Role for first zinc finger of WT1 in DNA sequence specificity: Denys–Drash syndrome-associated WT1 mutant in ZF1 enhances affinity for a subset of WT1 binding sites

Dongxue Wang[1], John R. Horton[2], Yu Zheng[3], Robert M. Blumenthal[4], Xing Zhang[2] and Xiaodong Cheng[1,2,*]

[1]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA, [2]Department of Molecular and Cellular Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA, [3]RGENE, Inc., 953 Indiana Street, San Francisco, CA 94107, USA and [4]Department of Medical Microbiology and Immunology, and Program in Bioinformatics, University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

ABSTRACT

Wilms tumor protein (WT1) is a Cys2-His2 zinc-finger transcription factor vital for embryonic development of the genitourinary system. The protein contains a C-terminal DNA binding domain with four tandem zinc-fingers (ZF1–4). An alternative splicing of *Wt1* can add three additional amino acids—lysine (K), threonine (T) and serine (S)—between ZF3 and ZF4. In the −KTS isoform, ZF2–4 determine the sequence-specificity of DNA binding, whereas the function of ZF1 remains elusive. Three X-ray structures are described here for wild-type −KTS isoform ZF1–4 in complex with its cognate DNA sequence. We observed four unique ZF1 conformations. First, like ZF2–4, ZF1 can be positioned continuously in the DNA major groove forming a 'near-cognate' complex. Second, while ZF2–4 make base-specific interactions with one DNA molecule, ZF1 can interact with a second DNA molecule (or, presumably, two regions of the same DNA molecule). Third, ZF1 can intercalate at the joint of two tail-to-head DNA molecules. If such intercalation occurs on a continuous DNA molecule, it would kink the DNA at the ZF1 binding site. Fourth, two ZF1 units can dimerize. Furthermore, we examined a Denys–Drash syndrome-associated ZF1 mutation (methionine at position 342 is replaced by arginine). This mutation enhances WT1 affinity for a guanine base. X-ray crystallography of the mutant in complex with its preferred sequence revealed the interactions responsible for this affinity change. These results provide insight into the mechanisms of action of WT1, and clarify the fact that ZF1 plays a role in determining sequence specificity of this critical transcription factor.

INTRODUCTION

Wilms tumor suppressor protein (WT1) is a Cys2-His2 zinc-finger (ZF) transcription factor vital to embryonic development of the genitourinary system (reviewed in (1–6)). Human WT1 contains an N-terminal region responsible for transcriptional regulation and for protein dimerization (7,8) and a C-terminal ZF array comprising four tandem fingers (Figure 1A). Alternative splicing between exons 9 and 10 of *Wt1* can add three amino acids—lysine (K), threonine (T) and serine (S)—between ZF3 and ZF4 of WT1 (Figure 1A). The last three fingers (ZF2–4) of the −KTS isoform are thought to be primarily responsible for DNA sequence-discrimination and target binding; ZF1 is reported to contribute to DNA binding affinity, but only in a relatively non-specific manner (9–11). The evidence is more clear that ZF1 is required for interaction with RNA by WT1, particularly in the +KTS isoform (12).

The consensus DNA binding sequence of WT1 is ∼10 bp, based on chromatin immunoprecipitation (ChIP) (13,14) (Supplementary Figure S1A). Most importantly, the ChIP studies did not reveal the consensus sequence motif for ZF1, except for the 5′ base of its putative triplet, suggesting that ZF1 might be flexible when WT1 is bound to DNA. Structural investigations of WT1 confirmed that ZF2–4 interacts DNA specifically, with each finger recognizing a 3-bp triplet sequence (11,15); again suggesting that ZF1 is a non-specific binder. However, ZF1 has all the features of a regular zinc

---

*To whom correspondence should be addressed. Tel. +1-713-834-6274; Email: xcheng5@mdanderson.org

Figure 1. The DDS mutant M342R has highest affinity for guanine in the 3′ triplet. (**A**) Human WT1 contains a C-terminal ZF DNA binding domain comprising four fingers in tandem. For the study described here, we used a fragment of WT1 containing ZF1–4 without KTS (the -KTS isoform). (**B**) DDS mutations in ZF1–3 that alter either the Cys2-His2 structural amino acids that coordinate the zinc ions and hydrophobic core (in red letter), or the sequence-recognition amino acids at the protein–DNA interface in each ZF (in black, blue and green). The information is extracted from the Human Gene Mutation Database (HGMD). (**C** and **D**) Binding affinities of normal WT1 (panel C) and M342R mutant (panel D) with oligos containing various base pairs in the 3′ triplet (position 'XYZ' in panel C). (**E** and **F**) Binding affinities of normal WT1 (panel E) and M342R mutant (panel F) with oligos containing a various base pair in the middle position of the recognition triplet for ZF2 (position 'X' in panel E).

finger unit and is predicted to bind a specific DNA sequence by a computational algorithm (16) (Supplementary Figure S1A). We sought to explore this apparent discrepancy.

Individuals with aberrant WT1 are invariably heterozygous, with copies of both normal and mutated *Wt1* genes and they exhibit a spectrum of unusual features typically early in life. Truncations of WT1 due to frame-shift or chain-termination mutations leads to pediatric renal malignancies termed Wilms tumors (17,18). Another class of mutations in *Wt1* causes Denys–Drash Syndrome (DDS) (4,5). These are predominantly missense mutations in the first three ZFs, most often clustered in ZF2 and ZF3, that

alter either the Cys2-His2 structural amino acids that coordinate the zinc ions, or the sequence-recognition amino acids at the protein–DNA interface (Figure 1B). DDS mutations result in an array of severe problems, including a high probability of Wilms tumor, mesangial sclerosis and early renal failure, gonadal dysgenesis and ambiguous or female genitalia in 46XY males (19). ZF3 is where the most common DDS mutation is found, changing arginine 394 to tryptophan (R394W) (20), but many additional DDS singletons have been documented (Figure 1B). Arginine 394 of ZF3 recognizes the conserved (3′) Gua of the central triplet. The R394W mutation was reported to abolish DNA bind-

ing (21–23), precluding structural analysis of its interactions with DNA. In ZF2, we recently examined another DNA base-interacting residue, glutamine 369 (Q369) and its three mutations to positively charged arginine, histidine or lysine (24). Unlike R394W, Q369 mutations are rare—each has been reported only once (25–27). These mutations alter the sequence-specificity of ZF2, increase its affinity for guanine (instead of adenine by the wild-type WT1) and for 5-carboxylated cytosine (an epigenetic form of cytosine) (24).

The role of ZF1 has been harder to define. In the +KTS isoform of WT1, ZF1 appears to play a significant role in interactions with specific RNA molecules (28), while in the more DNA-specific −KTS isoform, ZF1 has been characterized as playing no role in sequence specificity (see above). The C-terminal four-finger ZF array of WT1 is essentially identical in mammals, birds, amphibians and fish, whereas numerous differences occur in the N-terminal regulatory regions. More specifically, all four ZFs are highly conserved, in all vertebrate classes save the one most distant from Mammalia (Supplementary Figure S2). This high degree of conservation implies that all four ZFs are essential for normal vertebrate development and are essentially non-permissive for missense mutations.

The identification of DDS missense mutations located in the first ZF (29,30) suggest that it may participate in DNA binding and sequence-discrimination, and there is additional evidence to support this possibility. *In vitro* selection approaches revealed that ZF1 has a preference for Thy as the first base of its putative triplet (10) and, in a separate study, ZF1 was found to preferentially bind the ambiguous triplet (T/G)-(G/A/T)-(T/G) (9) in which cytosine was excluded as the central base. Nevertheless, previous NMR and X-ray crystallographic analyses of ZF1–4 in complex with an oligonucleotide failed to reveal sequence-specific contacts by ZF1, leading to the conclusion that ZF1 '…does not contribute significantly to binding specificity' (11).

We are investigating whether this discrepancy could be due to the oligonucleotide used in the previous structural analysis (T*C*T), which was based on the Egr1 consensus rather than the WT1 consensus [the three zinc fingers of Egr1 are nearly identical to ZF2–4 of WT1 (11,15)]. Here, we suggest that WT1 ZF1 might have higher affinity for the sequence of triplet 5′-T*G*T-3′ than for 5′-T*C*T-3′ used by Stoll *et al.* (11), and we further investigated a DDS-associated methionine 342 to arginine (M342R) mutant in ZF1 (30) (Figure 1B). M342 is one of the DNA base-interacting residues of ZF1, and the gain-of-function mutation from a hydrophobic residue to a positively charged arginine resulted in an increased DNA binding affinity, and an altered sequence preference for a 5′ Gua in the triplet 5′-*G*GT-3′, again supporting a role for ZF1 in determining WT1 DNA sequence specificity.

MATERIALS AND METHODS

Site-directed mutagenesis

GST-tagged human WT1 C-terminal four-finger array corresponding to residues 319–437 (UniProt: P19544.2) for the −KTS isoform (pXC1296) was mutated to generate M342R (pXC1586) by QuikChange site-directed mutagenesis kit (Stratagene). The corresponding wild-type (pXC1593) and

the mutant (pXC1598) for the +KTS isoform were generated in parallel (Supplementary Figure S1B). The mutants were verified by sequencing.

Protein expression and purification

Proteins used in this study were affinity-purified from *Escherichia coli* recombinants using a GST-tag, which was subsequently removed. All proteins were expressed in the *E. coli* strain BL21-CodonPlus(DE3)-RIL (Stratagene) and purified as described (15,24,31). Typically, 2 L of cultures were grown at 37°C to log phase (OD_{600} = 0.5–0.8) and then shifted to 16°C. $ZnCl_2$ was added to a final concentration of 25 μM, expression was induced by the addition of isopropyl-ß-D-1-thiogalactopyranoside to 0.2 mM and the cultures were incubated overnight at 16°C. Cells were harvested by centrifugation, resuspended in lysis buffer containing 20 mM Tris–HCl (pH 7.5), 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 1 mM phenylmethylsulphonyl fluoride (PMSF) and 25 μM $ZnCl_2$, and then lysed by sonication. Lysates were mixed with polyethylenimine (Sigma) to a final concentration of 0.3% (w/v) and clarified by centrifugation at 16 500 rpm (31). Cleared extracts were loaded onto a glutathione-Sepharose 4B column (GE Healthcare) pre-equilibrated with the lysis buffer. The GST fusion proteins were eluted with 20 mM glutathione (GSH) in the elution buffer containing 100 mM Tris-HCl (pH 8.0), 5% (v/v) glycerol, 25 μM $ZnCl_2$ and 250 mM NaCl. The GST tag was removed using PreScission protease (purified in-house), leaving five additional N-terminal residues (GPLGS). The proteins were loaded onto HiTrap-SP column (GE Healthcare) and were eluted using a linear gradient of NaCl from 250 mM to 1 M. Finally, the pooled protein was concentrated and loaded onto a size exclusion column (Superdex 200 10/300 GL; GE Healthcare) and eluted as a single peak in 500 mM NaCl, 20 mM Tris-HCl (pH 7.5), 5% (v/v) glycerol, 0.5 mM TCEP and 25 μM $ZnCl_2$. Final protein concentrations were estimated by absorbance at 280 nm (absorbance coefficient of 9970 for 1 mM WT1). The protein yields were estimated to be ∼4 mg per liter of culture for the −KTS isoform.

Fluorescence-based DNA binding assay

Fluorescence polarization was used to measure the dissociation constant (K_D) between these binding domains and double-stranded oligonucleotides (oligos) bearing fluorescent 5′-FAM labels. Fluorescence polarization measurements were carried out at 25°C on a Synergy 4 microplate reader (BioTek). The 6-carboxy-fluorescein (FAM)-labeled dsDNA probe (5 nM) was incubated for 10 min with increasing amounts of protein in 20 mM Tris–HCl (pH 7.5), 5% (v/v) glycerol, 0.5 mM TCEP and 300 mM NaCl. Curves were fit individually using GraphPad Prism 5.0 software (GraphPad Software, Inc.). Binding constants (K_D) were calculated as [mP] = [maximum mP] × [C]/(K_D + [C]) + [baseline mP] and saturated [mP] was calculated as saturation = ([mP] − [baseline mP])/([maximum mP] − [baseline mP]), where mP is millipolarization and [C] is protein concentration. Curves were normalized as percentage of bound

and reported is the mean ± SEM of the interpolated K_D from two independent experiments performed in duplicate.

Crystallography

We crystallized WT1 in the presence of DNA by the sitting-drop vapor diffusion method at 16°C using equal amounts of protein–DNA mixtures (0.5 mM) and well solution (Supplementary Tables S1 and S2). Protein–DNA mixtures in equimolar ratios were incubated for 30 min at 16°C before crystallization. Crystals were cryo-protected by soaking in mother liquor supplemented with 20% (v/v) ethylene glycol before plunging into liquid nitrogen. X-ray diffraction data were collected at 100K at the SER-CAT beamlines (22BM-D and 22ID-D) at the Advanced Photon Source, Argonne National Laboratory and processed using HKL2000 (32). Initial crystallographic phases were determined by molecular replacement using the coordinates of the three-finger WT1 ZF2–4 (PDB ID: 4R2E) as a search model. Phasing, molecular replacement, map production, and model refinement were performed using PHENIX (33,34); manual manipulation and any additional building was completed with the program COOT (35,36). The statistics were calculated for the entire resolution range (Supplementary Tables S1 and S2). The R_{free} and R_{work} values were calculated for 5% (randomly selected) and 95%, respectively, of the observed reflections. Molecular graphics were generated using PyMol (DeLano Scientific, LLC).

Analysis of ChIP profiles within WT1-binding sites

To build WT1-binding consensus sites from the published Chip-chip study in mouse embryonic kidney tissue (13), genomic coordinates from 1663 WT1-binding regions were extracted from their Supplementary Table S4. To identify the genomic locations of the WT1 consensus sites, genomic sequences of those regions from UCSC mouse mm9 genome were used as input to the motif-finding program MEME Suite (37). Similarly, genomic coordinates of WT1-binding regions from the Chip-seq study in embryonic kidneys (GSE58073) (14) were downloaded. Genomic sequences of these regions from mouse mm10 genome were used as input to MEME to locate WT1 consensus binding sites.

RESULTS

The M342R mutant has highest affinity for guanine in the 3′ triplet

In conventional C2H2 ZF proteins, each finger comprises two β strands and a helix (38). Characteristically, two histidines in the helix together with one cysteine in each of the β strands coordinate a zinc ion, forming a tetrahedral C2–Zn–H2 structural unit that confers rigidity to fingers. The amino acids occupying key 'canonical' positions of the helix and the preceding loop specify a DNA target sequence of three adjacent DNA base pairs (39,40), which we call a triplet element. The potential base-interacting residues in ZF1 are Lys336, His339 and Met342 (from N to C termini in Figure 1B), which could recognize DNA in a linear polarity from 3′ to 5′, i.e. corresponding to the 3′ base, middle

and 5′ base of putative triplet element (as illustrated in Figure 1C). In a simple code of DNA recognition specificity, described based on C2H2-ZF selection and structural data (39,41), Met342 has a thymine (T) preference and His339 has a guanine (G) preference. This is in general agreement with previous experimental observations that ZF1 has affinity for thymine as the 5′ base (10) and cytosine was excluded as the middle base of its putative triplet (9). Histidine also occurs naturally at the corresponding position in ZF3 of WT1 (H397), which also recognizes G at the center of triplet 2. We first compared the binding affinities of the −KTS isoform of WT1 to duplex oligos containing GGT, TGT or TCT, or CCC in the 3′ triplet (Figure 1C). WT1 showed approximately equal binding affinity to guanine as the middle base (GGT and TGT), ∼4-fold weaker binding to cytosine (TCT) and ∼12-fold weaker binding to CCC (a nonselective sequence by ZF1 in a binding site selection assay (9)).

In further binding-affinity experiments, we examined the M342R mutant associated with DDS. While the WT protein preferred Thy, the M342R protein bound most strongly to G as the 5′ base, preferring GGT to TGT by a factor of approximately 8× (Figure 1D). This is expected because juxtaposition of Arg with Gua is a common mechanism for guanine recognition (41,42). We further validated that the mutation in ZF1 does not affect the preference of neighboring ZF2 for adenine as the middle base of its GAG triplet, as GAG binds slightly (1.5–2×) better than when it is GCG (Figure 1E and F). This finding is in agreement with previous reports (10,22,24) as well as the anti-WT1 ChIP studies showing that A and C are the most frequent bases in the middle position of the ZF2-recognition triplet (13,14) (Supplementary Figure S1A).

Multiple conformations of ZF1 from wild-type WT1 in protein–DNA interactions

To illustrate the potentially different conformations in DNA binding between wild-type ZF1 and the DDS-associated mutant M342R, we determined the co-crystal structures of the −KTS isoform of ZF1–4 of normal WT1 and M342R mutant in complex with oligos containing either TGT or GGT in the 3′ triplet (Supplementary Table S1). The structures were solved at the resolution range of 1.55–2.7 Å (Supplementary Table S1). The protein and DNA components of all complexes involving ZF2–4 were structurally similar, with a root-mean-squared deviation (RMSD) of 0.3 Å. Here, we focus our discussion on ZF1 and its interaction with the 3′ triplet (Figure 2).

We crystallized the wild-type WT1 (−KTS) in three crystallographic space groups (*P*1, *P*2₁2₁2 and *P*6₃). First, in the *P*1 space group, the crystallographic asymmetric unit contains two WT1–DNA complexes (Mol A and Mol B in Figure 2A). Besides ZF1, the two complexes have non-crystallographic 2-fold symmetry (Figure 2B), and are highly similar with RMSD of less than 0.9 Å when comparing 85 pairs of Cα atoms (Figure 2C). When the two complexes are superimposed, the conformations of ZF1 are ∼180° rotation apart (Figure 2C), which could be achieved via a rotation of the main-chain torsion angle between E350 and K351 in the 7-residue linker (Figure 2D). In molecule
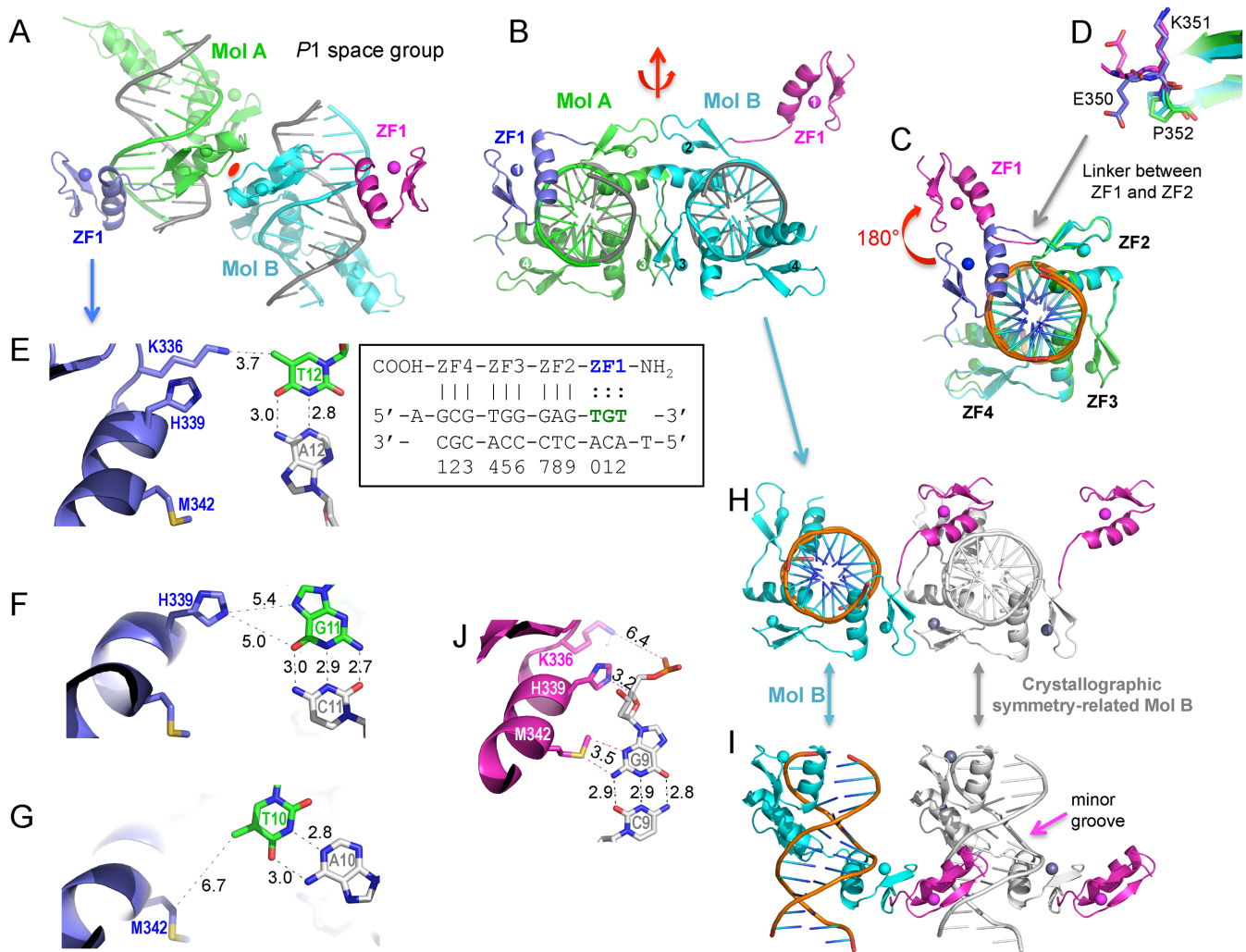
Figure 2. ZF1 could bind in DNA major groove or minor groove. (**A** and **B**) Two orthogonal views of two protei–DNA complexes (green and blue) in the *P*1 space group. ZF2–4 of Mol A (green) and Mol B (cyan) are related by a pseudo 2-fold symmetry as indicated by red symbols. (**C** and **D**) Superimposition of two molecules indicates conformational change of ZF1 by near 180° rotations along the backbone between E350 and K351 of the linker between ZF1 and ZF2 (panel D). (**E**–**G**) Three DNA base-interacting residues of ZF1 (blue) in Mol A form week interactions with TGT sequence in DNA major groove. (**H** and **I**) ZF1 of Mol B interacts with a symmetry-related, second DNA molecule via its minor groove. (**J**) The corresponding three residues of ZF1 (magenta) in Mol B form interactions with the G9 base, sugar moiety and phosphate group, respectively.

A, ZF1 is positioned in the DNA major groove, but the three side chains of canonical DNA interacting residues are a little too far away (>3.7 Å) to make base-specific hydrogen bonds of the conforming triplet TGT sequence (Figure 2E–G). We term the Mol A-DNA as the 'near-cognate' complex.

In molecule B, ZF1 swings completely away from the DNA molecule bound by ZF2–4, and reaches and inserts itself into the minor groove of the neighboring DNA molecule (Figure 2H and I). The same set of three 'base-interacting' residues switches to interact with guanine G_9 of the neighboring DNA molecule (Figure 2J): M342 makes van der Waals contacts to the guanine base from the minor groove side, the imidazole ring of H339 stacks with the ribose ring and K336 interacts weakly with the backbone phosphate group. In essence, Mol B associates with two DNA molecules: ZF2–4 assemble canonical base-specific interactions in the major groove of one DNA molecule,

whereas ZF1 interacts with a second DNA molecule via the minor groove side. If the two DNA molecules were connected, they could represent two different regions of a long DNA molecule, bridged by MolB.

Second, in the space group $P2_12_12$, the ZF1 of Mol C intercalates between the joint of two DNA molecules tail-to-head (Figure 3A and B). H339 stacks with A_{12} of one DNA molecule and K336 stacks with G_1 of the next DNA molecule, while M342 makes van der Waals contact with the A_{12} base (Figure 3B). If the intercalation does occur for a long, continuous DNA molecule, it would cause the DNA to be kinked at the ZF1 binding site. In the superimposition of Mol A and Mol B in the *P*1 space group and Mol C in the $P2_12_12$ space group, the ZF1 of Mol C has a conformation more closely related to that of Mol A (Figure 3C). The ability of ZF1 to stack between two DNA molecules had been noted previously (11), where two WT1 ZF1–4–DNA complexes stacked tail-to-tail (PDB ID: 2PRT; Supplemen-
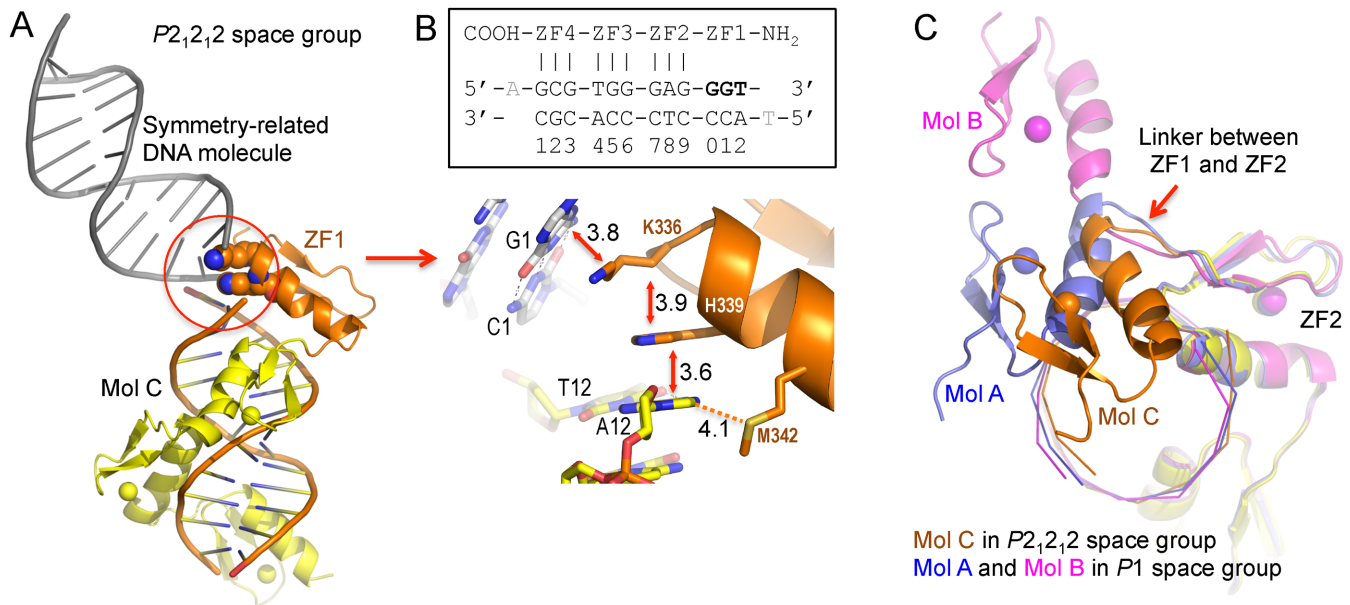
Figure 3. ZF1 stacks between two DNA molecules. (**A**) In the $P2_12_12$ space group, ZF1 of Mol C attaches at the joint of two tail-to-head DNA molecules. (**B**) The spacing distance between G1-K336-H339-A12 is equivalent to one DNA base pair helical rise. (**C**) Superimposition of Mol A, Mol B and Mol C indicates multiple conformations of ZF1 upon DNA binding.

tary Figure S3A and B). Comparing the structure of PDB ID: 2PRT to that of molecules A, B and C revealed yet another conformation of ZF1 with a switch point at the same linker residues between ZF1 and ZF2 (Supplementary Figure S3C). We note that, in the WT1 (−KTS) isoform, the three linker regions between ZFs (ZF1 and ZF2, ZF2 and ZF3, and ZF3 and ZF4) are identical in size (7 residues) and nearly conserved in composition (TG-E/V-KP-Y/F-Q/S). Thus we suggest that the ZF1 conformation in relation to ZF2–4 is not due to intrinsic features of inter-finger interactions, but rather due to the weakened binding of ZF1 to DNA that permits altered conformations.

ZF1 mediated self-association

Third, we co-crystallized ZF1–4 in complex with DNA in yet another space group $P6_3$, where two complexes were observed in the crystallographic asymmetric unit. Both complexes have their ZF1 swung out, like Mol B in the $P1$ space group. However, the two ZF1 units self associate (Figure 4A), with an interface of ∼510 Å², via expansion of two-strand β sheet into a four-strand β sheet (Figure 4B). In addition, a hydrophobic interface involving F323 of strand β1, L337 and L340 of the helix (Figure 4C), and M324 and A326 on the opposite side of β sheet (Figure 4B) further enhance the strength of inter-domain interaction. M324, A326 and L337 are not conserved and are substituted with polar and charged residues in the other ZFs (Figure 4D), implying that self-association might be unique to ZF1. Furthermore, M324 (just to the amino acid of ZF1) and L337 are conserved in all six classes of vertebrates, while A326 is conserved in four of them (the other two have a Val) (Supplementary Figure S2).

M342R DDS mutant allows ZF1 to form a cognate complex with DNA

Next, we crystallized the DDS-associated M342R protein with the GGT-containing oligo (Figure 5A), as M342R binds the 3′ GGT triplet with exceptionally high affinity (Figure 1D). With this mutant, ZF1 is now in the major groove as observed for ZF2–4 (Figure 5B), with its side chains in the DNA-interacting helix making base-specific interactions. R342 forms two hydrogen bonds with the *N7* and *O6* atoms of G_{10} (Figure 5C), a bonding pattern specific to guanine (42–44). The imidazole ring of H339 forms one hydrogen bond with *N7* atom of G_{11} via the Nε2 ring atom (Figure 5D). The aliphatic carbon atom Cε of K336 makes a van der Waals contact with the methyl group of T_{12}, while its terminal positively charged amino group interacts with a DNA backbone phosphate group (Figure 5D and E). Superimposition of this cognate structure of M342R-DNA complex with that of near-cognate complex, Mol A in the $P1$ space group, revealed a small (∼3 Å), but crucial shift of ZF1 toward DNA (Figure 5F and G). The largest conformational change lies in the side chains of M342-to-R substitution: M342 points away from the DNA, while R342 rotates its positively-charged guanidine group toward the paired guanine (Figure 5F).

The ability of the M342R mutant of WT1 (−KTS) to bind a 12-bp specific sequence, in contrast to the usual 9-bp sequence, raised the question of whether the mutant could bind more strongly to a subset of wild-type WT1 binding sites. Using previously published datasets of WT1 ChIP coupled to mouse promoter microarray (ChIP–Chip), with chromatin prepared from embryonic mouse kidney tissue (13) and ChIP-seq analysis of kidneys dissected from E18.5 mouse embryos (14), we extracted 295 and 1978 binding sites, respectively, based on stringent standards [enrichment
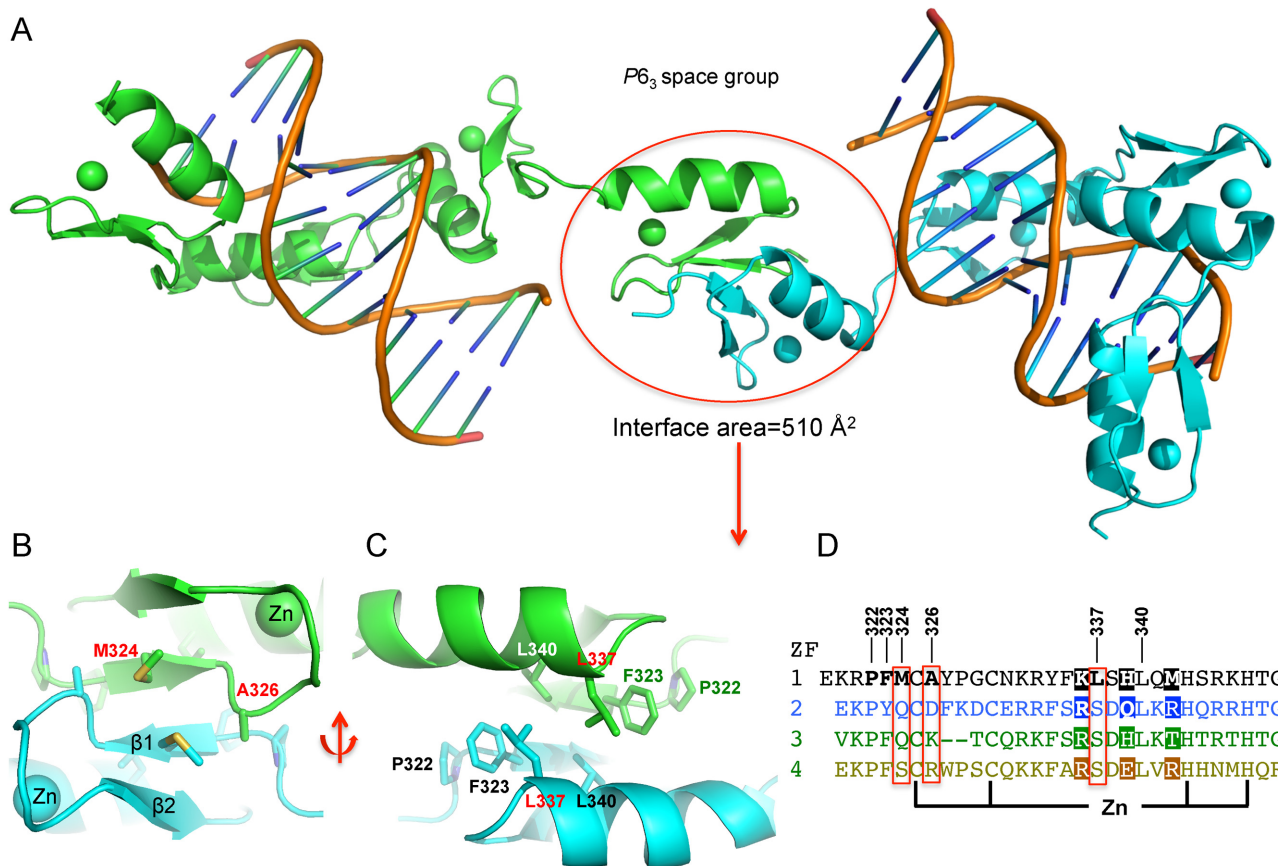
Figure 4. ZF1 mediated self-association. (**A**) In the *P*6₃ space group, two ZF1 units dimerize. (**B**) The dimerization resulted in a four-strand β sheet. (**C**) The additional hydrophobic interactions formed between two ZF1 units. (**D**) Sequence alignment of ZF1–4 indicates the hydrophobic residues are unique to ZF1.

fold >5 for the ChIP–chip dataset and irreproducibility discovery rate (IDR) <0.01 for the ChIP-seq dataset]. Motif discovery and searching analysis (37) yielded binding motifs that closely followed the published ones (Figure 6A and B). We next asked how many binding sites had GG at positions 10 and 11 (corresponding to the ZF1 binding triplet), finding 40% (120 sites) and 24% (471 sites) from the ChIP–chip and ChIP-seq datasets, respectively (Figure 6C and D). The frequency of GG at position 10 and 11 is significantly higher than expected (6.25%), which suggests a binding preference of ZF1 for these sites. However, only a small number (thirteen) of these binding sites overlap between the two datasets. Among the common 13 binding sites, 8 of them fall into the upstream 5 kb regions of Refseq genes (Figure 6E): one of these genes codes for transcription factor MafB, which is a confirmed WT1 target gene and is required for normal development of embryonic kidney in zebrafish (45), while three others specify a cluster of microRNAs.

M342R in the context of +KTS isoform

As mentioned, all known isoforms of WT1 include four ZFs at the C terminus with or without three extra amino acids (KTS) between ZF3 and ZF4 (46,47). Up to this point, analyses have focused exclusively on the −KTS isoform. Here,

we also expressed and purified the wild-type and M342R mutant in the context of the +KTS isoform.

First, we compared the binding affinities of the two wild-type isoforms. The +KTS isoform has ∼6-fold reduced binding affinity for the tested oligo, compared with the −KTS isoform (Figure 7A). This reduced affinity might result from increased linker flexibility due to the additional three amino acids between ZF3 and ZF4, leading to decreased DNA binding by ZF4 (48). In contrast, the three extra amino acids do not appear to affect the preference of neighboring ZF2 for adenine as the middle base of its triplet, with GAG bound slightly (∼2×) better than when it is GCG (Figure 7A), as was seen for the −KTS isoform (Figure 1E and F). Furthermore, the three amino acids do not appear to affect ZF1: as with the −KTS isoform (Figure 1C), ZF1 in the +KTS isoform demonstrated the strongest binding affinity for GGT, gradually decreasing affinity with TGT and TCT, and the weakest binding to CCC (Figure 7B).

Second, it seems plausible that the +KTS isoform has affinity for a length variant of the WT1 consensus, in which the corresponding triplets bound by ZF4 and ZF3 are separated by one or more additional base pairs to compensate for the increased length of the linker; for example, GCG-N₁₋₃-TGG-GAG-TGT (Figure 7C). There is precedent for such changes among the sequence-specificity sub-
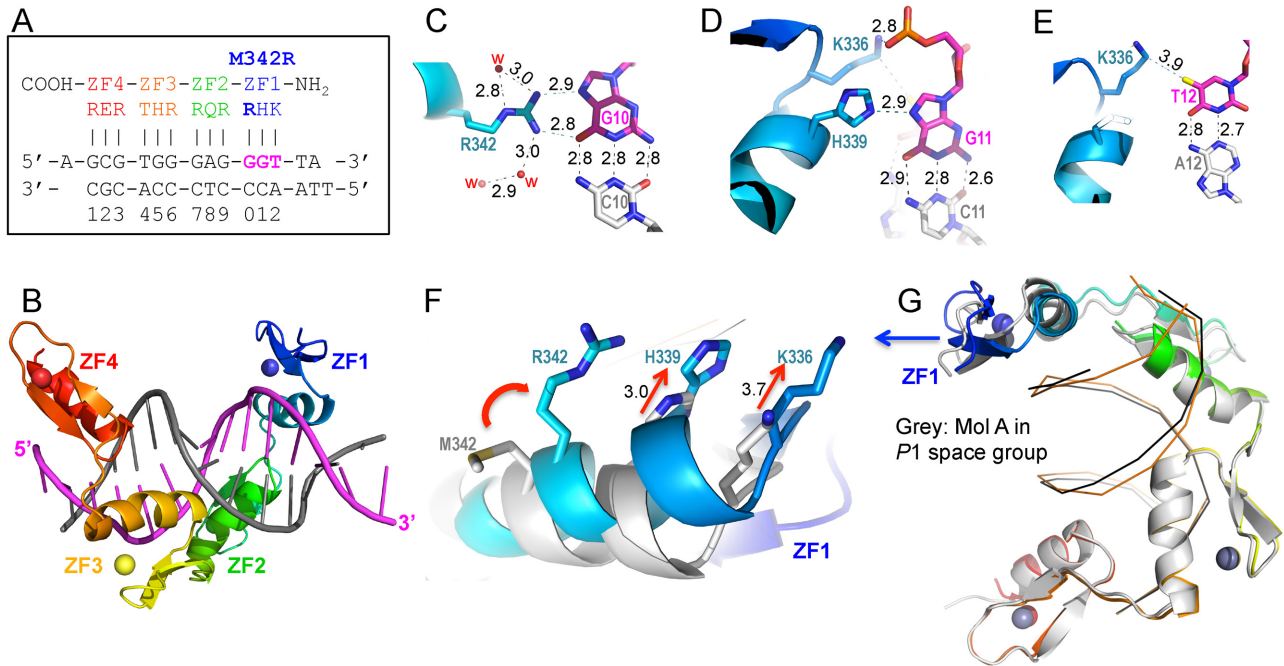
Figure 5. M342R DDS mutant allows ZF1 to form a cognate complex with DNA. (**A**) The oligo containing 3′ GGT triplet used in the co-crystallization. (**B**) Structure of ZF1–4 M342R in complex with cognate DNA. (**C**) R342 interacts with G10. (**D**) H339 interacts with G11 and K336 interacts with the phosphate group. (**E**) K336 interacts with T12. (**F** and **G**) Superimposition of M342 (wild-type) and R342 (mutant) shows the shift of ZF1 toward DNA and side chain conformation change of M342R mutation.
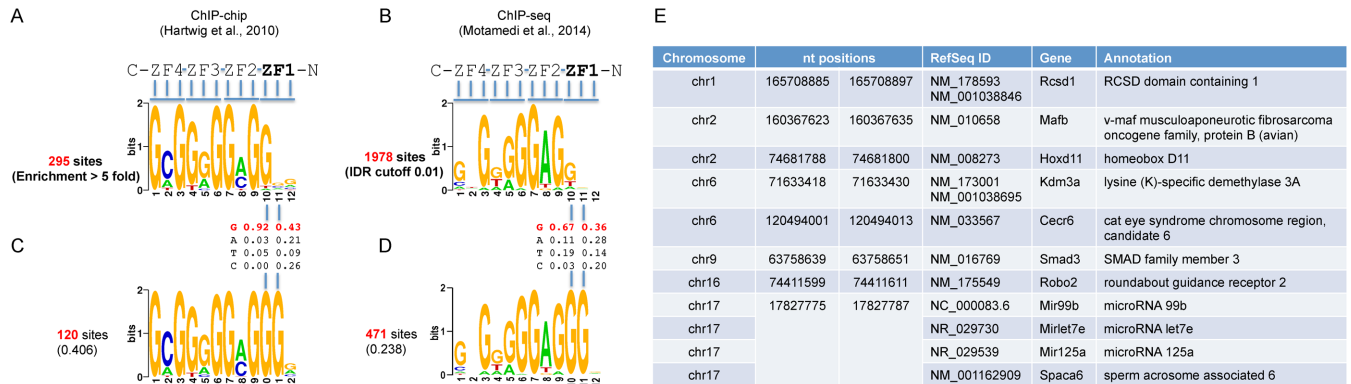

Figure 6. Examples of WT1 binding sites containing 3′ GG sequence for ZF1. (**A** and **B**) Number of WT1 binding sites, 295 and 1978, respectively, based on enrichment fold >5 for the ChIP–chip dataset (13) and IDR rate <0.01 for the ChIP-seq dataset (14). (**C** and **D**) Percentage of distribution of G, A, T and C at positions 10 and 11, and associated motifs with GG at positions 10 and 11. (**E**) Examples of genes associated with WT1 binding site containing 3′ GG sequence for ZF1.

units of certain restriction enzymes, where insertion or deletion of four amino acids between pairs of DNA-binding domains increases or reduces the separation between the sequences recognized by one base pair (49). Analyses of natural WT1+KTS binding sites in the insulin-like growth factor 2 (Igf2) gene (50), and recent ChIP-seq data of expressing biotinylated WT1+KTS isoform in leukemic K562 cells (51) (Supplementary Figure S5), cast some doubt on this possibility. Nevertheless, based on the analogy to the restriction enzymes noted above, we introduced 1–3 A:T base pairs between triplets 1 and 2 in the G:C rich sequence (Figure 7C). Interestingly, the +KTS has ∼3-fold increased affinity, whereas the −KTS isoform has more than 3-fold decreased affinity (as expected) (Figure 7C and D), regardless of the extent of increased length (1–3 nt) of the non-specific spacer. When we introduced the M342R mutant into the +/−KTS isoforms, and measured binding of the preferred triplet GGT sequence, the overall affinity again increased with increased spacer length, but to a lesser degree (<2-fold) (Figure 7E and F).

We next attempted to co-crystallize the wild-type and the mutant in the +KTS context with DNA oligonucleotides. We did not observe any indication of crystal formation for the wild-type +KTS form, probably because both ends (ZF1 and ZF4) are more flexible and thus resistant to crystallization. Fortunately, we did observe crystals of M342R
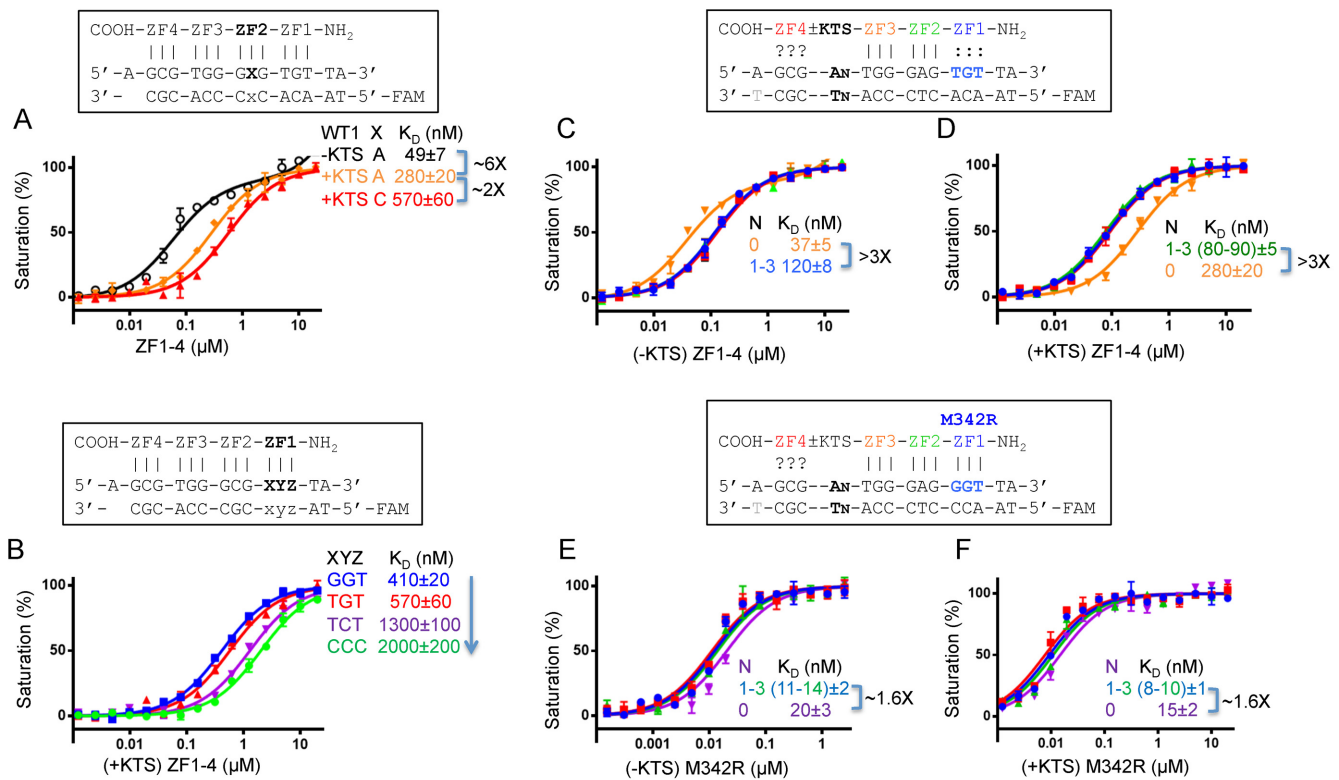
Figure 7. Binding affinities of wild-type ad M342R mutant in +KTS isoform. (**A** and **B**) Comparison of ±KTS isoforms with variation of sequence in the triplet recognized by ZF2 (panel A) or by ZF1 (panel B). (**C** and **D**) Binding affinities of –KTS isoform (panel C) and +KTS isoform (panel D) against oligos with 1–3 A:T base pair insertion between the two triplets recognized by ZF3 and ZF4, respectively. (**E** and **F**) Binding affinities of M342R mutant in –KTS isoform (panel E) and +KTS isoform (panel F) against oligos with 1–3 A:T base pair insertion.

in the +KTS form, screening against a set of oligos with increased spacer length by 1–3 A:T base pairs. Under three distinct conditions, the crystal diffracted X-rays sufficiently to permit us to collect usable diffraction data, resulting in three datasets with resolutions of 3.1 Å or lower and one of 1.85 Å (Supplementary Table S2). Further analysis revealed that the three low-resolution structures contained only ZF1–3 (residues immediately after Thr406 including ZF4 are disordered), while the high-resolution structure contained all four fingers (Figure 8A and B). As with the M342R mutant in the –KTS isoform, ZF1–3 are located in the DNA major groove and make base-specific contacts with their respective triplets, while ZF4 is located on the minor groove side with higher temperature-dependent atomic vibrations or static disorder in the crystal lattice. Superimposing the two M342R structures (±KTS) revealed that ZF4 swings nearly 180° from base-specific binding in the major groove (–KTS isoform) to the minor groove side (+KTS isoform; Figure 8C and D). This is reminiscent of the 180° rotation of ZF1 under some conditions (Figure 2C), indicating that the two terminal ZFs have some structural flexibility. Unfortunately, the seven-residue linker between ZF3 and ZF4, starting from the KTS sequence itself, was completely disordered in the current +KTS structure, further supporting the notion that the additional amino acids lead to increased linker flexibility resulting in the extreme C-terminal end ZF4 being more flexible than the other ZFs in this isoform.

DISCUSSION

Role of ZF1 in WT1 DNA specificity

Despite an enormous amount of work on the biochemistry, genetics, and physiology of WT1 spanning several decades, much remains to be discovered. Here we show that human ZF1, which has all the features of a regular zinc finger unit, could adopt multiple conformations, from DNA major groove binding, minor groove binding, and intercalation, to self-association. WT1 had been suggested to bind DNA as a dimer rather than a monomer (8,52,53) and recent ChIP-seq data found two similar motifs close to each other in many WT1-bound peaks (45). In contrast to the ZF1-mediated dimerization reported here (Figure 4), two regions N-terminal to ZF1 had previously been identified as being responsible for self-association (53). Additional data will be required to settle this point. We reanalyzed recently-published ChIP data in leukemic K562 cells expressing biotinylated WT1-KTS or WT1+KTS isoforms (51) (Supplementary Figure S5). While the consensus binding pattern of highly enriched WT1–KTS sites is in agreement with previously published motifs, the WT1+KTS binding sites revealed two weakly-related repeats, each containing six base pairs potentially occupied by ZF2 and ZF3 (Supplementary Figure S5). It is possible that the paired binding sites, which doubled the target sequence from 6 bp to 12 bp, are long enough to allow dimerized WT1+KTS to bind with sufficient affinity to be detected by ChIP.
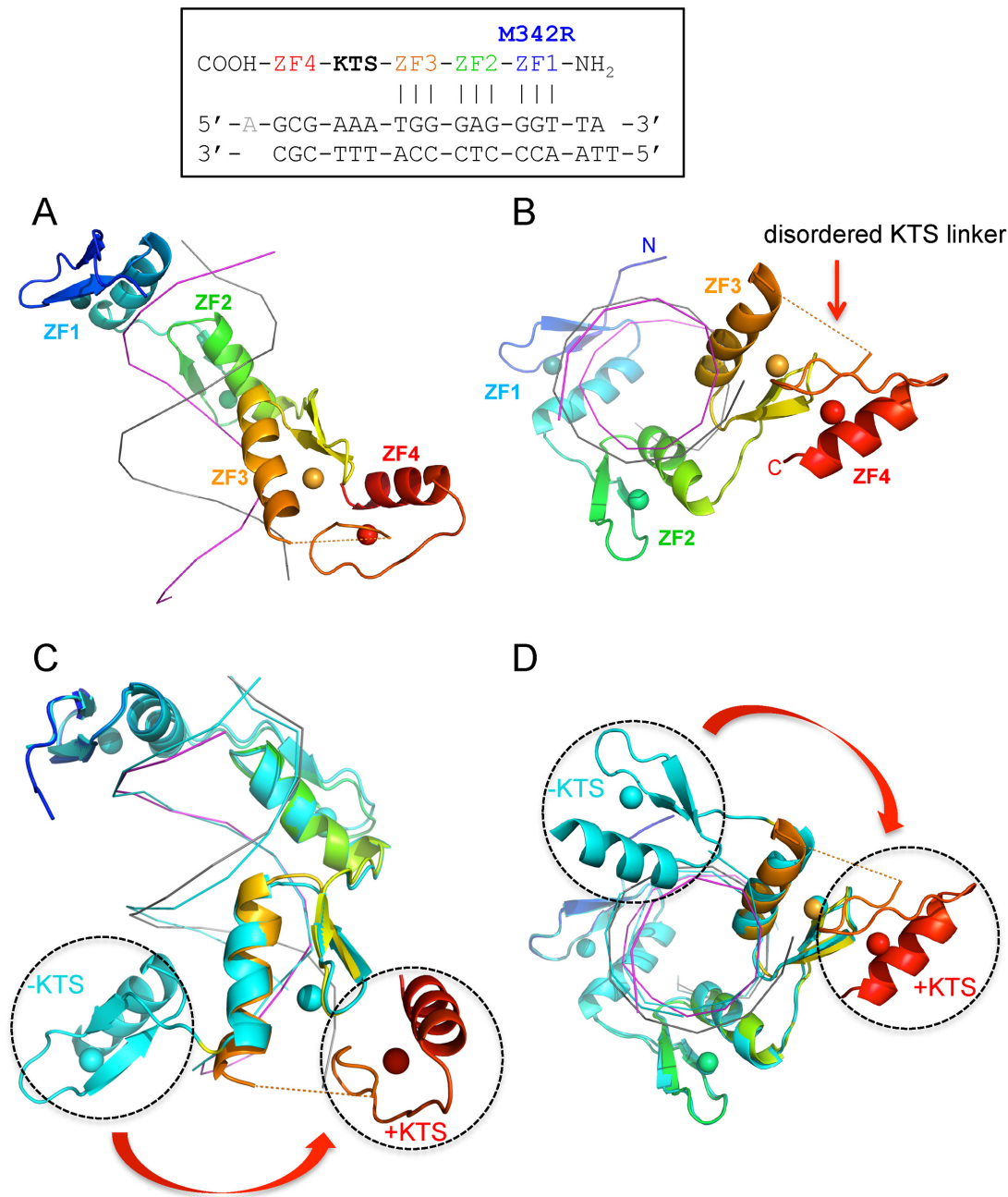
```
                              M342R
COOH-ZF4-KTS-ZF3-ZF2-ZF1-NH₂
              ||| ||| |||
5'-A-GCG-AAA-TGG-GAG-GGT-TA -3'
3'-  CGC-TTT-ACC-CTC-CCA-ATT-5'
```


Figure 8. Structure of M342R mutant in the +KTS isoform. (**A** and **B**) Two orthogonal views of ZF1–4. The +KTS linker region between ZF3 and ZF4 is disordered in the structure. (**C** and **D**) Superimposition of two M342R mutant structures in the −KTS isoform (uniformly cyan color) and +KTS isoform (colored in blue, cyan, green and red for ZF1–ZF4).

It is interesting that the same set of protein residues (K336, H339, and M/R342) could participate in base-specific H-bond interactions, non-specific van der Waals contacts, and phosphate interactions. A similar situation has been observed with *E. coli lac* repressor (LacI) DNA binding domain, and bacteriophage T4 DNA methyltransferase (T4Dam), where the same set of protein residues can switch, from an electrostatic interaction with the DNA backbone in a non-specific complex, to a specific binding mode with DNA base pairs in the cognate complex (54,55). These findings suggest that ZF1 detects local variations in

DNA shape (minor versus major grooves) and electrostatic potential.

WT1, particularly the −KTS isoform, has the ability to bind both DNA and RNA (56); for example, the gene promoter encoding Igf2 and specific Igf2 exonic RNA sequences (12). This is analogous to the case of transcription factor TFIIIA, which recognizes both the 5S rRNA gene promoter (57) and also binds its gene product (5S rRNA) (58). More recently, WT1 was shown to bind 3′ untranslated regions (UTRs) of developmental target RNAs (59). Interestingly, the DNA base-stacking interaction between

two DNA molecules by WT1 ZF1 (Figure 3) is somewhat similar to what had been observed for TFIIIA, where amino acid residues recognize individual RNA bases positioned in intricately-folded loop regions of the RNA (58) (Supplementary Figure S4).

DDS mutations are generally considered to abolish DNA binding by WT1 (11,22,23,26,53). However, we find that the substitutions of hydrophobic/polar to positively-charged residues at base-interacting positions 342 of ZF1 (M342R), and 369 in ZF2 (Q369R/H/K) (24), continue to bind DNA very well indeed; only now they bind to different sequences instead of, or in addition to, the original sequence. The predominant specificity of ZF1 DDS-associated mutant M342R is G-G-T, whereas ZF1 of normal WT1 has the more relaxed specificity T/G-G-T. This clearly supports a role for ZF1 in determining WT1 DNA sequence specificity in both the −KTS and +KTS variants.

Role of WT1 M342R in the pathology of Denys–Drash syndrome

The phenotypes associated with the changed specificity in the DDS-associated WT1 ZF1 mutant might be due to loss of binding to the subset of TGT sites, increased binding to the subset of GGT sites, or even something more subtle, such as the more stringent GGT specificity reducing generalized binding to DNA and raising the functional concentration of WT1 (60,61). Further, many apparently-opposing activities have been ascribed to WT1 (reviewed in (46)), including (but not limited to) transcriptional activation and repression (62); tumor suppression for Wilms tumor and an oncogenicity for adult tumors (63–66); a role in controlling both active and repressive histone modification marks (67); and a capacity to differentially bind to epigenetically modified DNA (15). It is possible that these opposing activities of WT1 maintain a balance and the potential to transition in either direction. Mutations in the *WT1* gene, or altered expression, could perturb this balance and lead to disease.

Among the small set of genes associated with the sequence preferred by the DDS-mutant M342R in ZF1 (Figure 6E), MafB is in the center of a WT1-dependent transcription factor network in control of podocyte gene expression (45,68) [podocytes are pericytes that help to form the glomerular basement membrane in developing kidneys (69)]. The *Rcsd1* gene codes for a protein kinase substrate (70) and is involved in cell ability to remodel actin filament assembly and remodeling of the actin cytoskeleton, which is an important step in mitosis. Interestingly, actin itself was identified as a WT1 interaction partner both in the nucleus and in the cytoplasm (71). Human *RCSD1* is fused to *ABL1* in a translocation-associated B-cell acute lymphoblastic leukemia (72). The resulting chimeric protein could result in an alteration of the cellular function by affecting cytoskeleton regulation, which could be an important step in leukemogenesis (73). HOXD11—Homeobox D11—is part of a developmental regulatory system that provides cells with specific positional identities on the anterior-posterior axis (74). In addition, the human HOXD11 and HOXD12 cluster is a Polycomb-dependent regulatory region important in embryonic stem cell differentiation (75). KDM3A (also known as JHDM2A or JMJD1A) is a histone demethylase that

physically interacts with the androgen receptor (AR) to up-regulate AR target gene expression through the demethylation of methylated histone H3 lysine 9 (76), a modification that is generally associated with transcriptional repression. More recent studies have revealed that histone lysine demethylases such as KDM3A play important roles in renal cell carcinoma, the most common kidney cancer, via hypoxia-mediated angiogenesis pathways (77–79). Smad3 is an intracellular signal transducer and transcriptional modulator activated by transforming growth factor (TGF)-β (80). On the molecular level, TGF-β/Smad3 signaling pathway plays a central role in fibrotic kidney disease (81).

Robo2 is a transmembrane receptor for Slit2, which are thought to act as molecular guidance cue in cellular migration, including axonal navigation at the ventral midline of the neural tube and projection of axons to different regions during neuronal development. Mouse mutants lacking either Slit2 or its receptor Robo2, molecules known primarily for their function in axon guidance and cell migration, develop supernumerary ureteric buds that remain inappropriately connected to the nephric duct. The Slit2/Robo2 signal is transduced in the nephrogenic mesenchyme, and thereby restricts precise positioning the site of kidney induction (82). Human ROBO2 mutations are known in 12 known dominant disease-causing genes in many congenital anomalies of the kidney and urinary tract (83). Finally, three microRNAs: Mir125a, Mirlet7e and Mir99b are bundled together on chromosome (17qE1.3–17qE3). Recent findings indicate the prominent role of microRNAs, small non-coding RNA molecules that inhibit gene expression through the post-transcriptional repression of their target mRNAs, in different pathologic conditions, including renal pathophysiology (81). Interestingly, like recent finding of WT1 binding 3′ UTR (59), microRNAs suppress target gene expression by binding to the 3′-UTR of mRNAs and inhibiting translation and/or promoting mRNA degradation (84). Specifically, Mir99a/b modulates the TGF-β pathway that alter SMAD3 phosphorylation (85) and the let-7 family—a family of tumor suppressor (86)—suppresses breast stem cell self-renewal, tumorigenesis and metastasis (87). The increased binding by DDS mutant M342R could alter the expression of these genes.

DATA AVAILABILITY

The X-ray structures (coordinates and structure factor files) of wild-type WT1(-KTS) ZF1–4 (PDB IDs: 6B0O in *P*1 space group, 6B0P in $P2_12_12$ space group and 6B0Q in $P6_3$ space group), M342R mutant in the −KTS isoform (PDB ID: 6B0R) and +KTS isoform (PDB ID: 6BLW) have been deposited to Protein Data Bank.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

performed X-ray data collection and structure determination; Y.Z. analyzed published WT1 ChIP data; X.Z. and X.C. organized and designed the scope of the study; R.M.B. performed data analysis and assisted in preparing the manuscript; and all were involved in analyzing data and preparing the manuscript.

REFERENCES

1. Lee,S.B. and Haber,D.A. (2001) Wilms tumor and the WT1 gene. *Exp. Cell Res.*, **264**, 74–99.
2. Wagner,K.D., Wagner,N. and Schedl,A. (2003) The complex life of WT1. *J. Cell Sci.*, **116**, 1653–1658.
3. Niaudet,P. and Gubler,M.C. (2006) WT1 and glomerular diseases. *Pediatr. Nephrol.*, **21**, 1653–1660.
4. Miller-Hodges,E. (2016) Clinical aspects of WT1 and the kidney. *Methods Mol. Biol.*, **1467**, 15–21.
5. Millar,A.J.W., Cox,S. and Davidson,A. (2017) Management of bilateral Wilms tumours. *Pediatr. Surg. Int.*, **33**, 737–745.
6. Hastie,N.D. (2017) Wilms' tumour 1 (WT1) in development, homeostasis and disease. *Development*, **144**, 2862–2872.
7. Wang,Z.Y., Qiu,Q.Q. and Deuel,T.F. (1993) The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains. *J. Biol. Chem.*, **268**, 9172–9175.
8. 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. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 11105–11109.
9. Hamilton,T.B., Barilla,K.C. and Romaniuk,P.J. (1995) High affinity binding sites for the Wilms' tumour suppressor protein WT1. *Nucleic Acids Res.*, **23**, 277–284.
10. 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. *Mol. Cell. Biol.*, **15**, 1489–1498.
11. Stoll,R., Lee,B.M., Debler,E.W., Laity,J.H., Wilson,I.A., Dyson,H.J. and Wright,P.E. (2007) Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. *J. Mol. Biol.*, **372**, 1227–1245.
12. 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. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 7562–7566.
13. Hartwig,S., Ho,J., Pandey,P., Macisaac,K., Taglienti,M., Xiang,M., Alterovitz,G., Ramoni,M., Fraenkel,E. and Kreidberg,J.A. (2010) Genomic characterization of Wilms' tumor suppressor 1 targets in nephron progenitor cells during kidney development. *Development*, **137**, 1189–1203.
14. Motamedi,F.J., Badro,D.A., Clarkson,M., Lecca,M.R., Bradford,S.T., Buske,F.A., Saar,K., Hubner,N., Brandli,A.W. and Schedl,A. (2014) WT1 controls antagonistic FGF and BMP-pSMAD pathways in early renal progenitors. *Nat. Commun.*, **5**, 4444.
15. Hashimoto,H., Olanrewaju,Y.O., Zheng,Y., Wilson,G.G., Zhang,X. and Cheng,X. (2014) Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev.*, **28**, 2304–2313.
16. Persikov,A.V. and Singh,M. (2014) De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins. *Nucleic Acids Res.*, **42**, 97–108.
17. Davidoff,A.M. (2012) Wilms tumor. *Adv. Pediatri.*, **59**, 247–267.
18. Dome,J.S., Fernandez,C.V., Mullen,E.A., Kalapurakal,J.A., Geller,J.I., Huff,V., Gratias,E.J., Dix,D.B., Ehrlich,P.F., Khanna,G. *et al.* (2013) Children's Oncology Group's 2013 blueprint for research: renal tumors. *Pediatr. Blood Cancer*, **60**, 994–1000.
19. Hu,M., Zhang,G.Y., Arbuckle,S., Graf,N., Shun,A., Silink,M., Lewis,D. and Alexander,S.I. (2004) Prophylactic bilateral nephrectomies in two paediatric patients with missense mutations in the WT1 gene. *Nephrol. Dial. Transplant.*, **19**, 223–226.
20. Royer-Pokora,B., Beier,M., Henzler,M., Alam,R., Schumacher,V., Weirich,A. and Huff,V. (2004) Twenty-four new cases of WT1 germline mutations and review of the literature: genotype/phenotype correlations for Wilms tumor development. *Am. J. Med. Genet. A*, **127A**, 249–257.
21. 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. *et al.* (1991) Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell*, **67**, 437–447.
22. Borel,F., Barilla,K.C., Hamilton,T.B., Iskandar,M. and Romaniuk,P.J. (1996) Effects of Denys-Drash syndrome point mutations on the DNA binding activity of the Wilms' tumor suppressor protein WT1. *Biochemistry*, **35**, 12070–12076.
23. Little,M., Holmes,G., Bickmore,W., van Heyningen,V., Hastie,N. and Wainwright,B. (1995) DNA binding capacity of the WT1 protein is abolished by Denys-Drash syndrome WT1 point mutations. *Hum. Mol. Genet.*, **4**, 351–358.
24. Hashimoto,H., Zhang,X., Zheng,Y., Wilson,G.G. and Cheng,X. (2016) Denys-Drash syndrome associated WT1 glutamine 369 mutants have altered sequence-preferences and altered responses to epigenetic modifications. *Nucleic Acids Res.*, **44**, 10165–10176.
25. Ohta,S., Ozawa,T., Izumino,K., Sakuragawa,N. and Fuse,H. (2000) A novel missense mutation of the Wt1 gene causing Denys-Drash syndrome with exceptionally mild renal manifestations. *J. Urol.*, **163**, 1857–1858.
26. Schumacher,V., Thumfart,J., Drechsler,M., Essayie,M., Royer-Pokora,B., Querfeld,U. and Muller,D. (2006) A novel WT1 missense mutation presenting with Denys-Drash syndrome and cortical atrophy. *Nephrol. Dial. Transplant.*, **21**, 518–521.
27. Patel,P.R., Pappas,J., Arva,N.C., Franklin,B. and Brar,P.C. (2013) Early presentation of bilateral gonadoblastomas in a Denys-Drash syndrome patient: a cautionary tale for prophylactic gonadectomy. *J. Pediatr. Endocrinol. Metabol.*, **26**, 971–974.
28. Ladomery,M., Sommerville,J., Woolner,S., Slight,J. and Hastie,N. (2003) Expression in Xenopus oocytes shows that WT1 binds transcripts in vivo, with a central role for zinc finger one. *J. Cell Sci.*, **116**, 1539–1549.
29. Bruening,W., Bardeesy,N., Silverman,B.L., Cohn,R.A., Machin,G.A., Aronson,A.J., Housman,D. and Pelletier,J. (1992) Germline intronic and exonic mutations in the Wilms' tumour gene (WT1) affecting urogenital development. *Nat. Genet.*, **1**, 144–148.
30. Takata,A., Kikuchi,H., Fukuzawa,R., Ito,S., Honda,M. and Hata,J. (2000) Constitutional WT1 correlate with clinical features in children with progressive nephropathy. *J. Med. Genet.*, **37**, 698–701.
31. Patel,A., Hashimoto,H., Zhang,X. and Cheng,X. (2016) Characterization of how DNA modifications affect DNA binding by C2H2 zinc finger proteins. *Methods Enzymol.*, **573**, 387–401.
32. Otwinowski,Z., Borek,D., Majewski,W. and Minor,W. (2003) Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A*, **59**, 228–234.
33. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
34. Adams,P.D., Grosse-Kunstleve,R.W., Hung,L.W., Ioerger,T.R., McCoy,A.J., Moriarty,N.W., Read,R.J., Sacchettini,J.C., Sauter,N.K. and Terwilliger,T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1948–1954.
35. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
36. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
37. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE:

tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.

38. Pavletich,N.P. and Pabo,C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science*, **252**, 809–817.

39. Wolfe,S.A., Nekludova,L. and Pabo,C.O. (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 183–212.

40. Choo,Y. and Klug,A. (1997) Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol.*, **7**, 117–125.

41. Persikov,A.V., Wetzel,J.L., Rowland,E.F., Oakes,B.L., Xu,D.J., Singh,M. and Noyes,M.B. (2015) A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Res.*, **43**, 1965–1984.

42. Luscombe,N.M., Laskowski,R.A. and Thornton,J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.*, **29**, 2860–2874.

43. Vanamee,E.S., Viadiu,H., Kucera,R., Dorner,L., Picone,S., Schildkraut,I. and Aggarwal,A.K. (2005) A view of consecutive binding events from structures of tetrameric endonuclease SfiI bound to DNA. *EMBO J.*, **24**, 4198–4208.

44. Patel,A., Horton,J.R., Wilson,G.G., Zhang,X. and Cheng,X. (2016) Structural basis for human PRDM9 action at recombination hot spots. *Genes Dev.*, **30**, 257–265.

45. Dong,L., Pietsch,S., Tan,Z., Perner,B., Sierig,R., Kruspe,D., Groth,M., Witzgall,R., Grone,H.J., Platzer,M. *et al.* (2015) Integration of cistromic and transcriptomic analyses identifies Nphs2, Mafb, and Magi2 as Wilms' tumor 1 target genes in podocyte differentiation and maintenance. *J. Am. Soc. Nephrol.*, **26**, 2118–2128.

46. Hohenstein,P. and Hastie,N.D. (2006) The many facets of the Wilms' tumour gene, WT1. *Hum. Mol. Genet.*, **15 Spec No 2**, R196–R201.

47. Ozdemir,D.D. and Hohenstein,P. (2014) Wt1 in the kidney–a tale in mouse models. *Pediatr. Nephrol.*, **29**, 687–693.

48. Laity,J.H., Dyson,H.J. and Wright,P.E. (2000) Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11932–11935.

49. Price,C., Lingner,J., Bickle,T.A., Firman,K. and Glover,S.W. (1989) Basis for changes in DNA recognition by the EcoR124 and EcoR124/3 type I DNA restriction and modification enzymes. *J. Mol. Biol.*, **205**, 115–125.

50. Drummond,I.A., Rupprecht,H.D., Rohwer-Nutter,P., Lopez-Guisa,J.M., Madden,S.L., Rauscher,F.J. 3rd and Sukhatme,V.P. (1994) DNA recognition by splicing variants of the Wilms' tumor suppressor, WT1. *Mol. Cell. Biol.*, **14**, 3800–3809.

51. Ullmark,T., Jarvstrat,L., Sanden,C., Montano,G., Jernmark-Nilsson,H., Lilljebjorn,H., Lennartsson,A., Fioretos,T., Drott,K., Vidovic,K. *et al.* (2017) Distinct global binding patterns of the Wilms tumor gene 1 (WT1) -KTS and +KTS isoforms in leukemic cells. *Haematologica*, **102**, 336–345.

52. Englert,C., Vidal,M., Maheswaran,S., Ge,Y., Ezzell,R.M., Isselbacher,K.J. and Haber,D.A. (1995) Truncated WT1 mutants alter the subnuclear localization of the wild-type protein. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 11960–11964.

53. Holmes,G., Boterashvili,S., English,M., Wainwright,B., Licht,J. and Little,M. (1997) Two N-terminal self-association domains are required for the dominant negative transcriptional activity of WT1 Denys-Drash mutant proteins. *Biochem. Biophys. Res. Commun.*, **233**, 723–728.

54. Kalodimos,C.G., Biris,N., Bonvin,A.M., Levandoski,M.M., Guennuegues,M., Boelens,R. and Kaptein,R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science*, **305**, 386–389.

55. Horton,J.R., Liebert,K., Hattman,S., Jeltsch,A. and Cheng,X. (2005) Transition from nonspecific to specific DNA interactions along the substrate-recognition pathway of dam methyltransferase. *Cell*, **121**, 349–361.

56. Zhai,G., Iskandar,M., Barilla,K. and Romaniuk,P.J. (2001) Characterization of RNA aptamer binding by the Wilms' tumor suppressor protein WT1. *Biochemistry*, **40**, 2032–2040.

57. Nolte,R.T., Conlin,R.M., Harrison,S.C. and Brown,R.S. (1998) Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 2938–2943.

58. Lu,D., Searles,M.A. and Klug,A. (2003) Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature*, **426**, 96–100.

59. Bharathavikru,R., Dudnakova,T., Aitken,S., Slight,J., Artibani,M., Hohenstein,P., Tollervey,D. and Hastie,N. (2017) Transcription factor Wilms' tumor 1 regulates developmental RNAs through 3′ UTR interaction. *Genes Dev.*, **31**, 347–352.

60. Mazur,S.J. and Record,M.T. Jr (1989) Association kinetics of site-specific protein-DNA interactions: roles of nonspecific DNA sites and of the molecular location of the specific site. *Biopolymers*, **28**, 929–953.

61. Esadze,A., Kemme,C.A., Kolomeisky,A.B. and Iwahara,J. (2014) Positive and negative impacts of nonspecific sites during target location by a sequence-specific DNA-binding protein: origin of the optimal search at physiological ionic strength. *Nucleic Acids Res.*, **42**, 7039–7046.

62. Roberts,S.G. (2005) Transcriptional regulation by WT1 in development. *Curr. Opin. Genet. Dev.*, **15**, 542–547.

63. Loeb,D.M., Evron,E., Patel,C.B., Sharma,P.M., Niranjan,B., Buluwela,L., Weitzman,S.A., Korz,D. and Sukumar,S. (2001) Wilms' tumor suppressor gene (WT1) is expressed in primary breast tumors despite tumor-specific promoter methylation. *Cancer Res.*, **61**, 921–925.

64. Koesters,R., Linnebacher,M., Coy,J.F., Germann,A., Schwitalle,Y., Findeisen,P. and von Knebel Doeberitz,M. (2004) WT1 is a tumor-associated antigen in colon cancer that can be recognized by in vitro stimulated cytotoxic T cells. *Int. J. Cancer*, **109**, 385–392.

65. Oji,Y., Yano,M., Nakano,Y., Abeno,S., Nakatsuka,S., Ikeba,A., Yasuda,T., Fujiwara,Y., Takiguchi,S., Yamamoto,H. *et al.* (2004) Overexpression of the Wilms' tumor gene WT1 in esophageal cancer. *Anticancer Res.*, **24**, 3103–3108.

66. Amini Nik,S., Hohenstein,P., Jadidizadeh,A., Van Dam,K., Bastidas,A., Berry,R.L., Patek,C.E., Van der Schueren,B., Cassiman,J.J. and Tejpar,S. (2005) Upregulation of Wilms' tumor gene 1 (WT1) in desmoid tumors. *Int. J. Cancer*, **114**, 202–208.

67. Essafi,A., Webb,A., Berry,R.L., Slight,J., Burn,S.F., Spraggon,L., Velecela,V., Martinez-Estrada,O.M., Wiltshire,J.H., Roberts,S.G. *et al.* (2011) A wt1-controlled chromatin switching mechanism underpins tissue-specific wnt4 activation and repression. *Dev. Cell*, **21**, 559–574.

68. Kann,M., Ettou,S., Jung,Y.L., Lenz,M.O., Taglienti,M.E., Park,P.J., Schermer,B., Benzing,T. and Kreidberg,J.A. (2015) Genome-wide analysis of Wilms' Tumor 1-controlled gene expression in podocytes reveals key regulatory mechanisms. *J. Am. Soc. Nephrol.*, **26**, 2097–2104.

69. Grahammer,F. (2017) New structural insights into podocyte biology. *Cell Tissue Res.*, **369**, 5–10.

70. Eyers,C.E., McNeill,H., Knebel,A., Morrice,N., Arthur,S.J., Cuenda,A. and Cohen,P. (2005) The phosphorylation of CapZ-interacting protein (CapZIP) by stress-activated protein kinases triggers its dissociation from CapZ. *Biochem. J.*, **389**, 127–135.

71. Dudnakova,T., Spraggon,L., Slight,J. and Hastie,N. (2010) Actin: a novel interaction partner of WT1 influencing its cell dynamic properties. *Oncogene*, **29**, 1085–1092.

72. De Braekeleer,E., Douet-Guilbert,N., Le Bris,M.J., Berthou,C., Morel,F. and De Braekeleer,M. (2007) A new partner gene fused to ABL1 in a t(1;9)(q24;q34)-associated B-cell acute lymphoblastic leukemia. *Leukemia*, **21**, 2220–2221.

73. De Braekeleer,E., Douet-Guilbert,N., Rowe,D., Bown,N., Morel,F., Berthou,C., Ferec,C. and De Braekeleer,M. (2011) ABL1 fusion genes in hematological malignancies: a review. *Eur. J. Haematol.*, **86**, 361–371.

74. Tschopp,P. and Duboule,D. (2011) A regulatory 'landscape effect' over the HoxD cluster. *Dev. Biol.*, **351**, 288–296.

75. Woo,C.J., Kharchenko,P.V., Daheron,L., Park,P.J. and Kingston,R.E. (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell*, **140**, 99–110.

76. Yamane,K., Toumazou,C., Tsukada,Y., Erdjument-Bromage,H., Tempst,P., Wong,J. and Zhang,Y. (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell*, **125**, 483–495.

77. Guo,X. and Zhang,Q. (2017) The emerging role of histone demethylases in renal cell carcinoma. *J. Kidney Cancer VHL*, **4**, 1–5.

78. Guo,X., Shi,M., Sun,L., Wang,Y., Gui,Y., Cai,Z. and Duan,X. (2011) The expression of histone demethylase JMJD1A in renal cell carcinoma. *Neoplasma*, **58**, 153–157.

79. Krieg,A.J., Rankin,E.B., Chan,D., Razorenova,O., Fernandez,S. and Giaccia,A.J. (2010) Regulation of the histone demethylase JMJD1A by hypoxia-inducible factor 1 alpha enhances hypoxic gene expression and tumor growth. *Mol. Cell. Biol.*, **30**, 344–353.

80. Fang,X., Gao,Y. and Li,Q. (2016) SMAD3 activation: a converging point of dysregulated TGF-beta superfamily signaling and genetic aberrations in granulosa cell tumor development? *Biol. Reprod.*, **95**, 105.

81. Loboda,A., Sobczak,M., Jozkowicz,A. and Dulak,J. (2016) TGF-beta1/Smads and miR-21 in renal fibrosis and inflammation. *Mediators Inflamm.*, **2016**, 8319283.

82. Grieshammer,U., Le,M., Plump,A.S., Wang,F., Tessier-Lavigne,M. and Martin,G.R. (2004) SLIT2-mediated ROBO2 signaling restricts kidney induction to a single site. *Dev. Cell*, **6**, 709–717.

83. Hwang,D.Y., Dworschak,G.C., Kohl,S., Saisawat,P., Vivante,A., Hilger,A.C., Reutter,H.M., Soliman,N.A., Bogdanovic,R., Kehinde,E.O. *et al.* (2014) Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. *Kidney Int.*, **85**, 1429–1433.

84. Fabbri,M., Croce,C.M. and Calin,G.A. (2008) MicroRNAs. *Cancer J.*, **14**, 1–6.

85. Turcatel,G., Rubin,N., El-Hashash,A. and Warburton,D. (2012) MIR-99a and MIR-99b modulate TGF-beta induced epithelial to mesenchymal plasticity in normal murine mammary gland cells. *PLoS One*, **7**, e31032.

86. Boyerinas,B., Park,S.M., Hau,A., Murmann,A.E. and Peter,M.E. (2010) The role of let-7 in cell differentiation and cancer. *Endocr. Relat. Cancer*, **17**, F19–F36.

87. Yu,F., Yao,H., Zhu,P., Zhang,X., Pan,Q., Gong,C., Huang,Y., Hu,X., Su,F., Lieberman,J. *et al.* (2007) let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell*, **131**, 1109–1123.