THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 1, pp. 647–656, January 5, 2007

Structure of Human Spindlin1

*TANDEM TUDOR-LIKE DOMAINS FOR CELL CYCLE REGULATION**

Received for publication, April 27, 2006, and in revised form, November 1, 2006 Published, JBC Papers in Press, November 1, 2006, DOI 10.1074/jbc.M604029200

Qiang Zhao[‡1], Lipeng Qin[§1], Fuguo Jiang[‡], Beili Wu[‡], Wen Yue[§], Feng Xu[‡], Zhili Rong[¶], Hongfeng Yuan[§], Xiaoyan Xie[§], Yanhong Gao[§], Cixian Bai[§], Mark Bartlam[‡], Xuetao Pei[§2], and Zihe Rao[‡3]

From the [‡]Tsinghua-Institute of Biophysics Joint Research Group for Structural Biology, Tsinghua University, Beijing 100084, China and National Laboratory of Biomacromolecules, Institute of Biophysics , Chinese Academy of Sciences, Beijing 100101, China, [§]Department of Stem Cell Biology, Beijing Institute of Transfusion Medicine, Beijing 100850, China, and [¶]Tsinghua Institute of Genome Research, School of Medicine, Tsinghua University, Beijing 100084, China

Spindlin1, a meiotic spindle-binding protein that is highly expressed in ovarian cancer cells, was first identified as a gene involved in gametogenesis. It appeared to be a target for cell cycle-dependent phosphorylation and was demonstrated to disturb the cell cycle. Here we report the crystal structure of human spindlin1 to 2.2 Å of resolution, representing the first three-dimensional structure from the spin/ssty (Y-linked spermiogenesis-specific transcript) gene family. The refined structure, containing three repeats of five/four anti-parallel β-strands, exhibits a novel arrangement of tandem Tudor-like domains. Two phosphate ions, chelated by Thr-95 and other residues, appear to stabilize the long loop between domains I and II, which might mediate the cell cycle regulation activity of spindlin1. Flow cytometry experiments indicate that cells expressing spindlin1 display a different cell cycle distribution in mitosis, whereas those expressing a T95A mutant, which had a great decrease in phosphorous content, have little effect on the cell cycle. We further identified associations of spindlin1 with nucleic acid to provide a biochemical basis for its cell cycle regulation and other functions.

Spindlin, a major maternal transcript in *Mus musculus*, was named for its association and co-migration with the meiotic spindle in the first meiotic cell cycle (1). It can be periodically phosphorylated during meiosis, which modulates its ability to associate with the meiotic spindle (2). The modification of spindlin depends at least partially on the Mos/mitogen-activated protein kinase pathway (2), which is controlled by meiotic checkpoint proteins cyclin B and Cdc2 (3, 4). As an ovarian-specific protein, its role in sperm development seems to be fulfilled by ssty (Y-linked spermiogenesis specific transcript) (1), a multicopy testis-specific spermatogenesis gene on the long arm of mouse Y chromosome whose dosage reduction was suggested to cause deformed sperm heads and infertility (5–7). These two genes share more than 50% identity in amino acid sequence (Fig. 1), and together they form a new spin/ssty gene family. Homologues of spin/ssty family genes are found in *Rattus norvegicus*, *Xenopus laevis*, and *Oryzias latipes* but not in *Drosophila melanogaster* or *Caenorhabditis elegans*, indicating that this gene family is restricted to vertebrates (8). Furthermore, bioinformatics analysis suggests that the spin/ssty family proteins are composed of three repeats of a new protein motif ~50 amino acids in length (8).

Spindlin1, a human homolog of spindlin, has been shown to be related to ovarian cancer (9). Its expression is up-regulated in ovarian cancer cells but not in normal tissues and was found to localize in cell nuclei. The transfected cells, which are prone to grow as cancer cells in nude mice, showed a complete morphological change (10). Overexpression of spindlin1 might lead to variations in cell cycle distribution during mitosis, which is consistent with previous reports for other spindlin1 homologs (11).

Although previous studies indicate that members of the spin/ssty protein family play important roles in tumorigenesis and early embryogenesis, their biochemical functions and mechanisms are largely unknown. Here we report the crystal structure of human spindlin1 at 2.2 Å of resolution, which represents the first crystal structure from the vertebrate spin/ssty gene family. The structure consists of three tandem repeats of Tudor-like domains, which represent a novel tandem repeat fold. We identify Thr-95 as important for the function of spindlin1, which is a key residue for the coordination of two phosphate ions. In addition, our preliminary biochemical data also indicate that spindlin1 associates with nucleic acid. These data provide a structural basis for further biochemical analysis of spindlin1 and the spin/ssty gene family.

EXPERIMENTAL PROCEDURES

Protein Expression, Purification, Crystallization, and Mutagenesis—The protocols for purification and crystallization of human spindlin1 have been described previously (12). The T95A mutant was generated using GCA to replace original codon by PCR and purified similar to the wild type protein. Both wild type spindlin1 and the T95A mutant were constructed into pCDNA3.1 myc/his(−) for cell cycle analysis.

* This study was supported by Ministry of Science and Technology Human Liver Proteomics Project Grant 2004CB520801, State 863 High-Tech Project Grants 2002BA711A12 and 2002AA205050 and 973 Project Grants G1999075602 and 2001CB509906, and the National Natural Science Foundation of China Grants 30221003 and 30271359. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2NS2) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] These authors made equal contributions.

[2] To whom correspondence may be addressed. E-mail: peixt@nic.bmi.ac.cn.

[3] To whom correspondence may be addressed. E-mail: raozh@xtal.tsinghua.edu.cn.

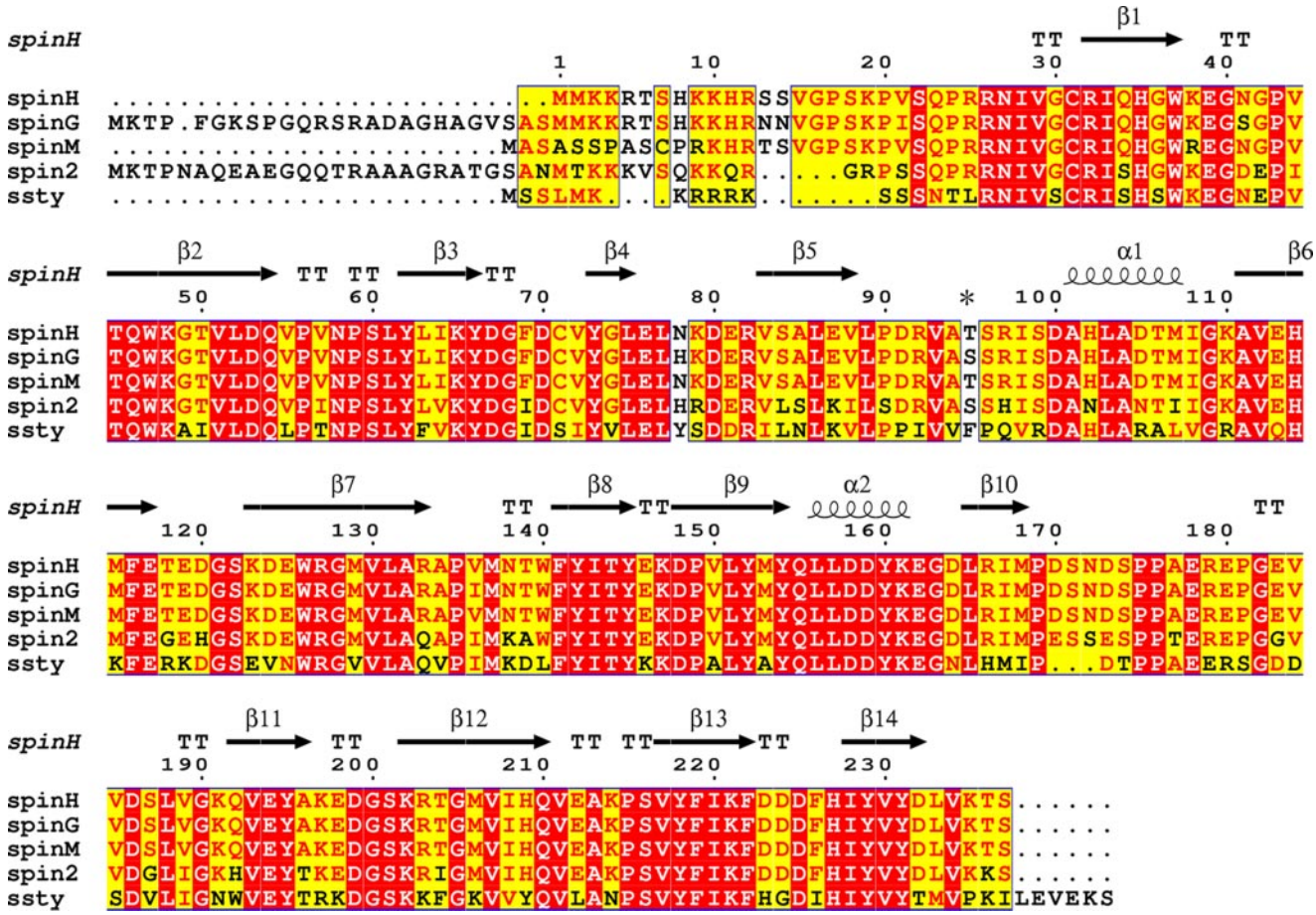

FIGURE 1. **Sequence alignment of spindlin1 with homologous proteins.** Spindlin genes from *Homo sapiens* (spinH, NCBI accession number AAG38112), *Gallus gallus* (spinG, NCBI accession number NP_989964), and *M. musculus* (spinM, NCBI accession number NP_035592) are aligned with Spin2 of *H. sapiens* (NCBI accession number NP_001006684) and ssty of *M. musculus* (NCBI accession number NP_033246). The secondary structure elements are labeled based on the crystal structure of spindlin1. The *asterisk* on the sequence represents Thr-95 in spindlin1. Sequences were aligned with ClustalW (28), and the alignment was drawn with ESPript (29).

Heavy Atom Derivatization—Crystals of spindlin1 were transferred into a solution containing 35% polyethylene glycol 6000, 100 mM Tris-HCl, pH 8.0, in a stepwise manner. The mercury derivative was obtained by soaking spindlin1 crystals in this buffer supplemented with 1 mM ethyl mercuric phosphate ($(C_2H_5HgO)HPO_2$) for 3 days.

Data Collection and Processing—Crystals of spindlin1 belong to the space group $P2_12_12_1$, with unit cell parameters $a = 40.8$ Å, $b = 84.9$ Å, $c = 136.6$ Å, $\alpha = \beta = \gamma = 90°$. The crystals contain two molecules per asymmetric unit. Native data up to 2.2 Å were collected from a flash-cryocooled crystal with 25% (v/v) glycerol used as a cryoprotectant. The single-wavelength anomalous diffraction data were collected using a Rigaku RU2000 rotating CuK_α anode source to 2.3 Å using a single mercury-spindlin1 crystal. Data were indexed and scaled using HKL2000 and SCALEPACK (13), and the unit cell dimensions were determined to be $a = 40.3$ Å, $b = 76.2$ Å, $c = 136.6$ Å, $\alpha = \beta = \gamma = 90°$.

Structure Determination and Refinement—Phases for the mercury-spindlin1 crystal were initially determined by the single-wavelength anomalous diffraction technique using CNS (14). The phasing power was calculated as 2.4, and three heavy atom sites were independently located by the heavy atom search routine, yielding an overall figure of merit of 0.41 after calculation of initial single-wavelength anomalous diffraction phases at 3.0 Å. The resolution was extended to 2.3 Å using CNS. After solvent flipping, the quality of the initial electron density maps was greatly improved. Initial manual model building and fitting were carried out using 2.2 Å-resolution native data in O (15). Positional refinement, B-factor refinement, and water molecules were added using CNS. Data collection, processing, phasing, and refinement statistics are given in Table 1.

Cell Cycle Analysis—The cell cycle phase distribution of HeLa cells was examined by flow cytometry using FACScan and Cell Quest software (BD Biosciences). 1×10^6 cells were co-transfected with pBB14 (green fluorescent protein) and pCDNA3.1 myc/his(−) vector, spindlin1 wild type, or spindlin1 point mutants. After incubation in full media for 20 h, HeLa cells were harvested and prefixed using 0.5% paraformaldehyde. The cells were washed with phosphate-buffered saline and fixed in phosphate-buffered saline/ethanol for 1 h. The cells were stained with propidium iodide (50 μg/ml) for 30 min after RNase digestion and analyzed.

Gel Shift Assays—For DNA binding assays, purified samples of about 30 bp of double-stranded DNA (dsDNA)[4] probe, 30 bp

[4] The abbreviation used is: dsDNA, double-stranded DNA.

TABLE 1
Crystallographic data collection and refinement statistics

	Native	Ethyl mercuric phosphate derivative
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell		
a (Å)	40.8	40.3
b (Å)	84.9	76.2
c (Å)	136.6	136.6
Resolution (Å)	50-2.2 (2.28-2.20)	50-2.3 (2.38-2.30)
Completeness (%)	99.9 (100)	97.9 (96.7)
Reflections		
Total	444,671	344,044
Unique	25,042	18,996
Redundancy	17.8 (17.5)	18.1 (17.9)
R_{merge}[a] (%)	5.6 (37.4)	9.0 (48.0)
$I/\sigma(I)$	30.8 (7.9)	30.9 (7.0)
Refinement statistics		
Resolution (Å)	2.2	
R-factor[b] (%)		
Working set	21.4	
Test set	26.8	
Root mean square deviation		
Bonds (Å)	0.012	
Angles (°)	1.72	
Ramachandran plot[c] (%)		
Most favored	86.5	
Allowed	12.6	
Generously allowed	0.3	
Disallowed	0.6	

[a] $R_{merge} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I_i(h)$, where $\langle I\rangle(h)$ is the mean of the observations $I_i(h)$ of reflection *h*.
[b] $R_{work} = \Sigma_h(||F_{obs}| - |F_{calc}||)/\Sigma_h|F_{obs}|$; R_{free} is the R factor for a subset (10%) of reflections that was selected before refinement calculations and not included in the refinement.
[c] Ramachandran plots were generated using PROCHECK.

of DNA primer, and a constructed spindlin1 vector of ~2.5 kilobases were employed. Binding reactions were conducted in 20 mM Tris, pH 8.0, 150 mM NaCl, 0.2 mM EDTA, 10%(v/v) glycerol, 1 mM dithiothreitol, 0.05 mg/ml bovine serum albumin, and 0.05% Triton-100 to final sample volumes of 10 μl. After 20 min of incubation at room temperature, 1 μl of loading buffer containing 0.05% bromphenol blue was added to the reaction mixtures. 1% agarose gels were run at room temperature at a constant voltage of 12 $V\cdot cm^{-1}$ in 1× Tris borate EDTA and afterward stained with ethidium bromide.

RESULTS

Overall Structure of Spindlin1 Folds into Three Similar Domains—The crystal structure of recombinant spindlin1 was determined by single-wavelength anomalous diffraction from a single crystal soaked with mercury. The initial model was built into an electron density map calculated to 2.3 Å resolution, and data from a native crystal allowed model refinement to 2.2 Å. The two spindlin1 molecules in one asymmetric unit were traced in two (residues 25–170 and 181–235 in molecule A) (Fig. 2*a*) and four (residues 27–91, 104–115, 125–168, and 184–234 in molecule B) fragments totaling 237 residues, respectively. No electron density was evident for the other residues.

Spindlin1, with dimensions of ~40 × 40 × 35 Å, exhibits an all-β structure that consists of 14 β-strands and two short α-helices. The structure is folded into three structural domains that are composed mainly of β-barrel-like structures. The three domains, which correspond to residues 25–90, 101–168, and 191–234, consist of β1–β5, β6–β10, and β11–β14, respectively. The α-helices are located in domain II, ahead of its first and fifth β-strands respectively (Fig. 2, *a* and *b*).

As shown in Fig. 2*c*, domain I, II, and III adopt a similar 5 (4) β-strand fold, and the main chain of their strands can be superimposed with root mean square deviations of 0.8 Å (from domain II to domain I) and 2.1 Å (from domain III to domain I). Each domain is composed of two strongly bent anti-parallel β-sheets: β1, β2, β5 and β2′, β3, β4 from domain I; β6, β7, β10 and β7′, β8, β9 from domain II; β11, β12 and β12′, β13, β14 from domain III. The two β sheets of each domain are roughly perpendicular, with the first β-sheet of each domain exposed to the solvent, and the second β-sheet of each domain buried inside. The long β strands β2, β7, and β11 are bent by ~90° around classical β-bulges located at amino acid positions 52, 53, 63 and 131, 132, 142 and 208, 209, 219, respectively. Strands β2′, β7′, and β12′ represent the posterior parts of strands β2, β7, and β11 after their respective β-bulges.

The two helices are located in domain II. Helix α1 flanks β7′, whereas helix α2 is above α1 and β8, and they compose the outer layer of domain II together with strands β6, β7, and β10. These two α-helices are tightly packed against domain II, forming hydrophobic interactions with Val-112 in β6, Val-130 in β7′, Phe-141 and Ile-143 in β8, Tyr-154 in β9, and Leu-165 and Ile-167 in β10 with a total buried area of 1089 $Å^2$. These two helices form a near vertical cross with hydrophobic interactions between Leu-103, Ala-104, and Met-107 of α1 and Leu-156, Leu-157, and Tyr-160 of α2.

Sequence alignment among the three domains shows a total sequence identity of 32%, with high conservation of some aromatic residues and nonpolar residues (Fig. 2*d*). These residues either contribute to the hydrophobic core of each domain or for the hydrophobic interactions among the three domains. The conserved glycines and some asparagine residues are important for the protein fold. Gly-30, Gly-109, and Gly-190 are significant for the formation of the first β-strand in each domain. Asn-35—Gly-36, Asp-120—Gly-121, and Asp-199—Gly-200 form β-turns between the first and second β-strands in each domain, whereas the flexibility of Gly-43, Gly-128, and Gly-205 should be crucial for the bends of strands β2, β7, and β11. There are a number of other conserved glutamate, aspartate, asparagine, glutamine, and serine residues that are distributed on the surface of the protein, but their structural roles are less clear (Fig. 2*d*).

Domains I, II, and III Possess a New Tudor-like Tandem Fold—A DALI search for structural similarity to the individual domains I, II, and III retrieved ~30 structural homologs for different fragments and suggested Tudor domains and malignant brain tumor repeat domains as closest matches. The spindlin domains and Tudor domains all compose two sheets that are made up of five anti-parallel β strands. The portions of the second strands of these domains all participating in the two sheets are delineated by a kink that changes the direction of the polypeptide (Fig. 3*a*). Furthermore, residues that assist in forming the Tudor-like domain fold are highly conserved between domains I, II, and III of spindlin1 and the Tudor domain (Fig. 3*b*). However, a DALI search did not yield any match for the overall structure, suggesting the spindlin1 structure is a novel arrangement of Tudor-like domains.

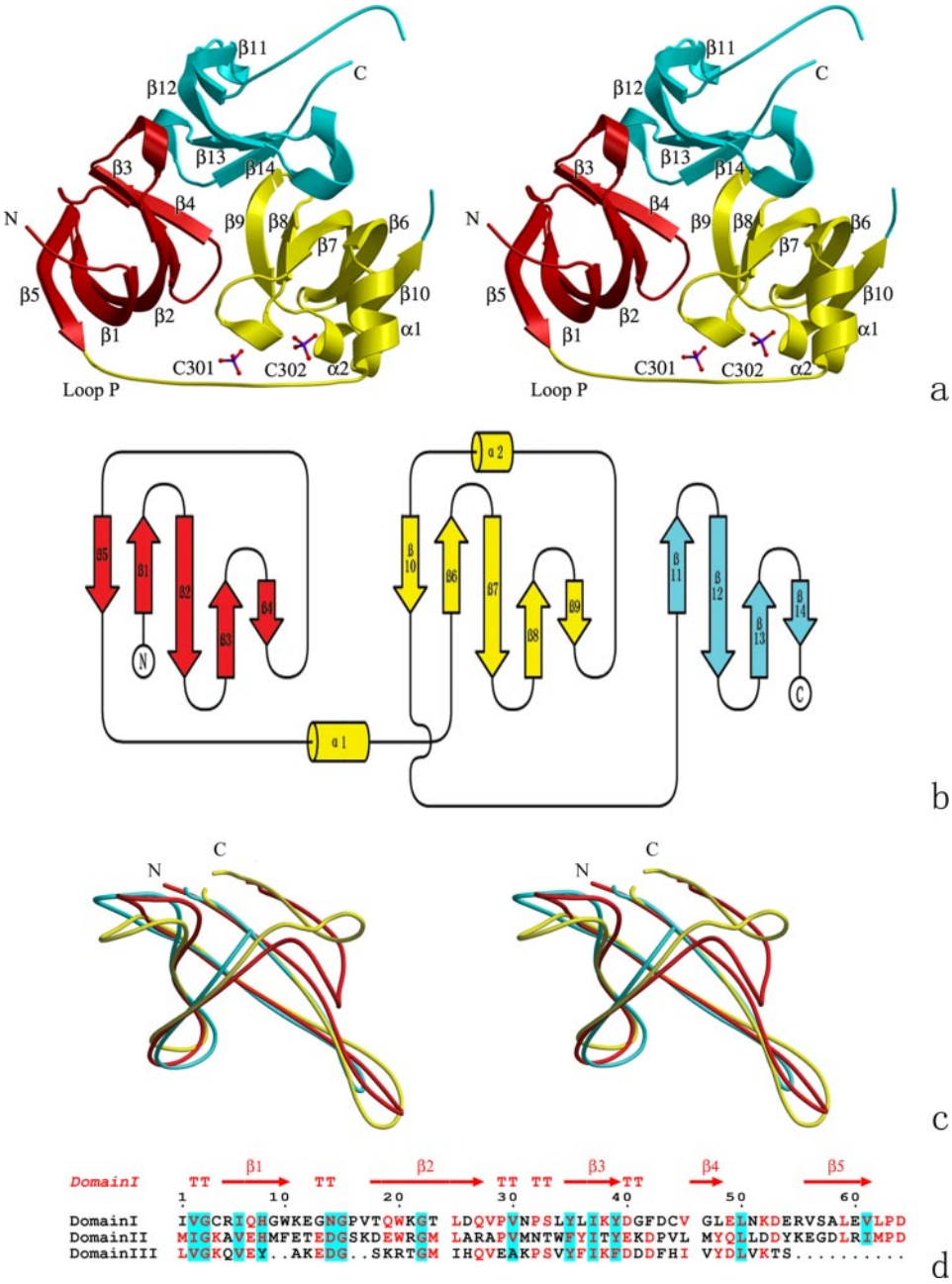

FIGURE 2. **Overall structure and topology of spindlin1.** The three domains are shown respectively in *red*, *yellow*, and *cyan*. *a*, stereo *ribbon diagram* of a spindlin1 monomer. The two phosphate groups are shown in *ball-and-stick* representation. All the secondary structures and the loop P are labeled. *b*, folding topology of spindlin1. The helices are drawn as *barrels*, the β-strands are drawn as *open arrows*, and both are numbered. *c*, a comparison of three domains based on the backbones of the first four β-strands. β1–5 of domain 1, β6–10 of domain 2, and β11–14 of domain 3 are shown, and the N and C termini are labeled. *d*, sequence alignment of the spindlin1 domains I, II, and III. The secondary structure elements are labeled based on the crystal structure of domain I. The residues colored in *red* are conserved among the three domains, and those *highlighted* with a *cyan* background are of structural significance for the protein fold.

Although spin/ssty repeats and Tudor domains share similar folds, there are some significant differences between them. First, superposition of these domains showed some significant disparity in the backbones of the five-strand structure. The orientation of the fifth strands of the spindlin1 domains is completely diverse from that of the Tudor domains. In addition, comparison of the anterior five β-strands gives a root mean square deviation of about 3.1 ~ 4.5 Å, indicating considerable differences between the two kinds of domains. Furthermore, previous studies have shown that the β1–β2 loop and the β3–β4 loop together with the regions between these two loops are of great importance for its function, but the residue organization in the spindlin domains is not essentially conserved with the known Tudor domains (Fig. 3, *a*, *c*, and *d*). All of these observations suggest that spin/ssty repeats might have a different function or at least function in a different mode than Tudor domains.

The Interactions between Spindlin1 Molecules in Crystal Structure and Solution—From the crystal structure, there are two molecules in one asymmetric unit that are essentially identical with a root mean square deviation of 0.4 Å for all C_α atoms. Examination of the crystal packing reveals the occurrence of a major interface between the two spindlin1 molecules. Each spindlin1 molecule offers two surfaces (surface A and surface B) for association with another molecule. Surface A is defined by strands β11, β14, and the loop β10–β11, whereas surface B is defined by strand β2 and the loops β1–β2, β3–β4, and β13–β14.

The two spindlin1 molecules associate through an interface that involves surface A of one molecule and surface B of another molecule, the interactions of which include van der Waals contacts, steric complementarily, and hydrogen bond contacts. First, the interaction between surface A and B buries about 624 $Å^2$ of solvent-accessible area, which would help to stabilize the dimer architecture. Meanwhile, the first strand (β11) of domain III in molecule A forms an anti-parallel β-sheet with strand β2 of domain I in molecule B. As a result, residues Pro-43 and Thr-45 in molecule B participate in the formation of the hydrophobic core of domain III of its neighboring molecule. It should be noted that domain III is composed of only four β strands, which is one strand less than domain I and II, and yet this β-barrel-like structure is closed by strand β2 from the neighboring molecule, in a similar way to strands β5 and β10 in the first two repeats. Thus, the "missing" fifth strand of this spin/ssty repeat is compensated at least in part by molecular interactions. The hydropho-

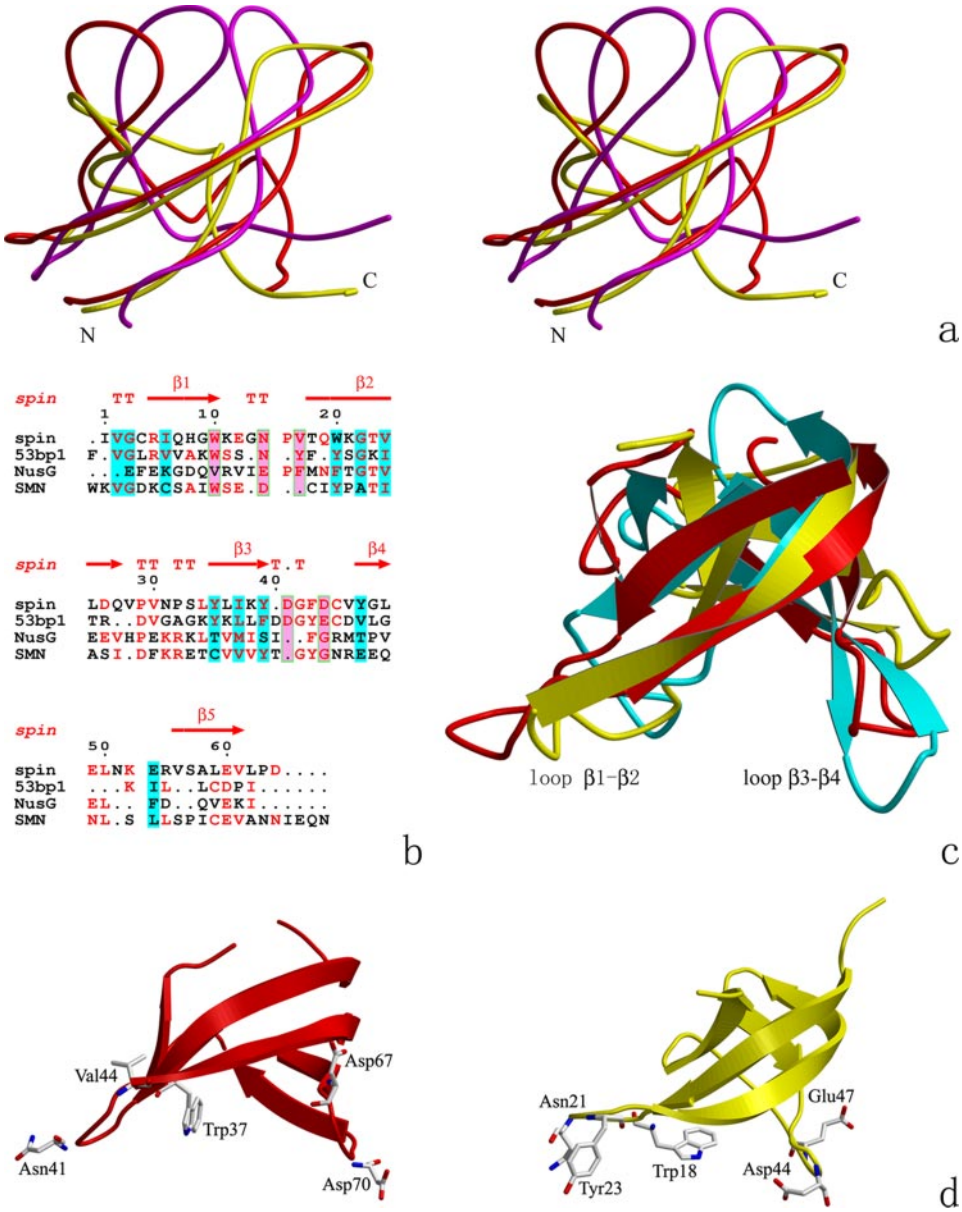

FIGURE 3. **Comparison of spindlin1 domain I with other Tudor-like domains.** *a*, superposition of domain I with two other Tudor-like domains. Domain I is shown in *red*, the Tudor domain in human Survival of Motor Neuron gene (SMN, PDB code 1G5V) is shown in *magenta*, and the Tudor domain in 53bp1 (PDB code 1SSF) is shown in *yellow*. The N and C termini of the three domains are labeled. *b*, structure-based sequence alignment of domain I with other Tudor-like domains shown in *a*. The secondary structure elements are labeled based on the crystal structure of spindlin1. The residues *highlighted* with a *cyan* background are the conserved residues for maintaining the protein fold, and residues *boxed* with a *green frame* and denoted in *pink* are essential for nucleic acid binding of 53bp1 and their counterparts in the other domains. *c*, comparison of domain I (*red*) with Tudor-like domains in microbial transcription modulator NusG (*cyan*) (PDB code 1M1G) and the Tudor domain in 53bp1 as shown in *a*. The loop β1–2 and loop β3–4, which have been shown to be of great importance of nucleic acid binding activity in NusG and 53bp1, are labeled. *d*, *ribbon diagrams* of domain I (*red*) and Tudor domain from human 53bp1 (*yellow*) (PDB code 1SSF). The residues denoted in *pink* in *b* are shown and labeled.

bic interface also involves Trp-47, Phe-69, Tyr-73, and Phe-226 in molecule B and Val-185 and Leu-188 in molecule A. In addition, the two anti-parallel β-strands between the two molecules also lead to the formation of at least eight hydrogen bonds around Asp-186 and Gln-192 (Glu-194A—Val-44B, Gln-192A—Gln-46B, Asp-186A—Tyr-73B, Lys-191A—Asp-67B, Lys-191A—Wat-78–Lys-48B, Gln-192A—Wat-29-Gln-46B, and Pro-181A—Wat-179—Asp-225B; Wat is water), which should also contribute to the interaction (Fig. 4, *a* and *b*).

Consistent with our structure, spindlin1 maintains a dimeric state in solution, as confirmed by dynamic light scattering and size exclusion chromatography (data not shown). Cross-linking assays using ethylene glycol succinate as linker also showed that spindlin1 exists largely as a homodimer in solution (Fig. 4*c*).

Thr-95 Is an Important Residue for Phosphate Ion Binding—Two phosphate ions maintained by hydrogen bonds in the loop between domain I and II (P loop) were found in the refined structure of spindlin1 and named c301 and c302, and their position in the overall structure is shown in Fig. 2*a*. Phasing with heavy atom assigned as phosphorous gives strong anomalous difference electron density at the phosphate binding site, confirming the existence of phosphates (Fig. 5*a*). Atomic emission spectroscopy revealed that spindlin1 associates with phosphate ions in solution, thus giving a characteristic line for phosphorus in phosphate free buffers.

Several arginine and lysine residues, but no aspartate or glutamate residues, are located in the phosphate binding surfaces, which form a positively charged surface region. The electrostatic interactions between these cationic residues and anionic phosphate groups should contribute to the phosphate binding. For further stabilization, the phosphate groups also form a number of hydrogen bonds with spindlin1.

There are 10 hydrogen bonds between c301 and spindlin1. The O1 oxygen of phosphate c301 forms three hydrogen bonds; two hydrogen bonds are formed directly with the OG1 and N atoms of Thr-95, and another hydrogen bond is mediated by the water molecule S40 with the main chain oxygen of Val-93. The O2 oxygen also involves three hydrogen bonds; two hydrogen bonds with the NH1 and NH2 atoms of Arg-92 and another hydrogen bond mediated by water molecule S40 with Val-93. The O3 and O4 oxygens of phosphate c301 both form two hydrogen bonds with the polypeptide. The OG1 and N of Thr-139 form two hydrogen bonds with O3, whereas the last two hydrogen bonds are generated between O4

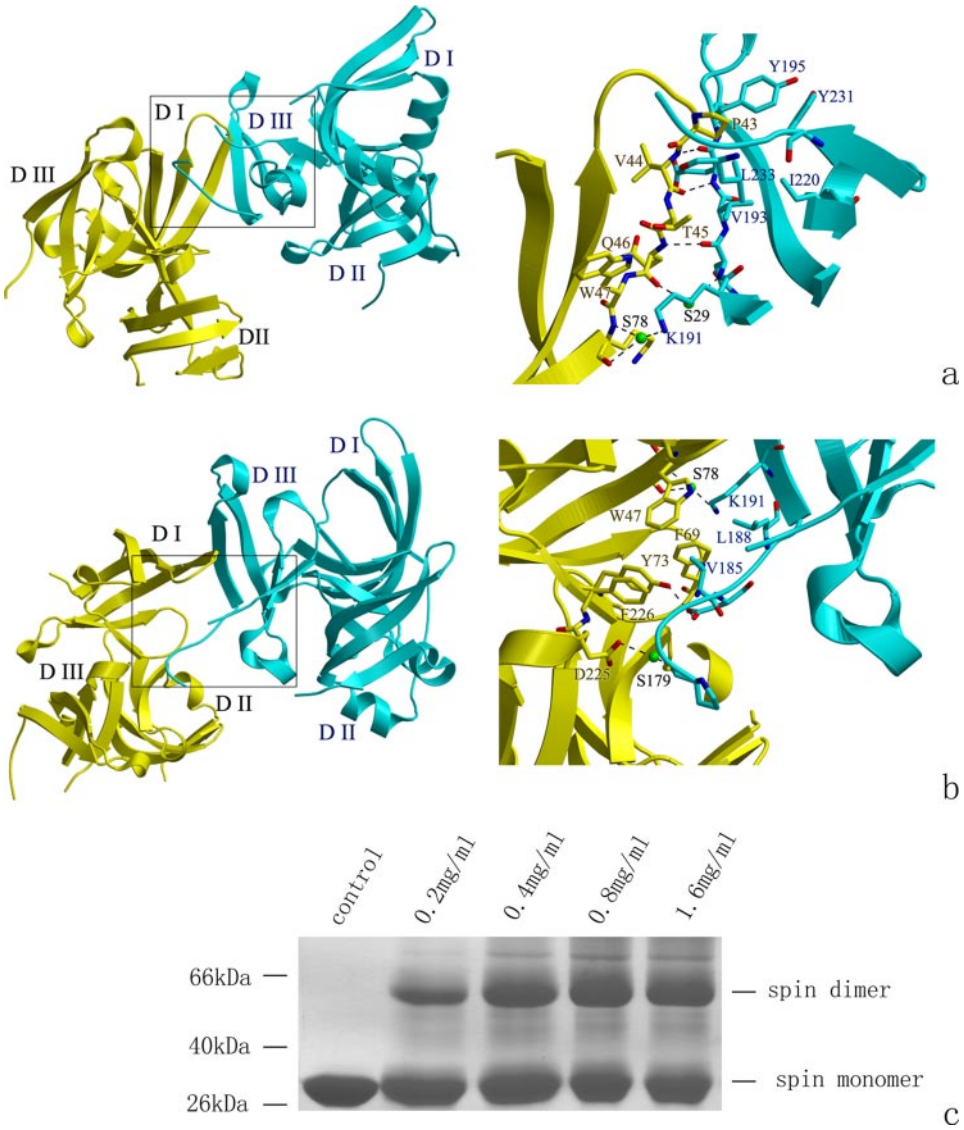

FIGURE 4. **The architecture of the spindlin1 homodimer.** *a*, interface around residue Gln-192. *Left*, an overall view of the dimer architecture, with molecule A shown in *cyan* and molecule B shown in *yellow*. The three domains of each molecule are labeled in *blue* and *brown*, respectively. *Right*, a detailed view of the boxed region, with residues of molecule A shown in *cyan*, residues of molecule B shown in *yellow*, and water molecules represented by *green spheres*. The residues from molecule A are labeled in *deep blue*, the residues from molecule B are labeled in *dark brown*, and the water spheres are labeled in *black*. Hydrogen bonds are denoted by *dashed lines*, and one residue of each hydrogen bond is labeled. Besides those involved in hydrogen bonds, the residues Pro-43 and Thr-45 in molecule B and Val-193, Tyr-195, Ile-220, Tyr-231, and Leu-233 in molecule A that form a hydrophobic core between different molecules are also labeled. *b*, the interface around residue 186. The labeled residues Trp-47, Phe-69, Tyr-73, Phe-226 in molecule B and Val-185, Leu-188, Lys-191 form another hydrophobic core. *c*, SDS-PAGE gel showing the results of chemical cross-linking of purified spindlin1. The cross-linker is ethylene glycol succinate, with a final concentration of 1 mM. Spindlin1 at a concentration of ~0.2, 0.4, 0.8, and 1.6 mg/ml was used for cross-linking, and the same amount of protein was loaded onto the SDS-PAGE gel. The spindlin1 dimer bands were clearly observed for all concentrations, whereas a small portion of protein is exhibited with even higher oligomer state as the increasing amount of protein.

and the N atoms of Thr-95 and Asn-138, respectively (Fig. 5*b*). The detailed length of each hydrogen bond is listed in Table 2.

Compared with c301, the interactions for the c302 phosphate ion are weaker with a total of eight hydrogen bonds. The O1 oxygen of c302 involves the formation of three hydrogen bonds; two hydrogen bonds with the NH2 and NE atoms of Arg-97 and one mediated by water molecule S4 with the O atom of Ile-98. The O2 oxygen of c302 forms two hydrogen bonds with NH1 of Arg-133 and OD1 of Asn-138. The O3 oxygen contributes one hydrogen bond with NH2 of Arg-133, whereas O4 forms one hydrogen bond with NH2 of Arg-97 and one mediated by water S98 with the ND2 atom of Asn-138 (Table 1, Fig. 5*c*).

From our structure, the long P loop (residues 90–100) is relatively stable (Fig. 2*a*) with an average B factor of 29.7 for main chain atoms. There are five residues in this loop that interact with the phosphate ions either directly (Arg-92, Thr-95, and Arg-97) or indirectly via hydrogen bonds mediated by water molecules (Val-93 and Ile-98). Thus, the presence of phosphate c301 and c302 should greatly benefit the stabilization of this loop.

Of these hydrogen bonds to the two phosphate ions, the interactions associated with the phosphate O1 atom are the strongest, as evidenced by the electron density between the phosphate O1 atom and OG1 of Thr-95 seen clearly even at a contour level of 3σ. From atomic emission spectroscopy experiments, the phosphorus content in wild type spindlin1 is 62 ± 2 g/mol, whereas only 24 ± 3 g/mol could be measured for the T95A mutant, indicating a greater decrease of the phosphate content in this mutant. This mutant could not be crystallized with the native crystallization conditions, which indicates that the structure of spindlin1 is at least partially influenced by the presence of phosphate. Thus, Thr-95 is a critical residue for the binding of phosphate ions.

The Loop between Domain I and II Is the Key Site for Cell Cycle Regulation Activity—To examine the potential effects of the phosphate ions on the function of spindlin1, the cell cycle distribution of control HeLa cells were compared with fluorescence-activated cell-sorted green fluorescent protein-positive HeLa cells expressing either wild type spindlin1 or the T95A mutant. All cells were incubated at 37 °C for 24 h after transfection with their respective vectors and were then applied to fluorescence-activated cell sorter to analyze the proliferation states of these samples.

Analysis of the cell cycle distribution of exponentially growing cells by propidium iodide staining revealed that 21% of cells transfected with wild type spindlin1 entered the G_2/M phase, whereas only 11% of HeLa cells and 10% of cells transfected with vectors were in G_2/M, indicating that overexpression of

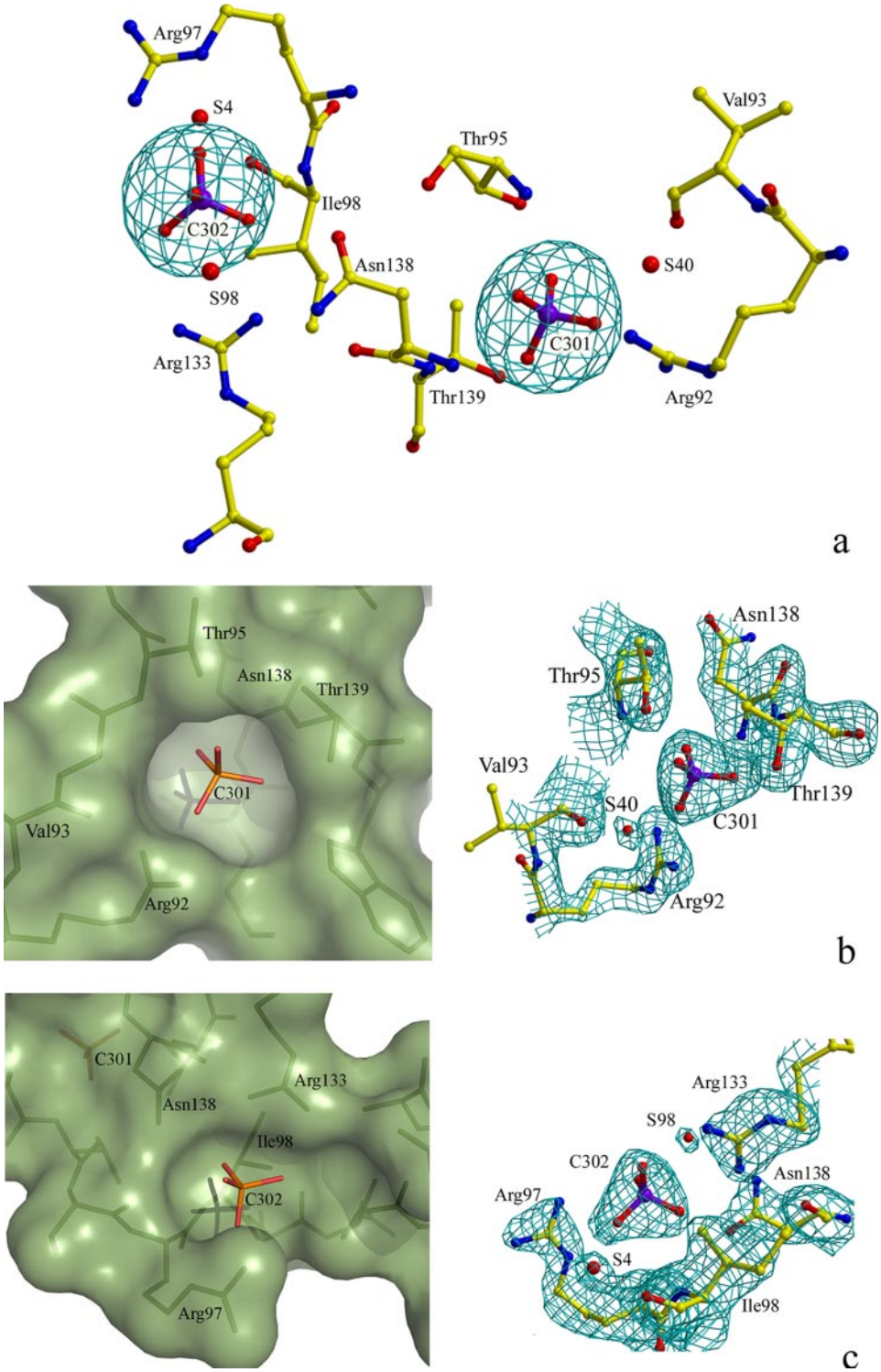

FIGURE 5. **The phosphate binding pocket and its influence on the function of spindlin1.** *a*, the anomalous difference map calculated on phosphorus, contoured at 8σ, superimposed on the refined model drawn in *stick* representation. The *red spheres* are water molecules and are labeled together with residues involved in phosphate binding. *b*, the binding site for the phosphate ion c301. *c*, the binding site for phosphate ion c302. *Left*, the surface areas that are involved in binding of phosphate groups are shown as transparent surfaces. Residues that interact with the phosphate groups are shown in *stick* representation and labeled, and the phosphate groups are shown in *stick* representation with the phosphorus atoms colored *orange*. *Right*, a portion of composite omit electron density map around the phosphate ion c301 and c302, contoured at 1.5σ, superimposed on the refined model drawn in *stick* representation. The residues and waters involved in phosphate binding are also shown and labeled.

spindlin1 would lead to a marked increase in the percentage of cells in G_2/M ($p = 0.001$). The results of fluorescence-activated cell sorter analysis also showed a slight increase of S stage cells when spindlin1 was overexpressed (34% of spindlin1 wild type, 26% of control cells, and 26% of cells transfected with vector, with a Student's *t* test difference of $p = 0.05$). These results were in agreement with previous studies (10, 11). Notably, cells expressing the spindlin1 T95A mutant showed a significant decrease in the percentage of cells in G_2/M (9%) and S (29%) phase compared with those expressing wild type spindlin1 and were nearly identical to those of the control cells (p values are 0.51 and 0.44, respectively) (Fig. 6, *a* and *b*).

Thus, our data suggest that Thr-95 is crucial for the function of spindlin1. From our structure the loss of the phosphate ions, which are maintained by hydrogen bonds, is expected to result in the distortion of the loop P. Thus, we propose that mutation of Thr-95 should break the interaction between the loop P and strands β7 and β8 of domain II, destroying its stability. Thus, the P loop should be related to the cell cycle regulation of spindlin1, although further work is required to confirm this hypothesis.

Spindlin1 Molecules Possess DNA Binding Activity—Prompted by the observation that some Tudor-like domains possess nucleic acid binding activity together with the potential cell cycle regulation, tumorigenesis, and anti-apoptosis functions of spindlin1, we performed gel shift assays to assess the binding of spindlin1 to nucleic acid.

In this assay, mixtures of different concentrations of protein and nucleic acids were analyzed on agarose gels. Our initial binding data showed that increasing the concentration of spindlin1 would lead to a shift of the band corresponding to 30 bp of dsDNA, which suggests that spindlin1 could interact with dsDNA. This interaction was not affected by altering the dsDNA

TABLE 2
Hydrogen bonds between phosphate and spindlin1

Phosphate atom	Phosphate c301 hydrogen bond partner			Distance X . . . Y	Phosphate atom	Phosphate c302 hydrogen bond partner			Distance X . . . Y
	Domain	Residue	Group			Domain	Residue	Group	
				Å					Å
O1	Hinge	Thr-95	N	3.2	O1	Hinge	Arg-97	NE	3.0
	Hinge	Thr-95	OG1	2.6		Hinge	Arg-97	NH_2	2.6
	Water (Hinge)[a]	S40 (Val-93)[a]	O (O)[a]	3.2 (2.8)[a]		Water (Hinge)[a]	S4 (Ile-98)[a]	O (O)[a]	2.6 (2.6)[a]
O2	Hinge	Arg-92	NH1	3.0	O2	Domain II	Arg-133	NH1	2.8
	Hinge	Arg-92	NH_2	2.5		Domain II	Asn-138	OD1	3.0
	Water (Hinge)[a]	S40 (Val-93)[a]	O (O)[a]	2.9 (2.8)[a]					
O3	Domain II	Thr-139	OG1	2.6	O3	Domain II	Arg-133	NH_2	3.0
	Domain II	Thr-139	N	2.8					
O4	Hinge	Thr-95	N	2.9	O4	Water (Domain II)[a]	S98 (Asn-138)[a]	O (ND2)[a]	2.8 (3.2)[a]
	Domain II	Asn-138	N	2.9		Hinge	Arg-97	NH_2	3.2

[a] The water-mediated hydrogen bonds. Values in parentheses refer to the corresponding coordinate of the bonds between respective water and peptide.

probe, indicating that the association between spindlin1 and dsDNA is nonspecific. However, similar assays using 30 bp of single-stranded DNA (ssDNA) as a probe did not show an obvious band shift, implying that spindlin1 does not bind ssDNA (Fig. 6*c*). Furthermore, spindlin1 also exhibited higher affinity for super helical than for open cycle dsDNA, as the super helical bands shifted much earlier than the open cycle bands when mixing spindlin1 with dsDNA vectors (Fig. 6*d*). Because no metal ions are found in the structure and the addition of EDTA did not inhibit the protein/DNA interaction, metal ions are not likely to be required for nucleic acid binding activity (data not shown). Our data suggest the existence of protein-nucleic acid interactions, providing a basis for investigating DNA/spindlin1 interactions and the opportunity to explore the biological consequences of this interaction in further detail.

DISCUSSION

Functional Implications of Other Tudor-containing Proteins—Although not shown in the previous bioinformatics analysis, the spin/ssty repeats and Tudor domains exhibit similar folds. Interestingly, like spin/ssty repeats, Tudor domains are usually presented in a tandem repeat manner. TUD, a protein with 11 Tudor domain repeats, is also specifically expressed during oogenesis and early embryogenesis (16). It is essential for the germ cell formation of embryos but without an obvious somatic function (17). All of these observations imply some functional relationship between spin/ssty repeat and Tudor domains.

Recent work has revealed that the Tudor domain might be a methylated protein binding domain. 53bp1, a conserved checkpoint gene for DNA double-stranded breaks (18) that contains two tandem Tudor domains, was found to bind directly to Lys-79-methylated histone H3 (19, 20). Furthermore, the Tudor domain in survival motor neuron, a gene responsible for autosomal recessive proximal spinal muscular atrophy, is required for its methylated partner binding activity (21, 22). All of these domains display a similar five anti-parallel β-strand structure with spin/ssty repeats (Fig. 3*a*), and thus, spindlin, the gene family that contains three spin/ssty repeats, might also possess a similar binding activity, although further evidence is required to confirm this.

Some other Tudor-like domains are also involved in nucleic acid binding. Research on Tudor domains in 53bp1 showed that the residues between loop β1–2 and loop β3–4 are necessary for dsDNA binding, whereas the residues Trp-18, Asn-21, Tyr-23, Asp-43, and Glu-47 should be responsible for nucleic acid binding (19). Although sequence alignment showed that these residues are largely conserved in spin/ssty repeat as well, their spatial positions are not necessarily the same (Fig. 3, *b* and *d*). Trp-37, Asn-41, and Asp-70 in spin/ssty repeat adopt similar positions in the structure, whereas Asp-67 is located in a different strand, and the side chain of Val-44 extends in an opposite direction compared with Tyr-23 of 53bp1. In another study, Steiner and co-workers (23) implied that the nucleic acid binding surface of Tudor-like domain in NusG is composed by loop β1–2 and loop β3–4 and portions of β1, β2, and β4. However, superposition of these structures shows that spindlin1 domains are greatly diverse in these loop regions, particularly in the β1–2 loop (Fig. 3, *a* and *c*). This loop is much longer in spin/ssty repeats than in other Tudor domains, leading to a greater distance between loop β1–2 and loop β3–4. It is not yet clear whether or not the spatial variations would give rise to different binding modes. Thus, whether spindlin1 interacts with nucleic acid in a similar manner to other Tudor domains awaits further resolution.

The Oligomer State Might Be Important for the Function of Spindlin1—Spindlin1 is a dimeric protein from gel filtration chromatography and dynamic light scattering experiments, and we observe in the crystal structure an interface between two spindlin1 molecules that buries 620 $Å^2$ from solvent. Based on our structure, the dimerization occurs by association of the surface A (strand β11, β14, loop β10–β11) and surface B (strand β2, loops β1–β2, β3–β4, β13–β14), which results in the formation of two equivalent surfaces that are not involved in dimer formation. These two surfaces may, therefore, associate with other tandems in a strictly identical way, as occurs in the crystal. Based on this, we speculate that spindlin1 has the potential to form higher order oligomers. Chemical cross-linking studies show that spindlin1 has a strong tendency to form a homodimer, yet a small portion of trimer can also be observed (Fig. 4*c*). Thus, the two vacant sites are likely to be the binding sites for other spindlin1 molecules.

Similar dimeric architectures are also found in other structures. The homodimer of the phosphorylation domain in PhoP, which is maintained mainly by hydrogen bonds of the acidic residues, involves an asymmetric association similar to the spindlin1 homodimer. Birck and co-workers (24, 25) also believe that the two free interacting surfaces on the tandem

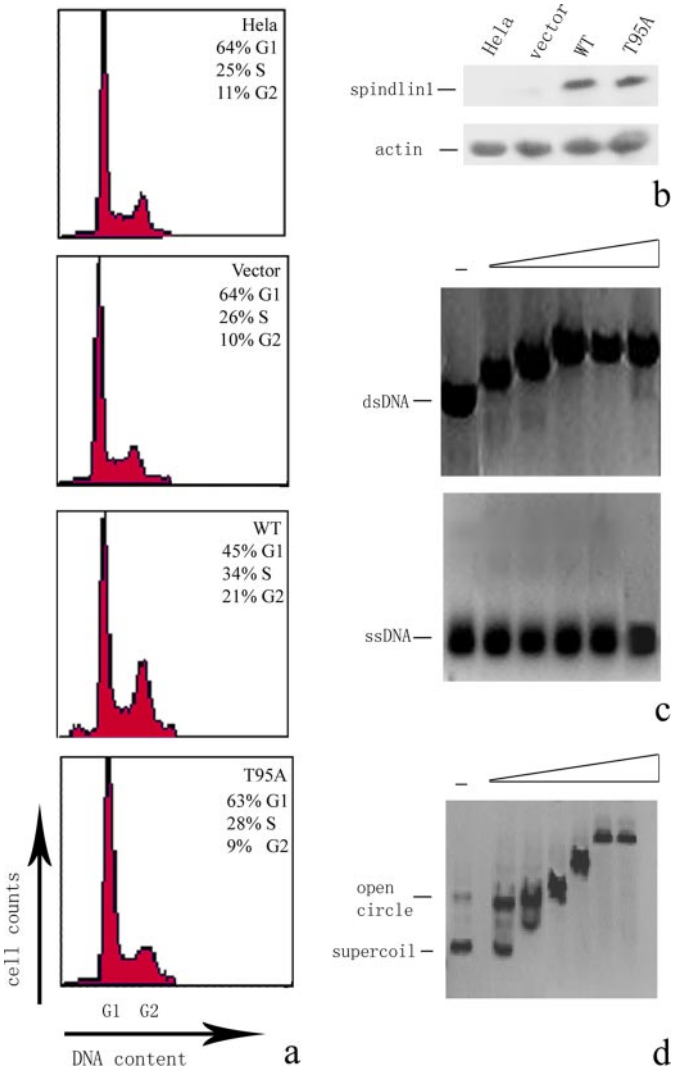

FIGURE 6. **Spindlin1 affects the cell cycle and possesses nucleic acid binding activity.** *a*, effects of spindlin1 or T95A overexpression on HeLa cell cycles. Fluorescence-activated cell sorter profiles of HeLa cells overexpressing green fluorescent protein alone, spindlin1 wild type, and T95A are presented. The cells of G_1 and G_2 are labeled. Percentages of G_1, S, and G_2 refer to green fluorescent protein-positive cells only. The cells overexpressing spindlin1 wild type protein display significant changes in the content of G_2/M cells. *b*, Western blotting of cells shown in *a*. The expression levels of spindlin1 were checked by anti-Myc antibody, and actin was used as a control for total protein. The cells transfected with WT spindlin1 and the T95A mutant have similar expression profiles. *c*, gel shift assay showing the binding of spindlin1 and dsDNA and single stranded DNA (*ssDNA*). The dsDNA and single-stranded DNA probes were labeled, and the *dash* represents control nucleic acid probe. The final concentration of spindlin1 was 2–6 mg/ml. *d*, gel shift assay showing the binding of spindlin1 and vector DNA (dsDNA). The *open circle* and *supercoil form* of control DNA was labeled, and the *dash* represents lane of vector DNA only. The final concentration of spindlin1 was 2–7 mg/ml, respectively.

repeats appear as an unprecedented unit of further oligomerization, which should be important for its function.

Spindlin1 is highly expressed at both the mRNA and protein level. Of all the maternal mRNA, about 0.35% of transcripts encode the *Spin* gene (1), whereas its protein product is repeatedly found in oocytes and early embryos and has been used as a marker protein for the cell cycle study (2, 26, 27). Based on the accumulation of spindlin1 to the cell nuclei and its potential for oligomerization, it is possible that spindlin1 is present in the oligomeric state under certain circumstances, which should be important for its function.

In summary, we have reported the three-dimensional structure of spindlin1, which is the first to be characterized from the spin/ssty gene family. The spindlin1 structure possesses a novel arrangement of the Tudor repeat domain fold. Thr-95 is a functionally important residue that plays roles in the binding of phosphate ions and stabilizing the loop between domain I and II and which is essential for regulating cell cycle progression. We have also identified double-stranded nucleic acid as its molecular binding partner, thus providing a structural and biochemical basis for further functional investigation of spindlin1.

Acknowledgments—We thank Sheng Ye for help with structure determination and Fang Chen and Wei Huang for help with electrophoretic mobility shift assay experiments.

REFERENCES

1. Oh, B., Hwang, S. Y., Solter, D., and Knowles, B. B. (1997) *Development* **124,** 493–503
2. Oh, B., Hampl, A., Eppig, J. J., Solter, D., and Knowles, B. B. (1998) *Mol. Reprod. Dev.* **50,** 240–249
3. Oita, E., Harada, K., and Chiba, K. (2004) *J. Biol. Chem.* **279,** 18633–18640
4. Frank-Vaillant, M., Haccard, O., Ozon, R., and Jessus, C. (2001) *Dev. Biol.* **231,** 279–288
5. Burgoyne, P. S., Mahadevaiah, S. K., Sutcliffe, M. J., and Palmer, S. J. (1992) *Cell* **71,** 391–398
6. Toure, A., Grigoriev, V., Mahadevaiah, S. K., Rattigan, A., Ojarikre, O. A., and Burgoyne, P. S. (2004) *Genomics* **83,** 140–147
7. Toure, A., Szot, M., Mahadevaiah, S. K., Rattigan, A., Ojarikre, O. A., and Burgoyne, P. S. (2004) *Genetics* **166,** 901–912
8. Staub, E., Mennerich, D., and Rosenthal, A. (2002) *Genome Biol.* **3,** RESEARCH0003
9. Yue, W., Sun, L. Y., Li, C. H., Zhang, L. X., and Pei, X. T. (2004) *Ai Zheng* **23,** 141–145
10. Gao, Y., Yue, W., Zhang, P., Li, L., Xie, X., Yuan, H., Chen, L., Liu, D., Yan, F., and Pei, X. (2005) *Biochem. Biophys. Res. Commun.* **335,** 343–350
11. Fletcher, B. S., Dragstedt, C., Notterpek, L., and Nolan, G. P. (2002) *Leukemia* **16,** 1507–1518
12. Jiang, F., Zhao, Q., Qin, L., Pang, H., Pei, X., and Rao, Z. (2006) *Protein Pept. Lett.* **13,** 203–205
13. Otwinowski, Z., and Minor, W. (1997) *Macromolecular Crystallography, Part A*, Vol. 276, pp. 307–326, Academic Press, Inc., New York
14. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
15. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr. A* **47,** 110–119
16. Bardsley, A., McDonald, K., and Boswell, R. E. (1993) *Development* **119,** 207–219
17. Thomson, T., and Lasko, P. (2004) *Genesis* **40,** 164–170
18. Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., and Halazonetis, T. D. (2004) *Nature* **432,** 406–411
19. Charier, G., Couprie, J., Alpha-Bazin, B., Meyer, V., Quemeneur, E., Guerois, R., Callebaut, I., Gilquin, B., and Zinn-Justin, S. (2004) *Structure* **12,** 1551–1562
20. Alpha-Bazin, B., Lorphelin, A., Nozerand, N., Charier, G., Marchetti, C., Berenguer, F., Couprie, J., Gilquin, B., Zinn-Justin, S., and Quemeneur, E. (2005) *Protein Sci.* **14,** 1827–1839
21. Friesen, W. J., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275,** 26370–26375
22. Friesen, W. J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001)

Mol. Cell **7,** 1111–1117
23. Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R., and Wahl, M. C. (2002) *EMBO J.* **21,** 4641–4653
24. Birck, C., Chen, Y., Hulett, F. M., and Samama, J. P. (2003) *J. Bacteriol.* **185,** 254–261
25. Chen, Y., Birck, C., Samama, J. P., and Hulett, F. M. (2003) *J. Bacteriol.* **185,** 262–273
26. Howlett, S. K. (1986) *Cell* **45,** 387–396
27. Szollosi, M. S., Kubiak, J. Z., Debey, P., de Pennart, H., Szollosi, D., and Maro, B. (1993) *J. Cell Sci.* **104,** 861–872
28. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
29. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15,** 305–308