Nucleic Acids Research, 1995, Vol. 23, No. 2 277–284

High affinity binding sites for the Wilms' tumour suppressor protein WT1

Tatyana B. Hamilton, Kathleen C. Barilla and Paul J. Romaniuk*

Department of Biochemistry and Microbiology, University of Victoria, PO Box 3055, Victoria, BC V8W 3P6, Canada

Received September 8, 1994; Revised and Accepted December 13, 1994

ABSTRACT

The Wilms' tumour suppressor protein (WT1) is a putative transcriptional regulatory protein with four zinc fingers, the last three of which have extensive sequence homology to the early growth response-1 (EGR-1) protein. Although a peptide encoding the zinc finger domain of WT1[-KTS] can bind to a consensus 9 bp EGR-1 binding site, current knowledge about the mechanisms of zinc finger–DNA interactions would predict a more extended recognition site for WT1. Using a WT1[-KTS] zinc finger peptide (WT1-ZFP) and the template oligonucleotide GCG-TGG-GCG-NNNNN in a binding site selection assay, we have determined that the highest affinity binding sites for WT1[-KTS] consist of a 12 bp sequence GCG-TGG-GCG-(T/G)(G/A/T)(T/G). The binding of WT1-ZFP to a number of the selected sequences was measured by a quantitative nitrocellulose filter binding assay, and the results demonstrated that these sequences have a 4-fold higher affinity for the protein than the nonselected sequence GCG-TGG-GCG-CCC. The full length WT1 protein regulates transcription of reporter genes linked to these high affinity sequences. A peptide lacking the first zinc finger of WT1[-KTS], but containing the three zinc fingers homologous to EGR-1 failed to select any specific sequences downstream of the GCG-TGG-GCG consensus sequence in the binding site selection 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.

INTRODUCTION

Wilms' tumor, or nephroblastoma, is a renal malignancy that is associated with aniridia, mental retardation and urogenital malfunction (the WAGR syndrome), affecting approximately 1 in 10 000 children under 16 years of age [for reviews see (1,2)]. A loss of heterozygosity can be detected at 11p13 in Wilms' tumor tissues but not normal tissues from the same individual (3–8). The smallest 11p13 deletion overlap found in affected individuals was resolved into a telomeric locus for aniridia (*AN2*) and a centromeric locus containing the putative Wilms' tumor gene *WT1* (9).

WT1 encodes a protein that has several features characteristic of transcription factors: it has a negatively charged amino terminus which contains a high concentration of proline and glutamine residues, and a putative DNA binding domain located at the carboxyl terminus containing four contiguous zinc finger motifs (9,10). The last three zinc fingers share a high degree of amino acid sequence homology with the three zinc fingers of the early growth response proteins EGR-1 and EGR-2, which are involved in regulating cell proliferation (11,12). Four cDNAs, encoding isoforms of the WT1 protein, arise from differential splicing of the primary transcript, and have been found in various amounts in certain tissues. Alternative splicing can result in the in-frame insertion of 17 amino acids into the P/Q domain (9), or an in-frame insertion of three amino acids between the third and fourth fingers of the DNA binding domain (13), or both insertions. Recent studies have demonstrated that the form of the WT1 zinc finger polypeptide lacking both of these insertions (WT1[-KTS]) can bind to DNA sequences similar to a consensus site for the EGR-1 protein, and can regulate promoters containing such sequences (14,15).

EGR-1 (also called NGFI-A, Zif268 and Krox-24) binds to the 9 bp sequence GCG(G/T)GGGCG (16,17), a sequence motif present in the promoter regions of *EGR-1* itself and other immediate-early genes (17). Other promoters which contain EGR-1 regulatory sites include platelet-derived growth factor A chain (18), insulin-like growth factor II (IGF-II) (19), adenosine deaminase (20) and α-myosin heavy chain (21).

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 (22) 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 9 bp EGR-1 recognition sequence (Fig. 1).

* To whom correspondence should be addressed

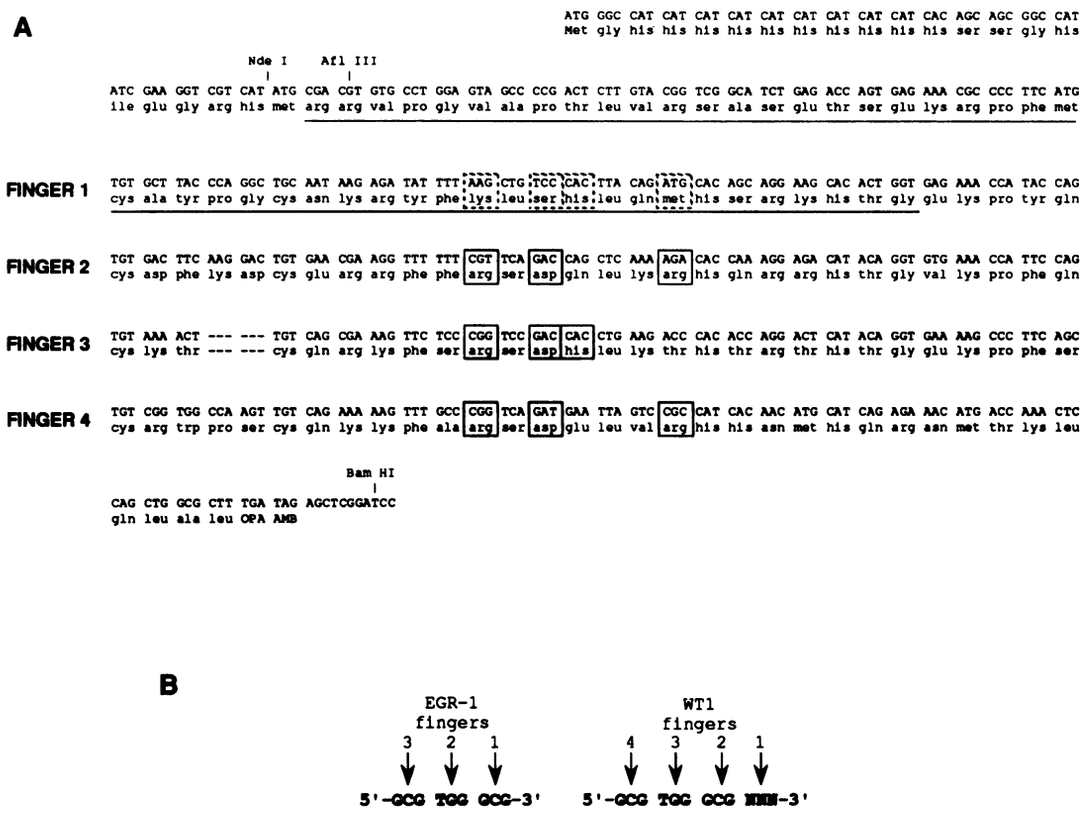

Figure 1. (**A**) DNA sequence of the peptide encoding insert in plasmid pET-WT1ZFP, with the amino acid sequence of the peptide. The region of sequence underlined is deleted in plasmid pET-WT1ΔF1, the peptide product lacking the first zinc finger of the WT1 protein. Amino acids in fingers 2–4 that are 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. Amino 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 of 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.

The crystallographic data indicates that each of the three fingers of EGR-1 interacts with a specific 3 bp subsite within the 9 bp consensus binding site (Fig. 1B). The yeast MIG1 repressor protein has two zinc fingers corresponding in sequence homology to fingers 2 and 3 of EGR-1. This protein recognizes a 6 bp G rich sequence (23). This trend predicts that WT1[-KTS] will recognize a specific 12 bp sequence, the first 9 bp of which will correspond to the EGR-1 consensus sequence. The effects on WT1-ZFP binding of mutating sequences within and flanking an EGR-1 consensus site suggested that this zinc finger peptide derived from WT1[-KTS] recognizes a longer sequence (15,24).

To determine the role of the first zinc finger of WT1[-KTS] in DNA binding, we have used an *in vitro* binding site selection assay (SAAB) (25). 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 3–4 bp immediately downstream of the EGR-1 consensus site (Fig. 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 4-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.

MATERIALS AND METHODS

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 *Afl*III site to the termination codon (Fig. 1A). PCR products were initially cloned into pUC18. The desired clones were identified by DNA sequencing, 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 (Fig. 1A). This PCR product was cloned into pUC18 and the correct clone identified by DNA sequencing. Once the sequences were verified, the zinc finger cassettes were cloned into the *NdeI*/*Bam*HI restriction sites of the T7 expression vector pET-16b (26) 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 (Fig. 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 h later by centrifugation at 7500 r.p.m. in a Beckman JA-14 rotor. The cell pellets were washed with buffer A (10 mM Tris–HCl, 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 (8×15 s pulses on ice with 1.5 min intervals between pulses). Cell debris was pelleted by centrifugation in 15 ml corex tubes at 16 000 r.p.m. for 15 min 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 of buffer B (buffer A containing 5 M 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 (27). Protein yield averaged 10 mg per litre of bacterial culture. On SDS–PAGE, the purified WT1-ZFP migrated with an apparent molecular weight of 22 000 kilodaltons (kDa), and the WT1ΔF1-ZFP with an apparent molecular weight of 17 000 kDa.

Selection amplification and binding (SAAB) assay

The SAAB assay was used according to the published procedure (25) with several minor modifications. A 69 bp oligonucleotide was synthesized that included the target sequence -GCGTGGGCGNNNNN- flanked by *Bam*HI and *Eco*RI restriction sites and by binding sites for M13 universal forward and reverse sequencing primers. Primer extension was used to convert the template oligonucleotide into a mixture of 1024 double stranded DNA sequences. 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.

Labeled DNA (10–20 000 c.p.m.) was incubated with either WT1-ZFP or WT1ΔF1-ZFP for 20 min 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 μg/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. 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 0.5 M ammonium acetate, 1 mM EDTA at 37°C. Labeled DNA was then prepared for the next selection round using the eluted DNA as a template in a PCR reaction (25). 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 *Bam*HI and *Eco*RI for cloning into pUC19. After transformation of *E.coli* JM109, 36–50 clones were randomly selected, and plasmid DNA prepared. The base sequences of the *Eco*RI/*Bam*HI inserts in these constructs were determined using an ABI 373A automated DNA sequencer.

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 (28).

Transient transfection assays

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 (29), 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 (30), was a kind gift from Dr R. Evans. The CAT reporter plasmids were constructed by cloning oligos containing either one or three repeats of the indicated WT1 binding motif into the unique *Hin*dIII 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 calf 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 (31). Cell extracts were prepared in 1× Reporter Lysis Buffer (Promega) 48 h after transfection. β-galactosidase activity was measured (32) 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) (33).

RESULTS

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

The selection/amplification assay (25) 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 (22), that fingers 2–4 of WT1 would interact with the EGR consensus sequence

```
                  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 2. 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 1.

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 sequences of 50 clones chosen from the final pool of DNAs selected by WT1-ZFP are shown in Figure 2. The percentage frequencies of the four bases at each of the five randomized positions are shown in Table 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 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 3 bp sequence was selected for the WT1-ZFP finger 1 binding subsite, while the nucleotides at positions 13 and 14 remained random. This finding is consistent with the paradigm of a 3 bp subsite recognition observed for a number of EGR family proteins.

Table 1. Frequencies of nucleotides selected by WT1-ZFP and WT1ΔF1-ZFP

Position:	10	11	12	13	14
WT1-ZFP[a]					
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]					
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.

Quantitative binding of WT1-ZFP to various DNA sequences

A nitrocellulose filter binding assay (28) 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 3A 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 $8.40\pm1.21\times10^8$ M^{-1}, while the nonconsensus CCC element has an apparent K_a of $2.10\pm0.41\times10^8$ M^{-1}. In comparison, the TRE sequence binds WT1-ZFP with approximately two orders of magnitude less affinity (K_a estimated to be ca. 3×10^6 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 2). In comparison, the WT1 element containing the CCC finger 1 subsite significantly reduced the binding of the WT1-ZFP, decreasing the affinity 4-fold compared to the selected, high affinity binding sites. The affinities of the TGT element ($K_a = 1.67\pm0.05\times10^8$ M^{-1}) and the CCC element ($K_a = 0.92\pm0.09\times10^8$ M^{-1}) for the mutant peptide WT1ΔF1-ZFP, which lacks the first zinc finger of WT1, are almost equivalent (Fig. 3B). 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

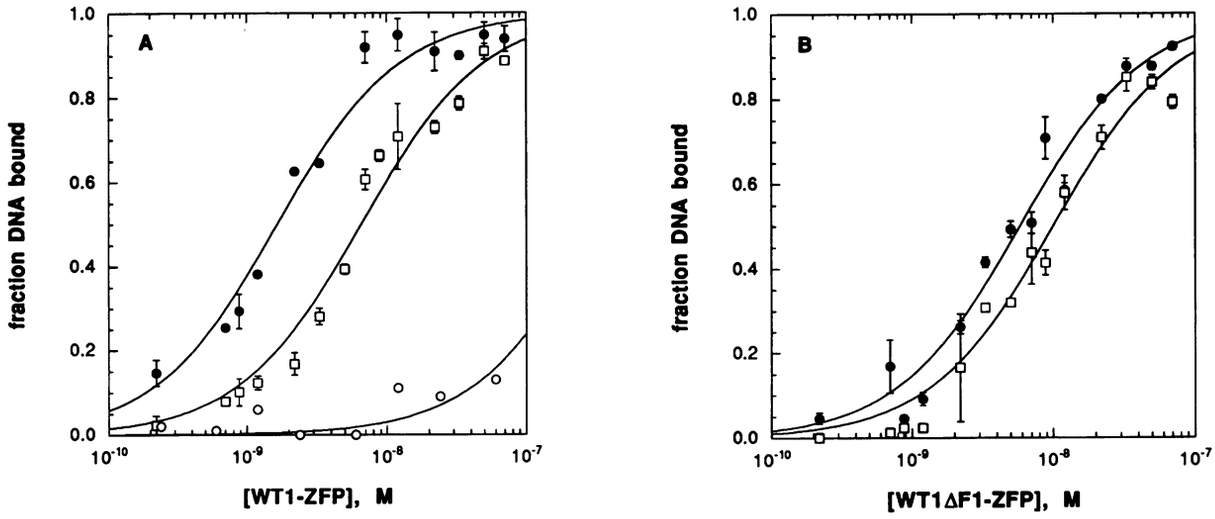

Figure 3. (**A**) Results of an nitrocellulose filter binding assay measuring the equilibrium binding of WT1-ZFP to the WT1 element with a TGT subsite for finger 1 (closed circles), the WT1 element with a CCC subsite for finger 1 (open squares) or a thyroid hormone receptor element (open circles). Curves 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).

subsite defined by the SAAB experiments (Table 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 4-fold.

Table 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.

WT1 binding sites have been identified in a number of eukaryotic promoters (19,34–40). To determine the biological significance of the WT1 recognition sequences selected in the 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 (19), with the affinity of the protein for the high affinity TGT target sequence identified by SAAB analysis. As the results in Figure 4 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 one consensus base but it is in the critical 10 position, and IGF-A has only one consensus base but it is in the less critical 11 position.

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 (14,29,38). 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 gene construct has been used to demonstrate the gene regulatory function of thyroid hormone reponsive elements cloned into the unique *Hin*dIII site (30). Similar tests of the regulatory activity of other WT1 binding sites have utilized an analogous tk-CAT reporter system (14). 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 h, cell extracts were prepared and after normalization for transfection efficiency, were assayed for CAT activity. As the results in Figure 5 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.

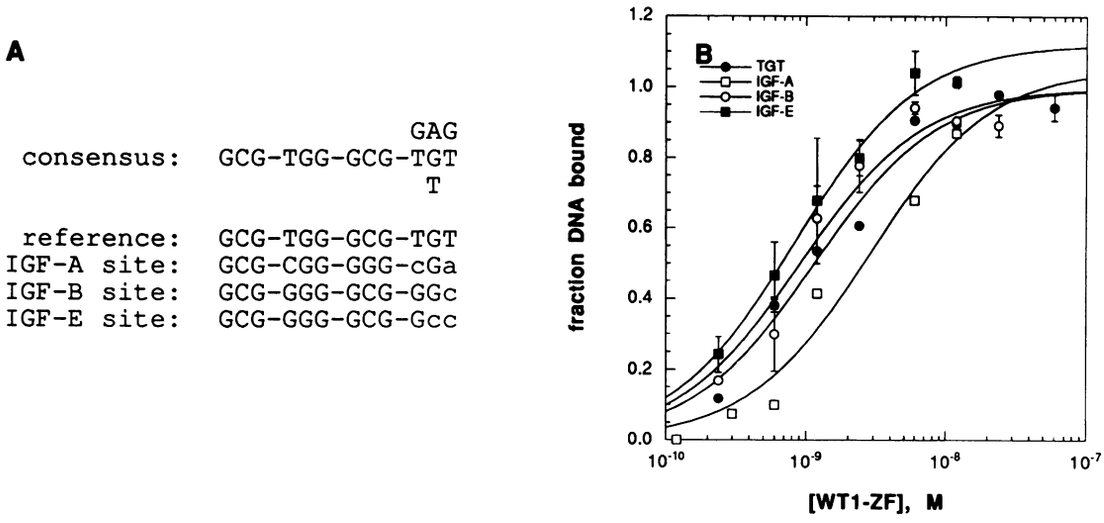

Figure 4. (**A**) Sequences of WT1 elements identified in the fetal promoter of the *IGF-II* gene. (**B**) Results of an nitrocellulose filter binding assay measuring the equilibrium binding of WT1-ZFP to the WT1 element with a TGT subsite for finger 1 (closed circles), the IGF-A element (open squares), the IGF-B element (open circles) or the IGF-E element (closed squares). Curves represent the best fit of the data in each case to a simple bimolecular equilibrium.

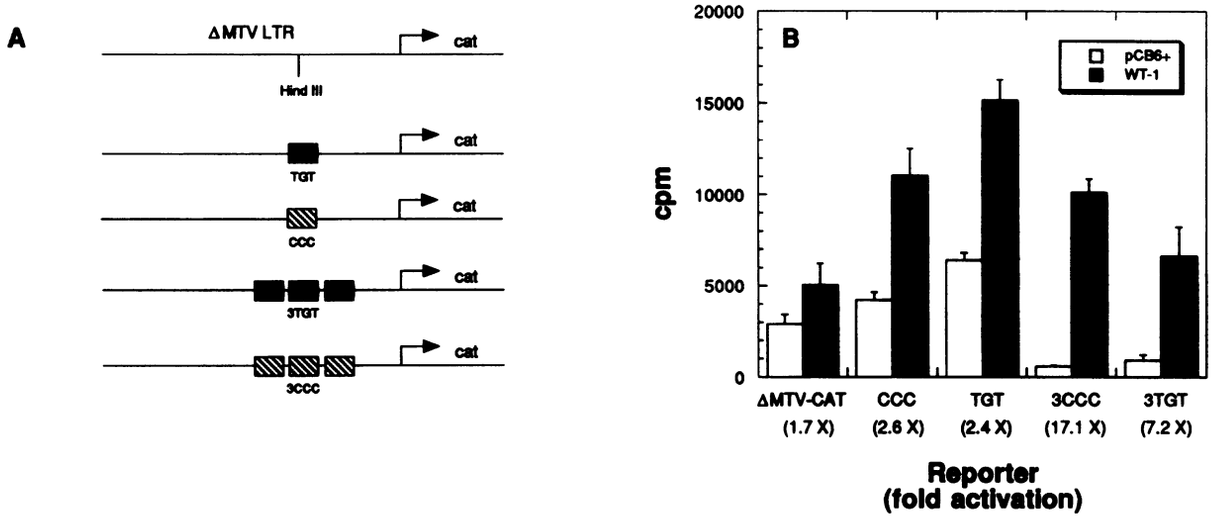

Figure 5. (**A**) CAT reporter gene constructs used in transient transfection assays to measure WT1 responsiveness. Boxes represent the number of WT1 elements with the indicated finger 1 subsite sequence inserted into the unique *Hin*dIII site of the ΔMTV-CAT plasmid. (**B**) Assay of chloramphenicol 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. C.p.m. values are reported as the mean ± standard deviation from five or more independent experiments. Fold activation is reported as the ratio of c.p.m. obtained in transfections with pWT1 to the c.p.m. obtained in transfections with pCB6+.

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

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 (22). Each zinc finger of EGR-1 interacts with a 3 bp subsite on the DNA, the amino acid residues in the α helix of the zinc finger interacting specifically with guanine residues along one strand. 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 3 bp subsite contiguous to the EGR-1 consensus site, implying that WT1[-KTS] recognizes a 12 bp sequence in target promoters.

Although this is a reasonable hypothesis, other possibilities must be taken into consideration. Two other zinc finger–DNA complexes have been studied by X-ray crystallography: the SWI–DNA complex (41) and the GLI–DNA complex (42). The results of these investigations demonstrated that zinc finger interactions with double stranded DNA are not restricted to one specific strand, nor to contiguous 3 bp subsites. Interactions can span both strands of the DNA, subsites can be larger than 3 bp 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 5 bp immediately downstream of the 9 bp EGR-1 consensus site. Based upon the antiparallel binding of EGR-1 to its 9 bp 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 2–4 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 3 bp subsite within a 12 bp 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. 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 manuscript was in preparation, a report was published describing the results of a SAAB experiment conducted with a recombinant peptide encompassing fingers 1–3 of WT1[-KTS] (WT1ΔF4-ZFP) (43). 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 3 bp subsite with a GT(C or T) consensus sequence. In the work we report here, 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 4-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 versus 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 bp WT1 selected sequence to be 7.1×10^6 M^{-1} versus a value of 4.5×10^6 M^{-1} for a non-selected sequence including the 9 bp EGR-1 consensus site (43). 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 $8.40\pm1.21\times10^8$ M^{-1} versus $2.10\pm0.41\times10^8$ 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 1.7×10^8 M^{-1} for the interaction of the zinc finger domain of EGR-1with its consensus DNA sequence (22). Thus the zinc finger domain of WT1 binds with high affinity to a 12 bp 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] (19,34–40). 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 (29). 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 signficantly 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 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. Although the high affinity consensus sequence can confer WT1 responsiveness to a promoter, the responsiveness may not be the optimal that can be achieved. 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.

ACKNOWLEDGEMENTS

The authors thank Dr T. Deuel for providing the pCB6+ and pWT1 plasmids and Dr R. Evans for providing the ΔMTV-CAT plasmid. The authors also thank C. Nelson and Q. You for helpful discussions. This work was supported by a grant from the National Cancer Institute of Canada.

REFERENCES

1 Stanbridge, E.J. (1990) *Annu. Rev. Genet.* **24**, 615–657.
2 Matsunaga, E. (1981) *Hum. Genet.* **57**, 231–246.
3 Glaser, T., Lewis, W.H., Bruns, G.A.P., Watkins, P.C., Rogler, C.E., Shows, T.B., Powers, V.E., Willard, H.F., Goguen, J.M., Simola, K.O.J. and Housman, D.E. (1986) *Nature* **321**, 882–887.
4 Fearon, E.R., Vogelstein, B. and Feinberg, A.P. (1984) *Nature* **309**, 176–178.
5 Koufos, A., Hansen, M.F., Lampkin, B.C., Workman, M.L., Copeland, N.G., Jenkins, N.A. and Cavenee, W.K. (1984) *Nature* **309**, 170–172.
6 Orkin, S.H., Goldman, D.S. and Sallan, S.E. (1984) *Nature* **309**, 172–174.
7 Reeve, A.E., Housiaux, P.J., Gardner, R.J.M., Chewings, W.E., Grindley, R.M. and Millow, L.J. (1984) *Nature* **309**, 174–176.
8 van Heyningen, V., Boyd, P.A., Seawright, A., Fletcher, J.M., Fantes, J.A., Buckton, K.E., Spowart, G., Porteous, D.J., Hill, R.E., Newton, M.S. and Hastie, N.D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8592–8596.
9 Gessler, M., Poustka, A., Cavenee, W., Neve, R.L., Orkin, S.H. and Bruns, G.A.P. (1990) *Nature* **343**, 774–778.
10 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) *Cell* **60**, 509–520.
11 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) *Cell* **53**, 37–43.
12 Joseph, L.J., le Beau, M.M., Jamieson, G.A.J., Acharya, S., Shows, T.B., Rowley, J.D. and Sukhatme, V.P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7164–7168.
13 Haber, D.A., Buckler, A.J., Glaser, T., Call, K.M., Pelletier, J., Sohn, R.L., Douglass, E.C. and Housman, D.E. (1990) *Cell* **61**, 1257–1269.
14 Madden, S.L., Cook, D.M., Morris, J.F., Gashler, A., Sukhatme, V.P. and Rauscher, F.J. (1991) *Science* **253**, 1550–1553.
15 Rauscher, F.J., Morris, J.F., Tournay, O.E., Cook, D.M. and Curran, T. (1990) *Science* **250**, 1259–1262.
16 Christy, B. and Nathans, D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8737–8741.
17 Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Frank, R. and Charnay, P. (1990) *Mol. Cell. Biol.* **10**, 3456–3467.
18 Takimoto, Y., Wang, Z., Kobler, K. and Deuel, T. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1686–1690.
19 Drummond, I.A., Madden, S.L., Rohwernutter, P., Bell, G.I., Sukhatme, V.P. and Rauscher, F.J. (1992) *Science* **257**, 674–678.
20 Ackerman, S.L., Minden, A.G., Williams, G.T., Bobonis, C. and Yeung, C.Y. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7523–7527.
21 Gupta, M.P., Gupta, M., Zak, R. and Sukhatme, V.P. (1991) *J. Biol. Chem.* **266**, 12813–12816.
22 Pavletich, N.P. and Pabo, C.O. (1991) *Science* **252**, 809–817.
23 Nehlin, J.O. and Ronne, H. (1990) *EMBO J.* **9**, 2891–2898.
24 Morris, J.F., Madden, S.L., Tournay, O.E., Cook, D.M., Sukhatme, V.P. and Rauscher, F.J. (1991) *Oncogene* **6**, 2339–2348.
25 Blackwell, T.K. and Weintraub, H. (1990) *Science* **250**, 1104–1110.
26 Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods Enzymol.* **185**, 60–89.
27 Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248–254.
28 Romaniuk, P.J. (1990) *J. Biol. Chem.* **265**, 17593–17600.
29 Wang, Z.Y., Qiu, Q.Q. and Deuel, T.F. (1993) *J. Biol. Chem.* **268**, 9172–9175.
30 Thompson, C.C. and Evans, R.M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3494–3498.
31 Chen, C. and Okayama, H. (1987) *Mol.Cell. Biol.* **7**, 2745–2752.
32 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
33 Sleigh, M.J. (1986) *Anal. Biochem.* **156**, 251–256.
34 Dey, B.R., Sukhatme, V.P., Roberts, A.B., Sporn, M.B., Rauscher, F.J. and Kim, S.J. (1994) *Mol. Endocrinol.* **8**, 595–602.
35 Harrington, M.A., Konicek, B., Song, A., Xia, X.L., Fredericks, W.J. and Rauscher, F.J. (1993) *J. Biol. Chem.* **268**, 21271–21275.
36 Rupprecht, H.D., Drummond, I.A., Madden, S.L., Rauscher, F.J. and Sukhatme, V.P. (1994) *J. Biol. Chem.* **269**, 6198–6206.
37 Wang, Z.Y., Qiu, Q.Q., Enger, K.T. and Deuel, T.F. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8896–8900.
38 Werner, H., Re, G.G., Drummond, I.A., Sukhatme, V.P., Rauscher, F.J., Sens, D.A., Garvin, A.J., Leroith, D. and Roberts, C.T. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5828–5832.
39 Wang, Z.Y., Madden, S.L., Deuel, T.F. and Rauscher, F.J. (1992) *J. Biol. Chem.* **267**, 21999–22002.
40 Gashler, A.L., Bonthron, D.T., Madden, S.L., Rauscher, F.J., Collins, T. and Sukhatme, V.P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10984–10988.
41 Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T. and Rhodes, D. (1993) *Nature* **366**, 483–487.
42 Pavletich, N.P. and Pabo, C.O. (1993) *Science* **261**, 1701–1707.
43 Drummond, I.A., Rupprecht, H.D., Rohwernutter, P., Lopezguisa, J.M., Madden, S.L., Rauscher, F.J. and Sukhatme, V.P. (1994) *Mol. Cell. Biol.* **14**, 3800–3809.