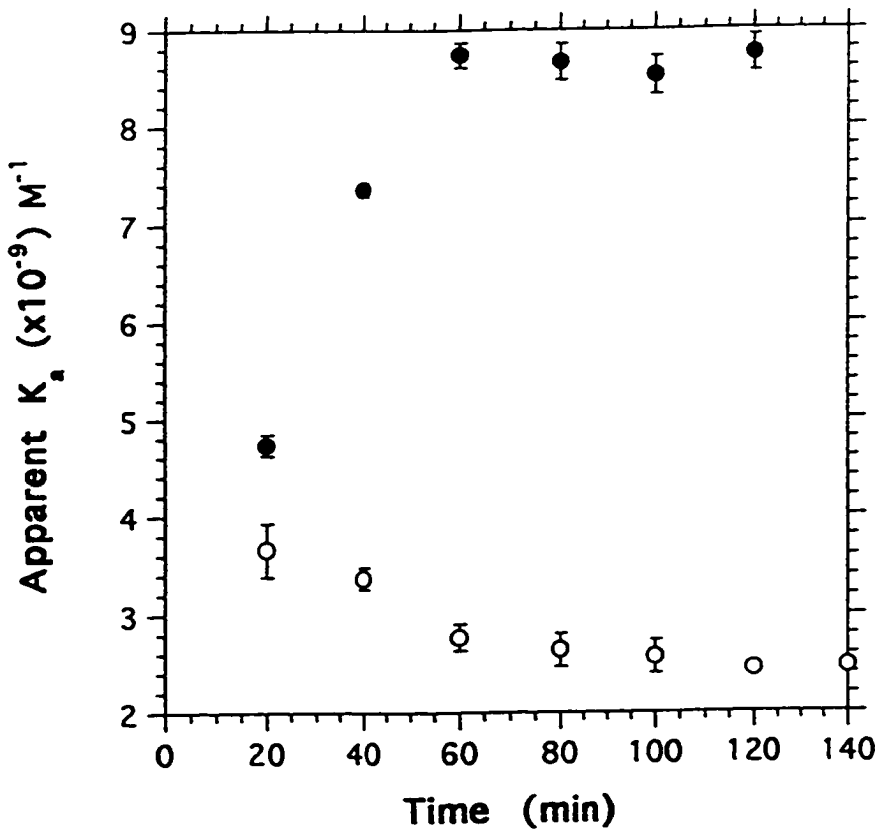


Figure 5.4 Time dependence of the binding of WT1-ZF (closed circles) and EGR1-ZF (open circles) to DNA consensus sequences. Apparent K_a s were determined using the nitrocellulose filter binding assay in TMK buffer. Each data point is the mean of 3 or more independent determinations, with the associated standard deviations indicated with the error bars.

in each complex. The number of ion pairs formed in a protein-DNA interaction can be determined using an analysis of the salt dependence of K_a based upon ion displacement (von Hippel & Schleich, 1969; Record et al., 1976). The number of ion pairs can be obtained using the equation: $\ln K_{obsd} = \ln K^\circ - Z_y \ln[M^+] - Z \ln(0.5\ (1+(1+4\ K^{Mg}_{obsd}\ [Mg^{2+}])^{0.5}))$ where K° is the apparent K_a at 1M salt, Z is the number of ion pairs formed, y is the fractional counterion bound per phosphate in the DNA (assumed to be 0.88 for double-stranded DNA), and K^{Mg}_{obsd} is the observed binding constant for the Mg^{2+} - DNA interaction. From the results presented in Figure 5.5 we calculated that there are 7±1 ion pairs formed in the WT1-ZF-DNA complex, and 6±1 ion pairs formed in the EGR1-ZF-DNA complex. This comparison indicates that WT1-ZF makes more ionic bonds to DNA than does EGR1-ZF.

The values of electrostatic and nonelectrostatic contributions to the observed free energy of binding can be determined using an analysis of the salt dependence of K_a based upon ion displacement (von Hippel, P.H. and Schleich, T. 1969; Record, M.T. et al. 1976). However, the determination of these values was not possible since both proteins demonstrate the presence of anionic binding sites.

5.3.3 pH dependence of K_a

The pH dependence of the apparent K_a for DNA binding was determined using appropriate buffers ranging from pH 5.5 to 9.0. As shown in Figure 5.6, both the WT1-ZF and EGR1-ZF-DNA interactions have a broad pH optimum from pH 5.5 to 8. Both proteins demonstrate a striking similarity in the pH dependence of the binding affinity. The affinity decreases at pH values above 8, which suggests that at low proton concentrations some of the functional groups on either protein (such as ε-NH_2 groups of arginines)

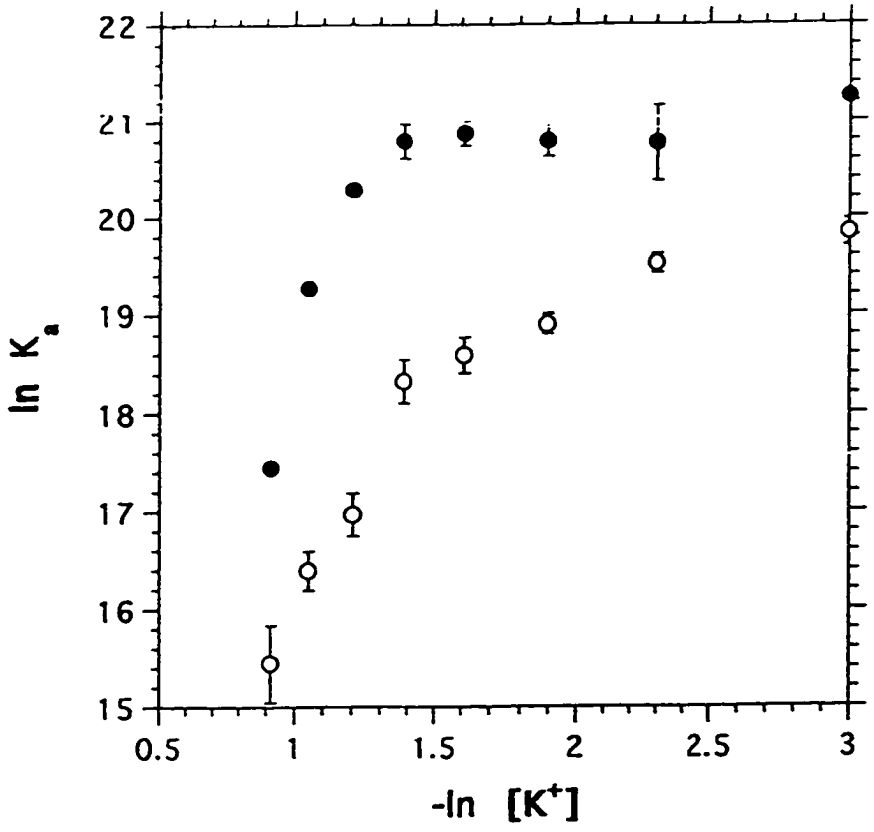


Figure 5.5. KCl concentration dependence of the binding of WT1-ZF (closed circles) and EGR1-ZF (open circles) to consensus sequences. The effect of increasing the KCl concentration in the binding buffer on the K_a was measured using the nitrocellulose filter binding assay. Each data point is the mean of 3 or more independent determinations, with the associated standard deviations indicated with the error bars. The experimental data were fit to equation 5 of Lohman et al (1980) by a regression method which varied the number of ion pairs and the apparent K_a at 1M salt.

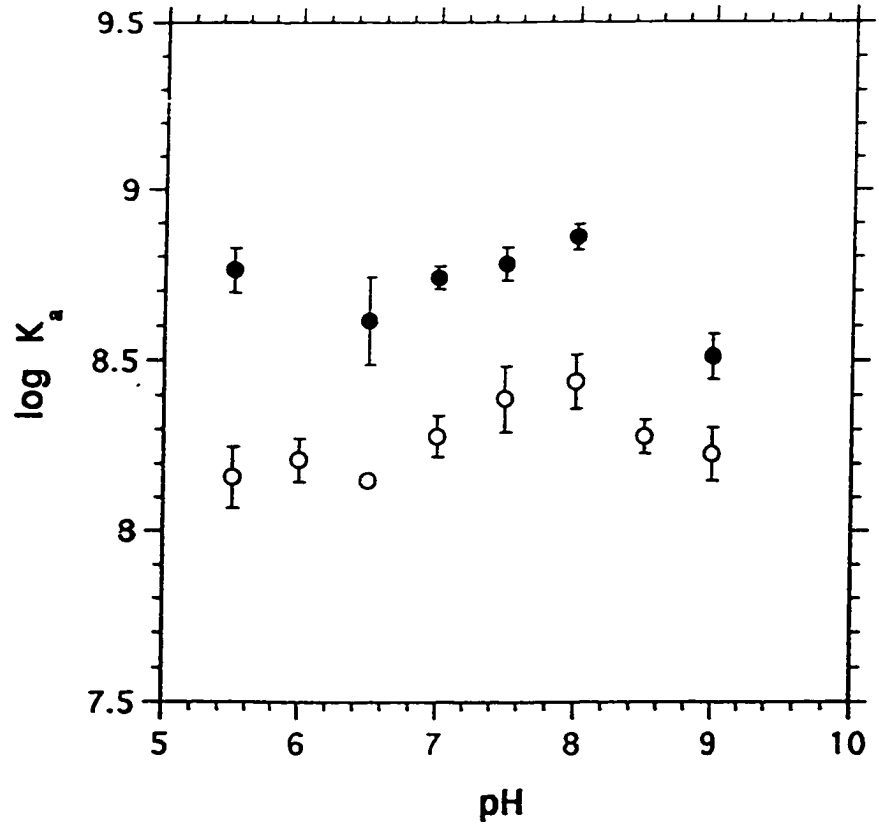


Figure 5.6. pH dependence of the binding of WT1-ZF (closed circles) and EGR1-ZF (open circles) to DNA consensus sequences. The effect of changing the pH of the binding buffer from 5.5 to 9.0 on the K_a was measured using the nitrocellulose filter binding assay. Each data point is the mean of 3 or more independent determinations, with the associated standard deviations indicated with the error bars.

or DNA get deprotonated. This reduces the amount of ion release that accompanies complex formation, and therefore the observed K_as decrease with increasing pH.

5.3.4 Effect of divalent metal ion concentration on DNA binding

The effect of the divalent metal ion concentration on the efficiency of binding was measured by varying the Mg^{2+} concentrations over a broad range: from 0 mM to 20 mM. The data in Figure 5.7 illustrates that WT1-ZF and EGR1-ZF binding have an optimum at 4 to 5 mM Mg^{2+}, gradually decreasing at either lower or higher Mg^{2+} concentrations.

5.3.5 Temperature dependence of the K_a value

The effects of temperature on the binding affinities of WT1-ZF and EGR1-ZF proteins have been determined using a range of temperatures from 4°C to 37 °C. The data is shown in Figure 5.8, and demonstrates the opposing effects of temperature on the WT1-ZF - DNA and EGR1-ZF - DNA complex formation. For WT1-ZF, the affinity of the interaction is highest at 37°C, while EGR1-ZF - DNA binding favors the lower values of the temperature range: the binding affinity is highest at 4°C.

A number of thermodynamic parameters can be calculated from these data: the free energy of the interactions can be calculated using the equation: $\Delta G^\circ = - RT\ln K_a$ (Lohman, T.M. & Maskotti, D.P. 1992), assuming that the apparent K_a value derived from the filter binding experiments is equivalent to the association constant for the WT1-ZF-DNA and EGR1-ZF-DNA equilibrium. The values of ΔH° can be calculated from van't Hoff analysis, and the values of ΔS° can be derived using the equation: $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$.

At 22°C, the ΔG° of the WT1-DNA interaction has a value of -12.1 kcal

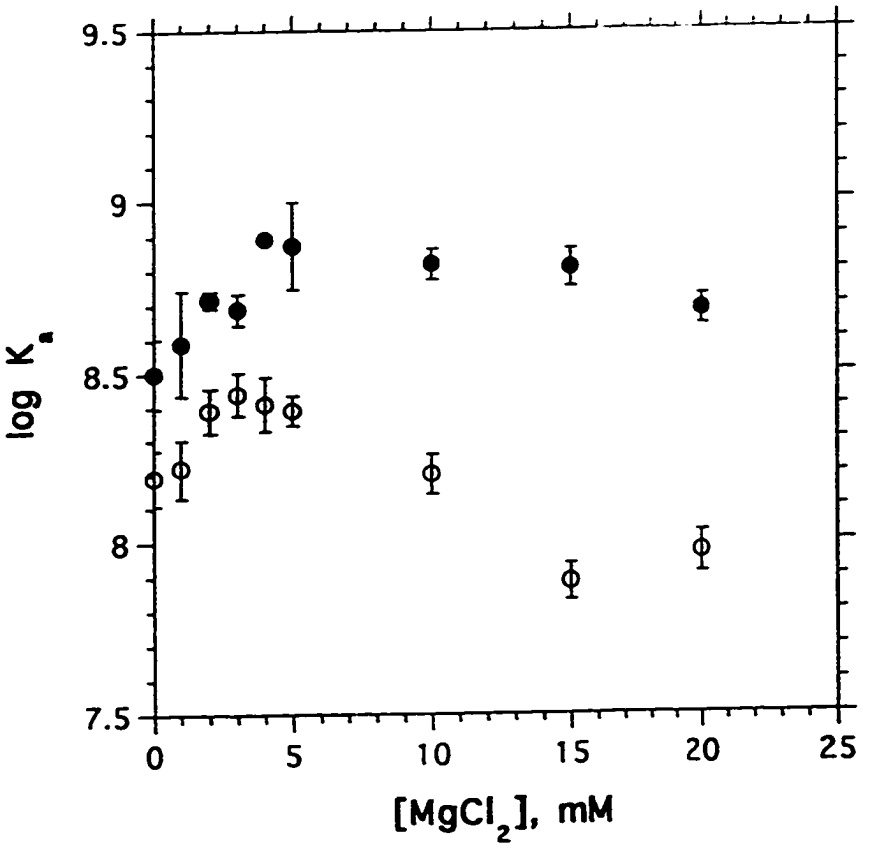


Figure 5.7. Effect of the magnesium ion concentration on the binding of WT1-ZF (closed circles) and EGR1-ZF (open circles) to DNA consensus sequences. The K_a values were determined using the nitrocellulose filter binding assay in TMK buffer supplemented with Mg^{2+} concentrations shown. Each data point is the mean of 3 or more independent determinations, with the associated standard deviations indicated with the error bars.

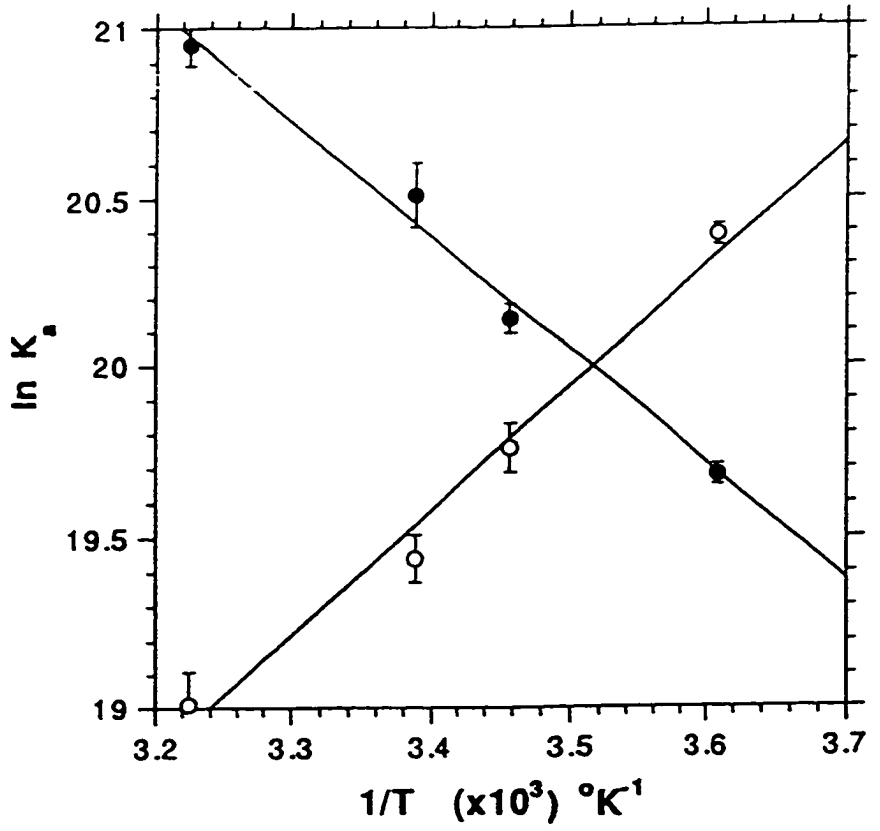


Figure 5.8. Temperature dependence of the binding of WT1-ZF (closed circles) and EGR1-ZF (open circles) to DNA consensus sequences. The K_a values were determined at temperatures ranging from 4°C to 37°C, using the nitrocellulose filter binding assay. The van't Hoff plot was used to calculate the enthalpy of these interactions by fitting each data set by linear least-squares analysis. The two correlation coefficients are greater than 0.99. Each data point is the mean of 3 or more independent determinations, with the associated standard deviations indicated with the error bars. The enthalpy values were determined from the slope of each line using the equation:

$$\frac{d \ln Ka}{d(1/T)} = \frac{-\Delta H^\circ}{R}$$

mol^{-1}; ΔH° has a value of +6.6 kcal mol^{-1}, and ΔS° has a value of +63.3 cal mol^{-1} deg^{-1}. This indicates that the WT1-ZF-DNA complex formation is an entropy driven process. The corresponding parameters for the EGR1-ZF-DNA interaction are: ΔG°= -11.4 kcal mol^{-1}; ΔH°= -6.9 kcal mol^{-1}; and ΔS°= 15.3 cal mol^{-1} deg^{-1}, and indicate that, in this case, the complex formation is an enthalpy driven process. The entropic binding process may reflect the release of counter ions and ordered water molecules as well as possible conformational changes that accompany protein-DNA complex formation. The enthalpic contribution to the free energy change indicates the amount of heat transferred in the course of the reaction.

5.3.6 Cation and anion effects on binding

We measured the effects of substituting potassium chloride by various anions and cations in the binding buffers. The varied cations and anions were kept at 0.1 M concentration. Replacement of potassium by sodium, lithium, or ammonium didn't have a significant effect on the measured K_as for either protein. The substitution of chloride by anions like acetate, nitrate, and iodide resulted in the decrease of K_a observed in the following order: acetate > chloride > nitrate > iodide (Table 5.2). While acetate and nitrate produced only modest changes in the K_as, the largest effect was observed by substituting chloride for iodide (about 100 - fold reduction of K_as for both proteins). The dependence of the binding on the identity of the anionic species indicates the presence of anionic binding sites on the proteins.

Table 5.2 Effect of the different monovalent salts on the binding of WT1-ZF and EGR1-ZF to DNA consensus sequences

	WT1-ZF		EGR1-ZF	
Salt	K_d (nM)[a]	Relative affinity	K_d (nM)	Relative affinity
KCl	1.14 ± 0.09	1.00	3.55 ± 0.02	1.00
NaCl	0.76 ± 0.01	1.5 ± 0.75[b]	2.76 ± 0.08	1.29 ± 0.06
LiCl	0.86 ± 0.16	1.33 ± 0.20	2.47 ± 0.1	1.44 ± 0.03
NH_4Cl	0.58 ± 0.14	1.96 ± 0.35	3.21 ± 0.11	1.11 ± 0.04
$NaNO_3$	1.47 ± 0.03	0.78 ± 0.06	4.99 ± 0.13	0.71 ± 0.03
$NaCH_3COO$	1.2 ± 0.14	0.95 ± 0.04	2.71 ± 0.03	1.31 ± 0.11
NaI	68.50 ± 6.61	0.017 ± 0.002	–[c]	>0.005

[a]Apparent dissociation constants determined by nitrocellulose filter binding assay. The salt concentrations were kept constant at 0.1M. Each value reported represents the mean of 3 or more independent determinations with the associated standard deviations.

[b]The errors for relative affinities are given by the expression $s = \{(s_1/M_1)^2 + (s_2/M_2)^2\}^{1/2} \times M_1/M_2$, where M_1 and M_2 are the respective dissociation constants for wild-type and mutant DNA and the s values are the corresponding standard deviations for these determinations.

[c]Binding affinity too low to be determined accurately.

5.3.7 Identification of relative contributions of each base pair in the consensus site to the binding specificities of WT1-ZF and EGR-1 proteins

The relative contribution of each base pair in the WT1-ZF consensus sequence 5' GCG TGG GCG TGT 3' to the binding affinity has been tested by point mutational analysis. The binding affinity of each mutant DNA was determined using the nitrocellulose filter binding assay. An identical analysis has been carried out for the EGR1-ZF protein interacting with the same consensus DNA sequence. These experiments have yielded a detailed picture of the interaction of each base pair in the DNA with the proteins under investigation. The results are presented in Table 5.3, and provide valuable information in assessing the relative importance of putative WT1 and EGR-1 binding sites as they are identified in genomic DNA.

Our analysis revealed that the guanine to cytosine substitutions at positions 3, 5, 6, 7, and 9 resulted in a drastically reduced (over 500-fold) affinity of binding for both WT1-ZF and EGR1-ZF proteins. This corresponds to the pattern of guanines shown to be contacted in the EGR-1 - DNA co-crystal structure (Pavletich & Pabo, 1991; Elrod-Erickson et al., 1996). The dissociation constants for these substitutions are indicated as "undetectable" in Table 5.3, since we could not accurately measure the equilibrium constants for those DNA mutants. The first guanine of the consensus binding site has also been shown to be contacted in the X-ray structure. However, this position appears to have a less drastic effect on binding. When the guanine is substituted by a cytosine, the binding affinity is reduced by approximately 10-fold for WT1-ZF, and about 5-fold for EGR1-ZF. Replacement of guanine residue with cytosine at position 1 removes the hydrogen-bonding potential for arginines, since cytosine has no hydrogen-bond-accepting groups in the major groove. This results in diminished affinity for both proteins, however

Table 5.3 Dissociation constants for WT1-ZF and EGR1-ZF binding to wild-type and mutant DNA consensus sequences

DNA	Relative affinity for WT1-ZF	Relative affinity for EGR1-ZF
wild-type	1.00	1.00
G1C*	0.09 ± 0.014[b]	0.18 ± 0.03
C2A	0.01 ± 0.005	0.15 ± 0.02
C2G	0.015 ± 0.005	0.20 ± 0.04
C2T	0.88 ± 0.14	0.79 ± 0.13
G3C*	< 0.005	< 0.005
T4A	< 0.005	< 0.005
T4C	< 0.005	< 0.005
T4G	0.64 ± 0.09	0.69 ± 0.10
G5C*	< 0.005	< 0.005
G6C*	< 0.005	< 0.005
G7C*	< 0.005	< 0.005
C8A	4.54 ± 0.75	0.08 ± 0.01
C8G	1.32 ± 0.27	0.04 ± 0.01
C8T	0.99 ± 0.12	0.25 ± 0.04
G9C*	< 0.005	< 0.005
T10C	0.88 ± 0.14	--
G11C	0.28 ± 0.06	--
T12C	0.53 ± 0.01	--

[a]Apparent dissociation constants determined by nitrocellulose filter binding assay. Each value reported represents the mean of 3 or more independent determinations with the associated standard deviations.

[b]The errors for relative affinities are given by the expression $s = \{(s_1/M_1)^2 + (s_2/M_2)^2\}^{1/2} \times M_1/M_2$, where M_1 and M_2 are the respective dissociation constants for wild-type and mutant DNA and the s values are the corresponding standard deviations for these determinations.

[c]Binding affinity is too low to be determined accurately.

*The positions of contacted residues in the EGR-1/DNA co-crytsal (Pavletich and Pabo, 1991). In each case an arginine (positions 1,3,6,7,9) or a histidine (position 5) interact with guanine residues in the DNA consensus sequence.

still providing a significant level of binding affinity. While nucleotides 3, 6, and 9 are contacted by an arginine residue immediately preceding the α-helix (position -1), the first and seventh bases in the binding site are contacted by fingers 1 and 3 which use arginines at α-helical position +6 as identified in the EGR-1 - DNA crystal structure. The reduction of the binding strength observed for thhe mutant at position one may reflect nonequivalence of the contributions to the binding affinity by the different nucleotides of the first triplet, as well as different roles played by arginines depending on their position in the recognition α-helix. However, we observed a much greater drop in affinity due to the substitution at position 7 than at position 1. This is evident from the fact that both guanine at position 7 and its complementary mate (C7') on the opposite DNA strand are involved in hydrogen bond contacts from an Arg 24 and an Asp 48, respectively, whereas there is only one residue contacted at position 1, as revealed in the crystal structure. In addition, Arg +6 (finger 3) interaction with the first guanine is not stabilized by a buttressing contact with an aspartic acid residue as observed in all three zinc fingers for the arginines at position -1. It may also be that individual zinc fingers provide nonequivalent contributions to the binding affinity. Overall, there is an excellent correspondence with the X-ray structural data in the way both proteins respond to changes in positions identified as direct base-pair contacts.

Base pairs 2, 4 and 8 are not involved in direct contacts with the EGR-1 zinc fingers. However, substitutions at any one of them demonstrated strong preferences at each of these positions. Thus, substitution of a cytosine residue at position 2 with adenine or guanine residues results in approximately 100-fold decrease in binding affinity for WT1-ZF, but only a 5-6-fold drop in affinity for the EGR1-ZF protein, suggesting that this interaction is more

critical for the WT1-ZF protein than for the EGR1-ZF. On the other hand, introduction of another pyrimidine, thymine, does not appreciably affect the binding strength of either protein, resulting only in a slight (less than 15%) decrease in the binding affinity. There may be several ways of rationalizing the observed effects: the identity of the base on the G-rich DNA strand, the stereochemical nature of the base, the contacts to the secondary, C-rich DNA strand or a combination of these features contribute to the strong selective pressure at this position. Indeed, water-mediated and van der Waals contacts are made to the second cytosine as well as to the complementary guanine at position 2' of the secondary strand as detected in the EGR-1 - DNA crystal structure. It is not readily apparent, however, why the magnitude of the affinity change is different between the two proteins. It is possible that WT1-ZF contacts are more stereochemically constricted that those of EGR1-ZF. More mutagenesis and structural data will be necessary to determine why the presence of guanine or adenine at position 2 is more disruptive for the WT1-DNA interaction.

The thymine substitution for either an adenine or a cytosine at position 4 abolished the binding ability of both WT1-ZF and EGR1-ZF (over 500-fold reduction in affinity), while both proteins showed just a slightly reduced (about 30 %) preference for a guanine residue in this position. As identified from the X-ray analysis, there are no direct hydrogen-bonding contacts made at this position. The threonine which could provide a hydrogen from its hydroxyl group to the major-groove hydrogen-bond acceptors of thymine (O4) or guanine (N7 and O6) is too far away to make such a contact (Pavletiich & Pabo, 1991). However, as described in Chapter 4, His 49 makes van der Waals contacts to the T4 as detected in the refined structure. Additionally, Asp +2 of finger 3, which forms a residue pair with

Arg -1 through a salt bridge, was tentatively suggested to accept a hydrogen bond from the N6 amine group of the adenine A4' located on the opposite strand of the DNA. In the X-ray structure, this interaction did not have a perfect geometry, and therefore its importance for DNA binding was not clear. In this study, we show that this contact likely occurs, as it is very important for the overall affinity of binding of both proteins. This interaction would also explain a preference for a cytosine at position 4' (complementary to the guanine on the primary strand): a cytosine, analogous to an adenine, also has a hydrogen-bond donating amine group - N4 - available for interacting with aspartic acid. In a recent study, Swirnoff & Milbrandt (1995) also demonstrated a strong selection for a thymine at position 4, and less strongly - for a guanine, whereas adenine or cysteine were never selected. The authors also showed that a methyl group of thymine was important for the optimum binding affinity by introducing a uracil substitution at position 4, which reduced binding strength of EGR-1.

Further differences in the binding preferences for the WT1-ZF and EGR1-ZF proteins were revealed when the cytosine at position 8 of the consensus site had been substituted by an adenine. The binding affinity of the EGR1-ZF protein dropped by a factor of about 12, whereas the binding ability of the WT1-ZF protein was actually improved by a factor of 4.5. The two other substitutions at position 8 - 8G and 8T gave nearly a wild-type affinity for the WT1-ZF protein. However, in the case of EGR1-ZF protein, the 8G and 8T mutations had resulted in a 25-fold and a 4-fold reduction in binding affinity, respectively. Cytosines at positions 2 and 8 are symmetrical in terms of their location in the EGR-1 nonamer binding site. However, there are different degrees of tolerance to substitutions at these positions. In the case of EGR1-ZF binding, some of the observed effects are implicit in the crystal

structure. Thus, cytosine 8 and its complementary base guanine 8' are both involved in water-mediated and van der Waals interactions with a number of amino acids: Asp -1 of finger 1 makes a water-mediated contact to C8; at the same time, Glu +3 of finger 1 makes van der Waals contacts to the same base; while three amino acids - Ser 47 and Asp +2 of finger 2 and Arg +6 of finger 1 make a series of water-mediated interactions to G8'. As mentioned earlier, at position 2 there are two amino acids making water-mediated bonds to C2 and G2', and one amino acid making van der Waals contacts to C2. Clearly, there is nonequivalence between positions 2 and 8 which is reflected in varying degrees of tolerance to substitutions at these positions.

In the case of WT1-ZF protein, position 8 of the DNA recognition sequence clearly plays a different role than it does in EGR1-ZF binding. Finger 2 of WT1-ZF has a glutamine at α-helical position +3, which EGR1-ZF doesn't have in its corresponding finger 1. This glutamine has the potential to be either a hydrogen bond donor or acceptor, which would help explain the base preferences seen in our data (A > G > T or C). Therefore, WT1-ZF protein uses a distinct mode of DNA binding from that of EGR1-ZF. In addition, comparison of the binding preferences between fingers 4 and 2 of WT1 suggests that different fingers of WT1 provide nonequivalent contributions to the binding.

The previously identified positions 10-12 for the WT1-ZF finger 1 binding (Drummond et al., 1994; Hamilton et al., 1995) had a relatively high tolerance for sequence alteration (Table 5.1). The largest reduction (four-fold) in affinity was observed when a guanine was substituted for a cytosine at position 11 of the binding site. Finger 1 of WT1 has a histidine at position +3 of the α-helix, which could be involved in contacting guanine 11. More

structural and mutagenesis data will be required to unambiguously assign specific amino acid-base interaction to finger 1 of WT1.

5.4 Discussion

The demonstration that WT1 protein binds to the DNA sequences similar to a consensus site for the EGR-1 protein, and can regulate promoters (Madden et al., 1991; Rauscher et al., 1990) containing such sequences, suggested that WT1 and EGR-1 proteins may act as antagonists in developmental regulation. We conducted a detailed quantitative analysis to define how the affinity and specificity of WT1-ZF for common regulatory sites would compare with competing factors such as EGR1-ZF. We investigated the molecular details of the interaction by measuring the preferences of WT1-ZF and EGR1-ZF proteins for various buffer components, as well as effects of temperature and incubation times on the binding efficiency.

The equilibrium dissociation constants were determined to be $1.14 \pm 0.2 \times 10^{-09}$ M for WT1-ZF, and $3.55 \pm 0.4 \times 10^{-09}$ M for EGR1-ZF protein, assuming that the complex formation is the result of a simple bimolecular equilibrium. Scatchard analysis had confirmed that both proteins form a 1 : 1 complex with DNA. The association constants measured by the slope of the Scatchard plot were in close agreement with those measured using the standard filter binding assay. The Scatchard analysis also confirms that the protein preparations were 100% active.

Measuring the monovalent salt dependence of the K_a for DNA binding indicated that WT1-ZF interaction has an optimum between 0 mM and 200 mM potassium, whereas EGR1-ZF apparently requires lower monovalent salt concentrations (from 0 to 100 mM) for the optimum binding. Both proteins demonstrate a dramatic decrease in binding affinity at higher salt

concentrations. In addition, we estimated the number of cationic residues and phosphates that interact in each complex: WT1-ZF protein forms about 7 such interactions, whereas EGR1-ZF makes 6. This number of phosphate contacts exactly corresponds to the number of phosphates found to be directly contacted in the EGR-1 - DNA co-crystal structure (Pavletich & Pabo, 1991). Additional, weaker water-mediated phosphate contacts were also observed in the refined EGR-1 structure (Elrod-Erickson, 1996), which were not detected in our analysis.

Determination of pH dependence of the DNA binding activity showed that both WT1-ZF and EGR1-ZF proteins can achieve optimum binding at a very broad pH range, declining at higher pH. The study of the effects of divalent metal ion concentration on DNA binding had shown that WT1-ZF achieves its optimal binding at 5 mM Mg^{2+}, while EGR1-ZF prefers a more modest concentration of about 1 mM Mg^{2+} for its optimum binding.

A number of thermodynamic parameters were obtained and compared for the WT1-ZF - DNA and EGR1-ZF - DNA interactions. At 22° C, the WT1-ZF - DNA complex had the following thermodynamic values: ΔG°= -12.1 kcal mol^{-1}; ΔH°= +6.6 kcal mol^{-1}, and ΔS°= +63.3 cal mol^{-1} deg^{-1}. The corresponding parameters for EGR1-ZF - DNA interaction are: ΔG°= -11.4 kcal mol^{-1}; ΔH°= -6.9 kcal mol^{-1}, and ΔS°= +15.3 cal mol^{-1} deg^{-1}. The large positive value of ΔS° for the WT1-ZF - DNA interaction indicates that it is an entropy driven process, where ordered water molecules and counter ions may become displaced from the binding interface upon complex formation. In contrast, the EGR1-ZF - DNA interaction appears to be an enthalpy driven process in the temperature range studied. These differences in thermodynamic parameters provide us with important insights into the mechanisms of protein-DNA complex formation in each case. As noted by

Beaudette & Langerman (1980), entropy-driven processes are often a reflection of hydrophobic interactions that occur in the system, whereas a large enthalpic contribution usually results from the formation of hydrogen-bond and van der Waals contacts. Thus, WT1-ZF protein contains 7 hydrophobic amino acids in its recognition α-helices, compared to 2 hydrophobic amino acids found in the corresponding regions of EGR1-ZF, which is consistent with our finding of a large entropic contribution to the overall free energy of binding in the WT1-ZF - DNA complex. Refined EGR-1 structure, on the other hand, revealed an extensive network of van der Waals and water-mediated contacts, which occur to both the DNA bases and the backbone phosphates and provide additional contributions to the binding energy. This is also consistent with the enthalpic driving force leading to the EGR1-ZF - DNA complex formation.

We used a nitrocellulose filter binding assay to determine the affinities of the WT1-ZF and EGR1-ZF proteins for a number of DNA point mutants designed to test the relative contributions of each base pair to the high affinity binding. The crystallographic data obtained for the EGR-1 - DNA complex has provided a structural framework for understanding the molecular basis of the interaction. Since the critical amino acids in EGR-1-ZF which contact specific base pairs in the EGR-1 consensus site are completely conserved in the WT1-ZF protein, it was reasonable to predict that the last three zinc fingers of WT1-ZF will play a similar role in their binding to the 9 bp EGR-1 recognition sequence. However, this prediction had to be tested considering that WT1-ZF protein contains a number of structural differences. These include an additional zinc finger (the first zinc finger of WT1) which does not possess significant homology with any of the three EGR1-ZF zinc fingers at the amino acid level (Call, K.M. et al. 1990), and the approximately 50 % dissimilarity in

the amino acid sequences between the EGR1-ZF protein and fingers 2-4 of the WT1-ZF.

In addition, crystal structural data available for a number of other zinc finger proteins, including Tramtrack (Fairall et al., 1993), SWI (Fairall et al., 1993), and GLI (Pavletich & Pabo, 1993) reveal complex and nonuniform recognition mechanisms taking place in different protein - DNA systems. The details of these structures are discussed in Chapter 4. In the view of this diversity in the binding modes observed for various zinc finger proteins, and the structural differences existing between the WT1-ZF and EGR1-ZF proteins, we made a detailed investigation into the roles of individual base pairs in the binding of WT1-ZF and EGR1-ZF proteins.

Overall, our results confirm the main recognition themes established in EGR-1 crystal structural data, while revealing additional details and providing a quantitative view of the importance of individual bases in high-affinity recognition. In agreement with the crystal structure, the critical guanines of the primary strand (positions 3, 5, 6, 7, and 9) provide the most important energetic contributions to the high affinity binding for both WT1-ZF and EGR1-ZF proteins. Substitutions at any of these positions resulted in the loss of binding by both proteins. Position 1, albeit shown to be involved in contact in the EGR-1 - DNA co-crystal, provides a less important contribution to the overall binding: a guanine to cytosine substitution at this position resulted in a 10-fold decrease in binding for WT1-ZF, and a 5-fold decrease for EGR1-ZF. Previous methylation interference studies by Lamaire et al. (1990) indicated a weaker interaction made by EGR-1 at position 1. This smaller contribution from position 1 to the overall binding energy can be underscored by the fact that there is only one contact made at this position, and this arginine-guanine contact is not stabilized by the aspartic acid side

chain. A similar Arg-G contact made at the third position in the binding site resulted in the loss of binding by both protein. The additional water-mediated contact formed between Asp+2 of finger 3 of EGR-1 (or, by analogy, finger 4 of WT1) combined with the stabilized Arg-Asp-G contact likely contribute to the higher energetic cost of sequence disruptions at this position. Position 2 of the recognition sequence was found to be important for the binding of both proteins. The sequence preferences were observed to be in the following order: C ≥ T » A » G. However, substitution of cytosine with either purine had a more deleterious effect on WT1-ZF binding, reducing its affinity by about 100-fold, whereas the corresponding drop in EGR1-ZF affinity was about 5-fold. It is likely that the stereochemical nature of the base plays an important role in recognition. It is not clear, however, why the affinity for WT1-ZF was reduced to a greater extent than that for EGR1-ZF. The role of position 2 is discussed in more detail in the results section.

The refined EGR-1 crystal structures demonstrated a direct role for the aspartic acid residue in finger 2. However, it was not clear whether the corresponding aspartic acid residues in fingers 1 and 3 also contribute to specific recognition, since they did not have a perfect geometry (Elrod-Erickson, 1996). Our mutagenesis data supports the importance of these aspartic acid residues for sequence-specific recognition of the fourth and seventh bases of the binding site, which are in the third and second zinc finger binding subsites, respectively. Both WT1-ZF and EGR1-ZF proteins lost their DNA binding activity when substitutions were made at these positions. However, we didn't test the importance of position 10 for EGR1-ZF binding. This would have given us an insight as to the contribution of aspartic acid residue in finger 1 to EGR1-ZF recognition, and, possibly, would have extended the EGR1-ZF recognition sequence to 10 base pairs. WT1-ZF showed

just a slightly decreased binding upon introduction of a cytosine at position 10. There is evidence from the work of others which corroborates our findings. Thus, a similar contact was seen in the TTK/DNA crystal structure: an aspartic acid located at the tip of finger 2 makes a good hydrogen bond to a cytosine on the secondary strand (Fairall et al., 1993). Replacement of the homologous aspartic acid in the second zinc finger of ADR1 protein significantly reduced protein binding to its DNA recognition sequence (Thukral et al., 1991; Thukral et al., 1992). Binding site selection studies support the same view: only thymine or guanine were selected at position 4 and 10 of the DNA sequence (Swirnoff & Milbrandt, 1995). The latter authors also proposed that single zinc fingers of EGR-1 as well as related family proteins may specify overlapping, 4-base pair subsites of DNA. This notion is supported in the refined EGR-1 crystal structure (Elrod-Erickson, 1996), in the study by Isalan et al. (1997), as well as in our study. While both thymine and guanine at position 4 could have satisfied a contact from aspartic acid to the secondary DNA strand (providing amine groups of A and C), we observed a preference for a thymine over a guanine for the binding of both proteins. This preference for a thymine at position 4 can be explained by the van der Waals contacts made by histidine +3 of the second finger of EGR1-ZF or the third finger of WT1-ZF. A similar observation was made by Swirnoff & Milbrandt (1995), who suggested that histidine +3 of EGR1-ZF contributes to the recognition of two adjacent bases at positions 4 and 5. More discussion is provided in the results section of this chapter.

As discussed earlier, adenine at position 8 provides a significant contribution to the high affinity binding of the WT1-ZF (4.5-fold increase compared to cytosine at this position). In contrast, this substitution reduced the binding strength of EGR1-ZF by 12-fold. This demonstrates that the two

proteins share related, but not identical DNA recognition specificities. The importance of an adenine at position 8 of the DNA site has been previously demonstrated for WT1 protein by Nakagama et al. (Nakagama et al., 1995). An additional interaction provided by a glutamine residue (Q352) of the WT1-ZF, which EGR1-ZF does not possess, was proposed by Nakagama et al. (1995). The results of the previous high affinity binding site selection experiments obtained in our lab identified the same phenomena: the wild type WT1-ZF protein selected an adenine at position 8 with high frequency (Borel et al., 1996). In addition, our data demonstrate that EGR1-ZF has a distinct preference (10-fold higher affinity) for a cytosine versus an adenine at position 8.

In conclusion, the data presented in this study provides accurate measurements of equilibrium binding parameters for WT1-ZF and EGR1-ZF proteins. Quantitative analysis of the contribution of each base pair in the DNA consensus binding sequence to the overall binding affinity was defined using DNA point mutants. Every base pair in the binding sequence was found to contribute, with varying degrees, to the high affinity binding for both proteins. The analysis demonstrated important differences in the binding mechanisms that exist between the WT1-ZF and EGR1-ZF proteins, with respect to both the thermodynamics, and the sequence preferences. This, in turn, suggests that the WT1 and EGR-1 proteins may act on distinct promoter sequences, while being able to compete for common regulatory sites. The information obtained from these mutagenesis studies extends our knowledge of the mechanisms of binding by the zinc finger proteins. Further studies into the physiological roles of these two regulatory proteins should provide valuable information into the origins of specificity, as well as shed

light on the involvement of both proteins in normal development and neoplastic processes.

CHAPTER 6.0 TFIIIA IS A PROTOTYPE OF THE C_2H_2 FAMILY OF ZINC FINGER PROTEINS

6.1 The C_2H_2 zinc finger domain

The majority of biological processes involve specific interactions between macromolecular species. The specific interactions are mediated by a multitude of distinct structural domains that maintain an appropriate stereochemical fold. Typically, in order for the polypeptides to fold autonomously, they need to be at least 50 amino acids long.

About a decade ago, a novel class of domains consisting on average of 30 amino acids was discovered. These domains are too small to fold independently, but fold stably when coordinated zinc ions stabilize the conformation. The ability of zinc to be tetrahedrally coordinated within such domains, as well as the lack of redox activity of the zinc ion, led to the evolution of a wide range of zinc-stabilized structural domains now known to exist.

Zinc has been known to be an essential trace element for eukaryotes for over a century (Coleman, 1992). An average adult human body contains 2.3 g of zinc (compared with 4 g of iron) (Coleman, 1992). This makes zinc the second most abundant trace metal in eukaryotes. As new families of zinc containing proteins continue to emerge, specific functions can be assigned to zinc to justify its abundance.

The first discovery of a zinc metalloprotein involved in transcription was made in 1983 by Hanas et al., who demonstrated that transcription factor IIIA from *Xenopus* oocytes was a zinc protein, and that zinc coordination correlated with its ability to bind DNA (Hanas et al., 1983). Of all the C_2H_2 zinc fingers studied to date, this unique polyfinger DNA- and RNA-binding protein is, perhaps, the most extensively studied of all.

From the nucleotide sequence of the gene, Klug and coworkers (Miller et al., 1985) deduced that the amino acid sequence of the translated product could be arranged so that a pair of conserved cysteines and a pair of conserved histidines separated by a 12 - to 13 - residue spacer defined 9 repeat series of amino acid sequences of 28 residues within the TFIIIA protein as follows: - C $-X_{2-5}$ - C $-X_{12-13}$ - H $-X_{3-4}$ - H - , where X - any amino acid. There was an additional pair of conserved amino acids found within the X_{12-13} spacers: an aromatic amino acid, F or Y, and a branched aliphatic amino acid, commonly L. The pairs of C and H residues were proposed to tetrahedrally coordinate Zn^{2+} ions.

The term " zinc finger" was coined by Miller et al. (1985) who proposed that the 12 - 13 -residue loop or "finger" formed the DNA binding surface (Figure 6.1A). The model of tetrahedral Zn^{2+} coordination was confirmed by extended X-ray absorption fine structure (EXAFS) which showed that the zinc absorption edge in TFIIIA was fit to a tetrahedral coordination sphere with two sulfur and two nitrogen atoms as ligands (Diakun et al., 1986). The zinc is bound to conserved histidine $N^{\varepsilon 2}$ atoms at 2 Å distance, and to cysteine S^{γ} atoms at 2.3 Å distance, with an affinity varying from 10^{-09} M to 10^{-12} M, depending on a particular zinc finger sequence (Diakun et al., 1986).

The stoichiometry of zinc metal calculated for TFIIIA showed it to contain 7 to 11 zinc atoms, which was more consistent with the 9 zinc fingers of TFIIIA than the originally reported 2 zinc ions per TFIIIA molecule (Miller et al., 1985). Proteolysis studies, revealing periodic intermediates of a limited digest coupled with the analysis of cDNA sequence of TFIIIA (Miller et al., 1985) supported the general idea of a repeated structural unit, each of which bound a single zinc ion. Model structures for the zinc finger motif have been put forward by Berg (1988) and Gibson et al. (1988). The first NMR structure of

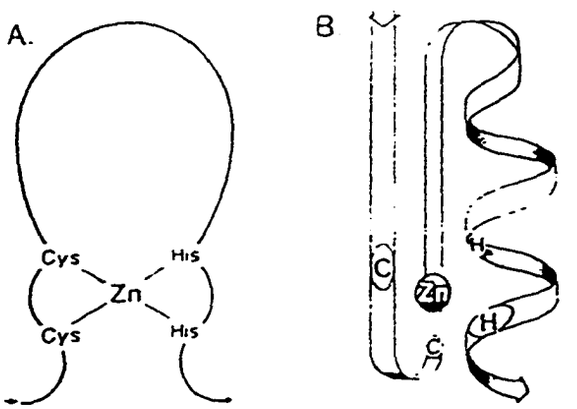


Figure 6.1 Schematic representation of a C_2H_2 zinc finger (Evans & Hollenberg, 1988).

a zinc finger was determined by Parraga et al. (1988) and Lee et al. (1989) and resembled more closely the model of Berg. Since then several NMR structures for zinc fingers have been published (Klevit et al., 1990; Omichinski et al., 1990, 1992; Kochoyan et al., 1991a, b; Neuhaus, et al., 1992). Model building, employing an insightful use of the structural data bases, two-dimensional nuclear magnetic resonance analysis and X-ray crystal structures have separately suggested an approximate structure for these zinc fingers.

The structure of each of the zinc finger domains consists of two antiparallel β-strands and an α-helix surrounding a centrally bound zinc ion (Figure 6.1 B). Three-dimensional structures now available for a number of zinc finger proteins very nicely accommodate the conserved sequence features. The cysteine and histidine side chains are involved in zinc coordination, and the three other conserved residues pack to form a hydrophobic core adjacent to the metal coordination unit. The use of hydrophobic amino acids by zinc fingers to stabilize local zinc protein structure differs from the folding strategy of metallothioneines which lack aromatic side chains.

As was proposed earlier (Berg, 1988), the α-helix in a zinc finger is directly involved in sequence-specific recognition, interacting in the major groove of the DNA. Thus, the zinc finger provides yet another way of inserting an α-helix - the most common mode of DNA recognition - into the DNA major groove.

The overall fold of an individual zinc finger has been shown to be maintained entirely by the zinc coordination as well as the packing of the hydrophobic framework residues. Michael et al. (1992) had used a minimalist finger peptide consisting of 26 amino acids, 16 of which were alanine residues,

and thus provided direct evidence for the importance of zinc coordination and the central hydrophobic core for maintaining the tertiary fold.

Once folded, zinc fingers are very stable structures, displaying little internal motion (Palmer et al., 1991), and resistance to protease digestion (Miller et al., 1985) or thermal denaturation (Frankel et al., 1987). Zinc fingers isolated under metal chelating conditions display no sequence specific DNA binding ability (Frankel et al., 1987; Lee et al., 1991a). Therefore, the tertiary structure of a protein is inseparable from its function.

Significant variation exists within the zinc requiring protein families, particularly within the Cys_2His_2 class of zinc finger proteins. The differences include such structural features as the nature and positioning of hydrophobic core residues, the spacing between the coordinating cysteine and histidine residues, the sequence of the interfinger linker regions, or the number of β-strands in the zinc fingers. Two proteins, SWI5 and Tramtrack, deviate from the classic ββα zinc finger fold. The amino terminal finger of SWI5 has an irregular β-sheet consisting of three strands as opposed to the canonical two (Nakaseko, 1992; Neuhaus, 1992). Interestingly, folding of finger 1 appears to be dependent on the presence of the additional strand. Another example of a three-stranded β-sheet was discovered in the amino-terminal finger of the *Drosophila* transcription repressor protein Tramtrack (Fairall et al., 1993). The third strand, while essential for DNA-binding activity, makes no direct contact with DNA in a co-crystal structure, suggesting it is somehow important for maintaining finger structure. It remains to be seen why these SWI5 and Tramtrack fingers differ from other known C_2H_2 fingers.

The extent of zinc finger occurrence in eukaryotes is astonishing. It was estimated that as much as 1% of the DNA in human cells specifies zinc fingers (Rhodes and Klug, 1993). In chromosome 19 the figure is as high as

8% (Rhodes and Klug, 1993). The zinc-finger containing proteins that have been identified so far carry from as few as two zinc fingers (MIG1), found in yeast, to as many as 37 tandem fingers (Xfin) in *Xenopus*. Well over 1300 zinc finger cDNA sequences have now been reported, suggesting that the C_2H_2 zinc finger domains represent the most abundant DNA binding motif in eukaryotic transcription factors.

These transcriptional regulators are involved in a multitude of functions, such as differential gene expression [MIG1 (yeast), MBP-1 (human)], developmental gene regulation [Kruppel and Tramtrack from *Drosophila* , EGR-1, mKr2, Krox-20 from mouse, Xfin, p43 and TFIIIA from *Xenopus*, tumor suppression [WT1 (identified in human and mouse)], human oncogenesis [GLI (human)], ubiquitous transcriptional regulation [Sp1], to name but a few.

The term "zinc finger" is sometimes loosely applied to other classes of zinc containing proteins, and has been used to refer to to almost any sequence that has a set of cysteines and/or histidines within a relatively short polypeptide chain. These other classes of proteins include large families such as the nuclear hormone receptor superfamily (Evans, 1988), the RING finger protein family (Freemont, 1993), which includes the breast and ovarian cancer susceptibility gene BRCA 1 (Miki et al., 1994) and the retroviral CCHC motifs represented by HIV-1 nucleocapsid protein (Blake and Summers, 1994). Recently, an NMR structure of human transcriptional elongation factor TFIIS has been reported (Quian et al. 1993a). This small domain binds nucleic acid and consists of a three-stranded β-sheet. The zinc is coordinated tetrahedrally by four cysteine residues. This type of structure is novel among the zinc-binding and nucleic acid-binding domains and thus represents a further striking example of diversity of zinc-containing motifs. Additionally, other

proteins have been shown to contain structural zinc ions, such as the tumor suppressor p53 (Cho et al., 1994) and the human growth hormone-prolactin receptor complex in which the zinc forms a bridge between the hormone and the receptor (Somers et al., 1994).

To give an appreciation for the diversity of zinc-binding motifs, the schematic representation in Figure 6.2 shows zinc modules from various protein families that differ in metal coordination strategy, secondary structure, and modularity, giving rise to distinctly different global folds. Each of these molecules exhibits nucleic acid recognition strategies that are unique among the zinc modules.

From an evolutionary point of view, it could be argued that the C_2H_2 class evolved from a single primordial ancestor, and that multifinger proteins are the result of internal gene duplication. This hypothesis is supported by the conservation of the zinc-coordinating, hydrophobic and H/C link residues, as well as the fact that most zinc fingers are encoded by separate exons (six of nine TFIIIA fingers are localized to separate exons (Tso et al., 1986)).

To rationalize the evolution of zinc fingers from the point of view of added diversity, one might argue that the zinc fingers have originally evolved as simple repeating motifs that could interact with RNA molecules (which inhabited the prebiotic soup before the DNA), and which later evolved to interact with DNA. Thus, the ability of TFIIIA to interact with both DNA and RNA suggests that perhaps it is an example of an original nucleic acid binding protein.

The reason for linking several zinc finger units together to create a functional protein can be underscored by the fact that isolated single fingers do not recognize DNA sequences specifically (Parraga et al., 1988). For

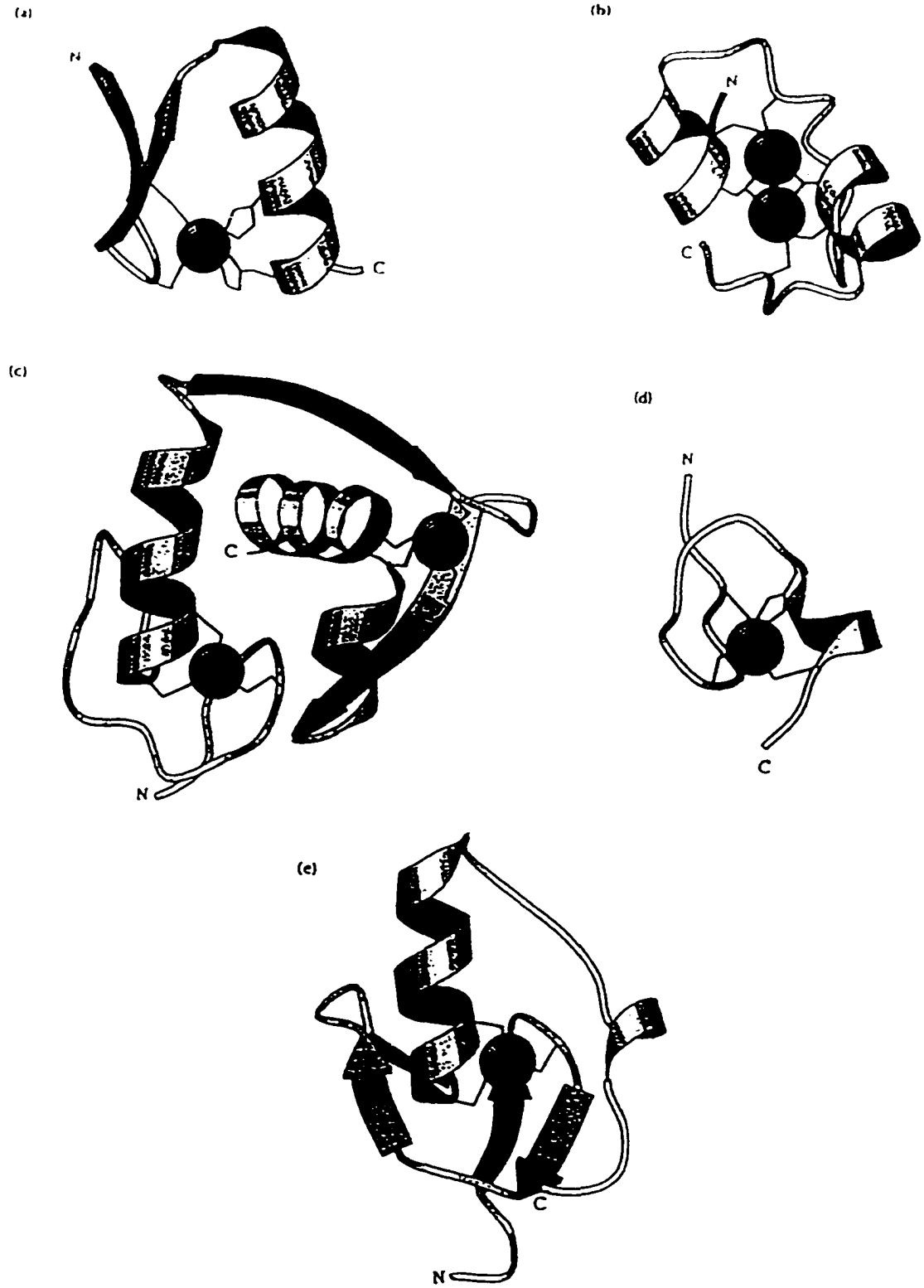


Figure 6.2 Representative structures from various zinc finger families, illustrating structural diversity among zinc-binding and nucleic acid-binding domains (Schmeideskamp & Klevit, 1994).

Figure 6.2 (a) EGR-1 finger 2 of C_2H_2 type; (b) the GAL4 Zn_2C_6 binuclear cluster; (c) the glucocorticoid receptor DNA-binding domain; (d) the amino-terminal retroviral-type CCHC finger of HIV-1 nucleocapsid protein; (e) the carboxy-terminal metal-binding domain of GATA-1.

example, in a three-finger protein SWITCH 5 (SWI5) two linked fingers had to be present for the protein to be able to interact with DNA with a reasonable affinity (Nakaseko et al., 1992). By then applying NMR to the first two zinc finger motifs of SWI5, it was confirmed that adjacent zinc fingers do not meld with each other, that they are truly independent "reading heads" joined together by flexible linkers.

This modular nature of zinc finger proteins is unique among the classes of DNA binding proteins. From an evolutionary point of view, producing a protein with multiple fingers could be advantageous for several reasons: such a protein could recognize longer stretches of DNA (important for an organism with a larger genome size); such a transcriptional regulator could bind DNA with higher specificity, and thus produce more stable interaction; finally, the specificities of individual fingers could be combined, giving rise to new binding specificities.

When speaking of "modularity" of zinc finger proteins, the independence of each zinc finger structure doesn't always imply that finger-finger interactions do not exist. For most zinc finger proteins examined, there are a few interfinger interactions taking place, and in general they are not conserved between different proteins.

For example, in the EGR-1 crystal structure there are conserved hydrogen bonds observed between zinc fingers 1 and 2 where an arginine located in the C-terminal portion of an α-helix of the first finger hydrogen bonds to the backbone carbonyl of a serine located in the turn between the β–sheet and the α-helix of the next finger. A similar interaction is seen for fingers 2 and 3, where arginine hydrogen bonds to the backbone carbonyl of an alanine residue (Pavletich & Pabo, 1991). A similar pattern of interfinger

interactions is observed in an NMR structure of MBP-1 protein (Omichinski et al., 1992): the two zinc fingers interact through hydrogen bonding of amino acids located in the same regions of the zinc fingers as seen in EGR-1 (C-terminal region of an α-helix of a first zinc finger and the "tip" region of the next finger) (Figure 6.3). The GLI - DNA co-crystal structure has revealed that both hydrogen bond and hydrophobic interactions occur, especially between zinc fingers 1 and 2 (Fairall et al. ,1993). As mentioned earlier, an NMR structure of SWI5 (Nakaseko et al., 1992) revealed no interfinger contacts. However, a possibility exists that the adjacent zinc fingers may come into contact when complexed with DNA.

Although the focus of this overview is the structure and function of zinc domains, it should be kept in mind that DNA binding is only the initial step in gene activation/repression. The majority of the polypeptide chain in the native zinc-containing transcription factors is taken up by the transactivating domains which are involved in interactions with other proteins, most notably the components of the basal transcriptional machinery. Figure 6.4 shows the schematic representations of complete zinc-containing transcription factors.

6.2 Structure/function analysis of TFIIIA

6.2.1 Purification and characterization of TFIIIA gene product

The fact that TFIIIA represents one of the best studied transcription factors is, perhaps, best explained by its great abundance in a naturally occurring system: *Xenopus* oocytes, which contain approximately 0.7-1 x 10^{12} molecules of TFIIIA per cell (Shastry et al., 1984). In the cytoplasm, TFIIIA predominantly exists as a component of a 7S ribonucleoprotein particle, which is composed of one molecule of TFIIIA and one molecule of 5S rRNA

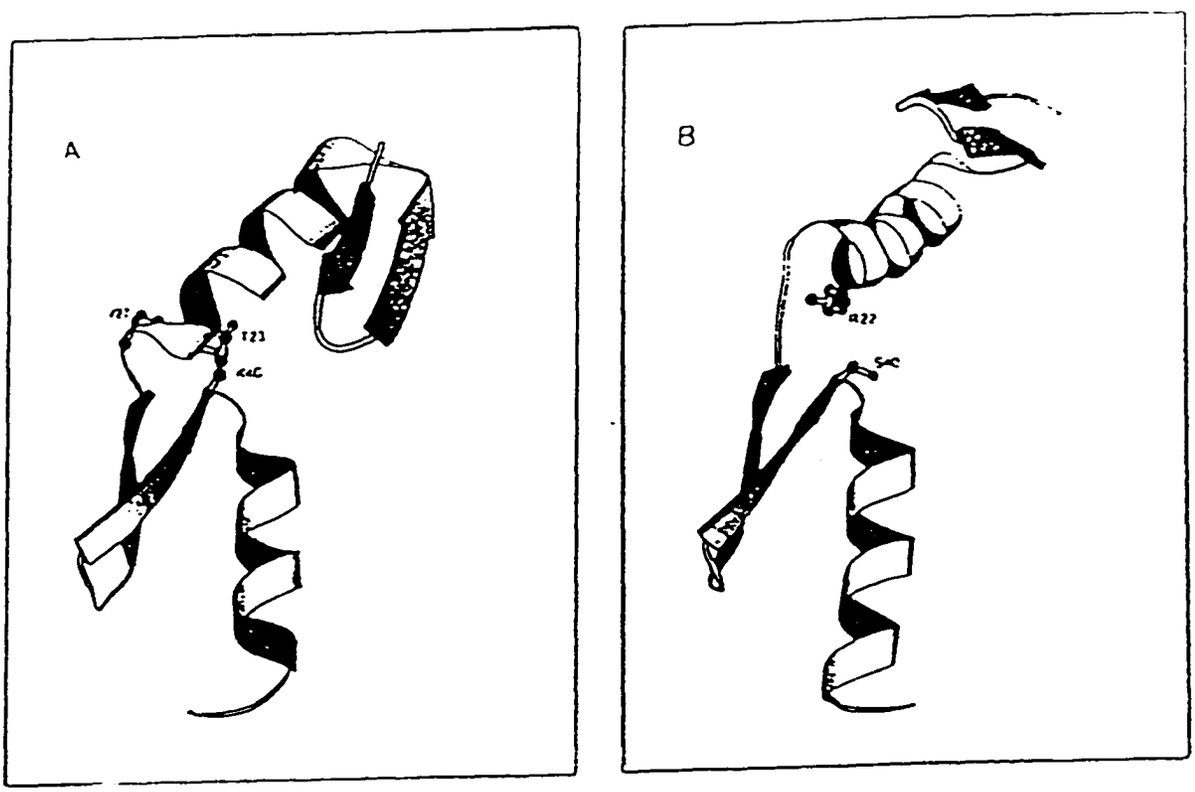


Figure 6.3 Schematic representation of the interfinger orientations of (**A**) MBP-1 zinc fingers and (**B**) EGR-1 fingers 1 and 2 (Omichinski et al., 1992).

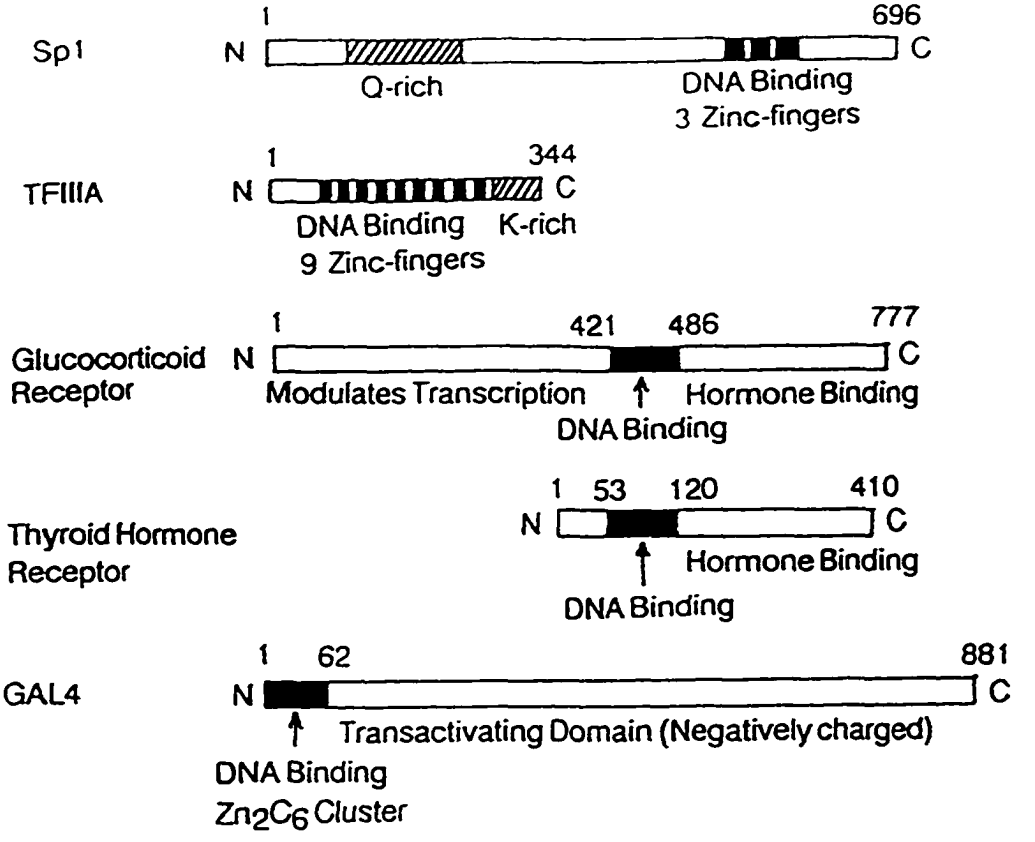


Figure 6.4 Diagrams of the functional domains of eukaryotic zinc-containing transcription factors (Coleman, 1992).

(Picard & Wegnez, 1979). Using conventional chromatography, TFIIIA was purified to homogeneity (Picard & Wegnez, 1979). Only later, when it was shown that TFIIIA could specifically bind DNA (Engelke et al., 1980) was it realized that one of the roles of TFIIIA is that of transcriptional regulation (Engelke et al., 1980; Pelham & Brown, 1980; Honda & Roeder, 1980). Other methods for TFIIIA purification have been developed, and include density-gradient centrifugation followed by DEAE cellulose chromatography (Smith et al., 1984), ion exchange chromatography (Romaniuk, 1985), as well as expression of recombinant TFIIIA in *E.coli* (Del Rio & Setzer, 1991).

The *Xenopus laevis* TFIIIA is a 344-amino acid, 38.5 kDa protein, containing nine C_2H_2 zinc finger motifs located at the aminoterminus (Brown et al., 1985; Miller et al., 1985). TFIIIA protein is proposed to be an extended rather than a globular protein (Reynolds & Gottesfeld, 1983; Hanas et al., 1984) based on its mode of DNA binding. First, TFIIIA binds a long stretch of DNA without compacting it into a tighter conformation or significantly changing the DNA linking number (Reynolds & Gottesfeld, 1983; Hanas et al., 1984). Second, proteolytic fragments of TFIIIA retain their ability to bind specific DNA regions, which suggests that TFIIIA must be in an extended conformation when interacting with DNA, to be able to cover the long internal control region (Ginsberg et al., 1984; Miller et al., 1985).

The amino acid sequence of the DNA-binding domain of TFIIIA is shown in Figure 6.5, and a schematic diagram showing TFIIIA alignment along the ICR is given in Figure 6.6, where the 30 kDa N-terminal DNA-binding domain, and a 10kDa C-terminal transcriptional activation domain are indicated (Smith et al., 1984). Amino acids 294 to 313 (Vrana et al., 1988), and 284 to 297 (Mao & Darby, 1993) within the C-terminus have been shown to be important for the transactivational function of TFIIIA. This region is

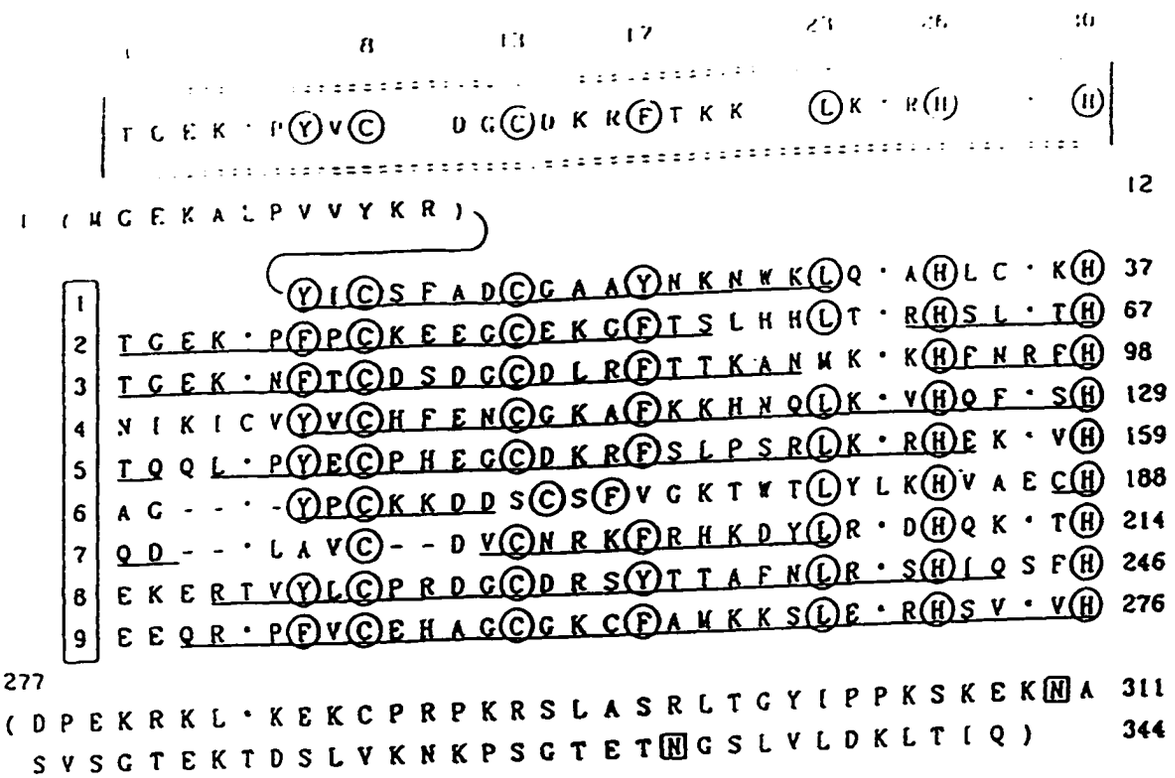


Figure 6.5 Amino acid sequence of the nucleic acid binding domain of TFIIIA (Miller et al., 1985).

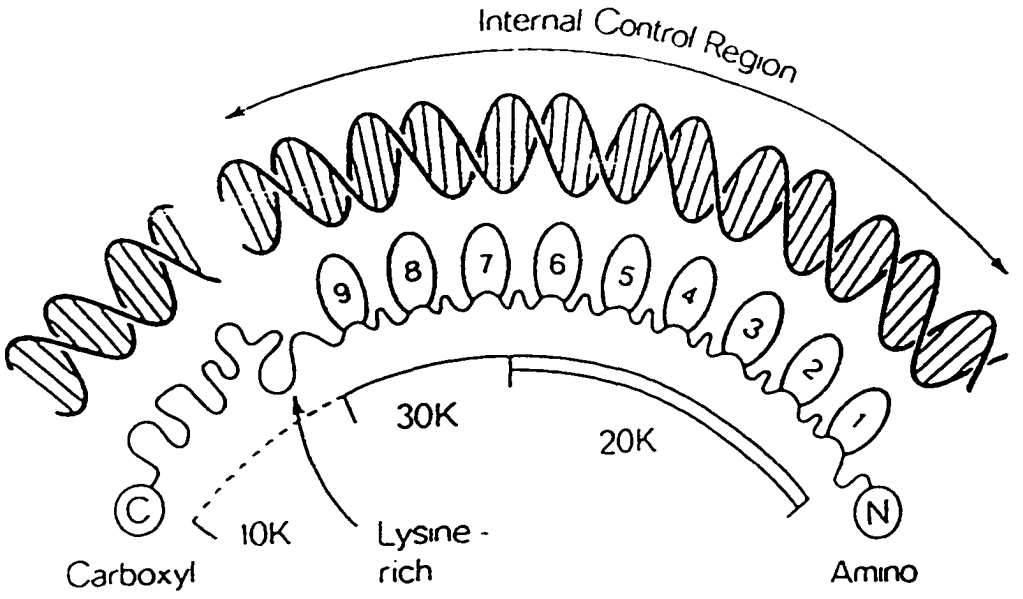


Figure 6.6 Schematic diagram illustrating alignment of TFIIIA zinc finger region along ICR (Miller et al., 1985).

rich in lysine residues which are proposed to mediate interactions with other components of the transcriptional machinery (Smith et al., 1984). Potential N-linked glycosylation sites were found in the C-terminal region of TFIIIA (Miller et al., 1985). However, the role of this post-translational modification is not yet clearly determined. The elucidation of the zinc finger structure of the TFIIIA DNA-binding domain is described in the previous section of the chapter.

6.2.2 TFIIIA gene expression

Xenopus laevis contains a single copy of the gene encoding TFIIIA per haploid genome (Ginsberg et al., 1984; Taylor et al., 1986, Tso et al., 1986; Scotto et al., 1989). The TFIIIA gene is approximately 11 kilobasepairs in length and contains nine exons separated by eight introns (Tso et al., 1986). The relative position of exons and introns, as well as the fact that 6 out of 9 zinc fingers of TFIIIA are encoded by separate exons suggest that TFIIIA evolved from a primordial ancestor specifying a zinc finger motif (Tso et al., 1986). The TFIIIA promoter was found located 5' to the transcriptional start site, and has features characteristic of many polymerase II transcribed genes (Tso et al., 1986). A consensus TATA box sequence is at position -32, and a CAAT box sequence is at position -96 (Tso et al., 1986). Further regulatory sequences include seven positive and four negative cis-acting elements, identified by deletion studies (Matsumoto & Korn, 1988; Hall & Taylor, 1989; Scotto et al., 1989; Pfaff et al., 1991; Pfaff & Taylor, 1992).

The levels of TFIIIA mRNA and protein reach their maximum in the early stages of oogenesis (approximately 10^{12} molecules/cell), and drop to about $4x10^5$ molecules per cell in later stages of a tailbud embryo (Shastry et al., 1984; Andrews & Brown, 1989). Since the expression of TFIIIA is tightly

developmentally regulated (Andrews & Brown, 1987), both housekeeping and stage-specific transcription factors are expected to interact with the TFIIIA promoter (Pfaff et al., 1991). A protein that binds to one of the positive regulatory sequences was identified and shown to recognize the core sequence CACGTG located at positions -269 to -264 (Scotto et al., 1989). This protein, termed the distal element factor (TDEF), is homologous to the adenovirus major late transcription factor, which belongs to the helix-loop-helix class of transcriptional regulators (Sawadogo & Roeder, 1985; Hall & Taylor, 1989; Scotto et al., 1989). The TDEF protein occurs in both oocyte and somatic cells. However, it was shown to regulate transcription only in oocytes, suggesting that it acts in a cell-specific fashion (Hall & Taylor, 1989).

Additional level of complexity in TFIIIA gene expression is introduced by the demonstration that the oocyte and somatic forms of TFIIIA are expressed from two different, overlapping promoters, utilizing distinct transcriptional start sites (Kim et al., 1990; Martinez et al., 1994). Moreover, it was recently demonstrated that both RNA polymerases II and III are involved in TFIIIA gene expression in somatic cells (Martinez et al., 1994). However, the significance of the involvement of the two RNA polymerases is not yet understood. Even though remarkable progress has been made toward understanding the regulatory mechanisms of TFIIIA gene expression, several questions still remain, such as the roles of other transcription factors, the fine details of the differential regulation of somatic versus oocyte TFIIIA genes, as well as the role of chromatin structure.

6.2.3 Roles of TFIIIA in transcription and storage of 5S rRNA

Transcription factor IIIA has a dual role as a positive transcription factor required for transcription of 5S rRNA genes (Honda & Roeder, 1980), as

well as a storage protein for 5S rRNA (Engelke et al., 1980). Association of TFIIIA with the 5S rRNA also serves to transport 5S rRNA from the nucleus into the cytoplasm where it is stored until it is required for ribosomal assembly in immature oocytes (Guddat et al., 1990). The binding of TFIIIA to the 5S rRNA gene nucleates an ordered assembly of a transcription preinitiation complex (Segall et al., 1980; Shastry et al., 1982; Bieker et al., 1985). First, TFIIIA binds to the internal control region of 5S rRNA gene, and forms a metastable complex (Engelke et al., 1980; Bieker & Roeder, 1984). Next, TFIIIC binds and stabilizes this complex (Lassar et al., 1983). Subsequently, TFIIIB binds (Bieker et al., 1985; Bieker & Roeder, 1986), and is responsible for maximal transcription rates (Kassavetis et al., 1990). The preinitiation complex is very stable, and persists through many cycles of transcription by RNA polymerase III (pol III) (Bogenhagen et al., 1982; Setzer & Brown, 1985). Transcription of the 5S rRNAs from *Xenopus* has become a model system for the study of the mechanisms that control eukaryotic RNA synthesis.

In eukaryotic cells, RNA polymerase III is responsible for the transcription of genes encoding 5S rRNA, tRNA, several cytoplasmic and nuclear RNAs (7S K and L, 4.5 S RNAs) and some virus-associated RNAs (adenovirus-associated VAI and VAII, EBERI and EBERII) (Ciliberto et al., 1983a). The most striking common feature shared by the genes transcribed by pol III (with several exceptions) is that promoters are contained in the coding sequence (Sakonju et al., 1980; Bogenhagen et al., 1980). Other similarities include the fact that Pol III genes are quite short compared to the average length of pol I or pol II genes (Ciliberto et al., 1983); the coding sequence is defined at the 3' end by at least one cluster of T residues, which are required for efficient transcriptional termination (Bogenhagen & Brown, 1981).

The binding of TFIIIA to 5S rRNA forms a 7S ribonucleoprotein (RNP) particle which serves to stabilize and transport the 5S RNA molecule (Guddat et al., 1990). Upon export from the nucleus to the cytoplasm, the ribosomal protein L5, which also binds 5S rRNA, displaces TFIIIA from its interaction with 5S rRNA, and the complex of L5 and 5S rRNA is incorporated into a ribosome in the nucleolus (Wormington, 1989).

6.2.4 Structure and function of 5S rRNA and its gene

The 5S rRNA is an essential component in all ribosomes, prokaryotic and eukaryotic, except those from mammalian mitochondria (Kjems et al., 1984). The first discovery of ribosomal 5S rRNA was made in 1963 (Erdmann et al., 1987). Since then, their structure and function have been studied extensively. The ribosomal RNA components are among the most highly conserved macromolecules in all living systems. Sequence comparisons of a large number of 5S rRNAs isolated from a wide range of organisms revealed a high degree of structural conservation throughout evolution (Hori & Osawa, 1987; Erdmann et al., 1987; Wolters & Erdmann, 1988). Due to this high conservation, the sequences of 5S rRNA genes and their transcripts were used to measure phylogenetic distances between distantly related organisms.

There are three classes of 5S rRNA genes: major-oocyte (Xlo), trace-oocyte (Xlt), and somatic (Xls) that are transcribed (Fedoroff & Brown, 1978; Peterson et al., 1980). The haploid *Xenopus* genome contains approximately 20,000 copies of the major-oocyte type genes which are transcribed only in oocytes , and about 400 copies of somatic 5S rRNA genes that are transcribed in both oocyte and somatic cells (Peterson et al., 1980). There are six nucleotide differences between the somatic and the oocyte gene sequences (Xing & Worcel, 1989). Trace oocyte genes are represented by approximately

2,000 copies per haploid genome, and combine some features of the oocyte-type genes, as well as somatic type (Brown et al., 1977). In addition, there are approximately 24,000 copies of 5S rRNA pseudogenes per haploid genome that have no known function since they do not produce a transcript (Fedoroff & Brown, 1978). The organization of the *Xenopus* 5S rRNA genes is shown in Figure 6.7. The main components of each gene copy are the GC-rich region which includes the gene and the pseudo-gene sequences, and a variable AT-rich block, called the spacer DNA (Miller et al., 1978; Fedoroff & Brown, 1978).

Based on the extensive sequence comparisons, a consensus model for a 5S rRNA molecule has been proposed, and is shown in Figure 6.8 (Delihas & Andersen, 1982). These molecules are 120 nucleotides long, have a molecular weight of about 40 kDa, and assume a Y-shape composed of five helical regions (I through V), and five loop domains (designated A through E) (Delihas & Andersen, 1982; Küntzel et al., 1983; Delihas et al., 1984).

6.2.5 Role of TFIIIA zinc fingers in RNA recognition

The dual function of TFIIIA as a positive transcription factor and a storage protein involved in regulation of 5S rRNA gene expression requires it to bind to two different nucleic acids - DNA and RNA (Engelke et al., 1980; Pelham & Brown, 1980). TFIIIA interacts with DNA and RNA in a mutually exclusive manner (Pelham & Brown, 1980). Except for the Wilms' tumor protein (Caricasole et al., 1996), TFIIIA is the only other C_2H_2 class protein capable of interacting with both DNA and RNA. In parallel to the study of TFIIIA-5S DNA interaction, the RNA binding function of the factor has been the subject of intense investigations.

TFIIIA - 5S rRNA equilibrium binding was characterized by Romaniuk (1985). The factor binds 5S rRNA with an apparent association constant of 1.4

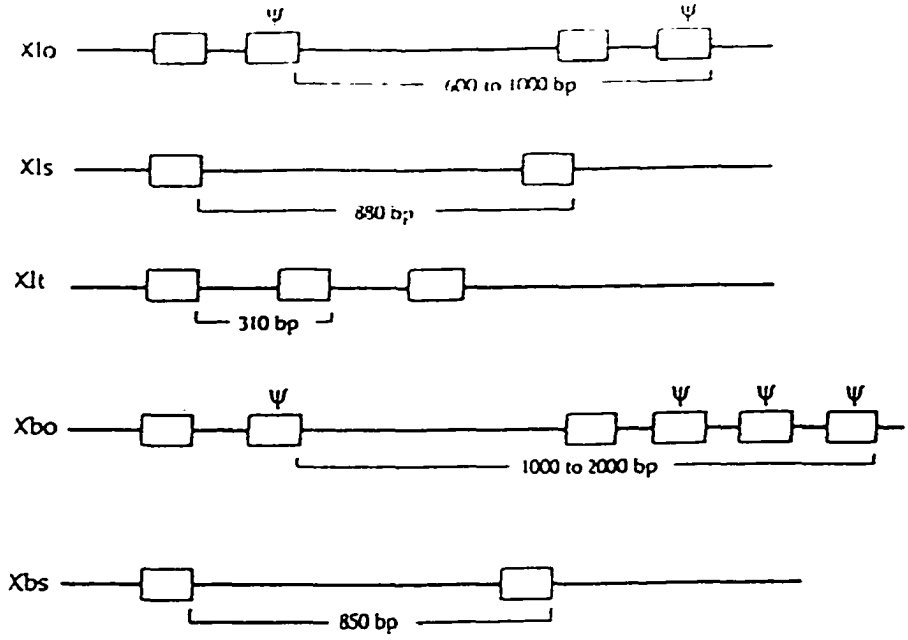


Figure 6.7 Organization of the *Xenopus* 5S rRNA genes. Pseudogenes are designated by a 'Ψ' (Korn, 1982).

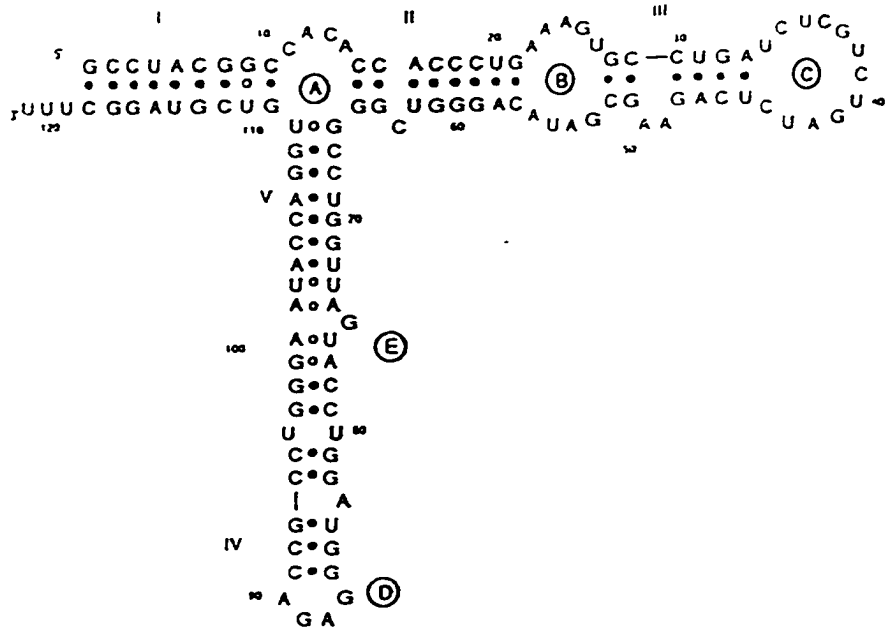


Figure 6.8 The secondary structure of 5S rRNA (Delihas & Andersen, 1982). 5S rRNA complex exhibits a broad pH optimum, and is favored by both entropy and enthalpy (Romaniuk, 1985).

x 10^9 M^{-01} under the conditions of 0.1 M salt, pH 7.5 and 20°C. The affinity of the nonspecific interaction with $tRNA^{Phe}$, on the other hand, is at least two orders of magnitude lower than for the 5S rRNA. Mutagenesis studies covering the entire 5S rRNA molecule complemented by nuclease digestion and hydroxyl radical footprinting cumulatively suggest the following features of TFIIIA - 5S rRNA interaction (Anderson & Delihas, 1986; Baudin et al., 1989; Christiansen et al., 1987; Pieler et al., 1984; Romaniuk, 1985; Romaniuk et al., 1988; Romaniuk, 1989; Baudin et al., 1989; You & Romaniuk, 1990; de Stevenson et al., 1991; You et al., 1991; McBryant et al., 1995; Zhang et al., 1995; Rawlings et al., 1996): almost all of the 5S rRNA molecule is involved in TFIIIA interaction, especially helices II, IV & V and loops A & E; the binding of TFIIIA is most notably altered by changes in the higher order structure, rather than base-specific contacts. The important role of RNA conformation for TFIIIA recognition is also evident from the experiments using 5S rRNAs isolated from different species (Hanas et al., 1984; Pieler et al., 1984b; Romaniuk, 1985; Andersen & Delihas, 1986). TFIIIA has comparable affinities for various eukaryotic and bacterial 5S rRNAs, supporting the notion that the overall three-dimensional fold of 5S rRNA molecules is the major determinant for specific interactions.

Since TFIIIA interacts with two different nucleic acids, it wasn't clear whether DNA and RNA interactions are mediated by different zinc fingers, and what features of individual zinc fingers were responsible for this differential activity. Hence, considerable effort went into defining high-affinity RNA-binding fingers of TFIIIA. It was soon found that while the amino-terminal three zinc fingers predominate in DNA binding, the central zinc finger triplet - fingers 4 to 7 - primarily mediate high-affinity RNA binding (Theunissen et al., 1992; Clemens et al., 1993). In fact, a polypeptide

consisting of fingers 4 to 7 has been shown to have higher affinity for 5S rRNA than the full length protein (McBryant et al., 1995). Similarly, Setzer et al. (1996) reported that polypeptide containing zinc fingers 4 to 9 bound 5S rRNA with a two-fold higher affinity than the full-length protein. This suggests that fingers 4 to 7 contain the necessary information for RNA recognition, as well as that there is a certain degree of interference between the zinc fingers. Combining the available RNA mutagenesis data (Sands & Bogenhagen, 1987; You et al., 1991; Theunissen et al., 1992) with the results using truncated TFIIIA mutants Clemens et al. (1993) proposed a model in which individual fingers are assigned to distinct structural elements of 5S rRNA (Figure 6.9). A peptide containing zinc fingers 5 to 7 was shown to have a 30-fold lower 5S rRNA affinity than the one with zinc fingers 4 to 7, indicating an important contribution for finger 4 (Clemens et al., 1993). Mutagenesis of finger 4 resulted in a four-fold reduction in affinity relative to the wild-type protein, while mutations in finger 6 showed a 30-fold drop in binding affinity (Clemens et al., 1993). Perhaps, the mutation introduced in finger 4 involved a residue (Gln121) which makes nonspecific backbone contacts, whereas a threonine (Thr176) in finger 6 may participate in a base-specific contact. This would explain the difference in the magnitude of affinity change for the two mutations. It is interesting to point out that all three mutations introduced in finger 6 produced significant reduction in binding affinity. These mutations were made in the amino acid triplet TWT which are the surface residues of finger 6 (Clemens et al., 1993). This triplet sequence is conserved in the same position in finger 6 of another 5S rRNA-binding protein, p43 (Joho, 1990). Like TFIIIA, p43 protein contains nine zinc fingers, seven of which are the same size as their counterparts in TFIIIA (Joho et al., 1990), (Figure 6.10). In addition, unlike TFIIIA, p43 only binds to 5S

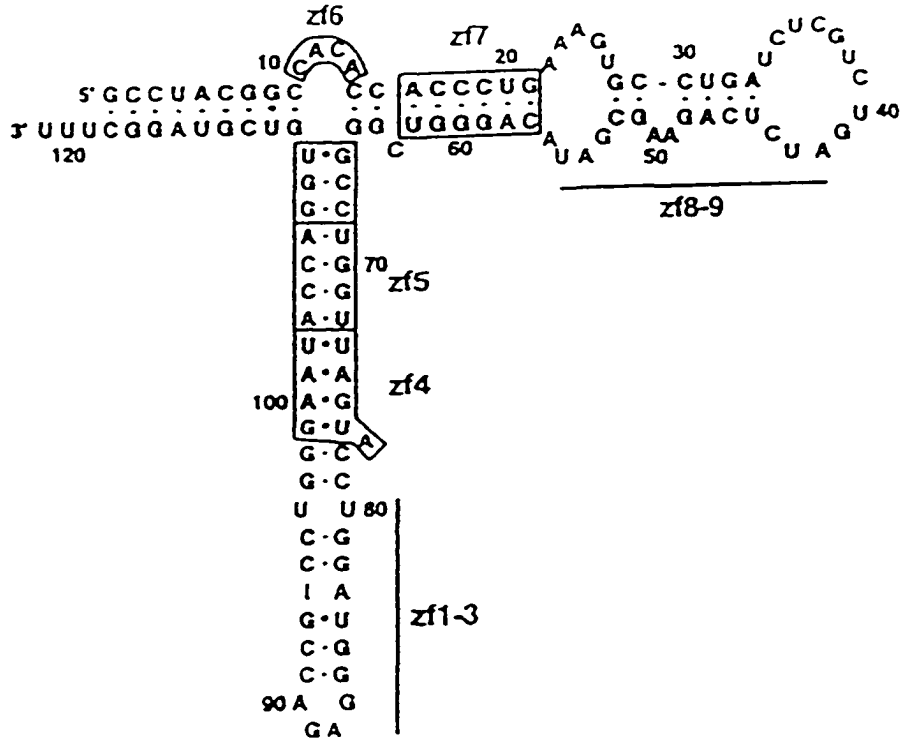


Figure 6.9 Model for the TFIIIA - 5S rRNA interaction. Proposed zinc finger positions are indicated. Nucleotides enclosed in boxes indicate that most important regions for recognition (Clemens et al., 1993).

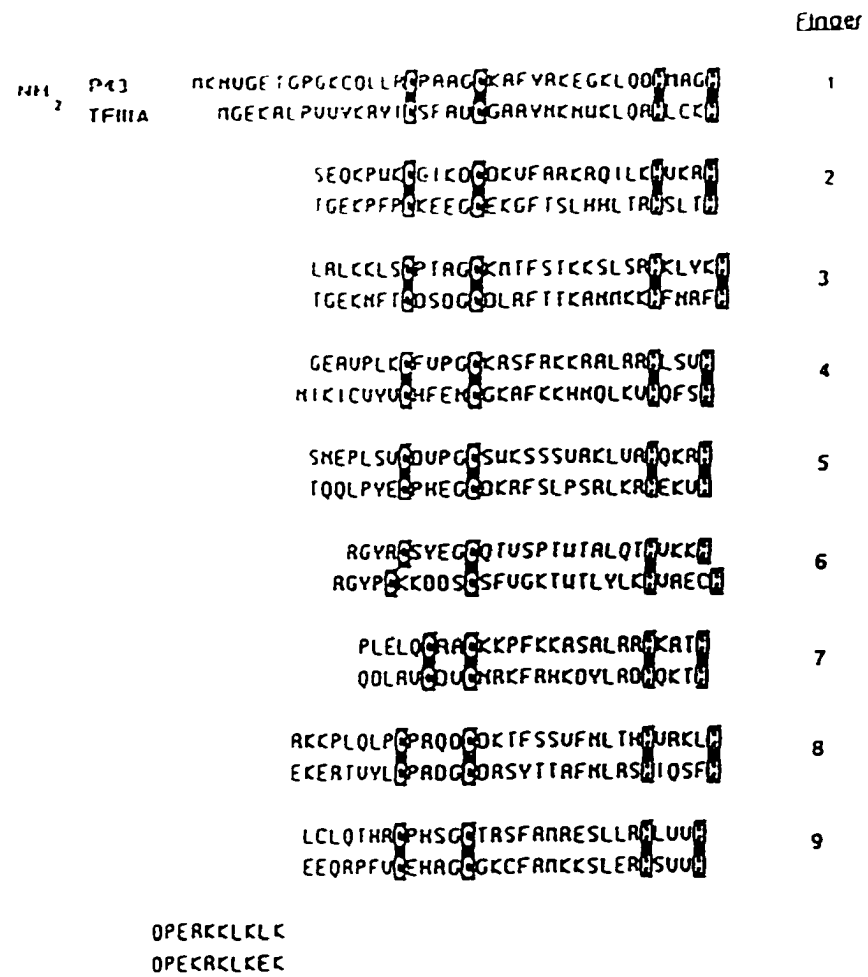


Figure 6.10 Alignment of zinc finger sequences of TFIIIA and p43 proteins. The shaded amino acids represent the zinc-coordinating cysteine and histidine residues (Joho et al., 1990).

rRNA, and has no DNA-binding ability. There is, however, only a limited amino acid sequence identity (33%) shared by the two proteins. This amino acid identity extends to the consensus zinc ligands (C & H), hydrophobic amino acids (Miller et al., 1985; Brown et al., 1985) and the TWT motif in finger 6 (Joho et al., 1990). The fact that only structure-determining amino acids are conserved between the two proteins suggests that it is not the overall three-dimensional fold of a zinc finger that confers specificity for either DNA or RNA.

The importance of fingers 4 and 6 for 5S rRNA binding is further demonstrated by the complementary 5S rRNA mutagenesis data. Fingers 4 and 6 were shown to bind to loop E and loop A regions of 5S rRNA, respectively (Clemens et al., 1993; McBryant et al., 1995). Loop A represents the central hinge region of the 5S rRNA, and is proposed to orient the adjacent arms of the molecule (Westhof et al., 1989). Likewise, loop E was also shown to provide important secondary and tertiary features necessary for specific TFIIIA interaction (Romaniuk, 1985; Huber & Wool, 1986; Christiansen et al., 1987; Wimberly et al., 1993). Thus, protection and missing nucleoside experiments emphasize the role of loop E as a major structural determinant for TFIIIA - 5S rRNA recognition. Similarly, substitutions in loop A result in reduction of TFIIIA binding affinity (Sands & Bogenhagen, 1987; Romaniuk, 1989; You et al., 1991). Although secondary and tertiary structure of 5S rRNA has been implicated as the major determinant of TFIIIA binding, it is still not clear whether loop A residues provide direct base-specific contacts or serve to preserve the overall integrity of the 5S rRNA molecule.

In a recently published study RNA binding by a zinc finger 4-7 peptide of TFIIIA was assayed using alanine substitutions made in the four positions

predicted to be α-helical in each of the RNA-binding fingers 4-7, as well as by phage display of zinc finger 4 (Friesen & Darby, 1997). Substitutions in fingers 4 and 6 resulted in the most significant reduction in RNA affinity: 77- and 30-fold, respectively. In contrast, substitutions in fingers 5 or 7 did not appreciably affect the RNA-binding ability. The effect of histidine to asparagine substitution in zinc fingers 4, 5 and 6 studied in the context of the full-length TFIIIA showed only a modest reduction in RNA binding: 2.6 - 4.1-fold (Setzer et al., 1996). This effect could probably be explained by supposing that the remaining fingers assume an alternative mode of binding, and compensate for the mutated ones. Other researchers observed a reduced affinity of fingers 5-9 compared to fingers 6-9, and a 60-fold decrease in affinity for a peptide containing fingers 1-6 relative to the 1-7 finger fragment (Darby & Joho, 1992; Clemens et al., 1993), suggesting that fingers 5 and 7 do provide an important contribution to the binding affinity. It is possible that amino acid determinants providing RNA specificity in fingers 5 and 7 are different from those mutated in the "alanine substitution" study (Friesen & Darby, 1997), and lie elsewhere in the zinc fingers. RNAse protection assays showed alterations visible only for mutant fingers 4, 5 and 6 (Setzer et al., 1996). Finger 5 and 6 mutants showed a loss of protection over helices II, IV & V (Setzer et al., 1996). In contrast, no protection could be detected for the 3-finger amino terminal fragment, even though it was shown to bind 5S rRNA with affinity that was only two-fold lower relative to the wild-type protein (Setzer et al., 1996). An explanation offered by the authors is that perhaps the finger 1-3 fragment - 5S RNA complex undergoes a very rapid dissociation, and the 5S RNA is cleaved by the CV1 RNAse shortly thereafter, which results in a lack of an RNAse footprint.

In an attempt to define amino acids in zinc finger 4 important for RNA binding, a phage-display method was used to screen a collection of mutant finger 4 sequences in which amino acids at positions -1, +2, +3 and +6 were randomized to generate the mutants with every possible amino acid substitution at these four positions (Friesen & Darby, 1997). These positions are used by EGR-1 protein for making direct hydrogen bond contacts to the DNA bases (Pavletich & Pabo, 1991), and were, by analogy, presumed important for TFIIIA - RNA recognition. Lysine at position -1 was selected in all the clones sequenced, and conserved in the wild-type protein, suggesting an important RNA contact or zinc finger structural role for this residue. Over half of the clones selected serine (versus asparagine in the wild-type) at +2 position. Serine was found at the +2 position in a number of other zinc-finger proteins (Pavletich & Pabo, 1991; Pavletich & Pabo, 1993; Kim & Berg, 1996). Only five other clones revealed identity with the wild-type: glutamine at position +3, and valine at position +6. Subsequent binding assays showed that Lys^{-1} —> Ala^{-1} substitution reduced zf 4-7 binding affinity 37-fold, while alanine substitutions at other positions resulted in a 7-fold lower affinity. Thus, some aspects of zinc finger - RNA interaction appear to be similar to DNA recognition, with some of the α-helical positions used for recognition of both nucleic acids.

Because different subsets of TFIIIA zinc fingers interact with different nucleic acids, a view emerged that TFIIIA is a fusion of a specific DNA-binding and an RNA-binding modules. However, this is an oversimplified view which does not acknowledge the contributions to the binding made by the remaining zinc fingers in each complex (Theunissen et al., 1992; Darby & Joho, 1992; Del Rio et al., 1993; Setzer et al., 1996). In addition, information obtained from studies using isolated sets of zinc fingers or zinc finger

mutants, while providing useful insights, may not necessarily reflect the binding mechanism of the full-length protein. Setzer et al. (1996) quantified thermodynamic contributions of different "broken finger" mutants of TFIIIA as well as N- and C-terminal TFIIIA fragments to RNA binding, and revealed that, similar to DNA binding, there exists a thermodynamic interference between zinc fingers at opposite ends of the protein. The sum of the free-energy change ($\Delta G°$) observed for the N- and C-terminal fragments is greater than the $\Delta G°$ of the full-length TFIIIA. If the two fragments were binding independently, the sum of $\Delta G°$ for the two fragments should be less than or equal to the $\Delta G°$ of the full-length protein.

Only two proteins, TFIIIA and Wilms' tumor protein have been demonstrated to bind both DNA and RNA. Another *Xenopus* protein, p43, (Joho et al., 1990) is able to bind RNA. However, p43 does not appear to bind DNA. A possibility exists that, in addition to these proteins, the function of other C_2H_2 type zinc finger proteins may also be through specific interactions with RNA. This possibility needs to be further investigated. Furthermore, it was recently reported that some zinc finger proteins may function through sequence-specific recognition of DNA-RNA hybrids (Shi and Berg, 1995). The basal transcription factor Sp1 and a designed finger protein ZF-QQR were shown to recognize DNA-RNA hybrids in a sequence-dependent manner. Strikingly, the binding affinities of these proteins for the DNA-RNA hybrids were equal to or even higher than those for the consensus DNA binding sites. This DNA-RNA binding activity may be a general property common to many other members of the C_2H_2 protein class. Although at present there are no known physiologically relevant examples of such interactions, DNA-RNA heteroduplexes are formed in such biological processes as gene transcription, DNA replication, and the reverse transcription of retroviruses. Therefore,

the potential of zinc finger proteins to specifically interact with DNA-RNA heteroduplexes, and, perhaps, macromolecules other than nucleic acids, exists, and requires further investigation.

6.2.7 Role of TFIIIA zinc fingers in DNA recognition

TFIIIA interacts specifically with the intragenic control region (ICR) of the 5S rRNA gene (Engelke et al., 1980) which overlaps the promoter region required for accurate transcription initiation (Sakonju et al., 1980; Bogenhagen et al., 1980). The stoichiometry of the interaction is 1 : 1 measured in titration experiments (Smith et al., 1984; Bieker & Roeder, 1984; Romaniuk, 1990). From DNA mutagenesis and *in vitro* transcription studies, the sequence comprising the ICR was defined and includes a region from base pairs +45 to +97 (Pieler et al., 1985a; McConkey & Bogenhagen, 1987). The functional domains of the *Xenopus* 5S rRNA gene promoter consist of three elements: box A (bp +50 to +64), box C (bp +80 to +97), and an intermediate spacer element (IE) (bp +67 to +72) (Pieler et al., 1985b; Pieler et al., 1987; You et al., 1991; Hayes & Tullius, 1992) (Figure 6.11). The box A component is common to all pol III genes that are regulated via intragenic promoter elements (Pieler & Theunissen, 1993). Box C and the intermediate element, on the other hand, are specific to 5S rRNA genes (Sakonju & Brown, 1982; Smith et al., 1984; Pieler et al., 1987). The 5S DNA ICR appears to assume a $B_{\text{enlarged groove}}$-DNA conformation (Fairall et al., 1989) which is found in a number of DNA binding sites of other zinc finger proteins (Fairall et al., 1993; Nekludova & Pabo, 1994). Studies using metal probes which recognize DNA on the basis of its shape (Huber et al., 1991) indicate that the structure of the ICR is not uniform throughout the sequence, but rather contains distinct segments of predominantly A-like, or B-like conformation, reflecting the

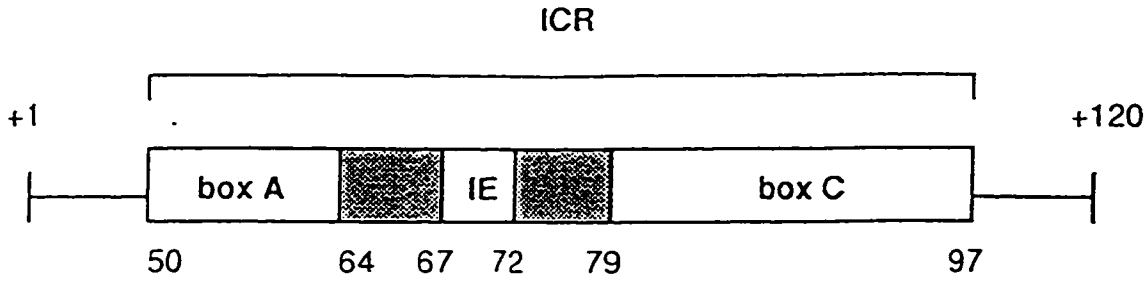


Figure 6.11 A schematic representation of the *Xenopus laevis* 5S rRNA gene internal control region.

periodicity with which TFIIIA interacts with the ICR. In addition, the binding of TFIIIA induces bending of the 5S gene promoter (Fairall et al., 1989; Schroth et al., 1989; Bazett-Jones & Brown, 1989).

Extensive studies of the interaction of TFIIIA with the 5S promoter DNA sequence indicate that three elements of the ICR provide nonequivalent contribution to the overall binding affinity. Chemical modification experiments revealed that methylation of eight guanine residues clustered at the 3' end of the ICR (box C) abolished TFIIIA binding (Sakonju & Brown, 1982). The same dramatic drop in TFIIIA binding affinity was observed upon ethylation of the eight phosphates in the same region of the TFIIIA binding site (Sakonju & Brown, 1982). The importance of the box C element for TFIIIA interaction was demonstrated in mutagenesis experiments (Sakonju et al., 1981; Pieler et al., 1985a; Pieler et al., 1987; You et al., 1991) and in methylation protection studies (Fairall et al., 1986). The latter showed a gradual increase in TFIIIA protection of the DNA binding site in the 5' to 3' direction (toward the C box) (Fairall et al., 1986). In contrast, box A and the intermediate element provide modest contributions to the overall strength of TFIIIA binding (Sakonju et al., 1981; Pieler et al., 1985a; McConkey & Bogenhagen, 1987; You et al., 1991). Studies using substitution mutagenesis within the box C and IE (Veldhoen et al., 1994) have further refined the regions as well as individual base pairs that contribute to the high affinity binding. Thus, base pairs 78-79 and 92-96 are not essential for the high-affinity interaction. Substitution at position 70 has a large effect, while substitution at position 71 has a moderate effect, and positions 67 to 69 have no effect. Substitutions at positions 81, 85, 89 and 91 have a dramatic effect, whereas positions 80, 82, 83, 86-88, and 90 have varying degrees of influence.

Therefore, almost every base pair within the box C region plays an important role in binding of TFIIIA.

A quantitative measurement of the TFIIIA-5S DNA equilibrium has been carried out by Romaniuk (1990). TFIIIA binds 5S DNA with an apparent association constant of 1.90×10^{9} M^{-01} in 0.1 M salt, pH 7.5, at 22°C (Romaniuk, 1990). Analysis of the thermodynamic parameters obtained from measuring temperature dependence of the binding of TFIIIA to 5S DNA indicates that TFIIIA - 5S DNA interaction is largely enthalpy driven at temperatures above 19°C, and is entropy driven at lower temperatures (Romaniuk, 1990). The DNA binding activity of TFIIIA has a broad pH optimum between 6 and 8 units of pH. The optimal divalent ion concentration ($MgCl_2$) is 5mM. The binding is insensitive to the identity of the monovalent cation in the binding buffer. The analysis of the monovalent salt dependence of the binding allowed the determination of the number of ion pairs that form at the TFIIIA - 5S DNA interface. There are as many as eight ion pairs formed (Romaniuk, 1990), which is consistent with the previous chemical modification experiments which determined that eight phosphates are essential for the binding of TFIIIA (Sakonju & Brown, 1982). The binding shows dramatic dependence on the identity of the anionic species, decreasing 100-fold following the lyotropic series (Romaniuk, 1990). This indicates the presence of several anion-binding sites on TFIIIA.

Numerous studies using mutant TFIIIA proteins provide complementary information about TFIIIA - 5S RNA gene complex formation. The results of footprinting experiments using truncated TFIIIA demonstrate that the amino-terminal zinc fingers interact primarily with the C box of the ICR, while the carboxy-terminal fingers interact with the A box region (Vrana et al., 1988). DNase I and hydroxyl radical footprinting studies suggest

that TFIIIA makes contacts primarily along one side of the double helix (Engelke et al., 1980; Churchill et al., 1990; Vrana et al., 1988). The arrangement of the individual zinc fingers over the 5S RNA gene promoter has been dissected using a series of recombinant polypeptides. The studies revealed that the contributions from the zinc finger domains are not functionally equivalent. The three amino-terminal zinc fingers (zf 1-3) are necessary and sufficient for high-affinity DNA binding, providing about 95% of the binding free energy (Christensen et al., 1991; Liao et l., 1992; Clemens et al., 1992; Choo & Klug, 1993; Hansen et al., 1993; Zang et al., 1995). TFIIIA variants with altered fingers 2 or 3 have significantly reduced DNA binding affinities (Theunissen et al., 1992; Darby & Joho, 1992; Del Rio et al., 1993; Zang et al., 1995). Finger swap and amino acid substitution studies conducted in Dr. Romaniuk's laboratory had identified that TFIIIA zinc fingers 2 and 3 provide the majority of free energy of TFIIIA - DNA complex, whereas finger 1 provides energetically less important contacts (Zang et al., 1995). Furthermore, a number of amino acids in the α-helical regions of the zinc fingers were identified as the most important residues for establishing high affinity interaction. These residues are predicted to occur at positions within the α-helix involved in DNA base recognition, by analogy with the EGR-1 - DNA complex (Pavletich & Pabo, 1991). It was recently shown that alanine substitutions at four critical α-helical positions -1, +2, +3, and +6 in any one of the first three fingers of zf 1-3 of TFIIIA eliminated DNA binding (Friesen & Darby, 1997). These four α-helical positions were chosen by analogy with the contact amino acids identified in the available crystal structures of C_2H_2 zinc finger proteins. The results of these studies indicate that amino acids important for TFIIIA DNA-binding affinity may be at the same positions of the recognition α-helix as those of EGR-1 and other closely related proteins.

Interestingly, whereas alanine substitutions in fingers 1 and 2 abolished DNA binding, the same substitutions in finger 3 produced less dramatic reduction in DNA binding. This result is different from the "broken finger" mutants of TFIIIA, which displayed a relatively small change in affinity for DNA: 2.5- and 7-fold for fingers 1 and 2, respectively (Del Rio et al., 1993). It also differs from the data obtained in our lab (Zang et al., 1995). The discrepancy may be attributable to the fact that the "broken finger" mutants were created in the context of the full-length TFIIIA protein (Del Rio et al., 1993), whereas the alanine substitutions were made in the three-finger amino-terminal fragment (Friesen & Darby, 1997), where other zinc fingers are not present to compensate for the loss of DNA contacts. In addition, different sources of the proteins, as well as the methods employed could all have contributed to these differring sets of data.

The peptide containing fingers 1-3 of TFIIIA binds to the box C promoter element (Liao et al., 1992), covering the region of approximately nucleotide positions 79 to 92 (Hayes & Tullius, 1992). Judging from the pattern of the hydroxyl radical and DNase I footprinting experiments with full length TFIIIA, or truncation mutants, a change in the orientation of TFIIIA relative to the DNA occurs at positions where fingers 4 and 6 are shown to bind (Clemens et al., 1992; Hayes & Clemens, 1992; Fairall & Rodes, 1992). Therefore, fingers 4 and 6 are proposed to cross the minor groove. Finger 5 binds in the major groove of the intermediate element, and the C-terminal triplet - fingers 7-9 - interacts with the box A sequence (Hayes & Tullius, 1992; Clemens et al., 1994). The model in which fingers 1-3 and 7-9 wrap around the ends of the TFIIIA binding site, while fingers 4-6 span the center, crossing the minor groove twice was originally proposed by Hayes and Tullius based on the results of the missing-nucleoside and hydroxyl radical

footprinting experiments (Hayes and Tullius, 1992) (Figure 6.12A). A very similar model was proposed by Clemens and coworkers (1994). Interestingly, it was found that polypeptide containing zinc fingers 1 to 5 binds the DNA with higher affinity than the wild-type TFIIIA, but the addition of finger 6 weakens the binding to DNA. These effects may be due to an altered conformation that the peptides adopt, resulting in the change in the strength of DNA-binding activity (Clemens et al., 1994). Fingers 7, 8 and 9 were shown to be dispensable for DNA binding. It should be noted, however, that fingers 8 and 9 are absolutely required for transcriptional activity (Rollins et al., 1993; Del Rio & Setzer, 1993). Thus, it appears that different subsets of TFIIIA zinc fingers provide energetically different contributions to the DNA interaction, contact DNA regions of different size, and bind in both major and minor grooves of the DNA.

The previous models for the TFIIIA - 5S DNA interaction were based on the assumption that the nine fingers of TFIIIA would interact with DNA in essentially the same way as does EGR-1 protein (Fairall et al., 1986; Churchill et al., 1990), with all the zinc fingers providing approximately equal contributions to the binding (Figure 6.12C). In fact, only the first three zinc fingers of TFIIIA show the closest similarity with the consensus sequences of other zinc finger proteins (Gibson et al., 1988; Klug & Rhodes,1987; Jacobs, 1992). The overall TFIIIA - DNA interaction appears much more complex and asymmetrical than originally proposed in the early models. A number of recent studies of TFIIIA-DNA interaction cumulatively confirm the view that there is a high degree of non-uniformity and functional interdependency among the zinc fingers (Clemens et al., 1994; Rawlings et al., 1996; Rowland & Segall, 1996; McBryant et al., 1996). Perhaps, proteins with relatively large numbers of zinc fingers should be analyzed differently from the small and

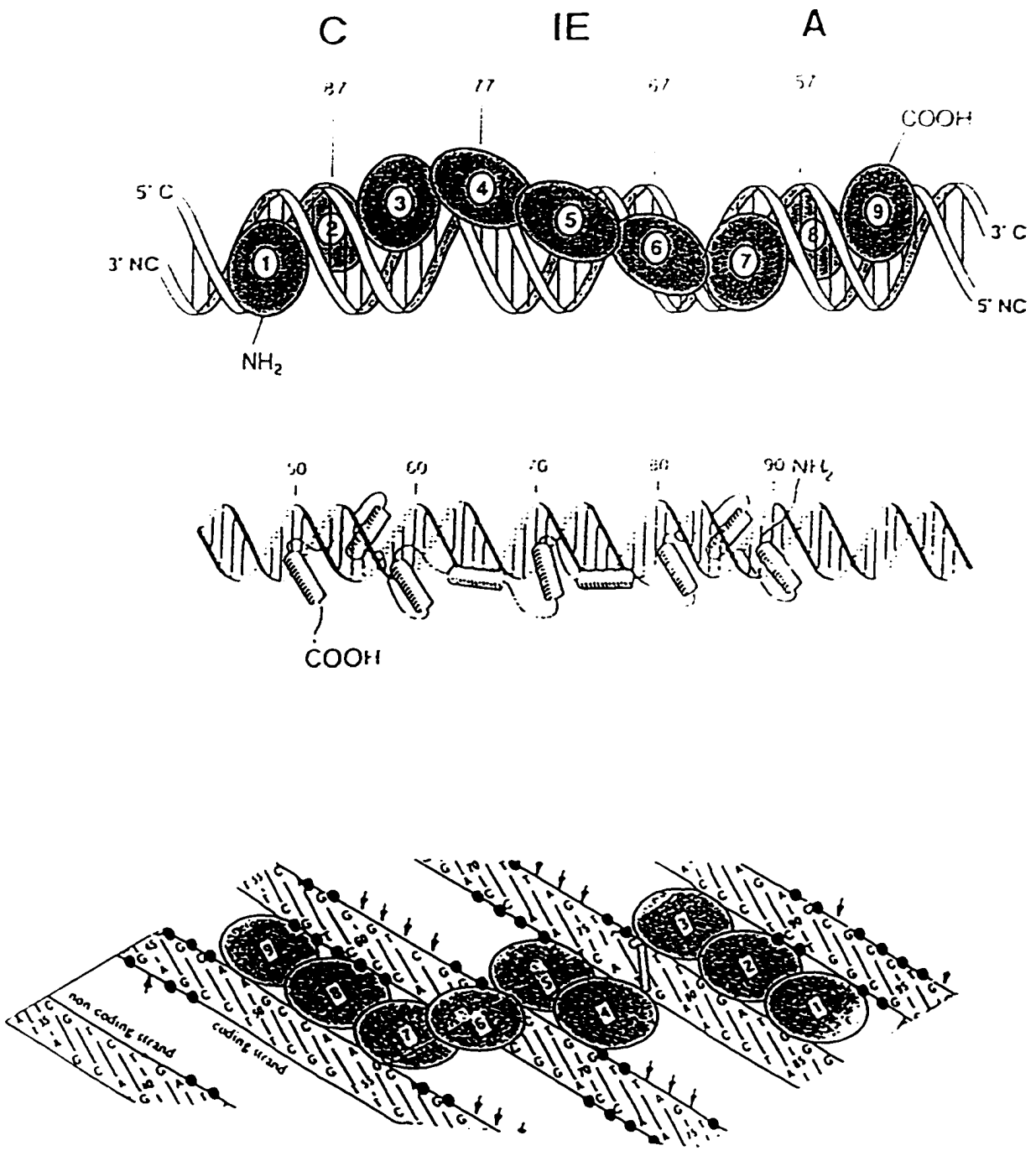


Figure 6.12 Models for TFIIIA - 5S DNA complex. (**A**) After Hayer & Tullius, 1992. (**B**) After Clemens et al., 1992. (**C**) After Fairall & Rhodes, 1992.

compact ones, like EGR-1. Therefore, simple extrapolation of an EGR-1-like mode of binding is not adequate when trying to understand proteins such as TFIIIA.

A series of "broken-finger" TFIIIA mutants were created, in which histidine residues were substituted by asparagines to introduce structural disruptions in each zinc finger (Del Rio et al., 1993). These experiments confirm that zinc fingers 1 and 9 are least important for DNA binding, while the most deleterious effect on DNA binding is produced when finger 3 is mutated (Del Rio et al., 1993). Curiously, while disruption of finger 4 did not result in any alterations of the footprint, it displayed the second largest reduction of the binding affinity (Del Rio et al., 1993). These results suggest that finger 4 plays an important role of correctly linking the first zinc finger triplet with the middle triplet. The finger swap experiments conducted in our lab (Zang et al., 1995) also identified fingers 2, 3, and 4 as important for TFIIIA-DNA interaction, and agree very well with the findings from the "broken finger" experiments (Del Rio et al., 1993).

The short sequences connecting zinc fingers, called linkers, are highly conserved in evolution (Choo & Klug, 1993). Therefore, it was proposed that they play an important functional role which until recently was not well understood. A linker of the sequence TGEKP is conserved between many zinc finger proteins, including EGR-1, but only in the linkers of the three amino-terminal fingers of TFIIIA (Gibson et al., 1988; Jacobs et al., 1992; Choo & Klug, 1993). The role of the interfinger linker sequences in the amino-terminal three zinc fingers has been assessed in a number of studies involving single and multiple amino acid substitutions (Smith et al., 1991; Choo & Klug, 1993; Clemens et al., 1994). These substitutions result in reduction of DNA-binding between 7- and 24-fold (Choo & Klug, 1993).

Interestingly, replacement of the first linker of TFIIIA with the linker connecting zinc fingers 3 and 4 abolished DNA binding (Choo & Klug, 1993). This suggests that different linkers provide nonequivalent contributions to the finger-finger relationships. Thus, linker sequences are important determinants of the overall structural integrity of TFIIIA and are required for high-affinity DNA binding. It remains to be established whether the function of the linker sequences is to correctly position the zinc fingers relative to the other domains of a protein, provide flexibility/rigidity, or make DNA contacts (the latter is less likely), or any combinations of the above.

Recently, a solution structure of the first three zinc fingers of TFIIIA complexed to its cognate DNA (base pairs 79 to 93 in the ICR numbering scheme) has been reported (Wuttke et al., 1997). Nuclear magnetic resonance spectroscopy has been employed to solve the high resolution structure. Figure 6.13 shows the amino acid sequence of zf 1-3 and the sequence of the C block of the 5S DNA element. The structure reveals that the three zinc finger peptide wraps around the DNA major groove, much like the zinc fingers of EGR-1. However, there are several new features: each zinc finger contacts 4 to 5 base pairs of the DNA, novel amino acids (such as tryptophan at position +2 of the finger 1 α-helix and arginine at position +10 of finger 3) are involved in making direct base contacts, while the entire α-helix of finger 3 is used for making sequence-specific contacts.

In summary, the protein-DNA recognition by the zf 1-3 fragment of TFIIIA is mediated by electrostatic and hydrogen bonding interactions, as well as structural complementarity at the protein-DNA interface. All contacts occur in the major groove of the DNA on both coding and noncoding DNA stands (with the majority of the contacts made to the non-coding strand). The contacts to the sugar-phosphate backbone, while not providing specificity,

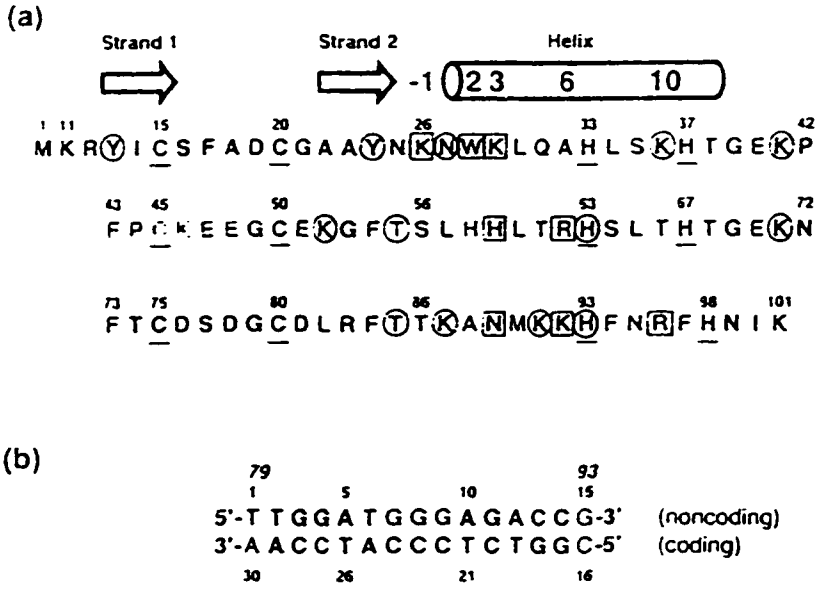


Figure 6.13 (a) Amino acid sequence of zf 1-3. The construct consists of residues 11 to 101 of TFIIIA. Conserved zinc ligands (Cys and His) are underlined; the β-strands are indicated by arrows, and the α-helix by a cylinder. Boxes indicate residues involved in making hydrophilic contacts to DNA bases, and circles denote those involved in contacting the DNA phosphate backbone. (b) DNA sequence of the C-block element of the 5S RNA gene. Both structural (nucleotides 1-30) and biochemical (nucleotides 79-93) numbering schemes are shown.

help anchor the zinc fingers to the DNA as well as properly orient them in the major groove.

Individual fingers appear to provide nonequivalent contributions to the overall binding. Thus, finger 3 contributes more than any other finger to the TFIIIA-DNA interaction. Indeed, unlike any other zinc finger-DNA structures characterized todate, the entire length of finger 3 of TFIIIA (i.e. positions -1 to +10) are involved in sequence-specific base contacts. Extensive mutagenesis data of finger 3 and corresponding DNA site also point to the critical role of finger 3 in the high affinity binding by TFIIIA (Zang et al., 1995; Veldhoen et al., 1994; Pieler et al., 1985). An unusual feature of finger 1 interaction is the extensive network of hydrophobic interactions formed by Trp28 (position+2 in the α-helix). It contacts four bases on both strands of the DNA, which has not been previously documented for any other zinc finger proteins.

The role of the interfinger linker structures are of particular interest as NMR structure determination provides a dynamic view of the interaction. It was shown previously that point mutations in the linker sequences can be deleterious to DNA binding (Choo & Klug, 1993; Clemens et al., 1994). The present NMR structure shows that linker sequences adopt a highly ordered structure upon the DNA complex formation thus actively contributing to the overall stability of the complex. Similarly, protein-protein interactions between zinc fingers indirectly contribute to the high affinity DNA binding. Another striking feature of the present NMR structure is that a surprisingly large number of DNA-contacting side-chains exhibit conformational flexibility within the protein-DNA interface without loss of specificity.

In conclusion, the study gives a molecular level prospective for the accumulated mutagenesis and footprinting data available for TFIIIA.

CHAPTER 7.0 THE STUDY OF THE NUCLEIC ACID INTERACTIONS OF ZINC FINGER 4-7 REGION OF TFIIIA

7.1 Introduction

In a recently published study, Zang et al. (1995) observed that the substitution of the TFIIIA zinc fingers 4 to 7 by the equivalent zinc fingers of an RNA-binding protein p43 produced a 100-fold reduction in the DNA binding affinity of the resultant protein construct. This effect can be explained by either assuming that a number of energetically important contacts are lost, or that the improper alignment of the mutant protein over the DNA precludes high-affinity interaction. Replacement of the finger 7 alone, on the other hand, did not result in a significant change in the binding affinity (Zang et al., 1995). Therefore, the remaining fingers 4, 5, and 6 must possess important determinants of the DNA-binding specificity.

We have approached the problem relating to the functional role of the zinc finger 4 to 6 region in high-affinity nucleic acid binding by quantitatively analyzing the properties of wild-type TFIIIA and a set of the mutant forms of the protein containing substitutions of an entire zinc finger, or either halves of a given finger. The donor proteins used were p43 and Wilms' tumor proteins. Since these donor proteins belong to the same zinc finger family as does TFIIIA, the stereochemical fold of the resultant mutant proteins is expected to be maintained. Thus, the differences in the measured protein-nucleic acid binding affinities are presumed to be attributable to the lost or gained contacts. p43 protein was chosen as a donor since it does not have a DNA-binding activity, and, therefore, would not interfere with the analysis of the DNA-binding properties of zinc fingers under investigation. Similarly, the Wilms' tumour protein does not specifically interact with either the 5S rRNA nor its gene, and is thus also well suited for the purposes of the study.

The proteins were purified to homogeneity, and assayed for their DNA and RNA binding abilities using a nitrocellulose filter binding assay. This is a collaborative study with Dr. Paul J. Romaniuk, who synthesized the oligonucleotides for the PCR-mediated site-directed mutagenesis, and Kathleen Barilla, who constructed and purified the Tp series of proteins , as well as made constructs for the expression of the WTF and TFW series of mutants (see nomenclature in the results section of this chapter). I constructed and purified the TF(4-7)WT mutant, carried out the expression and purification of the WTF and TFW sets of the mutant proteins, as well as carried out all the DNA and RNA binding experiments.

7.2 Materials and Methods

7.2.1 Construction of TFIIIA finger swap mutants

The TFIIIA mutants were produced by site directed mutagenesis, which was either PCR-mediated, or carried out with the use of a transformer kit (Clontech Laboratories, Inc.). Plasmid pT7-TF containing the TFIIIA cDNA (Ginsberg et al., 1984) was a kind gift from Dr. J Tso. The *Nde*I/*Bam*HI fragment of pT7-TF was cloned into the same restriction sites of the expression plasmid pET-11b (Studier et al., 1990) to yield pTF4 (Veldhoen, Ph.D. thesis). To facilitate mutagenesis, zinc finger 4 to 7 region of TFIIIA was amplified from pTF4 using primers which incorporate restriction sites *Bgl* II and *Xho*I for insertion back into TFIIIA cDNA (Table 7.1). The amplified zinc finger 4-7 fragment of TFIIIA was cloned by blunt-end ligation into *Sma*I site of pUC19, and the sequence of the resulting plasmid was verified by sequencing. The resultant plasmid was designated pUC19-zf4-7 (Figure 7.1A). To create a TFIIIA mutant in which zinc fingers 4-7 are replaced with the zinc fingers 1-4 of WT1, specific primers were designed to amplify zinc finger

Table 7.1 Sequences of oligonucleotides used in construction of TFIIIA substitution mutants (see details in the text).

Construct name	Oligonucleotide sequence	Specific Comments
pUC19-zf4-7	CATATGAAGATCTGCGTCTATGTGT ACTAGTATCTCGAGGGCAGAGATA CA	Upstream; *Bgl II* site Downstream; *Xho I* site
pUC19-WT1-BX	AGATCTGCGTCTATGTGTGTGCT TACCCAGGCTGCAAT	Upstream; *Bgl II* site Downstream; *Xho I*
Tp4A	CATATGAAGATCTGCCCCTTAAAG TGCTTTGTTCCGGGCTGTAAGAGA AGCTTCAAGAAACACAAT	Replaces first 15 a.a. of TFIIIA finger 4 with first 15 a.a. of p43 finger 4
Tp4B	AACTGTGGCAAAGCTTTCAGGAAA AAGCGTGCATTAAGGCGTCATCTG TCCGTTCACAGTAATCAGCTGCCA TACGAA	Replaces last 15 a.a. of TFIIIA finger 4 with last 15 a.a. of p43 finger 4
Tp5A	AGTCACACACAGGAGCCGCTAT CCGTATGTGATGTTCCAGGCTG TTCCTGGAAGTCTTCTTTGCCT	Replaces first 15 a.a. of TFIIIA finger 5with first 15 a.a. of p43 finger 5
Tp5B	AAGCGGTTTTCTTCGGTTGCCAAG CTTGTAGCTCATCAAAAACGCCAT CGAGGCTATCCCTGC	Replaces last 15 a.a of TFIIIA finger 5with last 15 a.a. of p43 finger 5
Tp5C	TGGAAGTCTTCTTCGGTTGCCAAG CTTGTAGCTCATCAAAAACGCCAT CGAGGCTATCCCTGC	Swaps TFIIIA finger 5 with p43 finger 5

Construct name	Oligonucleotide sequence	Specific Comments
Tp6A	GTCCATGCAGGCTATCGCTGCAG TTATGAAGGTTGCCAAACTGTGA GTCCGACTTGGACATTATAC	Replaces first 15 a.a. of TFIIIA finger 6with first 15 a.a. of p43 f.6
Tp6B	GGAAAGACTTGGACAGCACTGCA GACACACGTGAAAAAACATCCGC TCCTAGCAGTATGTGAT	Replaces last 15 a.a. of TFIIIA finger 6with last 15 a.a. of p43
WTF4	AAGAGATATAAAAAGCATAACCAG CTGAAGGTGCACCAGTTCAGCCAC ACTCAAGAGAAACCATAC	Replaces last 15 a.a. of WT1 finger 1 with last 15 a.a. of finger 4 of TFIIIA
WTF5	AGGTTTTCTCTTCCATCCCGGCTTA AGAGACACGAAAAGGTACATGCAG GTGTGAAACCA	Replaces last 15 a.a. of WT1 finger 2 with last 15 a.a. of finger 5 of TFIIIIA
WTF6	CGAAAGTTCAAGACGTGGACCCTC TATCTTAAGCACGTCGCCGAGTGT CATCAAGATGAAAAGCCCTTC	Replaces last 15 a.a. of WT1 finger 3 with last 15 a.a. of finger 6 of TFIIIA
TFW4	AAAGCATTCTTTAAACTCAGTCAT TTACAGATGCATTCGCGCAAGCA CACAGGGCAGCTGCCA	Replaces last 15 a.a. of TFIIIA finger 4 with last 15 a.a. of finger 1 of WT1
TFW5	CGGTTTTCTAGATCTGACCAGTTA AAACGTCATCAAAGACGCCATACA GGCTATCCC	Replaces last 15 a.a. of TFIIIA finger 5 with last 15 a.a. of finger 2 of WT1
TFW6	GTGGGAAAGAGATCTGATCATCTC AAGACACACACAAGAACCCATACG GGCCTAGCAGTATGT	Replaces last 15 a.a. of TFIIIA finger 6 with last 15 a.a. of finger 3 of WT1

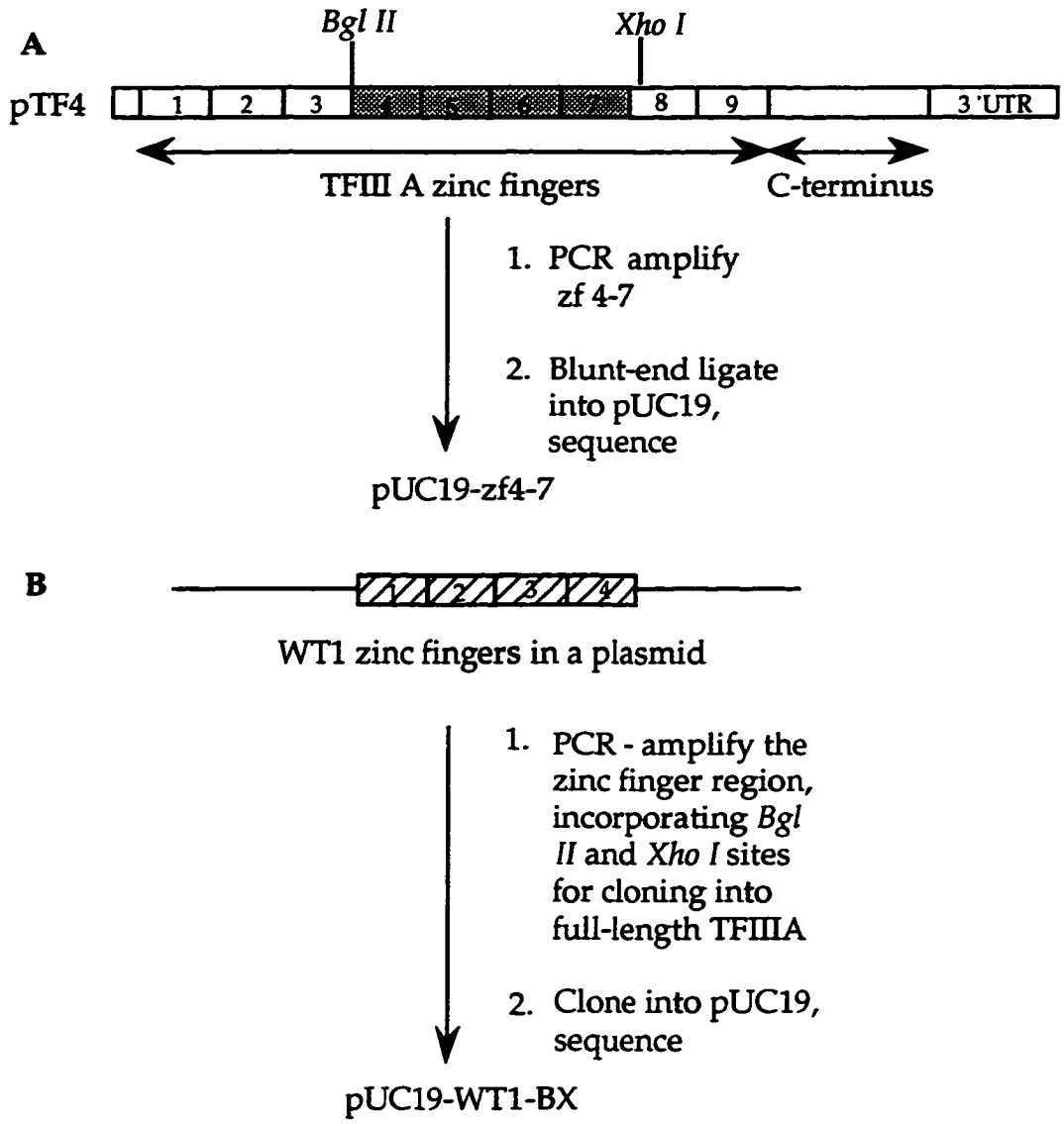


Figure 7.1 Schematic diagramm illustrating construction of (**A**) pUC19-zf4-7; (**B**) pUC19-WT1-BX

region of WT1, and incorporate *Bgl* II and *Xho* I restriction sites for subsequent cloning into full-length TFIIIA cDNA (Table 7.1). The general PCR protocol was used as described in Chapter 3 of this thesis. The PCR product was cloned into pUC19, sequenced, and designated pUC19-WT1-BX (Figure 7.1B).

Mutants Tp5A, Tp5B, TFW4, TFW5, WTF4, WTF5, and WTF6 were generated using PCR-mediated site directed mutagenesis procedure (Nelson & Long, 1989). The nomenclature 'Tp' refers to TFIIIA mutants in which the first or last 15 amino acids within a given zinc finger were replaced with the first or last 15 amino acids of the corresponding finger of p43. The nomenclature 'WTF' refers to constructs in which TFIIIA zinc fingers 4-7 were replaced by fingers 1-4 of the WT1 protein, and, subsequently, α-helical regions of TFIIIA fingers 4-6 were placed back. 'TFW' refers to TFIIIA substitution mutants in which α-helical regions of zinc fingers 4, 5, and 6 are replaced with those of the WT1 zinc fingers 1, 2, and 3.

pUC19-zf4-7 was used as a template to create mutants Tp5A, Tp5B, TFW4 and TFW5. pUC19-WT1-BX was used as a template to create mutants WTF4, WTF5, and WTF6. As the first step, the mutagenic sequence (Table 7.1) and a unique flanking primer site (RUMP) were introduced into the wild type template DNA (Figure 7.2). The 50 µl PCR reaction contained: PCR buffer (10 mM Tris-HCl pH 8.3 at 20°C, 1.5 mM $MgCl_2$, 25 mM KCl), 50 µg/ml gelatin, 2.5 nM each of dTTP, dGTP, dCTP, and dATP, 10 pmol FUMP and mutagenic primer, 1 ng template DNA, and 1.25 units Taq DNA polymerase. The reaction was overlaid with 50 µl mineral oil prior to thermocycling. Twenty four rounds of thermal cycling were carried out using the following conditions: denaturation temperature of 94°C (1.5 minutes), an annealing temperature of 55°C (1.5 minutes), and an extension temperature of 72°C (two

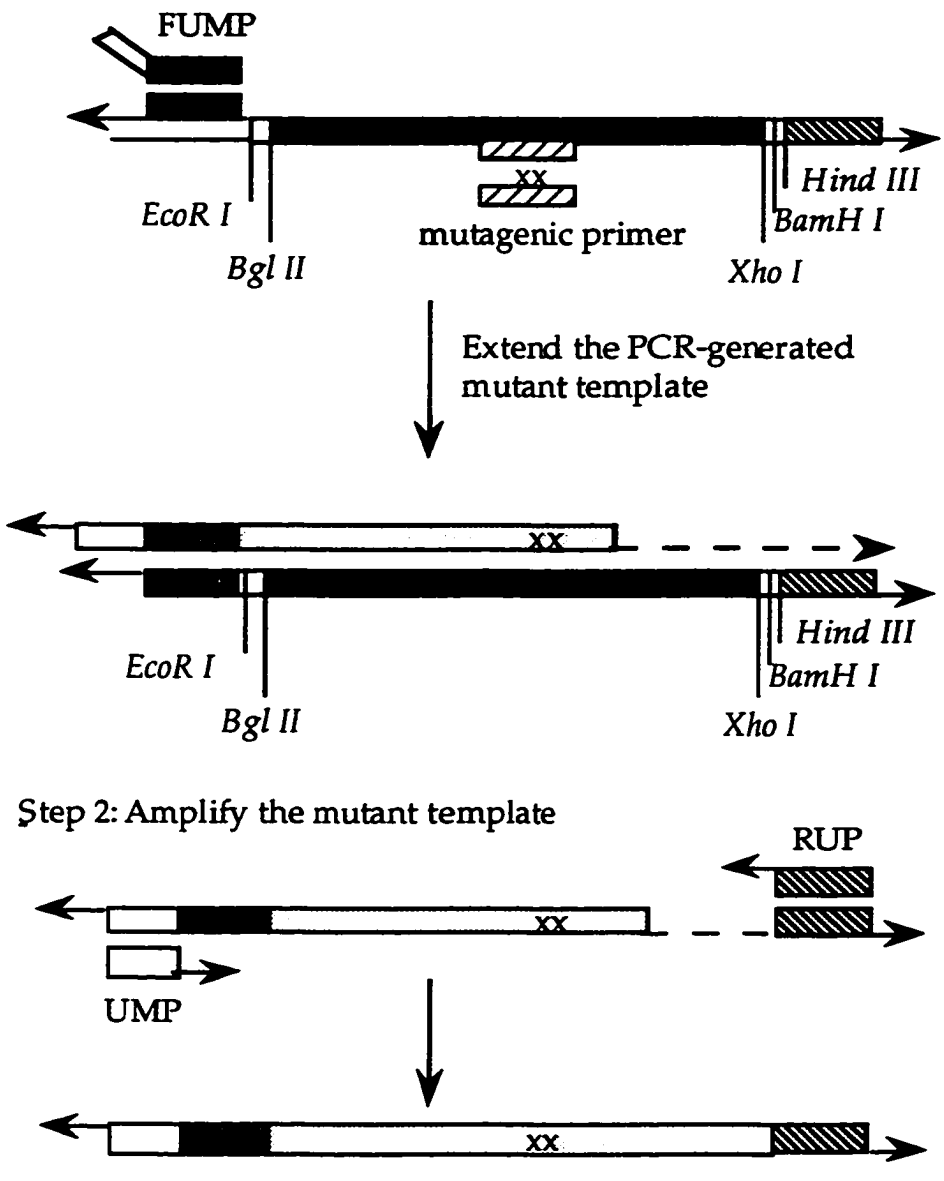


Figure 7.2 Schematic chart of the PCR-mediated site-directed mutagenesis protocol used to create TFIIIA substitution mutants

minutes). The second step reaction (80 µl) contained: 1 ng template DNA, PCR buffer (10 mM Tris-HCl pH 8.3 at 20°C, 1.5 mM $MgCl_2$, 25 mM KCl), 50 µg/ml gelatin, 5 nmol each of dTTP, dGTP, dCTP, dATP, 1-3 µl step 1 PCR, and 2.5 units Taq DNA polymerase. Initially, five rounds of thermocycling were carried out under the following conditions: denaturation temperature of 94°C (3 minutes), an annealing temperature of 55°C (1.5 minutes), and an extension temperature of 72°C (two minutes). 20 µl reaction buffer containing 20 pmol universal mutagenic (UMP) and M13 reverse universal primers (RUP) were added, and 25 additional cycles, identical to step 1, were carried out.

The PCR products were purified by phenol: $CHCl_3$ extraction, and ethanol-precipitated. The Tp series mutant constructs were cloned into *EcoR* I and *BamH* I sites of pUC19, and TFW & WTF constructs were cloned into *Eco*R I and *Hind* III sites of pUC19 for sequence verification. The correct sequences were subsequently subcloned into *Bgl* II and *Xho* I sites of pTF4 for protein expression.

Mutants Tp4A, Tp4B, Tp5C, Tp6A, Tp6B, and TFW6 were created with the use of the transformer site-directed mutagenesis kit (Clontech laboratories Inc., 1994) (Figure 7.3). Sequences of mutagenic oligonucleotides are given in Table 7.1. Sequences of selection primers are shown in Table 7.2. All the procedures were performed as per protocol (Clontech laboratories Inc., 1994). The resultant constructs were subcloned into *Bgl* II and *Xho* I sites of pTF4 plasmid for protein expression.

7.2.2 Expression and purification of TFIIIA proteins

Expression and purification of recombinant TFIIIA proteins was carried out as described by Del Rio & Setzer (1991). 100 ml LB medium supplemented

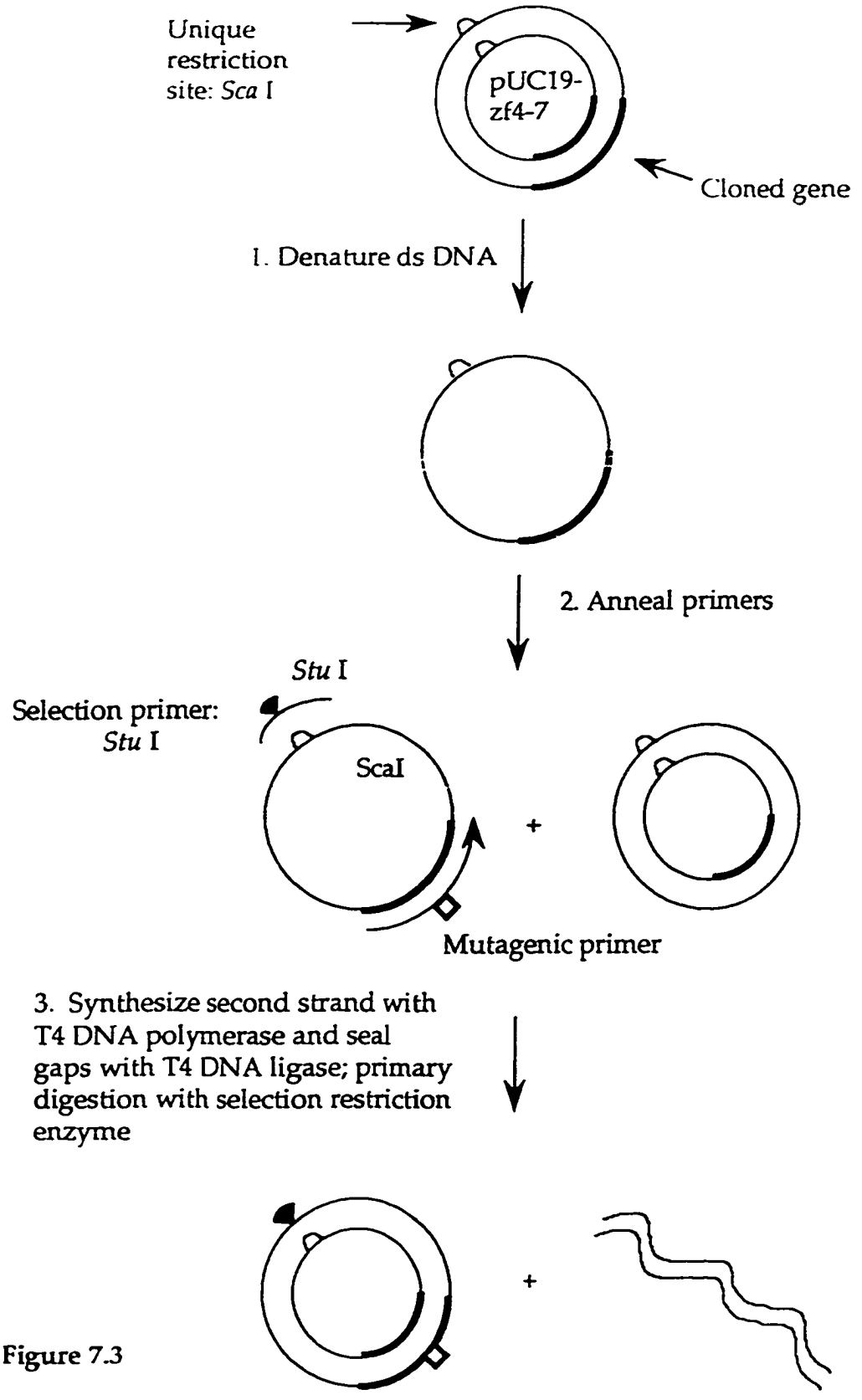

Unique
restriction
site: Sca I
pUC19-
zf4-7
Cloned gene
1. Denature ds DNA
2. Anneal primers
Stu I
Selection primer:
Stu I
ScaI
+
Mutagenic primer
3. Synthesize second strand with
T4 DNA polymerase and seal
gaps with T4 DNA ligase; primary
digestion with selection restriction
enzyme
+

Figure 7.3

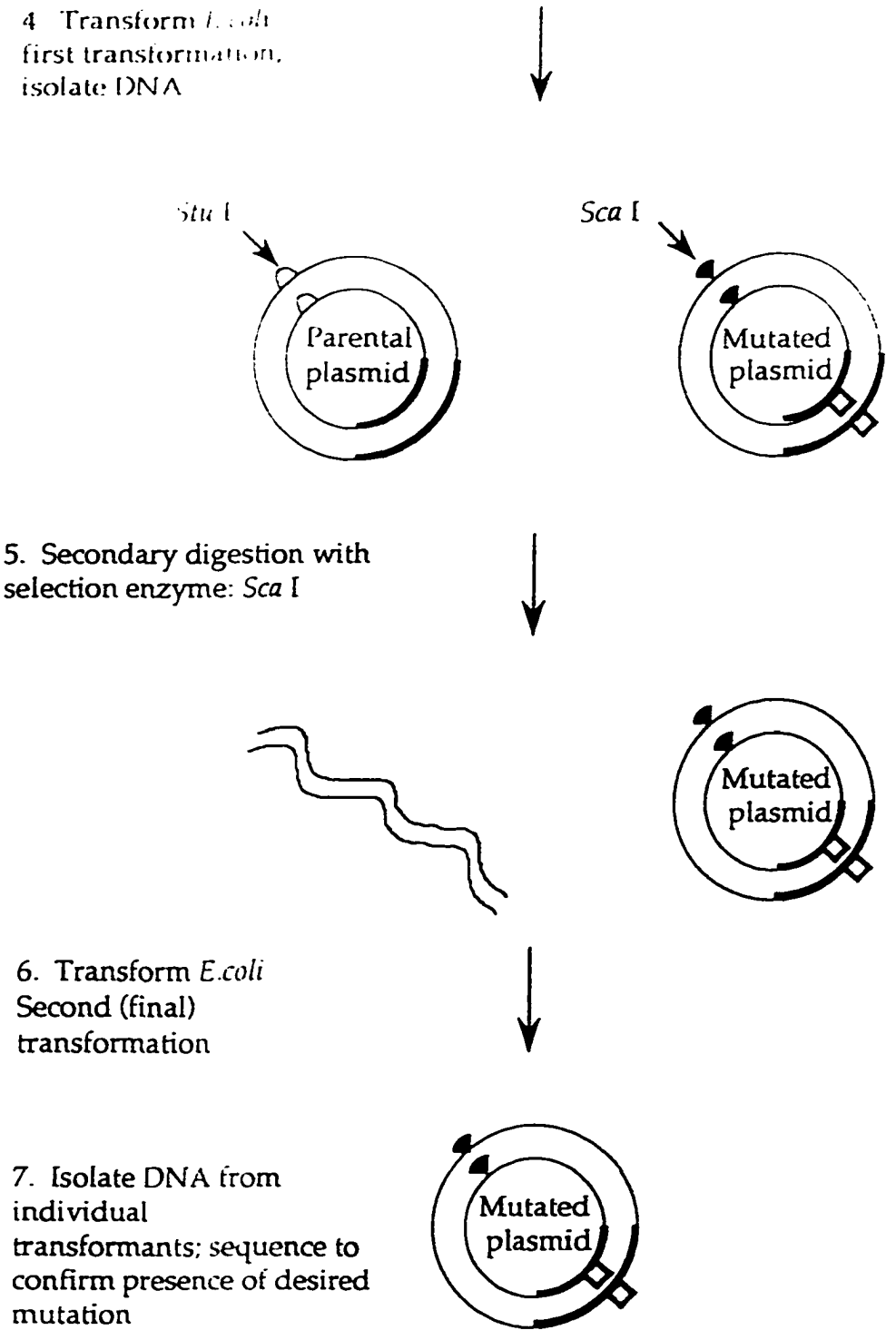


Figure 7.3 Schematic diagram showing the strategy for introducing mutations using the Transformer site-directed mutagenesis kit (Clontech Inc., 1994)

Table 7.2 Sequences of selection primers used in transformer mutagenesis protocol to construct substitution mutants of TFIIIA

Name of primer	Sequence	Specific comments
Trans Oligo Sca I/Stu I	GTGACTGGTG**AGG**	Turns wild type pUC19
	CCTCAACCAAGTC	Sca I site at 2177 to Stu I
Switch Stu I/Sca I	GTGACTGGTG**AGT**	Turns Stu I site in pUC19
	ACTCAACCAAGTC	back to wild type Sca I

with 100 μg/ml ampicillin was inoculated with 1 ml of overnight culture of *E. coli* strain BL21(DE3) harboring expression plasmid of interest (containing either wild type or mutant TFIIIA gene). The culture was grown at 37°C with shaking at 300 rpm to an $O.D._{600}$ of 0.5 to 0.7. Protein synthesis was initiated by addition of $ZnSO_4$ and IPTG to final concentrations of 50 μM and 1mM, respectively. After induction for three to four hours, cells were harvested by centrifugation at 3800 x g for 10 minutes at 4°C, in a Beckman JA-20 rotor, and washed with 10 ml of ice-cold TAB buffer (20 mM HEPES pH 7.4, 5 mM $MgCl_2$, 50 μM $ZnSO_4$, 5 mM DTT, 10% glycerol, 250 mM NaCl). All the following procedures were carried out on ice. The cells were pelleted once more, and resuspended in 4 ml of TABP buffer (TAB buffer + 1 mM PMSF). Cells were sonicated using a microtip sonicator (Heat Systems-Ultrasonics W-385) at setting four, 50% duty cycle for 6 x 20 second intervals with 1 minute cooling between the pulses. The sonicated cells were centrifuged for 20 minutes at 12000 x g in a Beckman JA-20 rotor. The pellet was resuspended in 1 ml of TABUP buffer (TAB + 5M urea + PMSF) and mixed by inversion overnight in a coldroom (4°C).

The solubilized extract was spun in a microcentrifuge for 20 minutes at 14000 x g, and the supernatant removed to a new tube. A series of ammonium sulphate cuts was started by addition of 0.666 ml of TABAS (TAB saturated with $(NH_4)_2SO_4$) to bring the sample to 40% saturation, and mixed by inversion for one hour. The proteins precipitated by the first round of salting-out were pelleted by centrifugation at 12000 x g for 20 minutes, and discarded. More $(NH_4)_2SO_4$ was added to the remaining supernatant to achieve 80% saturation. The sample was mixed by inversion for one hour, pelleted by centrifugation at 12000 x g for 20 minutes, and the pellet resuspended in 10 ml of TABU-250 (TAB + 5 M urea + 250 mM NaCl). This

protein solution was loaded onto BioRex 70 column (800 µl) pre-equilibrated with TABU-250. The column was then washed with 1 ml of loading buffer, followed by TABU-400 (TAB + 5M urea + 400 mM NaCl). The protein was eluted from the column in 800 µl of TABU-600 (TAB + 5M urea + 600 mM NaCl). Purified proteins were aliquoted and stored at -70°C. The purity of each protein preparation was checked on a 10% SDS-PAGE (Figure 7.4 & 7.5). Protein concentrations were determined by the method of Bradford (1976) using BSA as a standard.

7.2.3 Radiolabeling of 5S DNA

The 5S rRNA gene was released from plasmid pXlo (Romaniuk et al., 1987) using *Eco*RI and *Hind*III endonucleases of restriction and end-labeled with [α-P^{32}]-dATP and the Klenow fragment of DNA polymerase I (Sambrook et al., 1989). The details of the labeling protocol are the same as in Chapter 3 of this thesis.

7.2.4 Synthesis and Radiolabeling of 5S rRNA

A method of *in vitro* run-off transcription was used as described by Romaniuk et al. (1987). The pXlo plasmid containing 5S rRNA gene was digested with *Dra*I which cleaves the DNA at position +121 relative to the first nucleotide incorporated into RNA. The digested DNA was extracted with 100 µl of phenol: $CHCl_3$, and DNA precipitated with ethanol. The DNA pellet was dried in Speed Vac concentrator, resuspended in deionized water, and stored at -20°C.

Internally labeled 5S rRNA was made as described by Romaniuk et al. (1987). The labeling reaction had a final volume of 100 µl, and contained: 40 mM Tris-HCl pH 8.0, 15 mM $MgCl_2$, 5 mM DTT, 1mM spermidine, 100 µg/ml

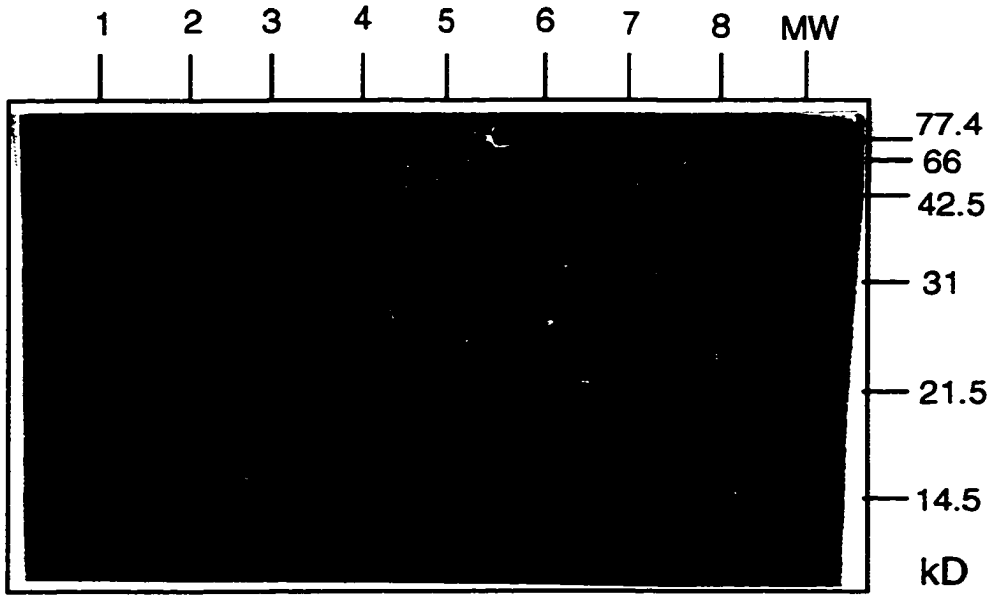


Figure 7.4 Coomassie blue-stained 15% SDS-polyacrylamide gel showing TFIIIA wild type and mutant proteins. Lane designated MW represents protein molecular weight marker; lane 1: wild-type TFIIIA; lanes 2-8: Tp4A, Tp4B, Tp5A, Tp5B, Tp5C, Tp6A, Tp6B.

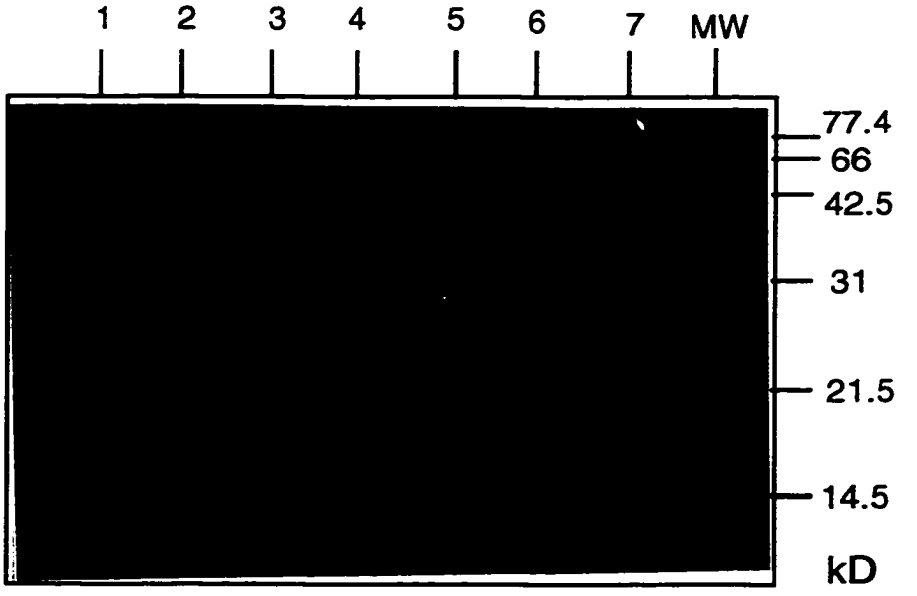


Figure 7.5 Coomassie blue-stained 15% SDS-polyacrylamide gel showing TFIIIA wild type and mutant proteins. Lane designated MW represents protein molecular weight marker; lane 1: wild-type TFIIIA; lanes 2-4: WTF4, WTF5, WTF6; lanes 5-7: TFW4, TFW5, TFW6.

BSA, 1000 U/ml RNasin, 0.5 mM each ATP, CTP, UTP, 0.0125 mM GTP, 50 μCi [α-P^{32}] GTP (600Ci/mmol), 1 μg of template DNA and 0.6 μg of T7 RNA polymerase. The reaction was incubated for 3 hours at 37°C, and purified on an 8M urea, 12% polyacrylamide gel. The gel band containing the labeled RNA was excised, eluted overnight in elution buffer and ethanol precipitated (Sambrook et al., 1989). The level of radiolabel incorporation was determined using an LKB 1214 Rackbeta Scintillation counter and toluene to which 0.4% PPO was added as a scintillant.

7.2.5 Nitrocellulose filter binding assays

The binding affinities were quantified using a nitrocellulose filter binding assay as described on page 88 of this thesis, except the binding reactions were supplemented with 28 μg/ml of nonspecific competitor polyA.

7.3 Results

7.3.1 Effects of zinc finger substitution mutagenesis on the DNA binding activity of TFIIIA

The first part of this study was to determine the contributions of individual zinc fingers 4, 5, and 6 to the high affinity DNA binding of TFIIIA. The employed strategy was to replace one half of a zinc finger at a time with the corresponding sequences of the p43 protein (Figure 7.6) to assess whether the second half (α-helical portion) of the zinc finger sequences, presumed to be the "recognition" helix by analogy with the EGR-1 protein, would provide a greater contribution to the binding than the first half. A nitrocellulose filter binding assay was used to measure the affinity for each mutant protein. The results of the experiments are given in Table 7.3 in a form of relative affinities to the wild-type TFIIIA.

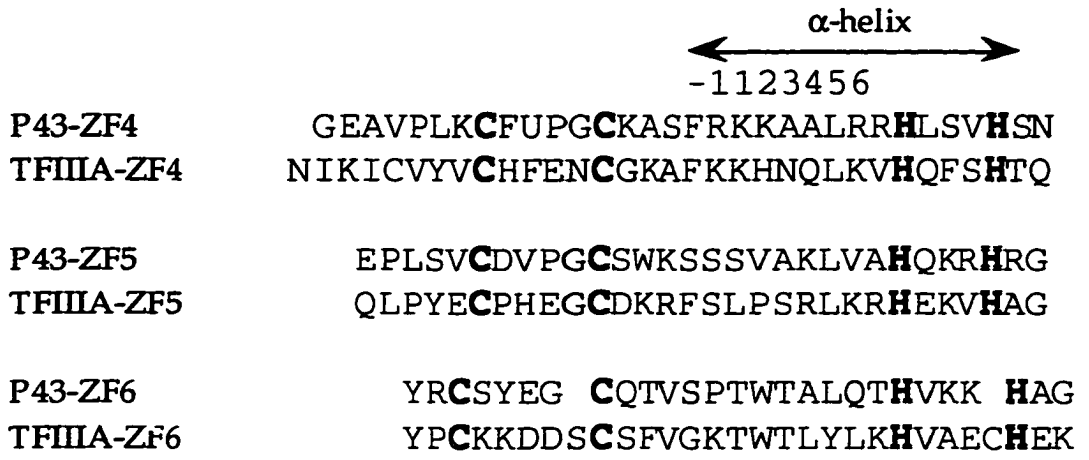


Figure 7.6 Comparison of the amino acid sequences of zinc fingers 4-6 of the donor p43 with the zinc fingers 4-6 of TFIIIA. The structural cysteines and histidines are outlined; the conserved 'TWT' triplet in finger 6 is shown in bold.

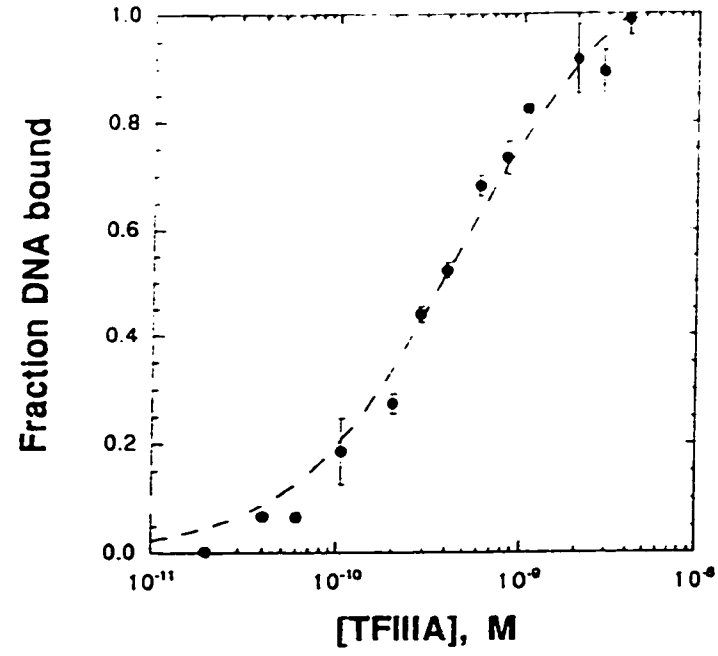
Table 7.3 The effects of TFIIIA zinc finger substitution mutations on the DNA and RNA binding of the factor

Mutant	Relative affinity for 5S rRNA gene	Relative affinity for 5S rRNA
Tp4A	0.72 ± 0.14	0.84 ± 0.09
Tp4B	0.59 ± 0.04	0.14 ± 0.02
Tp5A	0.29 ± 0.05	0.50 ± 0.06
Tp5B	0.23 ± 0.02	0.82 ± 0.11
Tp5C	0.09 ± 0.002	0.83 ± 0.10
Tp6A	1.02 ± 0.09	0.88 ± 0.02
Tp6B	1.56 ± 0.09	0.89 ± 0.02

The nomenclature 'Tp' refers to TFIIIA mutants in which the first or last 15 amino acids within a given zinc finger were replaced with the first or last 15 amino acids of the corresponding finger of p43. Letters A, B, and C refer to the first half, second half, or the whole zinc finger region replaced.
Relative affinities were calculated using the mean values of the association constants determined for the mutant proteins (each value representing the mean of three independent determinations), and compared against the value of the wild-type TFIIIA association constant (K_a=1.9 x 10^9 M^{-1} for DNA; K_a=1.4 x 10^9 M^{-1} for RNA).

The substitutions of either halves of finger 4 resulted in a less than a two-fold reduction in DNA binding, thus implying that likely no specific hydrogen-bond contacts were disrupted. Finger 5 mutants produced more substantial decreases in the binding ability: the entire finger 5 swap mutant (Tp5C) showed a 10-fold drop in binding, while either half resulted in about four-fold lower affinity. This suggests that some specific contacts between finger 5 and 5S DNA were disrupted by the mutations. Finally, finger 6 mutants showed an unchanged (in the case of the Tp6A mutant), and an actually improved binding (if only by a factor of 1.5) in the case of Tp6B mutant. It is interesting to note that no single mutant proved accountable for the dramatic reduction in binding seen when fingers 4-7 of TFIIIA are substituted with those of p43. Rather, it appears that the integrity of each TFIIIA zinc finger in the 4-6 finger region is required for the maximum binding. Of all the mutants tested, finger 5 seems to provide the most significant contribution to the overall binding affinity to DNA, while finger 4 provides the most significant contribution to 5S rRNA binding. The sample binding curves are shown in Figure 7.7.

The next two sets of mutants were created using Wilms' tumour zinc fingers as donors (Figure 7.8). Initially, we substituted the 4 to 7 region of TFIIIA with zinc fingers 1 to 4 of WT1. The resultant construct (TF(4-7WT)) did not have detectable DNA-binding activity, an effect similar to the one observed with the p43 as a donor of fingers 4-7 (Zang et al., 1995). Again, the explanation may be that this effect is due to the loss of specific protein-DNA contacts, or to the altered spatial arrangement of the zinc fingers in the context of the full-length protein. In the first set of mutants (WTF4-6), the α-helical portions of TFIIIA zinc fingers were reintroduced into the TF(4-7)WT one at a time to determine precisely which portions of zinc fingers 4-6 were



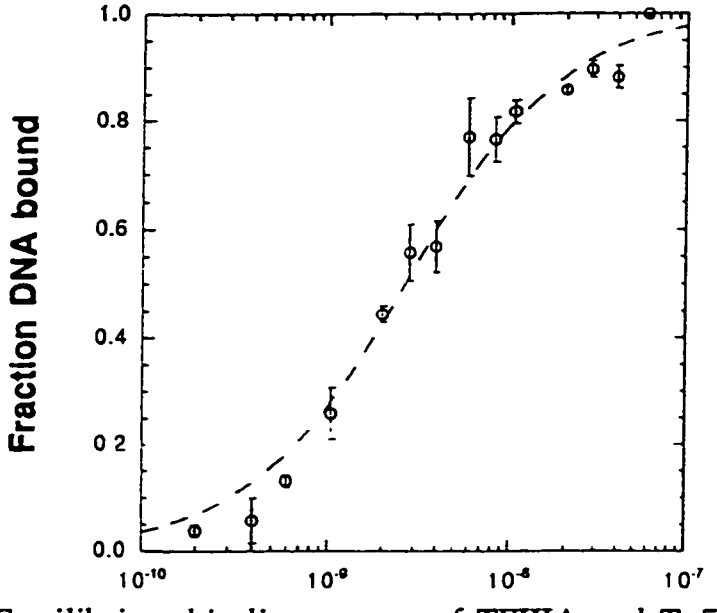


Figure 7.7 Equilibrium binding curves of TFIIIA and Tp5B mutant proteins interacting with 5S DNA. Each data point is the mean of three or more independent determinations, with the associated standard deviations indicated with the error bars.

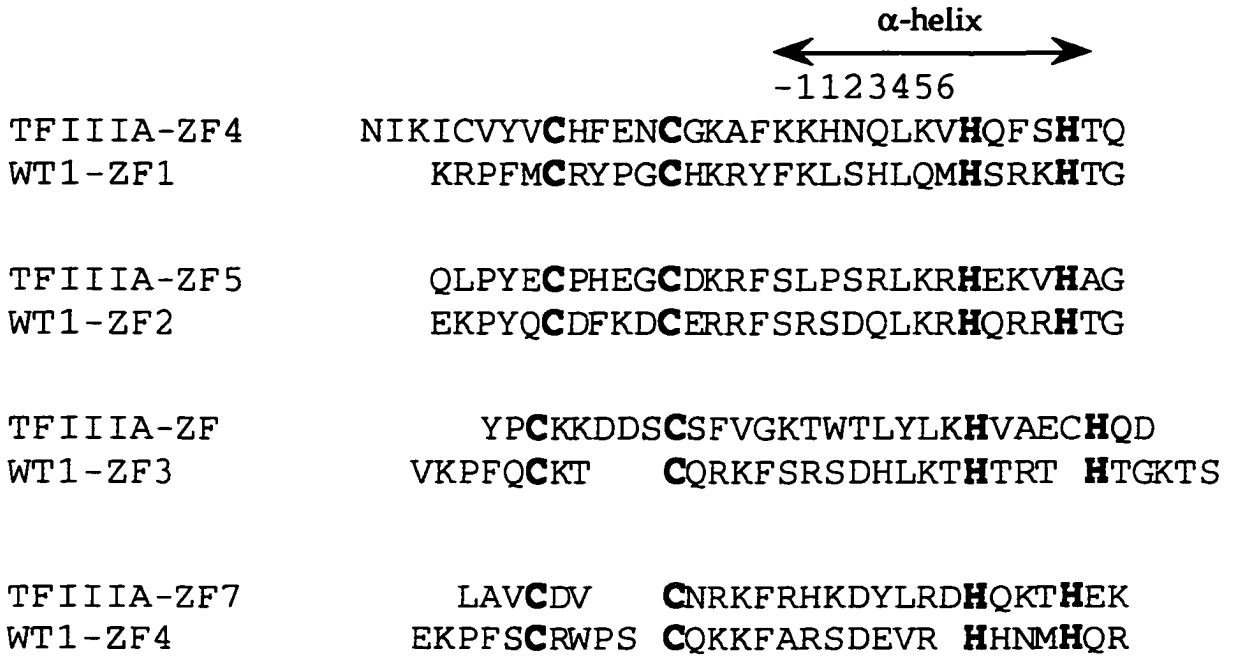


Figure 7.8 Comparison of the amino acid sequences of zinc fingers 1-4 of the donor WT1-ZF with the zinc fingers 4-7 of TFIIIA. The structural cysteines and histidines are shaded.

Table 7.4 The effects of TFIIIA zinc finger substitution mutations on the DNA and RNA binding of the factor

Mutant	Relative affinity for 5S rRNA gene	Relative affinity for 5S rRNA
WTF4	0.06 ± 0.01	<0.01
WTF5	0.04 ± 0.06	<0.01
WTF6	<0.01	<0.01
TFW4	0.34 ± 0.04	0.16 ± 0.02
TFW5	0.08 ± 0.008	0.48 ± 0.07
TFW6	0.55 ± 0.037	<0.01

The nomenclature 'WTF' refers to constructs in which TFIIIA zinc fingers 4-7 were replaced by fingers 1-4 of the WT1 protein, and, subsequently, α-helical regions of TFIIIA fingers 4-6 were placed back. 'TFW' refers to TFIIIA substitution mutants in which α-helical regions of zinc fingers 4, 5, and 6 are replaced with those of the WT1 zinc fingers 1, 2, and 3.

Relative affinities were calculated using the mean values of the association constants determined for the mutant proteins (each value representing the mean of three independent determinations), and compared against the value of the wild-type TFIIIA association constant (K_a=1.9 x 10^9 M^{-1} for DNA; K_a=1.4 x 10^9 M^{-1} for RNA).

necessary to restore the wild-type DNA-binding activity of TFIIIA. The results of filter-binding assays shown in Table 7.4 reveal that no single finger could rescue the binding of TF(4-7)WT to a significant degree. Mutant WTF6 did not produce any improvement in binding, suggesting that it either is the least important finger for DNA binding, or that it couldn't establish any contacts due to the imposed structural constraints. Mutants WTF4 and WTF5 did show somewhat improved binding: their affinities were about 17-fold and 22-fold lower relative to the wild-type TFIIIA, respectively. This small improvement is not surprising since more than one TFIIIA finger is replaced in a given construct.

In the second set of mutants, the α-helical portions of the WT1 zinc fingers were introduced into the fingers 4, 5, and 6 of the full-length TFIIIA. The resultant mutant proteins were designated TFW4, TFW5, and TFW6. Again, the largest reduction in binding affinity to 5S DNA was seen for the finger 5 mutant (about 13-fold), the next largest - for finger 4 (about 3-fold), followed by finger 6 (less that 2-fold). These results are consistent with the effects we observed using other mutants in this study.

7.3.2 Effects of zinc finger substitution mutagenesis on the RNA binding activity of TFIIIA.

In parallel to determining the effects of the zinc finger mutants on DNA binding, we also measured the affinities of these mutant proteins for 5S rRNA. The TF(4-7)WT did not specifically bind to 5S rRNA, although it exhibited a fairly high nonspecific binding for the molecule (Figure 7.9). The results of the RNA binding studies are summarized in Tables 7.3 and 7.4. Most of the mutations presented in Table 7.3 (mutants which incorporate p43 sequences) did not disrupt RNA binding appreciably, displaying almost wild-

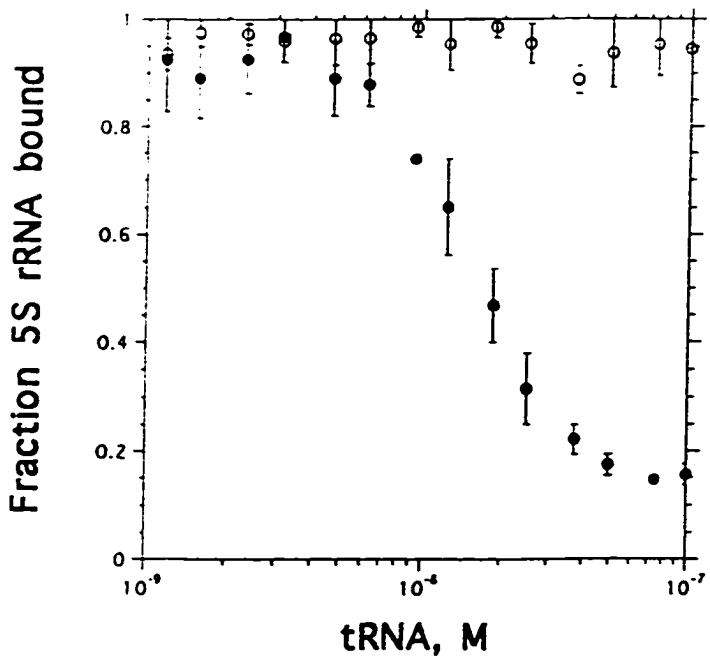


Figure 7.9 Competition assay for 5S rRNA binding using $tRNA^{Phe}$ as a competitor. Open circles represent TFIIIA; closed circles represent T F(4-7)WT proteins. Each data point is the mean of three independent determinations, with the associated standard deviations indicated with the error bars.

type affinities for RNA. The exception is mutant Tp4B, which showed an almost 10-fold drop in binding. The substitution of the first 15 amino acids in finger 5 produced a 2-fold reduction in binding, while swapping its α-helical 15 amino acids resulted in an almost wild-type affinity. The two categories of mutants in Table 7.4 (created using WT1 sequences as a donor) reveal the following: mutants WTF4, WTF5, and WTF6 did not recover any RNA-binding activity of the TF(4-7)WT protein. This is not surprising, since an extensive region of zf 4-7 has been replaced with exogenous (WT1) sequences in those constructs. The second category of mutants (TFW4-6) showed that while substitutions of the α-helical regions in fingers 4 and 5 produced approximately 6- and 2-fold reduction in binding, respectively, replacement of α-helical region in finger 6 abolished RNA binding. The reason why this result is so dramatically different from the one using a similar p43 substitution mutant of finger 6 is that unlike p43, WT1 protein doesn't have amino acid homology to TFIIIA in finger 6 sequence (except for the structural amino acids), Figure 7.6. p43, on the other hand, is equipped with amino acid triplet TWT, located at the tip of finger 6. This triplet amino acid sequence is conserved between TFIIIA and p43 (Joho et al., 1990; Clemens et al., 1993). In fact, these three amino acids in finger 6 are the only ones, apart from the zinc-coordinating cysteines and histidines and the structural aromatic residues that are conserved between the two proteins. The role of these amino acids in sequence-specific RNA recognition by TFIIIA and p43 needs to be investigated further.

7.4 Discussion

Since the demonstration that DNA- and RNA-binding activities of TFIIIA could be separated between the different zinc finger subsets of the

protein, TFIIIA began to be viewed as a fusion between two proteins that specifically evolved to interact with the two different nucleic acids.

However, the results of many recent studies conclusively demonstrate that all the zinc fingers are important for contributing to the overall DNA- and RNA-binding energy. Just as importantly, the zinc fingers do not provide equal contributions to the binding to either DNA or RNA. Although zinc fingers 1 to 3 were shown to dominate in DNA binding, substitution of zinc fingers 4-7 with their counterparts from p43 protein result in a 100-fold reduction in DNA-binding affinity (Zang et al., 1995), while maintaining wild-type RNA-binding affinity. We undertook this study to determine the contribution of each zinc finger in the central zinc finger triplet of TFIIIA to specific DNA and RNA binding. Several sets of TFIIIA mutants were generated, in which specific zinc finger sequences were replaced with those of p43 or Wilms' tumour proteins. Nitrocellulose filter binding assay was used to quantitatively determine affinities of the mutant proteins for both nucleic acids.

Analysis of the DNA-binding data for the Tp series of mutants indicated that the DNA-binding affinity correlated to that of the wild-type TFIIIA in the following order: Tp6B > wt TFIIIA = Tp6A > Tp4A > Tp4B > Tp5A > Tp5B > Tp5C. The largest reduction in DNA affinity was observed for mutant Tp5C (10-fold). Similarly, substitution of the finger 5 α-helix by the corresponding sequence of the WT1 protein (mutant TFW5) resulted in a 13-fold reduction of DNA-binding affinity. These findings are consistent with the current model of TFIIIA - DNA binding, in which finger 5 interacts with the major groove of the intermediate element (Clemens et al., 1992). In fact, it was shown that a zinc finger peptide containing fingers 1-5 bound DNA with higher affinity than full-length TFIIIA (Clemens et al., 1994). Additionally, a

strong footprint was obtained for finger 5 binding site (Clemens et al., 1994). 'Broken finger' experiments also identified a significant contribution of finger 5 to DNA binding (Del Rio et al., 1993). The structural disruption of the finger resulted in almost 7-fold decrease in DNA-binding affinity (Del Rio et al., 1993). The ten- to thirteen-fold reduction in DNA binding that we observed reflects a disruption of up to four hydrogen bonds that finger 5 makes in the intermediate element of the ICR (assuming that an energetic cost of a single hydrogen bond is reflected in a three-fold drop in binding affinity, as well as that energies of individual bonds are additive). Methylation interference experiments previously demonstrated that finger 5 makes two major groove contacts with guanines on the non-coding DNA strand (Clemens et al., 1992), which also agrees well with our findings. Complementary data from the DNA mutagenesis experiments conducted in our laboratory (You et al., 1991; Veldhoen et al., 1994) identified that the IE, and, in particular, two base pair positions within the IE are important for TFIIIA interaction (GC base pairs 70 and 71). These data agree well with the findings that finger 5 is important in promoter recognition by TFIIIA.

Finger 4 mutants produced a relatively modest decrease in the DNA binding: between 2- and 3-fold. Clemens et al. (1994) showed that addition of finger 4 to the three amino-terminal zinc fingers does not contribute significantly to DNA binding. Moreover, hydroxy-radical studies have shown that finger 4 makes closest contacts with the minor groove of the DNA (Hayes & Clemens, 1992). Interestingly, structural disruption of finger 4 in the 'broken finger' experiments lowered the DNA-binding affinity by as much as 9.3-fold (Del Rio et al., 1993). However, no alterations in the footprinting pattern were observed for the finger 4 TFIIIA mutant in the same study. It was thus assumed that finger 4 does not make any specific contacts with

DNA, but rather plays a role as a linker which positions adjacent fingers over the 5S DNA. In a recent study, McBryant et al. (1996) found that mutations in finger 4 DNA binding sequence resulted in a two-fold decrease in binding affinity, which is in concordance with our data. Interestingly, a binding site selection and amplification assay performed by the same group (McBryant et al., 1996) failed to recover wild-type finger 4 binding site sequence, but instead resulted in a selection of purine-rich sequence. This finding suggests that the overall DNA structure, rather than its particular sequence is the primary determinant of finger 4 binding affinity. Moreover, chemical modifications in the major groove of the finger 4 binding site did not affect high-affinity DNA binding of zf 1-5. These results further support the previously proposed model in which finger 4 binds across the DNA minor groove (Clemens et al., 1992; Hayes & Tillius, 1992).

Although both fingers 4 and 6 were proposed to interact with the minor groove of the DNA, the mutations in the two fingers produced nonequivalent results: while finger 4 mutants had an approximately two-fold reduction in binding, substitutions in finger 6 showed an improved DNA binding (Tp series of mutants). This may indicate that the two zinc fingers play distinct roles in DNA recognition by TFIIIA. There doesn't seem to be an unanimous agreement in the literature as to the energetic contribution of finger 6 to the overall DNA-binding affinity. Thus, Clemens et al. (1994) reported that addition of finger 6 to the 1-5 zf peptide weakened its binding to DNA by about three-fold. However, a loss of the binding affinity was detected in the 'broken-finger' experiments (Del Rio et al., 1993), which demonstrated a 3.4-fold drop for finger 6 mutant (versus a 9.3-fold drop for finger 4). It is of interest to note that the amino acid sequence of finger 6 of TFIIIA deviates most significantly from the consensus sequence established for fingers 1-3, as

well as from certain sequence elements of fingers 4 to 9 (Jacobs, 1992). Perhaps, due to this uniqueness finger 6 of TFIIIA adopts a distinct mode of DNA interaction in the context of the full-length protein.

With respect to the high-affinity RNA binding, most of the TFIIIA - p43 mutants (Tp4-6) retained the majority of the RNA-binding affinity. It is easily explained given that both proteins specialize in binding to the same RNA target, 5S rRNA. In addition, amino acids TWT located in the recognition α-helix of TFIIIA are conserved at the same position in finger 6 of p43, thus retaining the specific RNA-binding activity. Mutants that showed reduced RNA-binding activity include Tp4B mutant, which had about eight-fold lower affinity than the wild-type TFIIIA, and Tp5A mutant with a two-fold drop in affinity. The importance of finger 4 for RNA binding was previously demonstrated by many researchers. Thus, Clemens et al. (1993) showed that deletion of zinc finger 4 from the 4-7 zinc finger peptide resulted in a 20-fold drop in RNA binding. The fact that our substitution mutants produced smaller effects may be attributable to the use of the full-length TFIIIA protein, in which the loss of certain contacts may be compensated for by the rest of the protein. Different methods of assaying for affinities can also have an effect on the measured affinity values. A recent study by Friesen & Darby (1997) using a zf 4-7 zinc finger peptide of TFIIIA demonstrated that alanine substitutions in zinc fingers 4 or 6 of TFIIIA had a 77- and a 38-fold drop in affinity for RNA, respectively. In contrast, finger 5 alanine substitution mutant produced only a modest change in the RNA binding.

Finally, the most dramatic effect on RNA binding was revealed upon substitution of the 15 α-helical amino acids in zinc finger 6 with the corresponding region of WT1 protein (mutant TFW6). This substitution abolished specific RNA binding, while retaining almost wild-type affinity for

DNA. This suggests that the protein was active in our assays. A loss of RNA binding due to mutations in a single zinc finger was not previously observed. This may be due to the fact that other researchers used donor proteins that have extensive amino acid similarity with TFIIIA, thus creating mutants which retain significant levels of binding energy. As mentioned earlier, the three conserved amino acids at the tip of finger 6 may directly contribute to sequence-specific recognition of 5S rRNA. Thus, although the three-dimensional structure as opposed to a specific sequence of 5S rRNA is proposed to be the primary determinant for TFIIIA recognition, some sequence-specific contacts may also contribute to the interaction. A good candidate for such an interaction may be the loop A region of 5S rRNA, which was shown to provide direct contacts to the finger 6 (Romaniuk, 1989; Baudin et al., 1991; You et al., 1991; Setzer et al., 1996). The contacts of specific amino acids in zinc finger 6 of TFIIIA need to be assigned to individual residues in 5S rRNA.

In conclusion, this study provides quantitative analysis of the contributions of the central zinc fingers of TFIIIA to the high-affinity DNA and RNA binding. No single zinc finger in the central triplet proved to be responsible for the dramatic, 100-fold reduction in DNA binding affinity observed when all three zinc fingers were replaced. Rather, the central zinc fingers provide an integrated contribution to the DNA binding. In addition, we found that mutations in finger 6 α-helix result in a loss of specific RNA binding. Thus, finger 6 contains major determinants to support high-affinity RNA binding. The details of this phenomenon deserve further consideration.

8.0 Conclusions

Nucleic acid binding properties of three zinc finger proteins - WT1-ZF, EGR-1-ZF and TFIIIA were examined in the course of my Ph.D. work described in this thesis. Quantitative measurements of protein-nucleic acid equilibria were made using a nitrocellulose filter binding assay.

Wilms' tumour protein is a putative transcription factor of the C_2H_2 zinc finger family, and has four zinc finger motifs. EGR-1 - a prototypical member of the same family - has three zinc fingers for which consensus DNA binding site has been defined. Despite the fact that both WT1 and EGR-1 can bind to the same sequence, it was not clear what the preferred target sequence for WT1 was. In order to identify high affinity binding sites for WT1 and to determine the role of the additional, first zinc finger of WT1, we conducted a binding site selection assay (SAAB). The DNA target site had the sequence: GCG-TGG-GCG-**NNNNN**. The results of this experiment reveal that the high affinity binding site for WT1-ZF had the sequence: GCG-TGG-GCG-**(T/G)(A/G)(T/G)NN**. A WT1ΔF1-ZF protein which lacks the first zinc finger was used as a control, and showed no sequence preferences at any of the randomized positions. WT1-ZF protein had a four-fold higher affinity for the selected sequence versus the non-selected sequence GCG-TGG-GCG-**CCC**. The responsiveness of promoters containing either selected or non-selected sequences was measured in transient transfection assays using MTV-CAT reporter system. These experiments revealed that the lower affinity CCC site when present in three tandemly repeated copies confers greater responsiveness to WT1 than does the high affinity TGT site. Further experiments are necessary to clarify these phenomena.

Five point mutations in zinc fingers 2 and 3 of WT1 were isolated in DDS patients. It was therefore suggested that these mutations may interfere

with the normal function of WT1 by either lowering its affinity for DNA target sites, or conferring new DNA binding specificities. To test these possibilities, SAAB experiments were conducted for the DDS mutant proteins using two DNA templates: one with finger 2 randomized subsite, the other-with finger 3 randomized subsite. None of the five DDS mutant proteins selected new, high-affinity DNA binding sequences. This suggests that the observed DDS phenotype is likely not due to a change in gene networks regulated by WT1. Mutant R394W had no detectable DNA binding. The remaining mutants showed altered DNA binding selectivities. These mutant proteins had reduced DNA binding affinities, ranging from 1.4- to 14-fold. Thus, even small changes in DNA-binding affinity can result in a phenotype seen in DDS patients.

A number of equilibrium binding parameters for WT1-ZF and EGR1-ZF proteins were measured using a quantitative nitrocellulose binding assay. The equilibrium dissociation constants were determined to be $1.14 \pm 0.2 \times 10^{-09}$ M for WT1-ZF, and $3.55 \pm 0.4 \times 10^{-09}$ M for EGR1-ZF protein under the following conditions: pH 7.5, 0.1M KCl and incubation at 22°C. Point mutational anaysis was performed to measure the contribution of each base pair to high affinity binding for the two proteins. While exhibiting similarities, important differencies were revealed in the binding mechanisms of the two proteins, with respect to both thermodynamics and DNA sequence preferences. Thus, WT1-ZF-DNA interaction appears to be mostly an entropy driven process while EGR1-ZF-DNA interaction is driven by enthalpy. Placing adenine at position 8 of the consensus binding site resulted in 4.5-fold increase in WT1-ZF DNA-binding affinity while the EGR-1-ZF affinity dropped by a factor of 12 suggesting that the two proteins may act on distinct promoter sequences.

Transcription factor IIIA is necessary for the expression of the 5S rRNA genes and the storage of the 5S rRNA molecules. TFIIIA has thus both DNA and RNA binding activities. The roles of zinc fingers 4 through 7 in DNA and RNA binding were determined. The substitution of zinc fingers 4-7 of TFIIIA with those of WT1 resulted in the loss of DNA binding. The majority of mutations in individual zinc fingers produced modest effects on either DNA or RNA binding. The largest reduction of DNA binding affinity was observed when finger 5 of TFIIIA was mutated (10-fold), whereas substitution of an α-helical region of finger 6 abolished RNA binding activity. These findings suggest that the energetic contributions of TFIIIA zinc fingers for binding to either DNA or RNA are not equal. Thus, finger 6 is indispensable for high affinity RNA binding, while no single finger in the zf region 4-7 has the same effect on DNA binding of the factor.

9.0 Literature cited

Aggarwal, A.K., Rodgers, D.W., Drottar, M., Ptashne, M., Harrison, S.C. (1988). "Recognition of a DNA operator by the repressor of phage 434: a view at high resolution." Science **242**: 899-907.

Anant, S., Axenovich, S.A., Madden, S.L., Rauscher III, F.J. and Subramanian, K.N. (1994). "Novel replication inhibitory functions of the developmental regulator / transcription repressor protein WT1 encoded by the Wilms' tumor gene." Oncogene **9**: 3113-3126.

Andersen, J. and Delihas, N. (1986). "Characterization of RNA-protein interactions in 7S RNP from *Xenopus laevis* oocytes." The Journal of Biological Chemistry **261**: 2912-2917.

Andrews, M. T. and Brown, D. D. (1987). "Transient activation of oocyte 5S RNA genes in *Xenopus* embryos by raising the level of the trans-acting factor TFIIIA." Cell **51**: 445-453.

Baird, P.N., Groves, N., Haber, D.A., Housman, D.E. and Cowell J.K. (1992a). "Identification of mutations in the WT1 gene in tumours from patients with the WAGR syndrome." Oncogene **7**: 2141-2149.

Baird, P.N., Santos, A., Groves, N., Jadresic, L. and Cowell, J.K. (1992b). "Constitutional mutations in the WT1 gene in patients with Denys-Drash syndrome." Human Molecular Genetetics **1**: 301-305.

Baird, P.N. and Cowell, J.K. (1993). "A novel zinc finger mutation in a patient with Denys-Drash syndrome." Human Molecular Genetics **2**: 2193-2194.

Baudin, F., Romby, P., Romaniuk, P. J., Ehresmann, B. and Ehresmann, C. (1989). "Crosslinking of transcription factor TFIIIA to ribosomal 5S RNA from *X. laevis* by trans-diamminedichloroplatinum(II)." Nucleic Acids Research **17**: 10035-10046.

Bazett-Jones, D. P. and Brown, M. L. (1989). "Electron Microscopy reveals that transcription factor TFIIIA bends 5S DNA." Molecular and Cell Biology **9** (1): 336-341.

Beaudette, N.V., & Langerman, N. (1980). CRC Critical Reviews in Biochemistry. CRC Press, Cleveland, OH, pp 145-169.

Berg, J. M. (1988). "Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins." Proceeding of the National Academy of Sciences of the United States of America **85**: 99-102.

Berg, J. M. (1990). "Zinc Finger Domains: Hypotheses and Current Knowledge." Annual Review of Biophysics and Biophysical Chemistry **19**: 405-421.

Bhat, R., Worley, P., Cole, A., and Baraban, J. (1992). "Activation of the zinc finger encoding gene Krox-20 in adult rat brain: comparison with zif268." Molecular Brain Research **13**: 263-266.

Bickmore, W.A., Oghene, K., Little, M.H., Seawright, A., van Heyningen, V. and Hastie, N.D. (1992). "Modulation of DNA binding specificity by alternative splicing of the Wilms tumor wt1 gene transcript." Science **257**: 235-237.

Bieker, J. J. and Roeder, R. G. (1984). "Physical properties and DNA-binding stoichiometry of a 5S gene-specific transcription factor." The Journal of Biological Chemistry **259**: 6158-6164.

Bieker, J.J., Martin, P.L., and Roeder, R.G. (1985). "Formation of a rate-limiting intermediate in 5S RNA gene transcription." Cell **40**: 119-127.

Bieker, J. J. and Roeder, R. G. (1986). "Characterization of the nucleotide requirement for elimination of the rate-limiting step in 5S RNA gene transcription." The Journal of Biological Chemistry **261**: 9732-9738.

Blackwell, T. K. and Weintraub, H. (1990). "Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection." Science **250**: 1104-1110.

Blake, P.R. and Summers, M.F. (1994). "Probing the unusually similar metal coordination sites of retroviral zinc fingers and iron-sulfur proteins by nuclear magnetic resonance." Advances in Biophysical Chemistry **10**: 201-228.

Bogenhagen, D. F., Sakonju, S. and Brown, D. D. (1980). "A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region." Cell **19**: 27-35.

Bogenhagen, D. F. and Brown, D. D. (1981). "Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination." Cell **24**: 261-270.

Bogenhagen, D. F., Wormington, W. M. and Brown, D. D. (1982). "Stable transcription complexes of *Xenopus* 5S RNA genes: a means to maintain the differentiated state." Cell **28**: 413-421.

Bonetta, L., Kuehn, S.E., Huang, A., Law, D.J., Kalikin, L.M., Koi, M., Reeve, A.E., Brownstein, B.H., Yeger, H., Williams, B.R.G., and Feinberg, A.P. (1990). "Wilms tumour locus on 11p13 defined by multiple CpG island-associated transcripts." Science **250**: 994-997.

Bonventre, J.V., Sukhatme, V.P., Bamberger, M., Ouellette, A.J., and Brown, D. (1991). "Localization of the protein product of the immediate early growth response gene, *Egr-1* in the kidney after iscemia and reperfusion." Cell Regulation **2**: 251-260.

Bordereaux, D., Fichelson, S., Tambourin, P., and S. Gisselbrecht. (1990). "Alternative splicing of the Evi-1 zinc fingers gene generates mRNAs which differ by the number of zinc finger motifs." Oncogene **5**: 925-927.

Borel, F., Barilla K., Hamilton T., Iskandar, M., Romaniuk, P. (1996). "Effects of Denys-Drash syndrome point mutations on the DNA binding activity of the Wilms' tumour suppressor protein WT1." Biochemistry 35: 12070-12076.

Bradeesy, N., Zabel, B., Schmitt, K., and Pelletier, J. (1994). "WT1 mutations associated with incomplete Denys-Drash syndrome define a domain predicted to behave in a dominant-negative fashoin." Genomics **21**: 663-665.

Bradford, M. M. (1976). "A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding." Analytical Biochemistry **72**: 248-254.

Brodeur, G.M. (1990). "Neuroblastoma: clinical significance of genetic abnormalities." Cancer Surveys **9**: 673-688.

Brown, D. D., Carroll, D. and Brown, R. D. (1977). "The isolation and characterization of a second oocyte 5S DNA from *Xenopus laevis*." Cell **12**: 1045-1056.

Brown, R. S., Sander, C. and Argos, P. (1985). "The primary structure of transcription factor TFIIIA has 12 consecutive repeats." FEBS (Federation of European Biochemical Societies) Letters **186**: 271-274.

Bruening, W., Bardeesy, N., Silverman, B.L., Cohn, R.A., Aronson, A.J., Housman, D., and Pelletier, J. (1992). "Germline intronic and exonic mutations in the Wilms' tumor gene (WT1) affecting urogenital development." Nature Genetics **1**: 144-148.

Bruening, W., and Pelletier, J. (1996). "A non-AUG translational initiation event generates novel WT1 isoforms." The Journal of Biological Chemistry **271**: 8646-8654.

Call, K.M., Glaser, T., Ito, C.Y., Buckler, A.J., Pelletier, J., Haber, D.A., Rose, E.A., Kral, A., Yeger, H., Lewis, W.H., Jones, C. and Housman, D.E. (1990). "Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus." Cell **60**: 509-520.

Cao, X., Koski, R., and Gashler, A. McKiernan, M., Morris, C.F., Gaffney, R., Hay, R.V., and Sukhatme, V.P. (1990). "Identification and characterization of the EGR-1 gene product, a DNA-binding zinc finger protein induced by differentiation and growth signals." Molecular and Cellular Biology **10**: 1931-1939.

Caricasole, A., Duarte, A., Larsson, S.H., Hastie, N.D., Little, M., Holmes, G., Todorov, I. and Ward, A. (1996). "RNA binding by the Wilms tumor suppressor zinc finger proteins." Proceedings of the National Academy of Sciences of the United States of America **93**: 7562-7566.

Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R., and Charney, P. (1988). "A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells." The EMBO (European Molecular Biology Organization) Journal **7**: 29-35.

Cho, Y., Gorina, S. Jeffrey, P.D., N.P. Pavletich. (1994). "Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations." Science **265**: 346.

Choo, Y. and Klug, A. (1993). "A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA." Nucleic Acids Research **21** (15): 3341-3346.

Choo, Y. and Klug, A. (1994a). "Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage." Proceedings of the National Academy of Sciences of the United States of America **91**: 11163-11167.

Choo, Y. and Klug, A. (1994b). "Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions." Proceedings of the National Academy of Sciences of the United States of America **91**: 11168-11172.

Choo, Y., Sanchez-Garcia, I., and Klug, A. (1994c). "*In vivo* repression by a site-specific DNA-binding protein designed against an oncogenic sequence." Nature (London) **372**: 642-645.

Choo, Y., and Klug, A. (1995). "Designing DNA-binding proteins on the surface of filamentous phage." Current Opinion in Biotechnilogy **6**: 431-436.

Choo, Y., and Klug, A. (1997). "Physical basis of a protein-DNA recognition code." Current Opinion in Structural Biology **7**: 117-125.

Christensen, J. H., Hansen, P. K., Lillelund, O. and Thogersen, H. C. (1991). "Sequence-Specific Binding of the N-Terminal Three-Finger Fragment of *Xenopus* Transcription Factor-IIIA to the Internal Control Region of a 5S RNA Gene." FEBS (Federation of European Biochemical Societies) Letters **281** (1-2): 181-184.

Christiansen, J., Brown, R. S., Sproat, B. S. and Garrett, R. A. (1987). "*Xenopus* transcription factor IIIA binds primarily at junctions between double helical stems and internal loops in oocyte 5S RNA." EMBO (European Molecular Biology Organization) Journal **6**: 453-460.

Christy, B., Lau, L., and Nathans, D. (1988). "A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences." Proceedings of the National Academy of Sciences of the United States of America 85: 7857-7861.

Christy, B. and Nathans, D. (1989). "DNA binding site of the growth factor-inducible protein Zif268." Proceedings of the National Academy of Sciences of the United States of America **86**: 8737-8741.

Churchill, M. E. A., Tullius, T. D. and Klug, A. (1990). "Mode of interaction of the zinc finger protein TFIIIA with a 5S RNA gene of *Xenopus*." Proceeding of the National Academy of Sciences of the United States of America **87** (7): 5528-5532.

Ciliberto, G., Castagnoli, and Cortese, R. (1983a). "Transcription by RNA polymerase III." Current Topics in Developmental Biology **18**: 59-87.

Ciliberto, G., Raugei, G., Costanzo, F., Dente, L. and Cortese, R. (1983b). "Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase III." Cell **32**: 725-733.

Clemens, K. R., Liao, X. B., Wolf, V., Wright, P. E. and Gottesfeld, J. M. (1992). "Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex." Proceedings of the National Academy of Sciences of the United States of America **89** (22): 10822-10826.

Clemens, K. R., Wolf, V., McBryant, S. J., Zhang, P. H., Liao, X. B., Wright, P. E. and Gottesfeld, J. M. (1993). "Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein." Science **260** (5107): 530-533.

Clemens, K. R., Zhang, P., Liao, X., McBryant, S. J., Wright, P. E. and Gottesfeld, J. M. (1994). "Relative contributions of the zinc fingers of transcription factor IIIA to the energetics of DNA binding." Journal of Molecular Biology **244**: 23-35.

Coleman, J.E. (1992). "Zinc Proteins: enzymes, storage proteins, transcription factors, and replication proteins." Annual Review of Biochemistry. **61**: 897-964.

Coppes, M.J., Bonetta, L., Huang, A., Hoban, P., Chilton-MacNeill, S., Campbell, C.E., Weksberg, R., Yeger, H., Reeve, A.E., and Williams, B.R.G. (1992a). "Loss of heterozygosity mapping in Wilms' tumor indicates the involvement of three distinct regions and a limited role for non-disjunction or mitotic recombination." Genes Chromosomes Cancer **5**: 326-334.

Coppes, M.J., Liefers, G.J., Higuchi, M., Zinn, A.B., Balfe, J.W., and Williams, B.R.G. (1992b). "Inherited WT1 mutations in Denys-Drash syndrome." Cancer Research **52**: 6125-6128.

Coppes, M.J., Campbell, C.E. and Williams, B.R. (1993). "The role of WT1 in Wilms' tumorigenesis." FASEB (Federation of American Societies for Experimental Biology) Journal **7**: 886-895.

Crosby, S.D., Puetz, J.J., Simburger, K.S., Fahrner, T.J., and Milbrandt, J. (1991). "The early response gene NGFI-C encodes a zinc finger transcriptional activator and is a mamber of the GCGGGGGCG (GSG) element-binding protein family." Molecular and Cellular Biology **11**: 3835-3841.

D'Angio, G.J., Breslow, N., Bechwith, J.B., Evans, A., Baum, H., DeLorimier, A., Fernbach, D., Hrabovski, E., Jones, B., Kelalis, P., Othersen, B., Tefft, M., and Thomas, P.R.M. (1989). "Treatment of Wilms' tumor. Results of the third national Wilms' tumor study." Cancer **64**: 349-360.

Darby, M. K. and Joho, K. E. (1992). "Differential Binding of Zinc Fingers from *Xenopus* TFIIIA and p43 to 5S RNA and the 5S RNA Gene." Molecular and Cellular Biology **12** (7): 3155-3164.

Day, M., Fahrner, T., Aykent, S., and Milbrandt, J. (1990). "The zinc finger protein NGFI-A exists in both nuclear and cytoplasmic forms in nerve growth factor-stimulated PC12 cells." The Journal of Biological Chemistry **25**: 15253.

Del Rio, S. and Setzer, D. R. (1991). "High Yield Purification of Active Transcription Factor IIIA Expressed in *E Coli.*" Nucleic Acids Research **19** (22): 6197-6203.

Del Rio, S. and Setzer, D. R. (1993a). "The role of zinc fingers in transcriptional activation by transcription factor IIIA." Proceedings of the National Academy of Sciences of the United States of America **90** (1): 168-172.

Del Rio, S., Menezes, S. R. and Setzer, D. R. (1993b). "The function of individual zinc fingers in Sequence-Specific DNA recognition by transcription Factor IIIA." Journal of Molecular Biology **233** (4): 567-579.

Delihas, N. and Andersen, J. (1982). "Generalized structures of the 5S ribosomal RNAs." Nucleic Acids Research **10**: 7323-7344.

Delihas, N., Andersen, J. and Singhal, R. P. (1984). "Structure, function and evolution of 5S ribosomal RNAs." Progress in Nucleic Acid Research and Molecular Biology **31**: 161-190.

Denys, P., Malvaux, P., Van Den Berghe, H., Tanghe, W., and Proesmans, W. (1967). "Association d'un syndrome anatomo-pathologique de pseudohermaphroditisme masculin, d'une tumeur de Wilms, d'une nephropathie parenchymateuse et d'un mosaicisme XX/XY." Archives of French Pediatrics **24**: 729-739.

Desjarlais, J. R. and Berg, J. M. (1992). "Toward rules relating zinc finger protein sequences and DNA binding site preferences." Proceedings of the National Academy of Sciences of the United States of America **89** (16): 7345-7349.

Desjarlais, J. R. and Berg, J. M. (1993). "Use of a zinc finger consensus sequence framework and specificity rules to design specific DNA binding proteins." Proceedings of the National Academy of Sciences of the United States of America **90**: 2256-2260.

Dey, B.R., Sukhatme, V.P., Roberts, A.B., Sporn, M.B., Rauscher III, F.J. and Kim, S.J. (1994). "Repression of the transforming growth factor-β1 gene by the Wilms' tumor suppressor WT1 gene product." Molecular Endocrinology **8**: 595-602.

Diakun, G. P., Fairall, L. and Klug, A. (1986). "EXAFS study of the zinc-binding sites in the protein transcription factor IIIA." Nature (London) **324**: 698-699.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, Jr., C.A., Butel. J.S. and Bradley, A. (1992). "Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors." Nature (London) **356**: 215-221.

Drash, A., Sherman, F., Hartman, W.H. and Blizzard, R.M. (1970). "A syndrome of pseudohermaphroditism, Wilms' tumor, hypertension, and degenerative renal desease." Journal of Pediatrics **76**: 585-593.

Drummond, I.A., Madden, S.L., Rowher-Nutter, P., Bell, G.I., Sukhatme, V.P., and Rauscher III, F.J. (1992). "Repression of the insulin-like growth factor-II gene by the Wilms tumor suppressor WT1." Science **257**: 674-678.

Drummond, I.A., Rupprecht, H.D., Rohwer-Nutter, P., Lopez-Guiza, J.M., Madden, S.L., Rauscher III, F.J., and Sukhatme, V.P. (1994). "DNA recognition by splicing variants of the Wilms' tumor suppressor, WT1." Molecular and Cellular Biology **14**: 3800-3809.

Elbert, S.N., Balt, S.L., Hunter, J.P., Gashler, A., Sukhatme, V.P., and Wong, D.L. (1994). "Egr-1 activation in rat adrenal phenylethanolamine N-methyltransferase gene." The Journal of Biological Chemistry **269**: 20885-20898.

Elrod-Erickson, M., Rould, M.A., Nekludova, L., Pabo, C.O. (1996). "Zif268 protein-DNA complex refined at 1.6 A°: a model system for understanding zinc finger-DNA interactions." Structure **4**: 1171-1180.

Engelke, D. R., Ng, S.-Y., Shastry, B. S. and Roeder, R. G. (1980). "Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes." Cell **19**: 717-728.

Englert, C., Hou, X., Maheswaran, S., Bennett, P., Ngwu, C., Re, G.G., Garvin, A.J., Rosner, M.R., and Haber, D.A. (1995a). "WT1 suppresses synthesis of the epidermal growth factor receptor and induces apoptosis." EMBO (European Molecular Biology Organization) Journal **14**: 4662-4675.

Englert, C., Vidal, M., Maheswaran, S., Ge, Y., Ezzel, R.M., Isselbacher, K.J. and Haber, D.A. (1995b). "Truncated WT1 mutants alter the subnuclear localization of the wild-type protein." Proceedings of the National Academy of Sciences of the United States of America **92**: 11960-11964.

Erdmann, V.A., Wolters, J., Pieler, T., Digweed, M., Specht, T., Ulbrich, N. (1987). "Evolution of organisms and organelles as studied by comparative computer and biochemical analysis of ribosomal 5S RNA structure." Annals New York Academy of Sciences **503**: 103-123.

Evans, R.M. (1988). "The steroid and thyroid hormone receptor superfamily." Science **240**: 889-895.

Fairall, L., Rhodes, D. and Klug, A. (1986). "Mapping of the sites of protection on a 5S RNA gene by the *Xenopus* transcription factor IIIA. A model for the interaction." Journal of Molecular Biology **192**: 577-591.

Fairall, L., Martin, S. and Rhodes, D. (1989). "The DNA binding site of the *Xenopus* transcription factor IIIA has a non-B-form structure." EMBO (European Molecular Biology Organization) Journal **8** (6): 1809-1817.

Fairall, L. and Rhodes, D. (1992). "A new approach to the analysis of DNase I footprinting data and its application to the TFIIIA/5S DNA complex." Nucleic Acids Research **20** (18): 4727-4731.

Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T. and Rhodes, D. (1993). "The crystal structure of a two zinc finger-peptide reveals an extension to the rules for zinc-finger DNA recognition." Nature (London) **366** (6454): 483-487.

Fearon, E.R., and Vogelstein, B. (1990). "A genetic model for colorectal tumorigenesis." Cell **61**: 759-767.

Fedoroff, N. V. and Brown, D. D. (1978). "The nucleotide sequence of oocyte 5S DNA in *Xenopus laevis*. I. The AT-rich spacer." Cell **13**: 701-716.

Fraizer, G.E., Bowen-Pope, D.F. and Vogel, A.M. (1987). "Transcriptional regulation of the human Wilms' tumor gene (wt1)." Journal of Cellular Physiology **133**: 169-174.

Francke, U., Holmes, L.B., Atkins, L., and Riccardi, V.M. (1979). "Aniridia-Wilms' tumour association: evidence for specific deletion of 11p13." Cytogenetics Cell Genetics **24**: 185-192.

Frankel, A. D., Berg, J. M. and Pabo, C. O. (1987). "Metal-dependant folding of a single zinc finger from transcription factor IIIA." Proceeding of the National Academy of Sciences of the United States of America **84**: 4841-4845.

Freemont, P. S. (1993). "The RING finger. A novel protein sequence motif related to the zinc finger." Annals of New York Academy of Sciences **684**: 174-192.

Friesen, W.J. and Darby, M. K. (1997). "Phage display of RNA binding zinc fingers from transcription factor IIIA." The Journal of Biological Chemistry **272**: 10994-10997.

Gashler, A.L., Bonthron, D.T., Madden, S.L., Rauscher III, F.J., Collins, T., and Sukhatme, V.P. (1992). "Human platelet-derived growth factor A chain is transcriptionally repressed by the Wilms tumor suppressor WT1." Proceedings of the National Academy of Sciences of the United States of America **89**: 10984-10988.

Gashler, A.L., Swaminathan, S., and Sukhatme, V.P. (1993). "A novel repression module, an extensive activation domain, and a bipartate nuclear localization signal defined in the immediate-early transcription factor EGR-1." Molecular and Cellular Biology **13**: 4556-4571.

Gashler, A.L., and Sukhatme, V.P. (1995). "Early growth response protein 1 (EGR-1): Prototype of a zinc finger family of transcription factors." Nucleic Acid Research and Molecular Biology **50**: 191-221.

Gessler, M., Poustka, A., Cavenee, W., Neve, R.L., Orkin, S.H., and Bruns, G.A. (1990). "Homozygous deletion in Wilms' tumours of a zinc-finger gene identified by chromosome jumping." Nature (London) **343**: 774-778.

Gibson, T. J., Postma, J. P. M., Brown, R. S. and Argos, P. (1988). "A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger')

common to many eukaryotic transcriptional regulatory proteins." Protein Engineering **2** (3): 209-218.

Ginsberg, A. M., King, B. O. and Roeder, R. G. (1984). "*Xenopus* 5S Gene Transcription Factor, TFIIIA: Characterization of a cDNA Clone and Measurement of RNA Levels throughout Development." Cell **39**: 479-489.

Gogos, J.A., Hsu, T., Bolton, J., and F.C. Kafatos. (1992). "Sequence discrimination by alternatively spliced isoforms of a DNA binding zinc finger domain." Science **257**: 1951-1955.

Goodyer, P., Dehbi, M., Torban, E., Bruening, W. and Pelletier, J. (1995). "Repression of the retinoic acid receptor-alpha gene by the Wilms' tumor suppressor gene product, wt1." Oncogene **10**: 1125-1129.

Guddat, U., Bakken, A. H. and Pieler, T. (1990). "Protein-Mediated Nuclear Export of RNA: 5S rRNA Containing Small RNPs in *Xenopus* Oocytes." Cell **60** (4): 619-628.

Gupta, M.P., Gupta, G., Zak, R., and Sukhatme, V.P. (1991). "EGR-1, a serum-inducible zinc finger protein, regulates transcription of the rat cardiac alpha-myosin heavy chain gene." The Journal of Biological Chemistry **266**: 12813-12816.

Haber, D.A., Sohn, R.L., Buckler, A.J., Pelletier, J., Call, K.M., and Housman, D.E. (1991). "Alternative splicing and genomic structure of the Wilms tumour gene WT1. " Proceedings of the National Academy of Sciences of the United States of America 88: 9618-9622.

Haber, D. A., and Buckler, A. J. (1992a). "WT1: a novel tumor suppressor gene inactivated in Wilms' tumour." New Biology **4:** 97-106.

Haber, D.A., and Housman, D.E. (1992b). "The genetics of Wilms' tumor." Advances in Cancer Research **59**: 41-68.

Hall, R. K. and Taylor, W. L. (1989). "Transcription Factor IIIA Gene Expression in *Xenopus* Oocytes Utilizes a Transcription Factor Similar to the Major Late Transcription Factor." Molecular and Cellular Biology **9** (11): 5003-5011.

Hallahan, D., Sukhatme, V.P., Sherman, M., Virudachalam, S., Kufe, D., and Weichselbaum, R.R. (1991). "Protein kinase C mediates X-ray inducibility of

nuclear signal transducers EGR-1 and Jun." Proceedings of the National Academy of Sciences of the United States of America **88**: 2156-2160.

Hamel, P.A., Phillips, R.A., Muncaster, M., and Gallie, B. (1993). "Speculations on roles of Rb1 in tissue-specific differentiation, tumor initiation, and tumor progression." FASEB (Federation of American Societies for Experimental Biology) Journal **7**: 846-854.

Hamilton, T.B., Barilla, K.C., and Romaniuk, P.J. (1995). "High affinity binding sites for the Wilms' tumor suppressor protein WT1." Nucleic Acids Research **23**: 277-284.

Hamilton, T.B., Borel, F., and Romaniuk, P.J. (1998). "Comparison of DNA binding characteristics of the related zinc finger proteins WT1 and EGR-1." Biochemistry 37: 2051-2058.

Han, K., and Manley, J. (1993). "Transcriptional repression by the *Drosophila* even-skipped protein: definition of a minimal repression domain." Genes and Development **7**: 491-503.

Hanas, J. S., Hazuda, D. J., Bogenhagen, D. F., Wu, F. Y.-H. and Wu, C.-W. (1983). "*Xenopus* transcription factor A requires Zinc for binding to the 5S RNA gene." The Journal of Biological Chemistry **258**: 14120-14125.

Hanas, J. S., Bogenhagen, D. F. and Wu, C.-W. (1984). "Binding of *Xenopus* transcription factor A to 5S RNA and to single stranded DNA." Nucleic Acids Research **12**: 2745-2758.

Hansen, P. K., Christensen, J. H., Nyborg, J., Lillelund, O. and Thogersen, H. C. (1993). "Dissection of the DNA-Binding domain of *Xenopus laevis* TFIIIA - quantitative DNase-I footprinting analysis of specific complexes between a 5S RNA gene fragment and N-Terminal fragments of TFIIIA containing 3-Zinc Finger, 4-Zinc Finger or 5-Zinc Finger domains." Journal of Molecular Biology **233** (2): 191-202.

Harrington, M.A., Konicek, B., Song, A., Xia, X.-L., Fredericks, W.J. and Rauscher III, F.J. (1993). "Inhibition of colony-stimulating factor-1 promoter activity by the product of the Wilms' tumor locus." The Journal of Biological Chemistry **268**: 21271-21275.

Harris, H., Miller, O.J., Klien, G., Worst, P., and Tachibana, T. (1969). "Suppression of malignancy by cell fusion." Nature (London) **223**: 363-368.

Hastie, N.D. (1994). "The genetics of Wilms' tumor - a case of disrupted development." Annual Review of Genetics **28**: 523-558.

Hayes, J. J. and Clemens, K. R. (1992a). "Locations of contacts between individual zinc fingers of *Xenopus laevis* transcription Factor IIIA and the internal control region of a 5S RNA gene." Biochemistry **31** (46): 11600-11605.

Hayes, J. J. and Tullius, T. D. (1992b). "Structure of the TFIIIA-5S DNA complex." Journal of Molecular Biology **227** (2): 407-417.

Henry, I., Grandjouan, S., Coullin, P., Barichard, F., Huerre-Jeanpierre, C., Glaser, T., Philip, T., Lenoir, G., Chaussain, J.L., and Junien, C. (1989). "Tumor-specific loss of 11p15.5 alleles in del11p13 Wilms' tumour and in familial adrenocortical carcinoma." Proceedings of the National Academy of Sciences USA **86**: 3247-3251.

Hewitt, S.M., Fraizer, G.C., Wu, Y.-J., Rauscher III, F.J., and Saunders, G.F. (1996). "Differential function of Wilms' tumor gene WT1 splice isoforms in trnascriptional regulation." The Journal of Biological Chemistry **271**: 8588 - 8592.

Honda, B. M. and Roeder, R. G. (1980). "Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation." Cell **22**: 119-126.

Hori, H. and Osawa, S. (1987). "Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences." Molecular Biology and Evolution **4** (5): 445-472.

Houbaviy, H.B., Usheva, A., Shenk, T., Burley, S.K. (1996). "Cocrystal structure of YY1 bound to the adeno-associated virus P5 initiator." Proceedings of the National Academy of Sciences of the United States of America **93**: 13577-13582.

Hsu, T., Gogos, J.A., Kirsh, S.A., and F.C. Kafatos. (1992). "Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene." Science **257**: 1946-1950.

Huber, P. W. and Wool, I. G. (1986). "Identification of the binding site on 5S rRNA for the transcription factor IIIA: Proposed structure of a common binding site on 5S rRNA and on the gene." Proceeding of the National Academy of Sciences of the United States of America **83**: 1593-1597.

Huber, P. W., Morii, T., Mei, H. Y. and Barton, J. K. (1991). "Structural Polymorphism in the Major Groove of a 5S RNA Gene Complements the Zinc Finger Domains of Transcription Factor IIIA." Proceedings of the National Academy of Sciences of the United States of America **88** (23): 10801-10805.

Huff, V., Villalba, F., Strong, L.C., and Saunders, G.F. (1991). "Alteration of the WT1 gene in patients with Wilms' tumor and genitourinary anomalies." American Journal of Human Genetics **49**: 44.

Isalan, M., Choo, Y., and Klug, A. (1997). "Synergy between adjacent zinc fingers in sequence-specific DNA recognition." Proceedings of the National Academy of Sciences of the United States of America **94**: 5617-5621.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). "Tumor spectrum analysis of p53-mutant mice." Current Biology **4**: 1-7.

Jacobs, G. H. (1992). "Determination of the base recognition positions of zinc fingers from sequence analysis." EMBO (European Molecular Biology Organization) Journal **11** (12): 4507-4517.

Jamieson, A. C., Kim, S. H. and Wells, J. A. (1994). "*In vitro* selection of zinc fingers with altered DNA-binding specificity." Biochemistry **33** (19): 5689-5695.

Jamieson, A., Hongming, W., and Kim, S.-H. (1996). "A zinc finger directory for high-affinity DNA recognition." Biochemistry **93**: 12834-12839.

Joho, K.E., Darby, M.K., Crawford, E.T., and Brown, D.D. (1990). "A finger protein structurally similar to TFIIIA that binds exclusively to 5S RNA in *Xenopus*." Cell **61**: 293-300.

Joseph, L., Le Beau, M., and Jamieson, G. (1988). "Molecular cloning, sequencing, and mapping of EGR-2, a human early growth response gene encoding a protein with "zinc-binding finger" structure." Proceedings of the National Academy of Sciences of the United States of America **85**: 7164-7168.

Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. (1987). "Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain." Cell **51**: 1079-1090.

Kassavetis, G. A., Braun, B. R., Nguyen, L. H. and Geiduschek, E. P. (1990). "*S. Cerevisiae* TFIIIB Is the Transcription Initiation Factor Proper of RNA Polymerase III, While TFIIIA and TFIIIC Are Assembly Factors." Cell **60** (2): 235-245.

Kim, C.A., Berg, J.M. (1996). "A 2.2 Å resolution crystal structure of a designed zinc finger protein bound to DNA." Nature Structural Biology **3**: 940-945.

Kim, S. H., Darby, M. K., Joho, K. E. and Brown, D. D. (1990). "The Characterization of the TFIIIA Synthesized in Somatic Cells of *Xenopus laevis*." Gene & Development **4** (9): 1602-1610.

Kjems, J., Olesen, S. O. and Garrett, R. A. (1985). "Comparison of Eubacterial and Eukaryotic 5S RNA Structures: A chemical modification study." Biochemistry **24**: 241-250.

Klevit, R. E. (1991). "Recognition of DNA by Cys2, His2 zinc fingers." Science **253** (5026): 1367.

Klug, A. and Rhodes, D. (1987). "Zinc fingers: a novel protein motif for nucleic acid recognition." Trends in Biochemical Sciences **12**: 464-469.

Knudson, A.G. (1971). "Mutation and cancer: statistical study of retinoblastoma." Proceedings of the National Academy of Sciences of the United States of America **68**: 820-823.

Knudson, A.G., and Strong, L.C. (1972). "Mutation and cancer: a model for Wilms' tumor of kidney." Journal of National Cancer Institute **48**: 313-324.

Kochoyan, M., Havel, T. F., Nguyen, D. T., Dahl, C. E., Keutmann, H. T. and Weiss, M. A. (1991a). "Alternating zinc fingers in the human male associated protein ZFY - 2D NMR structure of an even finger and implications for "jumping linker" DNA recognition." Biochemistry **30** (14): 3371-3386.

Kochoyan, M., Keutmann, H. T. and Weiss, M. A. (1991b). "Alternating zinc fingers in the human male associated protein ZFY - refinement of the NMR structure of an even finger by selective deuterium labeling and implications for DNA recognition." Biochemistry **30** (29): 7063-7072.

Kochoyan, M., Keutmann, H. T. and Weiss, M. A. (1991c). "Alternating zinc fingers in the human male-associated protein ZFY - HX3H and HX4H motifs encode a local structural switch." Biochemistry **30** (39): 9396-9402.

Kochoyan, M., Keutmann, H. T. and Weiss, M. A. (1991d). "Architectural rules of the zinc finger motif: Comparative two-dimensional NMR studies of native and "aromatic-swap" domains define a "weakly polar switch"." Proceedings of the National Academy of Sciences of the United States of America 88: 8455-8459.

Koufos, A., Grundy, P., Morgan, K., Aleck, K.A., Hadro, T., Lampkin, B.C., Kalbakji, A., and Cavenee, W.K. (1989). "Familial Wiedemann-Beckwith syndrome and a second Wilms' tumor locus both map to 11p15.5." American Journal of Human Genetics 44: 711-719.

Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). "WT1 is required for early kidney development." Cell 74: 679-691.

Kriwacki, R. W., Schultz, S. C., Steitz, T. A. and Caradonna, J. P. (1992). "Sequence-specific recognition of DNA by zinc finger peptides derived from the transcription factor Sp1." Proceedings of the National Academy of Sciences of the United States of America 89 (20): 9759-9763.

Lamaire, P., Revelent, O., Bravo, R., and Charnay, P. (1988). "Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells." Proceedings of the National Academy of Sciences of the United States of America 85: 4691-4695.

Larsson, S.H., Chaliur, J.-P., Mijagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V. and Hastie, N.D. (1995). "Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing." Cell 81: 391-401.

Lassar, A. B., Martin, P. L. and Roeder, R. G. (1983). "Transcription of class III genes: formation of preinitiation complexes." Science 222: 740-748.

Lau, L.F., and Nathans, D. (1987). "Expression of a set of growth-related immediate-early genes in BALB/c 3T3 cells: coordinate regulation with c-fos or c-myc." Proceedings of the National Academy of Sciences of the United States of America 84: 1182-1186.

Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A. and Wright, P. E. (1989). "Three-dimensional solution structure of a single zinc finger DNA-binding domain." Science 245 (4918): 635-637.

Lee, M. S., Gottesfeld, J. M. and Wright, P. E. (1991). "Zinc is required for folding and binding of a single zinc finger to DNA." FEBS (Federation of European Biochemical Societies) Letters **279** (2): 289-294.

Lees-Miller ,S.P., and Anderson, C.W. (1991). "The DNA-activated protein kinase, DNA-PK: a potential coordinator of nuclear events." Cancer Cell **3**: 341-346.

Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Frank, R., and Charnay, P. (1990). "The serum-inducible mouse gene *Krox-24* encodes a sequence-specific transcriptional activator." Molecular and Cellular Biology **10**: 3456-3467.

Lemerle, J., Voute, P.A., Tournade, M.F., Rodary, C., Delemarre, J.F.M., Sirrazin, D., and Burgers, J.M.V. (1983) "Effectiveness of preoperative chemotherapy in Wilms' tumor: results of an international society of paediatric oncology (SIOP) clinical trial." Journal of Clinical Oncology **10**: 604-609.

Levine, M. and Manley, J.L. (1989). "Transcriptional repression of eukaryotic promoters." Cell **59**: 405-408.

Lewis, W.H., Yeger, H., Bonetta, L., Chan, H.L.S., Kang, J., Junien, C., Cowell, J., Jones, C., and Dafoe, L.A. (1988). "Homozygous deletion of a DNA marker from chromosome 11p13 in sporadic Wilms' tumor." Genomics **3**: 25-31.

Liao, X. B., Clemens, K. R., Tennant, L., Wright, P. E. and Gottesfeld, J. M. (1992). "Specific interaction of the first-three zinc fingers of TFIIIA with the internal control region of the Xenopus 5S RNA gene." Journal of Molecular Biology **223** (4): 857-871.

Licht, J., Grossel, M., Figge, J., and Hansen, U. (1990). "*Drosophila* Kruppel protein is a transcriptional repressor." Nature (London) **346**: 76-79.

Lim, R. W., Varnum, B. C., and Herschman, H. R. (1987). "Cloning of tetradecanoyl phorbol-ester-induced 'primary response' sequences and their expression in density-arrested Swiss 3T3 cells and a TPA non-proliferative variant." Oncogene **1**: 263-270.

Little, M.H., Prosser, J., Condie, A., Smith, P.J., Van Heyningen, V. and Hastie, N.D. (1992). "Zinc finger point mutations within the WT1 gene in Wilms' tumor patients." Proceedings of the National Academy of Sciences of the United States of America **89**: 4791-4795.

Little, M.H., Williamson, K.A., Mannens, M., Kelsey, A., Gosden, C., Hastie, N.D. and Van Heyningen, V. (1993). "Evidence that WT1 mutations in Denys-Drash syndrome patients may act in a dominant-negative fashion." Human Molecular Genetics **2**: 259-264.

Little, M.H., Williamson, K.A., Mannens, M., Kelsey, A., Gosden, C., Hastie, N.D., Heiningen, V. van. (1993). "Evidence that WT1 mutations in Denys-Drash syndrome patients may act in a dominant-negative fashion." Human Molecular Genetics **2**: 259-264.

Lohman, T., deHaseth, P. L., and Record, M.T. (1980). "Pentalysine-deoxyribonucleic acid interactions: A model for the general effects of ion concentrations on the interactions of proteins with nucleic acids." Biochemistry **19**: 3522-3530.

Lohman, T.M., and Mascotti, D.P. (1992). "Thermodynamics of ligand-nucleic acid interactions." Methods in Enzymology **212**: 400-424.

Lundin, M., Nehlin, J.O., and Ronne, H. (1994). "Importance of a flanking AT-rich region in target site recognition by the GC box-binding zinc finger protein MIG1." Molecular and Cellular Biology **14**: 1979-1985.

Madden, S.L., Cook, D.M., Morris, J.F., Gashler, A., Sukhatme, V.P. and Rauscher III, J.F. (1991). "Transcriptional repression mediated by the WT1 Wilms tumor gene product. " Science **253**: 1550-1553.

Madden, S.L., Cook, D.M. and Rauscher III, F.J. (1993). "A structure-function analysis of transcriptional repression mediated by the WT1, Wilms tumor suppressor protein." Oncogene **8**: 1713-1720.

Maheswaran, S., Park, S., Bernard, A., Morris, J.F., Rauscher III, F.J., Hill, D.E., and Haber, D.A. (1993). "Physical and functional interaction beween WT1 and p53 proteins." Proceedings of the National Academy of Sciences of the United States of America **90**: 5100-5104.

Maheswaran, S., Englert, C., Benett, P., Henrich, G., and Haber, D.A. (1995). "The WT1 gene product stabilizes p53 and inhibits p53-mediated apoptosis." Genes and Development **9**: 2143-2156.

Mao, X. Z. and Darby, M. K. (1993). "A Position-Dependent Transcription-Activating domain in TFIIIA." Molecular Cellular Biology **13** (12): 7496-7506.

Martinez, E., Lagna, G. and Roeder, R. G. (1994). "Overlapping transcription by RNA polymerase II and III of the *Xenopus* TFIIIA gene in somatic cells." The Journal of Biological Chemistry **269**: 25692-25698.

Martinez, J., Georgoff, I., and Levine, A.J. (1991). "Cellular localization and cell cycle regulation by a temperature-sensitive p53 protein." Genes and Development **5**: 151-159.

Matsumoto, Y. and Korn, L. J. (1988). "Upstream sequences required for transcription of the TFIIIA gene in *Xenopus* oocytes." Nucleic Acids Research **16**: 3801-3814.

Maw, M.A., Grundy, P.E., Millow, L.J., Eccles, M.R., Dunn, R.S., Smith, P.J., Feinberg, A.P., Law, D.J., Paterson, M.C., Telzerow, P.E., Callen, D.F., Thompson, A.D., Richards, R.I., and Reeve, A.E. (1992). "A third Wilms' tumor locus on chromosome 16q." Cancer Research **52**: 3094-3098.

McBryant, S.J., Veldhoen, N., Gedulin, B., Leresche, A., Foster, M., Wright, P.E., Romaniuk, P.J. and Gottesfeld, J.M. (1995). "Interaction of the RNA binding fingers of *Xenopus* transcription factor IIIA with specific regions of 5S ribosomal RNA." Journal of Molecular Biology **248**: 44-57.

McBryant, S.J., Gedulin, B., Clemens, K.R., Wright, P., and Gottesfeld, J.M. (1996). "Assessment of major and minor groove DNA interactions by the zinc fingers of *Xenopus* transcription factor IIIA." Nucleic Acids Research **24** (13): 2567-2574.

McConkey, G. A. and Bogenhagen, D. F. (1987). "Transition mutations within the *Xenopus borealis* somatic 5S RNA gene can have independent effects on transcription and TFIIIA binding." Molecular and Cellular Biology **7**: 486-494.

Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T. and Berg, J. M. (1992). "Metal binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide." Proceedings of the National Academy of Sciences of the United States of America **89** (11): 4796-4800.

Michalovitz, D., Halevy, O. and Oren, M. (1990). "Conditional inhibition of transformation and of cell proliferation." Cell **62**: 671-680.

Miki, Y. (1994). "A strong candidate for the breast and ovarian cancer susceptibility gene." Science **266**: 66-71.

Milbrandt, J. (1987). "A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor." Science **238**: 797-799.

Miller, J. R., Cartwright, E. M., Brownlee, G. G., Fedoroff, N. V. and Brown, D. D. (1978). "The nucleotide sequence of oocyte 5S DNA in *Xenopus laevis*. II. The GC-rich region." Cell **13**: 717-725.

Miller, J., McLachlan, A.D. and Klug, A. (1985). "Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes." EMBO (European Molecular Biology Organization) Journal **4**: 1609-1614.

Mitchell, P.J., and Tijian, R. (1989). "Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins." Science **245**: 371-378.

Moffett, P., Bruening, W., Nakagama, H., Bardeesy, N., Housman, D., Housman, D.E., and Pelletier, J. (1995). "Antagonism of WT1 activity by protein self-association." Proceedings of the National Academy of Sciences of the United States of America **92**: 11105-11109.

Molnar, G., Crozat, A., and Pardee, A.B. (1995). "The immediate-early gene Egr-1 regulates the activity of the thymidine kinase promoter at the G0 - to G1 transition of the cell cycle." Molecular and Cellular Biology **14**: 5242-5248.

Morris, J.F., Madden, S.L., Tournay, O.E., Cook, D.M., Sukhatme, V.P., and Rauscher, F.J. III. (1991). "Characterization of the zinc finger protein encoded by the WT1 Wilms tumor locus." Oncogene **6**: 2339-2348.

Muller, H.-J., Skerka, C., Bialonski, A., and Zipfel, P. (1991). "Clone pAT 133 identifies a gene that encodes another human member of a class of growth factor-induced genes with almost identical zinc-finger domains." Proceedings of the National Academy of Sciences of the United States of America **88**: 10079-10083.

Nakagama, H., Heinrich, G., Pelletier, J. and Housman, D.E. (1995). "Sequence and structural requirements for high-affinity DNA binding by the WT1 gene product." Molecular and Cellular Biology **15**: 1489-1498.

Nakaseko, Y., Neuhaus, D., Klug, A. and Rhodes, D. (1992). "Adjacent zinc-finger motifs in multiple zinc-finger peptides from SWI5 form structurally independent, flexibly linked domains." The Journal of Molecular Biology **228** (2): 619-636.

Nardelli, J., Gibson, T. J., Vesque, C. and Charnay, P. (1991). "Base sequence discrimination by zinc-finger DNA-binding domains." Nature (London) **349** (6305): 175-178.

Nardelli, J., Gibson, T., and Charnay, P. (1992). "Zinc finger - DNA recognition: analysis of base specificity by site-directed mutagenesis." Nucleic Acid Research **20**: 4137-4144.

Nehlin, J.O., and Ronne, H. (1990). "Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumor finger proteins." The EMBO (European Molecular Biology Organization) Journal **9**: 2891-2898.

Nekludova, L. & Pabo, C.O. (1994). "Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes." Proceedings of the National Academy of Sciences of the United States of America **91**: 6948-6952.

Nelson, R.M. and Long, G.L. (1989). "A general method of site-specific mutagenesis using a modification of the *Thermus aquaticus* polymerase chain reaction." Analitical Biochemistry **180**: 147-151.

Neuhaus, D., Nakaseko, Y., Schwabe, J. W. R. and Klug, A. (1992). "Solution structures of two Zinc-Finger domains from SWI5 obtained using 2-Dimensional H-1 nuclear magnetic resonance spectroscopy - a Zinc Finger structure with a 3rd strand of b-Sheet." The Journal of Molecular Biology **228** (2): 637-651.

Nguen, H.Q., Hoffman-Liebermann, B., Lieebermann, D.A. (1993). "The zinc finger transcription factor EGR-1 is essential for and restricts differentiation along the macrophage lineage." Cell **72**: 197-209.

Nordenskjold, A., Friedman, E. and Anvret, M. (1994). "WT1 mutations in patients with Denys-Drash syndrome: a novel mutation in exon 8 and paternal allele origin." Human Genetics **93**: 115-120.

O'Halloran, T. (1993). "Transition metals in control of gene expression." Science **261**: 715-725.

Ogawa, O., Eccles, M.R., Szeto, J., McNoe, L.A., Yun, K., Maw, M.A., Smith, P.J., and Reeve, A.E. (1993a). "Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour." Nature (London) **362**: 749-751.

Ogawa, O., Eccles, M.R., Yun, K., Mueller, R.F., Holdaway, M.D. and Reeve A.E. (1993b). "A novel insertional mutation at the third zinc finger coding region of the WT1 gene in Denys-Drash syndrome." Human Molecular Genetics **2**: 203-204.

Omichinski, J. G., Clore, G. M., Appella, E., Sakaguchi, K. and Gronenborn, A. M. (1990). "High-resolution 3-dimensional structure of a single zinc finger from a human enhancer binding protein in solution." Biochemistry **29** (40): 9324-9334.

Omichinski, J. G., Clore, G. M., Robien, M., Sakaguchi, K., Appella, E. and Gronenborn, A. M. (1992). "High-Resolution Solution Structure of the Double Cys2His2 Zinc Finger from the Human Enhancer Binding Protein MBP-1." Biochemistry **31** (16): 3907-3917.

Pabo, C.O., and Sauer, R.T. (1992). "Transcription factors: structural families and principles of DNA recognition." Annual Review of Biochemistry **61**: 1053-1095.

Palmer, A. G., Rance, M. and Wright, P. E. (1991). "Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance C-12 heteronuclear NMR spectroscopy." Journal of the American Chemical Society **113** (12): 4371-4380.

Park, S., Tomlinson, G., Nisen, P. and Haber, D.A. (1993). "Altered trans-activational properties of a mutated WT1 gene product in a WAGR-associated Wilms' tumor." Cancer Research **53**: 4757-4760.

Párraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T. and Klevit, R. E. (1988). "Zinc-dependent structure of a single finger domain of yeast ADR1." Science **241**: 1489-1492.

Patwardhan, S., Gashler, S., Siegel, M., Chang, L., Joseph, L., Shows, T., Le Beau, M., and Sukhatme, V. (1991). "EGR-3, a novel member of the Egr family of genes encoding immediate-early transcription factors." Oncogene **6**: 917-927.

Pavletich, N. P. and Pabo, C. O. (1991). "Zinc Finger-DNA Recognition - Crystal Structure of a Zif268-DNA Complex at 2.1Å." Science **252** (5007): 809-817.

Pavletich, N. P. and Pabo, C. O. (1993). "Crystal structure of a 5-Finger GLI-DNA complex - new perspectives on zinc fingers." Science **261** (5129): 1701-1707.

Pelham, H. R. B. and Brown, D. D. (1980). "A specific transcription factor that can bind either the 5S RNA gene or 5S RNA." Proceeding of the National Academy of Sciences of the United States of America **77**: 4170-4174.

Pelletier, J., Bruening, W., Kashtan, C., Mauer, M., Manivel, J.C., Striegel, J.E., Houghton, D.C., Junien, C., Habib, R., Fouser, L., Fine, R.N., Silverman, B.L., Haber, D.A., and Housman, D. (1991a). "Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denis-Drash syndrome." Cell **67**: 437-447.

Pelletier, J., Schalling. M., Buckler, A.J., Rogers, A., Haber, D.A. and Housman, D. (1991b). "Expression of the Wilms tumor gene WT1 in the murine urogenital system." Genes and Development **5**: 1345-1356.

Pelletier, J., Bruening, W., Li, F.P., Haber, D.A., Glaser, T., and Housman, D.E. (1991c). "WT1 mutations contribute to abnormal genital system development and hereditary Wilms' tumour." Nature (London) **353**: 431-434.

Peterson, R. C., Doering, J. L. and Brown, D. D. (1980). "Characterization of two *Xenopus* somatic 5S DNAs and one minor oocyte-specific 5S DNA." Cell **20**: 131-141.

Pfaff, S. L., Hall, R. K., Hart, G. C. and Taylor, W. L. (1991). "Regulation of the *Xenopus laevis* Transcription Factor IIIA Gene During Oogenesis and Early Embryogenesis - Negative Elements Repress the O-TFIIIA Promoter in Embryonic Cells." Developmental Biology **145** (2): 241-254.

Pfaff, S. L. and Taylor, W. L. (1992). "Characterization of a *Xenopus* Oocyte Factor That Binds to a Developmentally Regulated cis-Element in the TFIIIA Gene." Developmental Biology **151** (1): 306-316.

Picard, B. and Wegnez, M. (1979). "Isolation of a 7S particle from *Xenopus laevis* oocytes: a 5S RNA-protein complex." Proceeding of the National Academy of Sciences of the United States of America **76**: 241-245.

Pieler, T., Erdmann, V. A. and Appel, B. (1984). "Structural requirements for the interaction of 5S rRNA with the eukaryotic transcription factor IIIA." Nucleic Acids Research **12**: 8393-8406.

Pieler, T., Appel, B., Oei, S. L., Mentzel, H. and Erdmann, V. A. (1985a). "Point mutational analysis of the *Xenopus laevis* 5S gene promoter." EMBO (European Molecular Biology Organization) Journal **4** (7): 1847-1853.

Pieler, T., Oei, S.-L., Hamm, J., Engelke, U. and Erdmann, V. A. (1985b). "Functional domains of the *Xenopus laevis* 5S gene promoter." EMBO (European Molecular Biology Organization) Journal **4**: 3751-3756.

Pieler, T., Hamm, J. and Roeder, R. G. (1987). "The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing." Cell **48**: 91-100.

Pieler, T. and Theunissen, O. (1993). "TFIIIA: nine fingers - three hands?" TIBS (Trends in Biochemical Sciences) **18**: 226-230.

Quek, H.H., Chow, V.T., and Tock, E.P. (1993). "The third zinc finger of the WT1 gene is mutated in Wilms' tumor but not in a broad range of other urogenital tumors." Anticancer research **13**: 1575-1580.

Quian, X., Gozan, S.N., Yoon, H., Jeon, C., Agarwal, K., Weiss M.A. (1993a) "Novel zinc finger motif in the basal transcription machinery: Three-dimentional NMR studies of the nucleic acid binding domain of transcription elongation factor TFIIS." Biochemistry **32**: 9944-9959.

Quian, X., Jeon, C., Yoon, H., Agarwal, K., Weiss, M.A. (1993b) "Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS." Nature (London) **365**: 277-279.

Rackley, R.R., Flenniken, A.M., Kuriyan, N.P., Kessler, P.M., Stoler, M.H., and Williams, B.R. (1993). "Expression of the Wilms' tumor suppressor gene WT1 during mouse embryogenesis." Cell Growth and Differentiation **4**: 1023-1031.

Rainier, S., Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E., and Feinberg, A.P. (1993). "Relaxation of imprinted genes in human cancer." Nature (London) **362**: 747-749.

Rauscher, F.J., III, Morris, J.F., Tournay, O.E., Cook, D.M., and Curran, T. (1990). "Binding of the Wilms tumor locus zinc finger protein to the EGR-1 consensus sequence." Science **250**: 1259-1262.

Rauscher, F.J., III. (1993). "The WT1 Wilms tumor gene product: a developmentally regulated transcription factor in the kidney that functions as

a tumor suppressor." The FASEB (Federation of American Societies for Experimental Biology) Journal 7: 896-903.

Rawlings, S.L., Matt, G.D., and Huber, P.W. (1996). "Analysis of the binding of Xenopus transcription factor IIIA to oocyte 5S rRNA and to the 5S rRNA gene." The Journal of Biological Chemistry **271**: 869-877.

Read, D., and J.L. Manley. (1992). "Alternatively spliced transcripts of the Droosophila tramtrack gene encode zinc finger proteins with distinct DNA binding specificities." EMBO (European Molecular Biology Organization) Journal **11**: 1035-1044.

Rebar, E. J. and Pabo, C. O. (1994). "Zinc finger phage - affinity selection of fingers with new DNA-Binding specificities." Science **263** (5147): 671-673.

Record, M.T., Lohman, M.L., and Haseth, P. De (1976). "Ion effects on ligand-nucleic acid interactions." Journal of Molecular Biology **107**: 145-158.

Reddy, J.C.,Morris, J.C., Wang, J., English, M.A., Haber, D.A., Shi, Y., and Licht, J. (1995). "WT1-mediated transcriptional activation is inhibited by dominant negative mutant proteins." The Journal of Biological Chemistry **270**: 10878-10884.

Reddy, J.C., and Licht, J.D. (1996). "The WT1 Wilms' tumor suppressor gene: How much do we really know?" Biochimica et Biophysica Acta **1287**: 1-28.

Reeve, A.E., Eccles, M.R., Wilkins, R.J., Bell, G.I., and Millow, L.J. (1985). "Expression of insulin-like growth factor-II transcripts in Wilms tumour." Nature (London) **317**: 258-260.

Reynolds, W. F. and Gottesfeld, J. M. (1983). "5S rRNA gene transcription factor IIIA alters the helical configuration of DNA." Proceeding of the National Academy of Sciences of the United States of America **80**: 1862-1866.

Rhodes, D. and Klug, A. (1986). "An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA." Cell **46**: 123-132.

Rhodes, D. and Klug, A. (1993). "Zinc Fingers." Scientific American **February**: 56-65.

Riccardi, V.M., Sujansky, E., Smith, A.C., and Francke, U. (1978). "Chromosomal imbalance in the aniridia-Wilms' tumor association: 11p interstitial deletion." Pediatrics **61**: 604-610.

Rollins, M. B., Del Rio, S., Galey, A. L., Setzer, D. R. and Andrews, M. T. (1993). "Role of TFIIIA zinc fingers in vivo - analysis of Single-Finger function in developing *Xenopus* embryos." Molecular and Cellular Biology **13** (8): 4776-4783.

Romaniuk, P. J. (1985). "Characterization of the RNA binding properties of transcription factor IIIA of *Xenopus laevis* oocytes." Nucleic Acids Research **13** (14): 5369-5387.

Romaniuk, P. J. (1989). "The role of highly conserved single-stranded nucleotides of *Xenopus* 5S RNA in the binding of transcription factor IIIA." Biochemistry **28**: 1388-1395.

Romaniuk, P.J. (1990). "Characterization of equilibrium binding of *Xenopus* transcription factor IIIA to the 5S RNA gene." The Journal of Biological Chemistry **265**: 17593-17600.

Rowland, O., and Segall, J. (1996). "Interactionof wild-type and truncated forms of transcription factor IIIA from *Saccharomyces cerevisiae* with the 5S RNA gene." The Journal of Biological Chemistry **271**: 12103-12110.

Rupprecht, H.D., Drummond, I.A., Madden, S.L., Rauscher, F.J.III, and Sukhatme, V.P. (1994). "The Wilms' tumor suppressor gene WT1 is negatively autoregulated." The Journal of Biological Chemistry **269**: 6198-6206.

Russo, M.W., Mathny, C., and Milbrandt, J. (1993). "Transcriptional activity of the zinc finger protein NGFI-A is influenced by its interaction with cellular factor." Molecular and Cellular Biology **13**: 6858-6865.

Ryan, G., Steele-Perkins, V., Morris, J.F., Rauscher III, F.J. and Dressler, G.R. (1995). "Repression of Pax-2 by WT1 during normal kidney development." Development **121**: 867-875.

Sakai, A., Tadokoro, K., Yanagisawa, H., Nagafuchi, S., Hoshikawa, N., Suzuki, T., Kohsaka, T., Hasegawa, T., Nakahori, Y. and Yamada, M. (1993). "A novel mutation of the WT1 gene (a tumor suppressor gene for Wilms' tumor) in a patient with Denys-Drash syndrome." Human Molecular Genetics **2**: 1969-1970.

Sakonju, S., Bogenhagen, D. F. and Brown, D. D. (1980). "A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region." Cell **19**: 13-25.

Sakonju, S., Brown, D. D., Engelke, D., Ng, S.-Y., Shastry, B. S. and Roeder, R. G. (1981). "The binding of a transcription factor to deletion mutants of a 5S rRNA gene." Cell **23**: 665-669.

Sakonju, S. and Brown, D. D. (1982). "Contact points between a positive transcription factor and the *Xenopus* 5S RNA gene." Cell **31**: 395-405.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Second edition). Cold Spring Harbor, Cold Spring Harbor Laboratory Press.

Sands, M. S. and Bogenhagen, D. F. (1987). "TFIIIA binds to different domains of 5S RNA and the *Xenopus borealis* 5S RNA gene." Molecular and Cellular Biology **7**: 3985-3993.

Sawadogo, M. and Roeder, R. G. (1985). "Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region." Cell **43**: 439-448.

Schmiedeskamp, M., and Klevit, R. (1994). "Zinc finger diversity." Current Opinion in Structural Biology **4**: 28-35.

Schroth, G. P., Cook, G. R., Bradbury, E. M. and Gottesfeld, J. M. (1989). "Transcription Factor IIIA Induced Bending of the *Xenopus* Somatic 5S Gene Promoter." Nature (London) **340** (6233): 487-488.

Scott, J., Cowell, J., Robertson, M.E., Priestley, L.M., Wadey, R., Hopkins, B., Pritchard, J., Bell, G.I., Rall, L.B., Graham, C.F., and Knott, T.J. (1985) "Insulin-like growth factor-II gene expression in Wilms tumours and emryonic tissues." Nature (London) **317**: 260-262.

Scotto, K. W., Kaulen, H. and Roeder, R. G. (1989). "Positive and negative regulation of the gene for transcription factor IIIA in *Xenopus laevis* oocytes." Genes & Development **3**: 651-662.

Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976). "Sequence-specific recognition of double helical nucleic acids by proteins." Proceedings of the National Academy of Sciences of the United States of America **73**: 804-808.

Segall, J., Matsui, T. and Roeder, R. G. (1980). "Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III." The Journal of Biological Chemistry **255** (24): 11986-11991.

Setzer, D. R. and Brown, D. D. (1985). "Formation and stability of the 5S RNA transcription complex." The Journal of Biological Chemistry **260** (4): 2483-2492.

Setzer, D.R., Menezes, S.R., Del Rio, S., Hung, V.S., and Subramanya, G. (1996). "Functional interactions between the zinc fingers of *Xenopus* transcription factor IIIA during 5S rRNA binding." RNA **2**: 1254-1269.

Sharma, P.M., Bowman, M., Yu, B.F., and Sukumar, S. (1994) "A rodent model for Wilms' tumor: embryonal kidney neoplasms induced by N-nitroso-N'-methyurea." Proceedings of the National Academy of Sciences of the United States of America **91**: 9931-9935.

Shastry, B. S., Ng, S.-Y. and Roeder, R. G. (1982). "Multiple factors involved in the transcription of class III genes in *Xenopus laevis*." The Journal of Biological Chemistry **257**: 12979-12986.

Shastry, B. S., Honda, B. M. and Roeder, R. G. (1984). "Altered levels of a 5S gene-specific transcription factor (TFIIIA) during oogenesis and embryonic development of *Xenopus laevis*." The Journal of Biological Chemistry **259**: 11373-11382.

Shi, Y., and Berg, J. (1995). "Specific DNA-RNA hybrid binding by zinc finger proteins." Science **268**: 282-284.

Silver, P. A., Keegan, L. P., and Ptashne, M. (1984). "Amino-terminus of the yeast GAL4 gene product is sufficient for nuclear localization." Proceedings of the National Academy of Sciences of the United States of America **81**: 5951-5955.

Simmons, D. L., Levy, D. B., Yannoni, Y., and Erikson, R. L. (1989). "Identification of a phorbol ester-repressible *v-src*-inducible gene." Proceedings of the National Academy of Sciences of the United States of America **86**: 1178-1182.

Smith, D. R., Jackson, I. J. and Brown, D. D. (1984). "Domains of the positive transcription factor specific for the *Xenopus* 5S RNA gene." Cell **37**: 645-652.

Somers, W., Ultsch, M., De-Vos, A., Kossiakoff, A.A. (1994). "The X-ray structure of a growth hormone-prolactin receptor complex." Nature (London) **372:** 478-481

Stanbridge, E.J. (1990). "Human tumor suppressor genes." Annual Review of Genetics **24:** 615-657.

Steitz, T.A. (1990). "Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding." Quaternal Reviews of Biophysics **23 (3):** 205-280.

Steitz, T.A. (1993). "Structural studies of protein-nucleic acid interaction." Cambridge University Press, Cambridge.

Studier, W.F. and Moffat, B.A. (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." Journal of Molecular Biology **189:** 113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990). "Gene Expression Technology. Use of T7 RNA polymerase to direct expression of cloned genes." Methods of Enzymology **185:** 60-89.

Suggs, S., Katzowitz, J., Tsai-Morris, C., and Sukhatme, V. (1990). "cDNA sequence of the human early growth response gene Egr-1." Nucleic Acids Research **16:** 8835-8846.

Sukhatme, V.P., Kartha, S., Toback, F.G., Taub, R., Hoover, R.G., and Tsai-Morris, C.-H. (1987). "A novel earlly growth response gene rapidly induced by fibroblast, epithelial cell and lymphocyte mitogens." Oncogene Research **1:** 343-355.

Sukhatme, V.P., Cao, X., Chang, L.C., Tsai-Morris, C.H., Stamenkovich, D., Ferreira, P.C.P., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T., Le Beau, M.M., and Adamson, E.D. (1988). "A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization." Cell **53:** 37-43.

Suzuki, M. (1993). "Common features in DNA recognition helices of eukaryotic transcription factors." EMBO (European Molecular Biology Organization) Journal **12:** 3221-3226.

Suzuki, M., Brenner, S.E., Gerstein, M., and Yagi, N. (1995). "DNA recognition code of transcription factors." Protein engineering **8:**319-328.

Swirnoff, A.H., and Milbrandt, J. (1995). "DNA-binding specificity of NGFI-A and related zinc finger transcription factors." Molecular and Cellular Biology **15**: 2275-2287.

Taylor, W., Jackson, I. J., Siegel, N., Kumar, A. and Brown, D. D. (1986). "The developmental expression of the gene for TFIIIA in *Xenopus laevis*." Nucleic Acids Research **14**: 6185-6195.

TenHarmsel, A., Auston, R.J., Savenelli, N and Biggin, M.D. (1993). "Cooperative binding at a distance by even-skipped protein correlates with repression and suggests as mechanism of silencing." Molecular and Cellular Biology **13**: 2742-2752.

Theunissen, O., Rudt, F., Guddat, U., Mentzel, H. and Pieler, T. (1992). "RNA and DNA binding zinc fingers in *Xenopus* TFIIIA." Cell **71** (4): 679-690.

Thompson, C.C. and Evans, R.M. (1989). Proceedings of the National Academy of Sciences of the United States of America **86**: 3494-3498.

Thukral, S.K., Morrison, M.L., and Young, E.T. (1991). "Alanine scanning site-directed mutagenesis of the zinc fingers of transcription factor ADR1: residues that contact DNA and that transactivate." Proceedings of the National Academy of Sciences of the United States of America **88**: 9188-9192.

Thukral, S.K., Morrison, M.L., and Young, E.T. (1992). "Mutations in the zinc fingers of ADR1 that change the specificity of DNA binding and transactivation." Molecular and Cellular Biology **12**: 2784-2792.

Ton, C.C., Huff, V., Call, K.M., Cohn, S., Strong, L.C., Housman, D.E., and Saunders, G.F. (1991). "Smallest region of overlap in Wilms tumor deletions uniquely implicates an 11p13 zinc finger gene as the disease locus." Genomics **10**: 293-297.

Tso, J. Y., Van Den Berg, D. J. and Korn, L. J. (1986). "Structure of the gene for *Xenopus* transcription factor TFIIIA." Nucleic Acids Research **14**: 2187-2200.

Varanasi, R., Bardeesy, N., Ghahremani, M., Petruzzi, M.J., Nowak, N., Adam, M.A., Grundy, P., Shows, T.B., and Pelletier, J. (1994). "Fine structure analysis of the WT1 gene in sporadic Wilms' tumors." Proceedings of the National Academy of Sciences of the United States of America **91**: 3554-3558.

Veldhoen, N., You, Q. M., Setzer, D. R. and Romaniuk, P. J. (1994). "Contribution of individual base pairs to the interaction of TFIIIA with the *Xenopus* 5S RNA gene." Biochemistry **33** (24): 7568-7575.

von Hippel, P.H., and Schleich, T. (1969). in Structure and Stability of Biological Macromolecules, pp. 417-574, Marcel Dekker Inc., New York.

Vrana, K. E., Churchill, M. E. A., Tullius, T. D. and Brown, D. D. (1988). "Mapping functional regions of transcription factor TFIIIA." Molecular and Cellular Biology **8**: 1684-1696.

Wang, Z.-Y., and Deuel, T. (1992a). "An S1 nuclease-sensitive homopurine-homopyrimidine domain in the PDGF A-chain promoter contains a novel binding site for the growth factor-inducible protein EGR-1." Biochemical and Biophysical Research Communications **188**: 433-439.

Wang, Z.-Y., Madden, S.L., Deuel, T.F., and Rauscher III, F.J. (1992b). "The human platelet-derived growth factor A-chain (PDGF-A) gene is a target for repression by the WT1 Wilms tumor protein." The Journal of Biological Chemistry **267**: 21999-22002.

Wang, Z.-Y., Qiu, Q.-Q., Enger, K.T., and Deuel, T.F. (1993a). "A second transcriptionally active DNA-binding site for the Wilms tumor gene product, WT1." Proceedings of the National Academy of Sciences of the United States of America **90**: 8896-8900.

Wang, Z.-Y., Qiu, Q.Q., and Deuel, T.F. (1993b). "The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains." The Journal of Biological Chemistry **268**: 9172-9175.

Wang, Z.-Y., Qiu, Q.Q., Huang, J., Gurrieri, M., and Deuel, T.F. (1995). "Products of alternatively spliced transcripts of the Wilms' tumor suppressor gene, wt1, have altered DNA binding specificity and regulate transcription in different ways." Oncogene **10**: 415-422.

Watson, M.A. and Milbrandt, J. (1990). "Expression of the nerve growth factor - regulated NGFI-A and NGFI-B genes in the developing rat." Development **110**: 173-183.

Weissman, B.E., Saxon, P.J., Pasquale, S.R., Jones, G.R., Geiser, A.G., and Stanbridge, E.J. (1987). "Introduction of a normal human chromosome 11 into a Wilms tumor cell line controls its tumorigenic expression." Science **236**: 175-180.

Werner, H., Re, G.G., Drummond, I.A., Sukhatme, V.P., Rauscher III, F.J., Sens, D.A., Garvin, A.J., LeRoith, D., D., and Roberts, C.T., Jr. (1993). "Increased expression of the insulin-like growth factor I receptor (IGF-I-R) gene in Wilms tumor is correlated with modulation of IGF-I-R promoter activity by the WT1 Wilms' tumor gene product." Proceedings of the National Academy of Sciences of the United States of America **90**: 5828-5832.

Werner, H., Rauscher III, F.J., Sukhatme, V.P., Drummond, I.A., Roberts Jr., C.T., and LeRoith, D. (1994). "Transcriptional repression of the insulin-like frowth factor I receptor (IGF-I-R) gene by the tumor suppressor WT1 involves binding to sequences both upstream and downstream of the IGF-I-R gene transcription start site." The Journal of Biological Chemistry **269**: 12577-12582.

Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J.-P., Ehresmann, C. and Ehresmann, B. (1989). "Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5S rRNAs." Journal of Molecular Biology **207**: 417-431.

Wilson, T.E., Day, M.L., Pexton, T., Padgett, K.A., Johnson, M., and Milbrandt, J. (1992). "*In vivo* mutational analysis of the NGFI-A zinc fingers." The Journal of Biological Chemistry **267**: 3718-3724.

Wimberly, B., Varani, G. and Tinoco, I. (1993). "The conformation of Loop-E of eukaryotic 5S Ribosomal RNA." Biochemistry **32** (4): 1078-1087.

Witman, K.G. (1993). "The retinoblastoma gene: role in cell cycle control and cell differentiation." FASEB (Federation of American Societies for Experimental Biology) Journal **7**: 834-845.

Wolters, J. and Erdmann, V. A. (1988). "Compilation of 5S rRNA and 5S rRNA gene sequences." Nucleic Acids Research **16** (supplement): r1-r9.

Wormington, W. M. (1989). "Developmental Expression and 5S rRNA-Binding Activity of *Xenopus laevis* Ribosomal Protein L5." Molecular and Cellular Biology **9** (12): 5281-5288.

Wright, J., Gunter, K., Mitsuya, H., Irving, S., Kelly, K., and Siebenlist, U. (1990). "Expression of a zinc finger gene in HTLV-1 and HTLV-II-transformed cells." Science **248**: 588-591.

Wuttke, D.S., Foster, M.P., Case, D.A., Gottesfeld, J.M. and Wright, P. (1997). "Solution structure of the first three zinc fingers of TFIIIA bound to the

cognate DNA sequence: determinants of affinity and sequence specificity." Journal of Molecular Biology **273 (1)**: 183-206.

Xing, Y. Y. and Worcel, A. (1989). "The C-terminal domain of transcription factor IIIA interacts differently with different 5S RNA genes." Molecular and Cellular Biology **9**: 499-514.

Ye, Y., Raychaudhuri, B., Gurney, A., Campbell, C.E. and Williams, B.R.G. (1996). "Regulation of WT1 by phosphorylation:inhibition of DNA binding, alteration of transcriptional activity and cellular translocation." EMBO (European Molecular Biology Organization) Journal 15: 5606-5615.

You, Q., Veldhoen, N., Baudin, F. and Romaniuk, P. J. (1991). "Mutations in 5S DNA and 5S RNA Have Different Effects on the Binding of *Xenopus* Transcription Factor IIIA." Biochemistry **30** (9): 2495-2500.

Zambetti, G.P., and Levine, A.J. (1993). "A comparison of the biological activities of wild-type and mutant p53." FASEB (Federation of American Societies for Experimental Biology) Journal 7: 855-865

Zang, W-Q., Veldhoen, N. and Romaniuk, P.J. (1995). "The effect of substitution mutations of transcription factor IIIA on nucleic acid binding activity." Biochemistry **34**: 15545-15552.

Zhang, Y., and Tycko, B. (1992). "Monoallelic expression of the human H19 gene." Nature Genetics **1**: 40-44.