Alternative Splicing of Wilms' Tumor Suppressor Protein Modulates DNA Binding Activity through Isoform-Specific DNA-Induced Conformational Changes†

John H. Laity, John Chung, H. Jane Dyson, and Peter E. Wright*

Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10596 North Torrey Pines Road, La Jolla, California 92037

Received November 18, 1999; Revised Manuscript Received February 25, 2000

ABSTRACT: The Wilms' tumor suppressor protein (WT1) is a zinc finger transcription factor that appears to function differently according to the presence of a posttranscriptional modification that adds three amino acids into one of the linker regions between the zinc fingers. We have investigated the structural consequences of the insertion of the Lys-Thr-Ser (KTS) sequence by preparing recombinant protein constructs of the four zinc finger DNA-binding domain of WT1 corresponding to the two isoforms with (+KTS) and without (−KTS) the insertion, which is located in the linker region between the third and fourth zinc fingers. NMR resonance assignments were used to estimate the structural differences between the two isoforms both free in solution and in complex with a 14 base pair DNA duplex corresponding to the WT1 recognition element. The NMR spectra indicate that the two isoforms are nearly identical in structure in the absence of the DNA. Only the immediate region of the insertion showed any change in chemical shifts. Upon DNA binding, the NMR spectrum of each isoform changed to indicate greater structure formation in the linker regions. Significant differences were observed between the spectra of the DNA complexes of the +KTS and −KTS isoforms, with the −KTS construct forming a more stable complex, consistent with prior biochemical assays. The majority of the differences between the spectra of the two complexes occur in the immediate region of the insertion, which appears to be closer in structure to the free form of the protein in the case of the +KTS complex. The insertion of the KTS sequence disrupts important interactions of the linker region with the adjacent zinc fingers, thus lowering the stability of the complex. The "normal" (−KTS) sequence of the linker appears to be involved in a C-terminal helix-capping interaction with the helix of the preceding zinc finger, a stabilizing interaction which is abrogated in the +KTS isoform.

Transcription factors play an important role in the cascade of cellular signal transduction. Wilms' tumor suppressor protein (WT1)[1] has many properties that are characteristic of a transcription factor. It contains a C-terminal DNA-binding domain comprised of four Cys_2His_2 zinc fingers, a proline/glutamine-rich N-terminus (*1*, *2*), activation and repression domains, nuclear localization signals (*3*), and at least two self-association domains (*4*−*6*). The cellular expression pattern of WT1 is tissue-specific, and also depends on the growth stage of the organism. WT1 is essential for normal mammalian urogenital development, with the highest expression levels of WT1 observed in the specialized podocyte cells of the embryonic kidney (*7*, *8*). The *WT1* gene was originally identified through its predisposition for deletions or mutations in a subset of patients with the common pediatric malignancy, Wilms' tumor (*1*, *2*). Constitutional defects in the *WT1* gene are also associated with other diseases such as the Denys-Drash (*9*), WAGR (*10*), and Frasier syndromes (*11*), all of which result in malformations of the genito-urinary system, and an increased risk for the development of Wilms' tumor. WT1 mutations have also been implicated in several acute leukemias (*12*), while altered expression of WT1 has recently been detected in human breast cancer (*13*).

Alternative pre-mRNA splicing at two sites in the WT1 gene produces four protein isoforms. The first alternative splice site (exon 5) inserts or omits 17 amino acids N-terminal to the zinc finger domain, but has no effect on the DNA-binding activity of the protein. The second alternative splice site either inserts (+KTS) or omits (−KTS) a highly conserved Lys-Thr-Ser sequence in the linker between zinc-fingers 3 and 4. The binding affinity of WT1+KTS for putative GC-rich DNA sequences is dramatically reduced compared to that of the −KTS isoform (*14*−*16*). The +KTS isoform of WT1 has alternative functions; for example, it binds to a 25 nucleotide region of untranslated mouse Igf-2 mRNA with greater affinity than does WT1-KTS (*17*). Recently, in a study of Igf-2 gene regulation by WT1, the two isoforms were found to have opposite effects, with the −KTS isoform activating the Igf-2 P3 promoter and WT1+KTS repressing its activity (*18*). The subnuclear localization of WT1 also appears to be controlled by the

† This work was supported by Grant GM36643 from the National Institutes of Health. J.H.L. is supported by Fellowship CA79192 from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (858) 784 9721. Fax: (858) 784 9822. Email: wright@scripps.edu.

[1] Abbreviations: WT1, Wilms' tumor suppressor protein; KTS, amino acid sequence of insertion Lys-Thr-Ser; NMR, nuclear magnetic resonance; wt 1−4, 4-zinc finger domain of WT1; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactose; DTT, dithiothreitol; GdnHCl, guanidine hydrochloride; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

10.1021/bi9926678 CCC: $19.00
Published on Web 04/13/2000

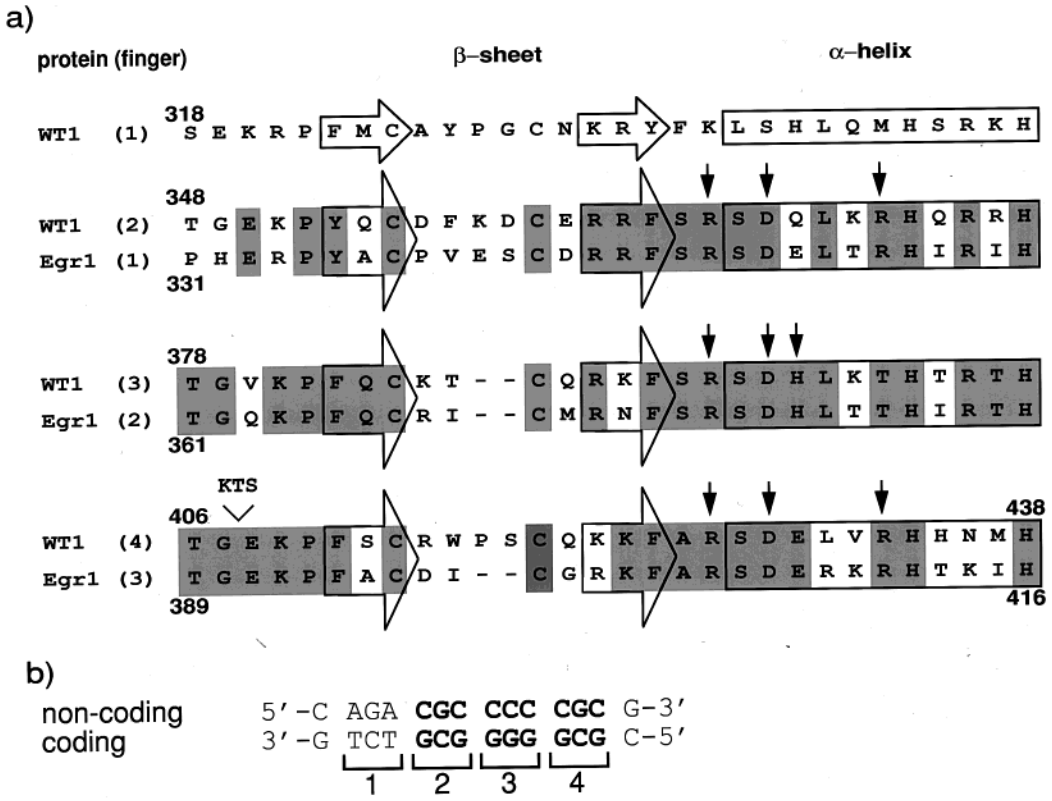

FIGURE 1: Sequence of (a) homologous WT1 and Egr1 zinc finger domains and (b) the corresponding DNA cognate. (a) WT1 and Egr1 sequences are shaded in regions with primary sequence identity. β-Sheet and α-helix secondary structures, predicted for WT1 based on the Egr1/DNA X-ray structure (*21*, *22*), chemical shift data for WT1 (present work), and conserved Cys_2His_2 secondary structure propensities (WT1 finger 1), are indicated. Egr1 residues that contact DNA bases in the crystal structure are indicated with arrows. Note: the numbering of amino acids in the WT1 protein proceeds according to the +KTS isoform; therefore, residues Lys408−Ser410 are not present in −KTS. (b) DNA sequence used for wt1−4/DNA complexes. The Egr1 DNA consensus sequence is shown in boldface text. The "antiparallel" alignment of the WT1 zinc fingers (1−4) with the DNA base pairs is predicted from homologous Egr1/DNA interactions (*21*, *22*).

alternative KTS splicing motif. WT1+KTS preferentially co-localizes with nuclear pre-mRNA splicing domains (spliceosomes and coiled bodies), and the −KTS isoform of WT1 associates predominantly with the transcription factor domains (DNA-rich regions) (*19*). Alterations in WT1 +KTS/−KTS isoform ratios, reported to be steady throughout normal development (+KTS/−KTS ∼ 2:1) (*20*), can result in severe developmental abnormalities. Collectively, these observations suggest that WT1 may regulate gene expression through both transcriptional and posttranscriptional processes, utilizing alternative ±KTS pre-mRNA splicing to differentiate, and subsequently regulate, DNA and RNA targets.

As shown in Figure 1a, zinc fingers 2−4 of WT1 have 67% amino acid identity to the three zinc fingers of the mouse immediate early protein (Egr1). The 3D structure of Egr1 bound to the DNA consensus sequence 5′-GCG-TGG-GCG-3′ has been reported (*21*, *22*); this structure identifies nine Egr1 residues which directly contact the DNA bases in the major groove. Interestingly, although all nine of these residues are conserved between Egr1 and WT1 (see arrows in Figure 1a), and the −KTS isoform of WT1 binds the Egr1 DNA consensus with high affinity, the +KTS binding is dramatically reduced, even though the insertion is distant from the protein−DNA interface.

To date, no structural studies of WT1−DNA complexes have been reported. Identifying the specific isoform-dependent changes that occur upon binding to DNA will provide insights into the molecular mechanism that governs the activity of WT1. To investigate the structural characteristics of both free and DNA-bound WT1, we have constructed a recombinant system to produce polypeptides that correspond to both isoforms of the WT1 zinc finger domain (wt1−4; Figure 1a). The 14 base pair region of DNA studied (Figure 1b) contains the core Egr1 consensus sequence, and was optimized to give high binding affinity for both +KTS and −KTS based on observed finger 1 DNA sequence preferences (*16*). Using NMR spectroscopy, we have identified regions of DNA-induced conformational changes in +KTS and −KTS that contribute to the stability of both protein−DNA complexes. It is clear that the lower affinity of the +KTS isoform for DNA is related to specific changes in structure of the linker between fingers 3 and 4, notably disruption of the capping and other protein−protein interactions that occur in this and other linker regions upon DNA binding.

MATERIALS AND METHODS

Gene Design and Plasmid Construction. A synthetic gene was designed to facilitate high-level expression of the +KTS isoform of human wt1−4 in *E. coli*. This *+KTS* gene encodes residues 318−438 of the +KTS isoform of WT1 (Figure 1a), and includes an alanine residue at the N-terminus (122 residues total) to facilitate cleavage of the N-terminal methionine (*23*). Synthesis of *+KTS* was carried out recursively (*24*) using PCR from 8 alternating coding/noncoding oligonucleotides of equal length (79−80 bases) which overlapped 20 base pairs. The DNA sequence for this gene was optimized to balance usage of codons for highly

expressed indigenous proteins in *E. coli*, while avoiding repetitive DNA sequences within the gene that could interfere with the PCR reaction. The DNA expression vector for +KTS (pWT1−4+) was constructed by subcloning the *+KTS* synthetic gene into the *Eco*RI/*Hin*dIII site of the T7-expression vector pET21d (Novagen). The pWT1−4+ DNA plasmid was then used as a template for PCR-based deletion of 9 base pairs to create the corresponding DNA expression vector pWT1−4− that encodes −KTS (119 residues). Both DNA constructs direct high-level intracellular expression of the corresponding wt1−4 protein isoform in *E. coli*.

Protein Expression and Purification. The DNA plasmids pWT1−4+ and pWT1−4− were transformed into competent *E. coli* BL21(DE3) cells, from which a single colony was used to inoculate 1 L of minimal medium (0.67% Na_2HPO_4, 0.3% KH_2PO_4, 0.05% NaCl, 1 mM $MgCl_2$, and 1 mM $CaCl_2$) containing 50 μg/L carbenicillin, and supplemented with vitamins (Gibco) and trace metals. Uniform (U) isotopic enrichment of the media was obtained by using 2 g/L U-[$(^{15}NH_4)_2SO_4$] and either 3 g/L unlabeled glucose or 2 g/L U-[^{13}C]glucose when producing U-[^{15}N]- or U-[^{15}N,^{13}C]-protein, respectively. Similarly, uniformly U-[^{2}H (nonexchangeable),^{15}N,^{13}C]-protein was produced in an analogous manner except ^2H_2O was used in place of ^1H_2O to make the cell growth media. All cultures were grown to OD_{600} ∼0.6, induced with IPTG (1 mM), and further incubated for 6−10 h at 37 °C to produce both soluble and insoluble wt1−4 in a ratio of ∼1:4. Cells were recovered by centrifugation, resuspended in lysis buffer (50 mM Tris, pH 8.0, containing 300 mM NaCl, 50 μM $ZnSO_4$, 10 mM DTT), and lysed by sonication. The insoluble lysate (inclusion bodies) was solubilized using lysis buffer containing 7 M GdnHCl and 100 mM DTT, and refolded by a 1:100 dilution into argon-saturated lysis buffer.

The filtered soluble lysate and refolded inclusion bodies were pooled, and loaded onto a HiTrap SP column (Pharmacia) preequilibrated with buffer (50 mM Tris, pH 8.0, containing 10 μM $ZnSO_4$). The wt1−4 was eluted using a linear NaCl gradient. This protein solution was purified further using a HYDROPORE SCX 21.4 mm ID column (Rainin) preequilibrated with 25 mM HEPES, pH 7.0, containing 10 μM $ZnSO_4$. The wt1−4 was eluted using another NaCl gradient. The purified protein was then desalted into 50 mM NH_4OAc using a Sephadex G25F column, lyophilized, and further analyzed by SDS−PAGE and electrospray mass spectrometry to determine purity and extent of isotopic enrichment. Uniform isotopic enrichment with ^{15}N and ^{13}C was essentially complete (>98%) for both +KTS and −KTS, and ^{2}H incorporation was 80% for proteins expressed in ^2H_2O and U-[^{1}H,^{13}C]glucose. The yield of +KTS and −KTS proteins (>95% pure) was 10−20 mg/L.

DNA Synthesis and Purification. The two complementary oligonucleotides (5′-CGCGGGGGCGTCTG-3′, 5′-CAGACGCCCCCGCG-3′), commercially synthesized (Operon, Keystone) using the phosphoramidite method, were obtained in lyophilized form with the trityl group on. Each oligonucleotide was resuspended in H_2O, and loaded on a preparative C18 HPLC column equilibrated in 100 mM TEAA, pH 6.5. The single-stranded DNA was eluted from the column by a linear CH_3CN gradient. The purified DNA was lyophilized, and then resuspended again in H_2O; the trityl group was removed spontaneously by hydrolysis due to the low pH (∼4−5) of the solution. Once the detritylation of the oligonucleotides was deemed complete by analytical HPLC, the pH was raised to ∼7−8 and KCl was added to 200 mM. Stoichiometric amounts of the two oligonucleotides were combined, heated to 85 °C in a water bath, and then slowly annealed by cooling the bath to room temperature. The DNA duplex was loaded onto a HiTrap Q column preequilibrated in 25 mM Tris buffer (pH 7.5), and eluted with a linear KCl gradient to separate the duplex from any remaining single stranded DNA in the original annealing mixture. The purity of the DNA duplex was confirmed by re-injection into an analytical HiTrap Q column and reversible thermal denaturation experiments. The DNA was concentrated and exchanged into NMR buffer.

WT1−4/DNA Complex Formation. Both +KTS and −KTS complexes with DNA were prepared as follows: an equal volume of 0.6 mM DNA solution was added in aliquots to a 0.6 mM wt1−4 protein solution. With each aliquot of DNA added, a small amount of precipitate formed initially in the solution containing the complex, but redissolved after gentle agitation. Once the stoichiometry of the wt1−4/DNA complex reached 1:1, additional aliquots of DNA added to the complex no longer resulted in reversible precipitation. Difference 1D-NMR spectra of the of the DNA imino region between subsequent DNA aliquot additions was used to confirm that the titration end-point was reached (protein/DNA stoichiometery was 1:1).

NMR Spectroscopy. All NMR samples were prepared in 10 mM d_{11}-Tris saturated with argon, pH 6.7, containing 5% (v/v) ^2H_2O, 20 mM KCl, 50 μM $ZnSO_4$, 2 mM NaN_3, and 0.2 mM DSS. Sample concentrations for all NMR experiments were 2−3 mM for both free wt1−4 isoforms, and 0.3 mM for the DNA-bound +KTS and −KTS complexes. The samples were placed in either a siliconized 5 mm or 8 mm (Shigemi) NMR tube. Spectra of free wt1−4 proteins were recorded with U-[^{15}N]- or U-[^{15}N,^{13}C]-wt1−4 at 293 K using a Bruker AMX 500. NMR data for the wt1−4/DNA complexes were acquired with U-[^{15}N]- or U-[^{2}H (80%, nonexchangeable), ^{15}N,^{13}C]-wt1−4 at 310 K using either a Bruker DRX 600 equipped with an 8 mm probe (Nalorac) or a Bruker DMX 750. Quadrature detection in the indirect dimensions was obtained by combining n- and p-spectra selected with pulse field gradients, or by using the States-TPPI method. For some of the NMR spectra acquired, the proton carrier was set to the water frequency to reduce the amplitude of spurious water echos. A summary of the relevant NMR parameters is listed in Table S1 of the Supporting Information.

NMR data were processed with the Felix software package (Molecular Simulations Inc.). Linear prediction (LP) was applied to the time domain data in the indirect dimensions, using mirror-image LP for all spectra recorded with constant time evolution. All dimensions were zero-filled once (except when t_1 was greater than 256 complex points, Table S1 of Supporting Information), and apodized with a sinebell window function prior to Fourier transformation. A large first-order phase correction was applied to all frequency domain spectra that were recorded using digital filters (on the DRX 600 and DMX 750).

The ^{1}H, ^{13}C, and ^{15}N chemical shifts were referenced with DSS as an internal standard (*25*). Sequential assignments of protein ^{1}H, ^{15}N, and ^{13}C backbone and $^{13}C^{\beta}$ side-chain resonances were determined manually by matching intraresi-

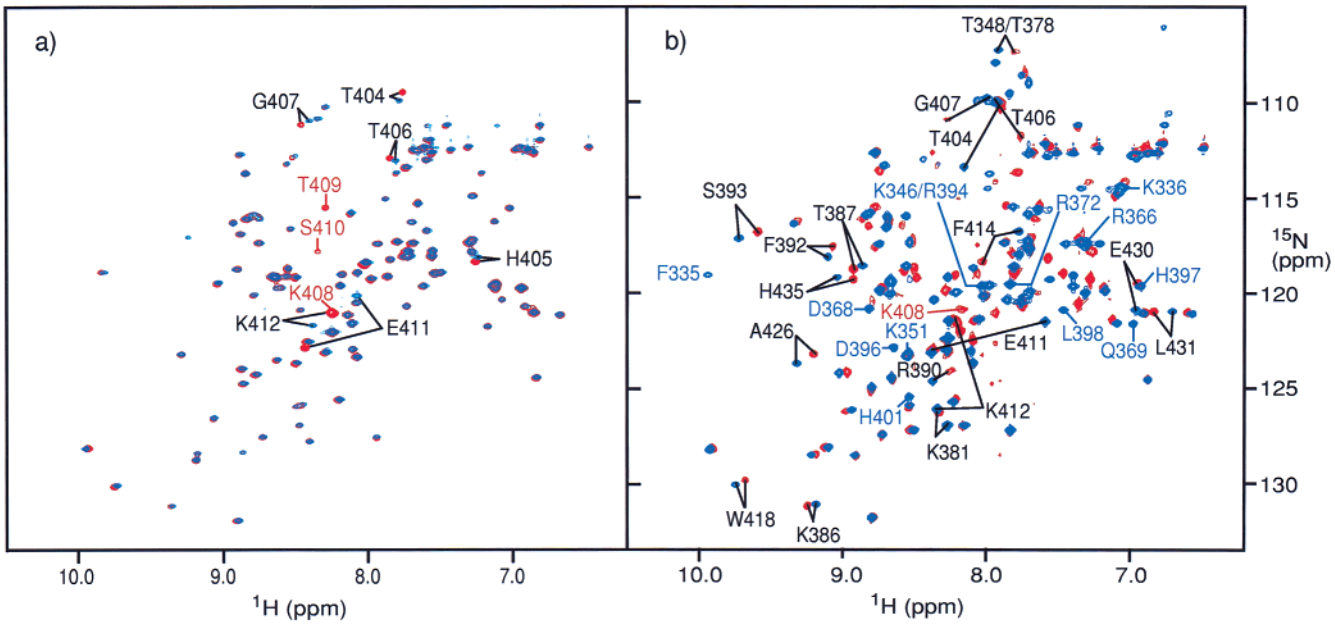

FIGURE 2: (a) Overlaid ^{1}H−^{15}N-HSQC spectra of free −KTS (blue peaks) and +KTS (red peaks). Alternatively spliced Lys408-Thr409-Ser410 residues and residues with different $^1H^N$ and $^{15}N^H$ shifts beween −KTS and +KTS are annotated, with the KTS splice site residues annotated in red. (b) Overlaid ^{1}H−^{15}N-HSQC spectra of DNA-bound −KTS (blue peaks) and +KTS (red peaks). The Lys408 residue in the alternative splice site (red annotation), residues that had resonances in the −KTS/DNA spectrum but not in the +KTS/DNA spectrum (blue annotation), and many residues with large $^1H^N$ and $^{15}N^H$ chemical shift differences beween DNA-bound −KTS and +KTS (black annotation) are indicated. Note: the Ser410 resonance was assigned in the +KTS/DNA complex, but was not visible at the contour level shown in the ^{1}H−^{15}N-HSQC spectrum presented above.

due and sequential $^{13}C^\alpha$ and $^{13}C^\beta$, $^{13}C'$, or $^1H^\alpha$ chemical shifts from the corresponding dimension of a 3D spectrum with sequential-only chemical shifts from the same dimension of a complementary 3D experiment. Amide resonances for each residue were assigned from the corresponding values in the amide $^1H^N$ and ^{15}N dimensions of the same set of 3D spectra. The resonance assignments have been deposited in the BioMagResBank (accession numbers 4707−4710).

RESULTS

NMR Spectra of Free +KTS and −KTS Isoforms. Nearly complete ^{1}H, ^{15}N, and ^{13}C backbone and $^{13}C^\beta$ side-chain resonance assignments were determined for the free +KTS and −KTS proteins. A summary of the total number of assigned residues for each nucleus of free and DNA-bound wt1−4 is given in Table S2 of the Supporting Information. A full description of the assignment procedure is also given in the Supporting Information.

The NMR spectra indicate that the structures of the free +KTS and −KTS isoforms are almost identical. Overlaid ^{1}H−^{15}N-HSQC spectra of free +KTS and −KTS are shown in Figure 2a, and a quantitative comparison of $^1H^N$, $^{15}N^H$, and $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shift differences between the two isoforms of wt1−4 is presented in Figure 3. The structural similarity of +KTS and −KTS is striking, as evidenced by the identical positions of >95% of the corresponding $^1H^N$−^{15}N peaks in the overlaid HSQC spectra (Figure 2a). Large differences in backbone amide ^{1}H and ^{15}N chemical shifts between +KTS and −KTS are observed only for the Glu411 residue immediately N-terminal to the alternative splice site. A few subtle differences in $^1H^N$ and $^{15}N^H$ shifts between wt1−4 isoforms are also evident in the region from Thr404 to Lys412 (Figures 2a and 3). Significantly, with the exception of the Thr406 $^{13}C^\alpha$ shift ($\Delta\delta_{[(+KTS) - (-KTS)]} = -0.35$ ppm), all the $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shifts of free isoforms of wt1−4 are superimposed (all $\Delta\delta_{[(+KTS) - (-KTS)]} < 0.2$ ppm). A qualitative characterization of the ϕ, ψ backbone conformations in proteins can be obtained from a comparison of $^1H^\alpha$, $^{13}C^\alpha$, and $^{13}C'$ chemical shift deviations from statistical coil values ($\Delta\delta_{coil}$). A summary of $\Delta\delta_{coil}$ values for +KTS is shown in Figure 4. The negative $^1H^\alpha$ and positive $^{13}C^\alpha$ and $^{13}C'$ $\Delta\delta_{coil}$ values in regions α1, α2, α3, and α4 of Figure 4 clearly demonstrate the presence of an α-helix in the "finger-tip" region of each zinc finger of WT1. The α-helix and β-sheet structures found in both isoforms of WT1 are characteristic of the Cys_2-His_2 zinc finger motif. It is noticeable, however, that there are small $^1H^\alpha$, $^{13}C^\alpha$, and $^{13}C'$ $\Delta\delta_{coil}$ values in the C-terminal residues of the first three α-helices (α1−α3) in the −KTS isoform as well as the +KTS shown in Figure 4. This indicates poor helical geometry and fraying of the helix in regions that are predicted to be well-structured from the Egr1/DNA Xray structure (*21*, *22*).

NMR Spectra of the DNA Complexes of the +KTS and −KTS Isoforms. Assignment of the ^{1}H, ^{15}N, and ^{13}C backbone and $^{13}C^\beta$ side-chain resonances for the DNA-bound complexes of wt1−4 proved more difficult than for the free protein. Two experimental enhancements were needed to circumvent short carbon T_2 times and extreme ^{1}H, ^{15}N, and ^{13}C line-broadening during triple resonance data collection for the 25 kDa protein−DNA complexes. First, all nonexchangeable H atoms in the +KTS and −KTS proteins were deuterated, resulting in sufficiently long (>100 ms) $^{13}C^\alpha$ T_2 times to record a set of deuterium-decoupled HNCACB and HN(CO)CACB spectra for each wt1−4/DNA complex. In addition, all NMR spectra were acquired at 310 K, compared to 293 K for the free wt1−4 proteins. The line-narrowing effect of deuteration and increased temperature on the DNA-

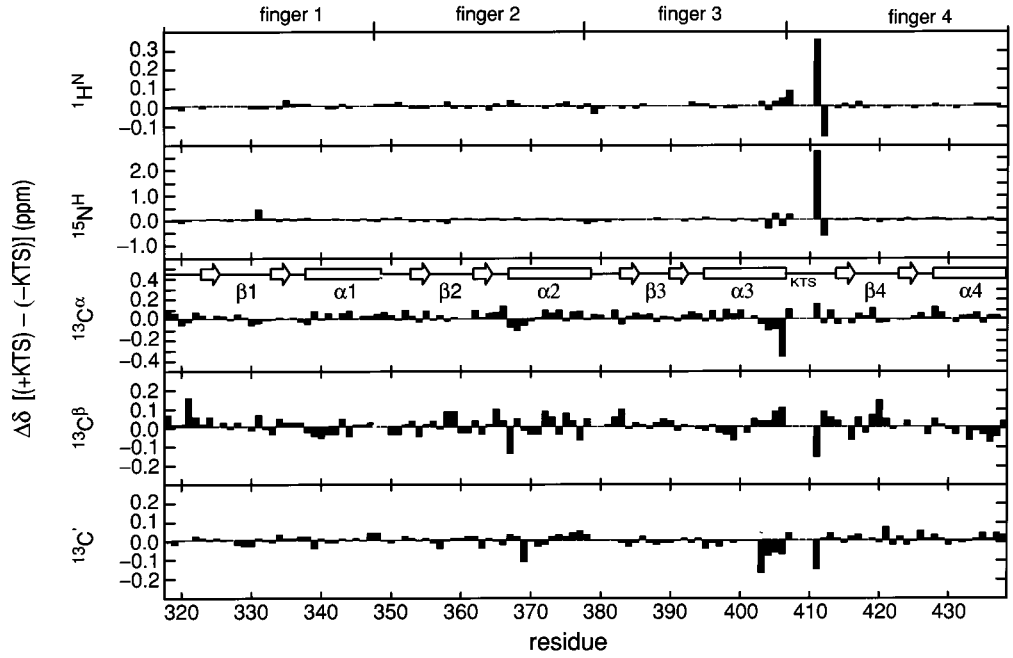

FIGURE 3: Comparison of backbone $^1H^N$, $^{15}N^H$, $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shift differences as a function of residue number between the −KTS and +KTS isoforms free in solution. Predicted α-helix and β-sheet structures are indicated (*21*, *22*).

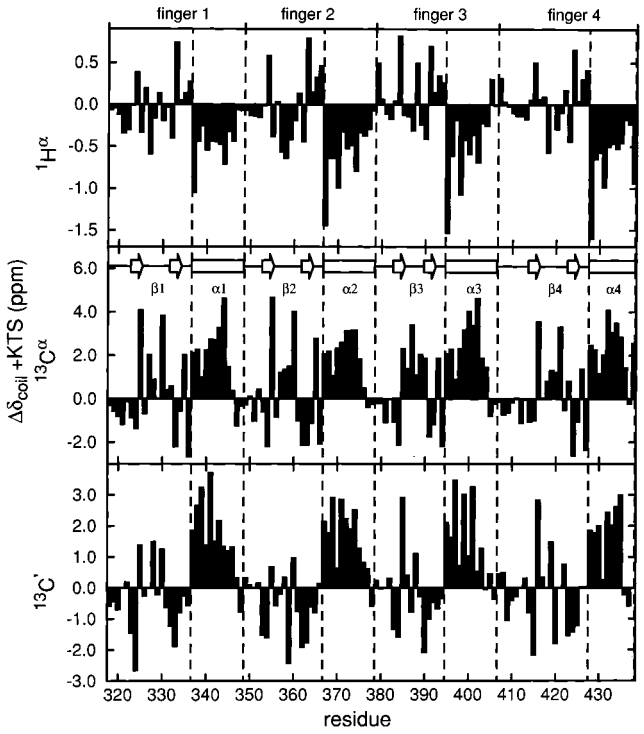

FIGURE 4: Chemical shift deviations in from random coil ($\delta\Delta_{coil}$) for $^1H^\alpha$, $^{13}C^\alpha$, $^{13}C'$ nuclei in free +KTS as a function of residue number. Predicted α-helix and β-sheet structures are indicated (*21*, *22*).

bound +KTS and −KTS complexes was significant. The $^1H^N$−^{15}N line widths of the two wt1−4/DNA complexes, shown in overlaid HSQC spectra in Figure 2b, are comparable to the corresponding $^1H^N$−^{15}N line widths of the free wt1−4 isoforms (Figure 2a).

Chemical Shift Changes upon DNA Binding. Comparison of Figures 2a and 2b and the $^1H^N$, $^{15}N^H$, and $^{13}C^\alpha$ shift differences between free and DNA-bound +KTS and −KTS presented in Figure 5 illustrates that there are greater differences between the +KTS and −KTS isoforms when they are bound to DNA. Due to slight differences in solution conditions for the bound and free proteins, there are some small systematic changes in the chemical shifts. In particular, there is a small change in $\Delta\delta$ due to the temperature difference between free (293 K) and DNA-bound (310 K) NMR experiments and a small $^2H/^1H$ isotope effect ($\Delta\delta_I$) on the $^{13}C^\alpha$ and $^{13}C^\beta$ shifts in the wt1−4 DNA complexes. Overall, these combined temperature and isotope effects are small for the $^{13}C^\alpha$ [$\Delta\delta_I+\Delta\delta_{(393\ K-293\ K)} \sim -0.1$ to -0.2 ppm], $^{15}N^H$ [$\Delta\delta_{(393\ K-293\ K)} \sim 0.3$ to -0.3 ppm], and $^1H^N$ [$\Delta\delta_{(393\ K-293\ K)} \sim -0.05$ to -0.15 ppm] shifts in wt1−4. Accordingly, only residues with DNA-induced chemical shift changes ($\Delta\delta_{complex\ -\ free}$) outside these ranges are considered to be significantly shifted. It is clear from Figure 5 that only a subset of the backbone resonances is significantly affected by DNA binding. The chemical shift changes upon DNA binding are of two types: (1) those that result from direct contact with the DNA at known contact sites, and (2) those that arise due to conformational changes at sites distant from the DNA binding site.

DISCUSSION

The recently published solution structure of the three N-terminal zinc fingers of TFIIIA (zf1−3) bound to nucleotides 79−93 of the 5S RNA gene (*26*) illustrates the importance of peptide linkers between adjacent zinc fingers in mediating sequence-specific DNA recognition. The peptide linkers in zf1−3 are nearly identical to the highly conserved TGEKP consensus sequence in WT1 and Egr1 (see Figure 1a). Upon binding to DNA, the zf1−3 linkers adopt well-defined structures involving multiple hydrogen-bonding interactions with the preceding finger, and pack against the adjacent fingers to bury substantial areas of the protein surface (*26*). Although we have not yet calculated three-dimensional structures of the WT1 isoform complexes, it is clear that similar mechanisms are operating for at least some of the zinc fingers of WT1. Interestingly, the chemical shift

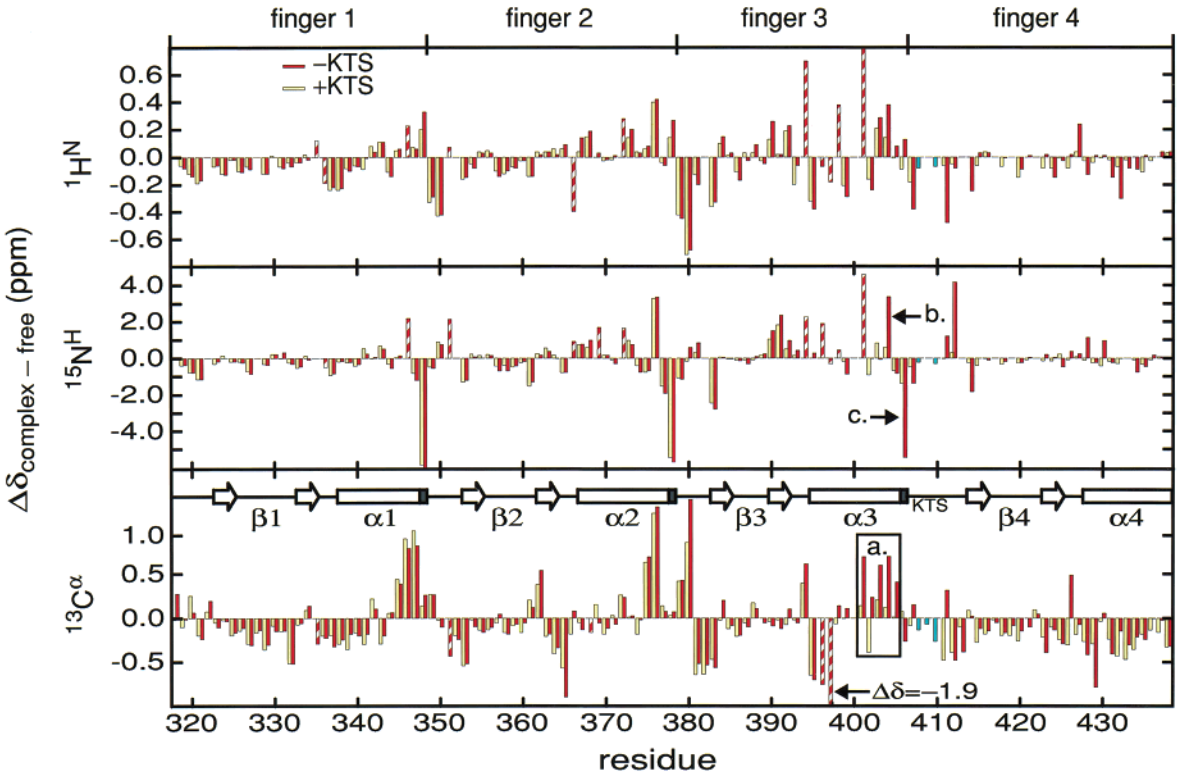

FIGURE 5: Comparison of DNA-induced $^1H^N$, $^{15}N^H$, and $^{13}C^\alpha$ chemical shift changes ($\delta\Delta_{complex\ -\ free}$) for −KTS (red bars) and +KTS (yellow bars) as a function of residue number. Red and white striped bars indicate $\delta\Delta_{complex\ -\ free}$ $^1H^N$, $^{15}N^H$, or $^{13}C^\alpha$ values of −KTS/DNA for which a corresponding value for the +KTS/DNA complex could not be determined. Residues in the C-terminal region of finger 3 with large $\delta\Delta_{complex\ -\ free}$ $^{13}C^\alpha$ (box a) and $^{15}N^H$ (arrows b and c) values for DNA-bound −KTS, but not +KTS, are indicated. Predicted α-helix and *β*-sheet structures are indicated (*21*, *22*). The solid black portions in the α-helices of fingers 1−3 indicate the locations of the DNA-induced C-capping threonine residues.

changes that accompany DNA binding by the TFIIIA zinc fingers (*27*) reflect the structural changes in the linker regions and helices, and parallel those observed in the present work.

Chemical Shift Mapping of DNA-Binding Interfaces. The DNA contact residues in fingers 2−4 can be inferred by analogy with the contacts observed in the highly homologous Egr1 complex (*21*, *22*) (see Figure 1). For the −KTS complex, significant chemical shift changes are observed, especially in the $^1H^N$ and $^{13}C\alpha$ shifts, for residues in fingers 2−4 in the immediate vicinity of the DNA base contacts. These include Ser365, Arg366, and Arg372 (finger 2); Arg394−His397 (finger 3); and Ala426, Arg427, and Asp429 (finger 4). The large $\Delta\delta_{complex\ -\ free}$ $^{13}C^\alpha$ values for Arg394−His397 in finger 3 of the −KTS/DNA complex (Figure 5) are particularly striking and suggest that the N-terminal portion of the α-helix in finger 3 may become less helical upon binding to DNA. Significantly, the three homologous residues (Arg377, Asp379, and His380) in Egr1 make the only direct hydrogen bonds to DNA bases in finger 2 of the Egr-1/DNA crystal structure (Figure 1a). Many of the WT1 mutations associated with Denys−Drash syndrome have been mapped to finger 3 (*28*−*30*). The most common of these mutations, Arg394 → Trp, is located at position −1 of the α-helix in finger 3 in a region that is shown by our data to be critical for site-specific DNA recognition and binding affinity. For the +KTS/DNA complex, similar DNA-induced chemical shift changes are seen for several residues in fingers 2 and 3 that are in the immediate vicinity of the DNA base contacts, including Arg366, Arg372, Arg394, and Ser395; resonances corresponding to Asp396 and His397 are too broad to assign for this complex.

For both the −KTS and +KTS complexes with DNA, there are very few chemical shift changes in the DNA interface of finger 1 (Figure 5). This suggests that finger 1 binding is weaker than for the other fingers, in agreement with binding studies with mouse Igf-2 exon 2 DNA, where it was determined that the deletion of finger 1 had no deleterious effect on DNA binding by either −KTS or +KTS. Indeed, finger 1 has been implicated in RNA binding (*17*).

Chemical Shift Mapping of DNA-Induced Conformational Changes for the −KTS Isoform. Significant chemical shift changes occur upon DNA binding (Figure 5) in regions that are distant from the DNA contact site. This observation exactly parallels that for the DNA complex of zinc fingers 1−3 from TFIIIA and reflects structure formation in the linkers induced by DNA binding (*26*). The most significant DNA-induced change in the −KTS/DNA complex is an increase in helical character in the C-terminal region of the first three zinc fingers (Figure 5). Figure 4 shows that the residues in these C-terminal regions are not helical in the free proteins. This increase in helical structure is observed generally in zinc finger proteins and can be accounted for by the formation of a DNA-induced C-capping interaction that acts as a "snap-lock" to hold the multifinger protein on the DNA (*31*). Figure 6 shows a schematic representation of the hydrogen bonding in the helices of the four fingers, deduced from the NMR spectra of the DNA complex of −KTS. The C-terminus of α4 is the same in the free and complexed −KTS protein, since there are no residues to form a C-cap for this helix. For fingers 1−3, DNA-induced helix termination motifs [α_L C-cap (*32*) and Schellman C-cap (*33*)] are observed that involve the first two or three residues of

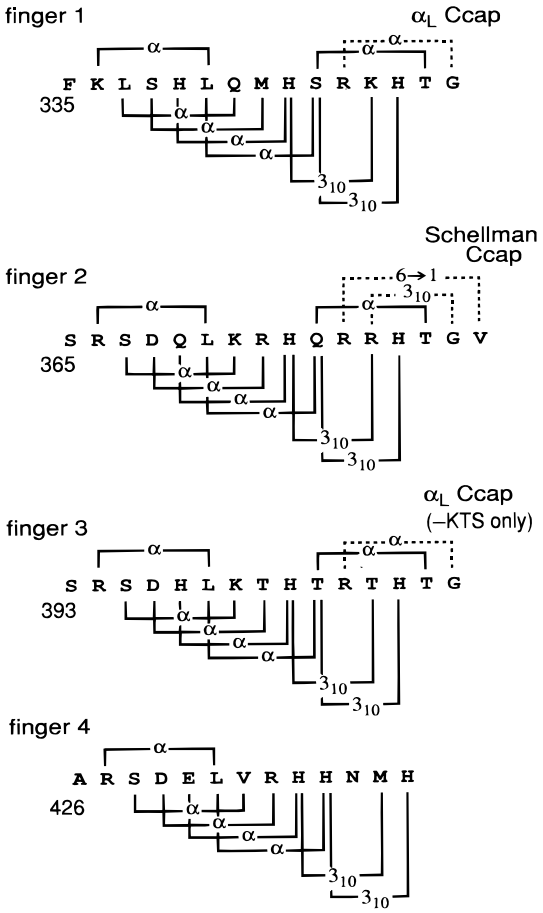

FIGURE 6: Model of backbone hydrogen bonding in the four α-helices of DNA-bound −KTS and +KTS. Hydrogen bonding was predicted from NMR data in Figures 4 and 6, and 3D structures of Egr1 (*21*, *22*) and zf1−3 (*26*) complexes with DNA. Brackets depicting hydrogen bonding originate at the residue that donates the amide proton and proceed to the residue with the carboxyl oxygen. Hydrogen bonding patterns indicative of an α-helix (*i*,*i*+4) and 3_{10}-helix (*i*,*i*+3) are indicated. C-capping hydrogen bonds use dotted brackets. α_L C-caps are characterized by one *i*,*i*+4 (α) hydrogen bond, while the Schellman C-cap has both an *i*,*i*+3 (3_{10}) and an *i*,*i*+5 (6→1) hydrogen bond. DNA-induced C-capping in finger 3 is observed in the −KTS/DNA complex only. Brackets below the sequence indicate that both of the residues that participate in the hydrogen bond have helical (ϕ,ψ) dihedral angles; brackets above the sequence indicate that one or both of the participating residues does not have helical dihedral angles.

the conserved linker sequence (consensus sequence TGEKP). The DNA-induced capping of the helix in fingers 1−3 stabilizes and extends the C-terminus of the helix, as observed for a number of other zinc finger−DNA complexes (*31*), and accounts for the positive changes in $^{13}C^{\alpha}$ chemical shifts of the −KTS/DNA complex compared to the free protein.

Influence of the KTS Insertion. Remarkably, in the +KTS/DNA complex, the same DNA-induced C-capping motifs are observed in fingers 1 and 2, but not in finger 3 (Figure 5). These observations suggest that C-terminal helix capping stabilizes the DNA complexes of both −KTS and +KTS isoforms, and that this stabilization is reduced in the +KTS/DNA complex because one of the three C-capping motifs is severely destabilized in the DNA-bound +KTS isoform. It is important to note that these DNA-induced helix termination motifs in both wt1−4/DNA complexes are distant from the protein−DNA interface.

The KTS insertion occurs between G407 and E408 in the sequence of WT1 −KTS, in the middle of the conserved TGEKP linker (Figure 1a). The role of the linker in increasing affinity for DNA has recently been elucidated (*26*, *31*). Since the T and G residues are present at the C-terminal end of the helix of finger 3 in the +KTS isoform as well as in the −KTS isoform, it is somewhat puzzling that the capping interaction is so thoroughly abrogated in the +KTS isoform.

A possible explanation for this resides in additional interactions, over and above those expected for classic α_L helix-capping motifs. It appears that the α_L C-caps in zinc finger−DNA complexes are supplemented by an interaction between the amide proton of the conserved glutamic acid residue of the linker and the O^{γ} of the conserved threonine (*31*). Such an interaction is present in the majority of published zinc finger−DNA structures. Given that this same TGEKP linker sequence is present between fingers 1−2 and fingers 3−4 in −KTS, but disrupted between fingers 3−4 in +KTS (Figure 1a), it seems likely that two *i*,*i*+2 Glu amide proton to Thr O^{γ} hydrogen bonds help to stabilize the −KTS/DNA complex; in contrast, only one such hydrogen bond (between fingers 1−2) is possible for the +KTS/DNA complex. The large difference in the Glu411 $^{13}C^{\alpha}$ chemical shift between DNA-bound −KTS and +KTS (Figure 5) is consistent with the disruption of the Glu411 to Thr406 O^{γ} hydrogen bond in the +KTS/DNA complex. In addition, the chemical shift of the Glu 411 amide proton is significantly shifted upon complexation of WT1 −KTS, consistent with the formation of the Thr O^{γ} hydrogen bond, but shows only a small shift on binding of the +KTS isoform. In fact, this additional set of interactions provides enthalpic compensation for the entropy loss incurred on going from the flexible free state of the protein to the DNA-bound form. The additional entropic cost of constraining the longer finger 3−finger 4 linker in the +KTS isoform is also deleterious to DNA binding. As a consequence, it appears that finger 4 remains largely unbound in the +KTS isoform (see later discussion).

It is still puzzling why the NH of the inserted lysine of the KTS sequence could not act equally well in the formation of the hydrogen bond with the Thr O^{γ}. It is doubtless significant that the residue at the third position of the linkers in DNA-binding zinc fingers is a conserved negatively charged amino acid (Glu, Asp, or in a few cases Gln), whereas the residue in the insertion is a highly hydrophobic (close to the backbone) or positively charged (toward the end of the side chain) amino acid. This conservation and its replacement with a different amino acid type in +KTS point to a specific charge interaction that needs to occur in the linker for efficient formation of the NH−O^{γ} hydrogen bond. These results imply that the "snap-lock" mechanism of stabilization of the zinc finger−DNA complex (*31*) involves many interactions, including but not limited to the C-capping of the helix of the zinc finger when it is followed by an appropriate linker sequence and another zinc finger.

We conclude that the presence of the extra residues in the +KTS isoform prevents formation of the required structure in the linker between fingers 3 and 4. The consensus TGEKP linker sequence gives rise to a highly specific structure in complexes with DNA (*26*, *31*), a structure that confers considerable stabilization on the complex by enabling specific protein−protein interactions between the adjacent

fingers. Structuring of the linker also positions the zinc fingers correctly for recognition of their target DNA sequences. These interactions are disrupted in the +KTS isoform because the additional linker residues cannot accommodate the required linker structure and because of the entropic cost of ordering the longer linker.

Does Finger 4 in the +KTS Isoform Bind to DNA? The DNA-binding properties in finger 4 appear to be isoform-specific. For the −KTS complex, a noticeable chemical shift difference, $\Delta\delta_{complex\ -\ free}$, corresponding to DNA base contact, is seen for the $^{1}H^{N}$ of R427, and significant DNA-induced chemical shifts are also observed for the $^{13}C\alpha$ of both Ala426 and Asp429 (Figure 5). However, these changes are absent for finger 4 of the +KTS/DNA complex. Other evidence points to a significantly lower affinity of finger 4 for the DNA: Figure 5 shows that the $^{15}N^{H}$ values for Thr404 and Thr406 in the +KTS/DNA complex are very similar to those of the free +KTS protein, which is characterized by flexible linkers between adjacent fingers. The absence of DNA-induced chemical shift changes suggests that finger 4 in the +KTS complex may in fact not be tightly bound to the DNA. This most likely accounts for the significantly lower affinity of the +KTS isoform for this particular DNA sequence (*14*−*16*).

CONCLUSION

WT1 may function as an activator or a repressor of gene expression (*28*), and, depending on the isoform and its localization, the protein may act both at the transcriptional and at the posttranscriptional level. The presence of different isoforms of the WT1 protein represents an efficient means of utilizing a single gene sequence for diverse, though related functions. We have identified numerous site-specific intramolecular and intermolecular interactions that occur both near and distant from the protein−DNA interface that stabilize the −KTS and +KTS complexes with DNA. Some of the interactions that occur in fingers 3 and 4 of the −KTS/DNA complex are abrogated, or at least greatly diminished, in the +KTS/DNA complex. This information sheds new light on the molecular basis for the different DNA-binding properties of the −KTS and +KTS isoforms of WT1, and can be used to advance understanding of their various roles in transcriptional regulation in vivo.

ACKNOWLEDGMENT

We thank Linda Tennant for technical assistance and Dr. Maria Yamout for assistance with cloning methods.

SUPPORTING INFORMATION AVAILABLE

Three tables providing data for both wt 1−4 and −XTS and +KTS complexes (3 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. 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.
2. Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H., and Bruns, G. A. P. (1990) *Nature 343*, 774−778.
3. Bruening, W., and Pelletier, J. (1996) *J. Biol. Chem. 271*, 8646−8654.
4. Reddy, J. C., Morris, J. C., Wang, J., English, M. A., Haber, D. A., Shi, Y., and Licht, J. D. (1995) *J. Biol. Chem. 270*, 10878−10884.
5. Moffett, P., Bruening, W., Nakagama, H., Bardeesy, N., Housman, D., Housman, D. E., and Pelletier, J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 11105−11109.
6. Englert, C., Vidal, M., Maheswaran, S., Ge, Y., Ezzell, R. M., Isselbacher, K. J., and Haber, D. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 11960−11964.
7. Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., and Housman, D. (1990) *Nature 346*, 194−197.
8. Armstrong, J. F., Pritchard-Jones, K., Bickmore, W. A., Hastie, N. D., and Bard, J. B. (1993) *Mech. Dev. 40*, 85−97.
9. Pelletier, J., Bruening, W., Kashtan, C. E., Mauer, S. 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 Houseman, D. (1991) *Cell 67*, 437−447.
10. Ricciardi, V. M., Sujansky, E., Smith, A. C., and Francke, U. (1978) *Pediatrics 61*, 604−610.
11. Barbaux, S., Niaudet, P., Gubler, M. C., Grunfeld, J. P., Jaubert, F., Kuttenn, F., Fekete, C. N., Souleyreau-Therville, N., Thibaud, E., Fellous, M., and McElreavey, K. (1997) *Nat. Genet. 17*, 467−470.
12. King-Underwood, L., Renshaw, J., and Pritchard-Jones, K. (1996) *Blood 87*, 2171−2179.
13. Silberstein, G. B., Van Horn, K., Strickland, P., Roberts, C. T., Jr., and Daniel, C. W. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8132−8137.
14. Rauscher, F. J., Morris, J. F., Tournay, O. E., Cook, D. M., and Curran, T. (1990) *Science 250*, 1259−1262.
15. Bickmore, W. A., Oghene, K., Little, M. H., Seawright, A., van Heyningen, V., and Hastie, N. D. (1992) *Science 257*, 235−237.
16. Drummond, I. A., Rupprecht, H. D., Rohwer-Nutter, P., Lopez-Guisa, J. M., Madden, S. L., Rauscher, F. J., III, and Sukhatme, V. P. (1994) *Mol. Cell. Biol. 14*, 3800−3809.
17. Caricasole, A., Duarte, A., Larsson, S. H., Hastie, N. D., Little, M., Holmes, G., Todorov, I., and Ward, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 7562−7566.
18. Duarte, A., Caricasole, A., Graham, C. F., and Ward, A. (1998) *Br. J. Cancer 77*, 253−259.
19. Larsson, S. H., Charlieu, J.-P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V., and Hastie, N. D. (1995) *Cell 81*, 391−401.
20. Haber, D. A., Sohn, R. L., Buckler, A. J., Pelletier, J., Call, K. M., and Housman, D. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9618−9622.
21. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
22. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) *Structure 4*, 1171−1180.
23. Meinnel, T., Mechulam, Y., and Blanquet, S. (1993) *Biochimie 75*, 1061−1075.
24. Casimiro, D. R., Toy-Palmer, A., Blake, R. C., II, and Dyson, H. J. (1995) *Biochemistry 34*, 6640−6648.
25. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
26. Wuttke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *J. Mol. Biol. 273*, 183−206.
27. Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L. L., Reymond, M. T., Chung, J., and Wright, P. E. (1998) *J. Biomol. NMR 12*, 51−71.
28. Reddy, J. C., and Licht, J. D. (1996) *Biochim. Biophys. Acta 1287*, 1−28.
29. Coppes, M. J., Campbell, C. E., and Williams, B. R. G. (1993) *FASEB J. 7*, 886−894.
30. Hastie, N. D. (1994) *Annu. Rev. Genet. 28*, 523−558.
31. Laity, J. H., Dyson, H. J., and Wright, P. E. (2000) *J. Mol. Biol. 295*, 719−727.
32. Aurora, R., Srinivasan, R., and Rose, G. D. (1994) *Science 264*, 1126−1130.
33. Schellman, C. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 53−61, Elsevier/North-Holland, New York.

BI9926678