Human Mismatch Repair Protein MSH6 Contains a PWWP Domain That Targets Double Stranded DNA[†]

Cédric Laguri,[‡,▽] Isabelle Duband-Goulet,[§] Nikolas Friedrich,[‡] Marianne Axt,[‡] Pascal Belin,[||] Isabelle Callebaut,[⊥] Bernard Gilquin,[*,‡] Sophie Zinn-Justin,[‡] and Joël Couprie[‡,#]

CEA Laboratoire de Biologie Structurale et Radiobiologie, iBiTec-Saclay, 91191 Gif sur Yvette, France, Laboratoire de Biologie Structurale du Noyau, Institut Jacques Monod CNRS-Université Paris 6 et 7, 2, place Jussieu 75251 Paris Cedex 05, France, CEA Service d'Ingénierie Moléculaire des Protéines, iBiTec-Saclay, 91191 Gif sur Yvette, France, and Département de Biologie Structurale, Institut de Minéralogie et de Physique des Milieux Condensés CNRS UMR7590/Uversités Paris 6 et Paris 7 Campus Boucicaut, 140 rue de Lourmel 75015 Paris, France

Received December 18, 2007; Revised Manuscript Received February 12, 2008

ABSTRACT: The eukaryotic mismatch repair (MMR) protein MSH6 exhibits a core region structurally and functionally similar to bacterial MutS. However, it possesses an additional N-terminal region (NTR), comprising a PCNA binding motif, a large region of unknown function and a nonspecific DNA binding fragment. Yeast NTR was recently described as an extended tether between PCNA and the core of MSH6 (*1*). In contrast, we show that human NTR presents a globular PWWP domain in the region of unknown function. We demonstrate that this PWWP domain binds double-stranded DNA, without any preference for mismatches or nicks, whereas its apparent affinity for single-stranded DNA is about 20 times lower. The S144I mutation, which in human MSH6 causes inherited somatic defects in MMR resulting in increased development of hereditary non polyposis colorectal cancer (*2*), is located in the DNA binding surface of the PWWP domain. However, it only moderately affects domain stability, and it does not perturb DNA binding *in vitro*.

Correction of DNA errors is an essential mechanism in maintaining genome stability. Mismatch repair (MMR[1]) contributes to the integrity of genetic material by correcting mismatches and small insertion and deletion loops occurring during replication (*3*). In the absence of MMR, these DNA errors remain uncorrected, which result in a mutator phenotype that is most often accompanied by microsatellite instability and, possibly, in cancer (*4*).

Molecular mechanisms underlying MMR are best described in *Escherichia coli*, where both the early steps in mispair detection and later steps in discrimination among the newly synthesized and template DNAs, strand nicking, and resynthesis are understood (*5*). However, the importance of MMR malfunction for human disease can only be appreciated by studying mammalian systems. It was shown that the bacterial MutS homodimer that recognizes and binds mismatches is conserved in eukaryotes in the form of two heterodimeric MutS homologue complexes with overlapping substrate specificity. The MSH2/MSH6 complex (MutSα) participates in the early recognition of mismatches and small insertion/deletion loops, while large insertion/deletion loops are mainly recognized by MSH2/MSH3 (MutSβ). MutSα binds directly to the mismatch and recruits the MLH1/PMS2 complex (MutLα). It is then postulated that MutSα, probably associated to MutLα, translocates along the DNA, until it finds a strand discontinuity, either nicked DNA or gapped DNA in the nascent DNA (*6*). Degradation of the error-containing strand by exonuclease ExoI involves several proteins (PCNA, RFC, and RPA) and follows a different pathway depending on the 5′ or 3′ location of nicked or gapped DNA relative to the mismatch (*7–10*). The resulting single-stranded DNA is then probably coated by RPA, DNA polymerase δ fills the gap, and DNA ligase I seals the remaining nick, thereby completing the repair process. Finally, eukaryotic MMR proteins are not only involved in this complex, not well-understood process but also implicated in other DNA-metabolic pathways, ranging from DNA-damage signaling to recombinogenic and mutagenic processes that are specific to multicellular organisms (*4*).

When aligned with *E. coli* MutS, the sequences of eukaryotic MSH6 proteins are characterized by an additional N-terminal region (NTR), comprising a PCNA interacting

[†] This work was supported by a grant from Association pour la Recherche sur le Cancer (number 4524).

* To whom correspondence should be addressed. Phone: 00 33 1 69 088153. Fax: 00 33 1 69 084712. E-mail: Bernard.Gilquin@cea.fr.

[‡] CEA Laboratoire de Biologie Structurale et Radiobiologie, iBiTec-Saclay.

[§] Laboratoire de Biologie Structurale du Noyau, Institut Jacques Monod CNRS-Université Paris.

[||] CEA Service d'Ingénierie Moléculaire des Protéines, iBiTec-Saclay.

[⊥] Département de Biologie Structurale, Institut de Minéralogie et de Physique des Milieux Condensés CNRS UMR7590/Uversités Paris.

[#] Present address: Laboratoire de Biochimie et Génétique Moléculaire, Faculté des Sciences et Techniques, Université de la Réunion, 15 av René Cassin, BP 7151, 97715 Saint-Denis Messagerie Cedex 9, France.

[▽] Present address: Laboratoire d'Enzymologie Moléculaire-Glycosaminoglycans, Institut de Biologie Structurale UMR 5075 CEA-CNRS-UJF, Grenoble, France.

[1] Abbreviations: MMR, mismatch repair; NTR, N-terminal region; HNPCC, hereditary nonpolyposis colorectal cancer; NOE, nuclear Overhauser effect; EMSA, electro mobility shift assay; ds, double stranded; ss, single stranded; CD, circular dichroism.

10.1021/bi7024639 CCC: $40.75
Published on Web 05/17/2008

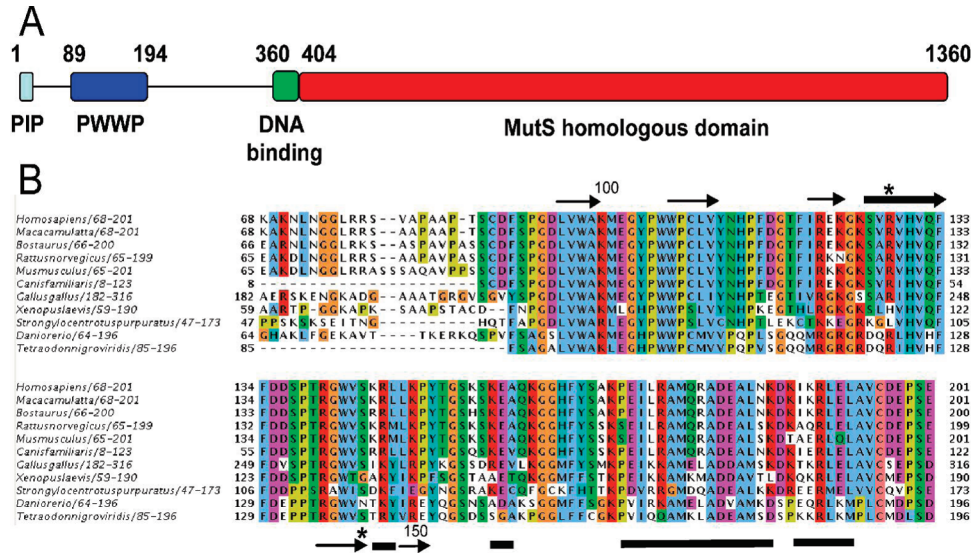

FIGURE 1: Domain arrangment of hMSH6 and alignment of eukaryotic hMSH6 PWWP domains. (A) hMSH2, hMSH3, and hMSH6 mismatch repair proteins are all highly similar to the bacterial MutS protein (in red). In hMSH3 and hMSH6, the MutS domain is extended by several hundred aminoacids in the N-terminus, but only hMSH6 possesses an additional domain predicted to be a PWWP domain (blue). The PCNA interaction box (4–11 PIP) is also represented. The hMSH6 region, similar to the yMSH6 N-terminal region and responsible for unspecific DNA binding, is also indicated (*11*). (B) Sequence alignment of N-terminal PWWP domains of MSH6 from metazoa with the hMSH6 construct used in this study (68–201). The lethal S144I and R128L mutation positions are indicated by asterisks. The secondary structure content of the hMSH6 PWWP domain is shown as lines (helices) and arrows (strands).

motif, a large region of unknown function and a nonspecific DNA binding sequence (Figure 1A (*11*)). Recently, Shell and co-workers (*1*) demonstrated that yeast NTR is an unstructured tether to PCNA. However, sequence analysis shows that the region located between the PCNA-interacting motif and the nonspecific DNA-binding region cannot be aligned between yeast and metazoan MSH6 proteins. Preliminary experiments also suggest that human MSH6 NTR is essential for MSH6 chromatin targeting (*12, 13*).

In this study, we have investigated the 3D structure and DNA binding properties of the large region of unknown function found in human MSH6 NTR. This region exhibits a PWWP domain with unique structural features between residues 89 and 194. It contains S144, whose mutation into isoleucine causes a cancer predisposition syndrome called HNPCC (hereditary nonpolyposis colorectal cancer) (*2*). It also contains R128, which is mutated into leucine in a patient suffering from endometrial cancer (*14*). We report the localization of these mutations in the PWWP 3D structure and relative to the PWWP DNA binding surface.

EXPERIMENTAL PROCEDURES

Expression and Purification of the Regions (68–201) and (68–240) of Human MSH6. Human MSH6 fragments (68–201) and (68–240) (wild type and Cys → Ser mutated) cDNAs were cloned into a pET M30 plasmid (EMBL, Protein Expression Facility). Thus, the resulting protein fragments were expressed as fusion proteins, with an N-terminal (6xHis)-tagged GST protein followed by a TEV protease cleavage site. The pET M30 plasmids were used to transform BL21(DE3) competent cells (Invitrogen). Cells were induced at an OD of 0.8 with 1 mM IPTG during 4 h at 30 °C in LB or minimal Medium supplemented with $^{15}NH_4Cl$ and ^{13}C-Glucose. Cell disruption was performed using an Eaton press. After centrifugation, disrupted cell supernatant was recovered and treated with benzonase endonuclease (Sigma-Aldrich) for 30 min at room temperature, then subjected to Glutathion Sepharose chromatography (GS4B, Amersham) and the GST-PWWP fusion was eluted with 20 mM GSH. The fusion protein was cleaved by TEV protease (overnight at 4 degrees) and the mixture injected on a Ni-NTA Column (Qiagen), which retains GST and uncleaved fusion proteins. MSH6 fragments were further purified on a SPsepharose Fast-Flow to reach purity greater than 98%, as observed on Coomassie blue stained SDS−PAGE.

The S144I and Cys → Ser mutations were introduced using the Quickchange site-directed mutagenesis kit (Stratagene) following the Quickchange standard protocol. The presence of the desired mutation(s) was checked by DNA sequencing using an ABI prism sequencer (Applied Biosystems). Expression and purification of the mutant proteins was performed in the same conditions as those for the wild type proteins.

NMR Experiments. NMR samples were prepared in Na-Phosphate buffer (50 mM) at pH 6.9, 225 mM NaCl, and 10% D_2O, with a protease inhibitor cocktail (SIGMA) and 1 mM 3-(trimethylsilyl)[2,2,3,3-2H4] propionate (TSP). Backbone and side-chain assignment experiments were performed at 27 °C on Bruker DRX-500 and DRX-600 spectrometers equipped with a triple resonance cryoprobe. NOESY experiments were recorded on the 800 MHz Varian spectrometer at IBS in Grenoble (^{15}N-HSQC NOESY, ^{13}C-HSQC-NOESY aliphatic and aromatic). Hydrogen bond restraints were derived from slowly exchanging amide protons, identified by dissolving the protein in D_2O and by recording successive $^{1}H,^{15}N$ HSQC experiments. Processing was performed with NMRPipe (*15*) and analysis with Felix (Accelrys) using macros provided by J. Craven (Sheffield University, U.K.).

NMR Resonance Frequency Assignment. Backbone resonance assignment was performed using standard triple resonance NMR experiments (*16*), and the vast majority of

the backbone resonances could be assigned. Five residues from the unstructured N-terminal part of the domain were not assigned because of overlaps in the central part of the ^{1}H-^{15}N HSQC spectrum (K68, N71, L72, N73, and R78). Several residues from region (113–128) had missing amide proton assignments, and their side-chain resonances were only partially assigned through the determination of NOE connectivities (H113, T118, F119, I120, E122, S126, V127, and R128). Residues 123 to 125 remained completely unassigned. Region (113–128) is probably involved in a conformational exchange process, and/or its amide protons are rapidly exchanging with water. Y151 aromatic ring resonances could not be assigned, whereas its side chain is buried in the calculated structures and lies at close proximity to the T152 and G153 backbone. The G153 amide proton remained unassigned too, probably as a result of exchange broadening due to motions of the Y151 ring.

Solution Structure Determination. The solution structure of the human MSH6 region (68–201) was solved on the basis of 2589 interproton distances deduced from 3800 NOE cross-peaks. A semiautomated iterative assignment procedure was applied for the NOE assignment and construction of the 3D structures (*17*). A force field adapted to NMR structure calculation (file parallhdg.pro in CNS 1.0 (*18*)) was used. 174 torsion angle (Φ and Ψ) values were deduced from the backbone ^{1}H, ^{15}N, and ^{13}C chemical shifts using the program TALOS (*19*). Twenty-four hydrogen bonds were imposed at a later stage of the structure calculation process. At the last iteration of the procedure, 1000 structures were calculated; the 40 best energy structures were retained and refined with a standard energy function (CHARMM22), including an electrostatic energy term. This term is calculated with no net charge on the side-chain atoms and with a distance-gated dielectric constant. The 20 structures of lowest overall energy were retained after the refinement process. Structural statistics (Table 1) were calculated using in-house scripts and Procheck-NMR (*20*) and the structure deposited in the RCSB database (PDB: 2GFU). All figures were created with the Pymol software (Delano Scientific, LLC).

DNA Preparation. The 146 bp DNA fragment was generated by polymerase chain reaction with a thermostable DNA polymerase (Promega) using a PTC-100 PCR System (MJ Research, Inc.). The 146 bp DNA fragment, obtained from the *Bam*HI and *Dra*I double digest of the plasmid pUC (357.4) (*21*), was used as the template and the oligonucleotides 5′-GATCCTCTAGAGTCCGGCTAC-3′ and 5′-AAAGGGTCAGGGATGTTATGACG-3′ as sense and antisense primers. The four double-stranded DNA fragments 35 base pairs long were obtained by annealing oligonucleotides (see Table 2 below) purchased from MWG. The 35 bp fragment was obtained by annealing oligonucleotides 1 and 2, the 35 bp fragment containing a G/T mismatch by annealing oligonucleotides 1 and 3, the 35 bp fragment containing a nick by annealing oligonucleotides 1, 4, and 5, and the 35 bp fragment containing a gap by annealing oligonucleotides 1 and 4. The single-stranded DNA of 35 nucleotides was oligonucleotide 1. Annealing was performed in 10 mM Tris HCl at pH 8.0, 1 mM EDTA (ethylene diamine tetraacetic acid), and 200 mM NaCl by heating at 80 °C for 5 min followed by slow cooling. Correct annealing was controlled on 8% native polyacrylamide gels. 5′-End labeling with ^{32}P-ATP and T4 polynucleotide kinase was performed at 25 °C for oligonucleotides and at 37 °C for the 146 bp fragment according to standard protocols (*22*).

Table 1: Structural Statistics for the HsMSH6 [68–201, Cys → Ser] PWWP domain (89–194)

number of violations	
NOE distance restraints >0.5 Å	0
dihedral restraints >10°	0
experimental restraints	
distance restraints	
total	2589 (rmsd 0.0247 +/−0.0003)
unambiguous	2196
ambiguous	393
hydrogen bonds	24
dihedral restraints (phi/psi)	174 (rmsd 0.87 +/− 0.08)
rms deviation from idealized covalent geometry	
bonds (Å)	0.0165 +/−0.0003
angles (deg)	3.23 +/−0.05
impropers (deg)	2.3 +/−0.2
energy (kcal/mol)[a]	
van der Waals	15 +/−20
electrostatic	−406 +/−25
Ramachandran plot (%)	
most favored regions	83.2
additionally allowed regions	14.2
generously allowed regions	2.3
disallowed regions	0.3
coordinate precision[b] (Å)	
backbone atoms	1.05 +/−0.13
heavy atoms	1.6 +/−0.14

[a] The van der Waals energy is calculated with a Lennard−Jones potential. The electrostatic energy is calculated with no net charge on side-chain atoms and a distance-gated dielectric constant. CHARMM22 parameters are used. [b] Precision compared to the best energy structure.

Table 2: Oligonucleotides Used To Prepare 35 nt DNAs for Gel Shift Assays

Oligo 1	5′-gggGCATGCCTGCAGGTCGACTCTAGAGgatcccc-3′
Oligo 2	3′-cccCGTACGGACGTCCAGCTGAGATCTCctagggg-5′
Oligo 3	3′-cccCGTACGGACGTCCAGtTGAGATCTCctagggg-5′
Oligo 4	3′-cccCGTACGGACGTCCAGCT-5′
Oligo 5	3′-GAGATCTCctagggg-5′

EMSA (Electro Mobility Shift Assay) Experiments. Proteins in 50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 1 mM 4-(2-aminoethyl) benzenesulfonyl fluoride, 1 mM TCEP, 50 mM NaCl, and 0.1% triton X-100 were incubated with the radioactive DNA fragments for 30 min at 22 °C. Each experiment contained 0.5 pmol of DNA (26 nM). Protein−DNA complexes were analyzed on 5.5% polyacrylamide gels (5% for 146 bp DNA interaction) at an acrylamide to bisacrylamide ratio of 29/1 (w/w), in 0.5× TGE [12.5 mM Tris-HCl (pH8.4), 95 mM glycine, and 0.5 mM EDTA]. After 30 min of pre-electrophoresis, samples were loaded onto the gels and resolved by electrophoresis at 70 V for 30 min to 1 h, depending on the size of the DNA. DNA retardation was detected by autoradiography of dried polyacrylamide gels at −80 °C using Biomax MR films (Kodak) and an intensifying screen. For affinity measurements, dried polyacrylamide gels were exposed to a phosphor screen, and measurements of the radioactive signals were performed with a STORM 860 scanner (Amersham) using image Quant software (Molecular Dynamics, Inc.).

NMR Protein−DNA Titration. The 35 bp DNA fragment used for NMR titration was prepared as described above,

exchanged with standard NMR buffer and concentrated with a vivaspin concentrating device (final concn, 900 μM). DNA was slowly added to the MSH6 fragment (68–201) mutated on its cysteine residues at 160 μM concentration. Dilution of the protein sample did not exceed 30% at the end of titration. Chemical shift variations were calculated as $\delta = ((\delta^1H)^2 + (\delta^{15}N/10)^2)^{1/2}$. The chemical shift variations considered as significant upon DNA interaction were at least equal to the mean value plus one standard deviation of all chemical shift variations.

CD and 1D NMR of Mutant Domains. Circular dichroism experiments were performed with a Jobin Yvon CD6 spectrometer in a 1 cm path length cell. Protein concentration was 10 μM in 50 mM NaH_2PO_4 at pH6.9, 11 mM NaCl, and 0.4 mM TCEP (tris-(2-carboxyethyl)-phosphine). Comparison of 1D ^{1}H NMR spectra of native, Cys → Ser, and S144I mutants was performed in the standard NMR buffer (see above) at 75 μM concentration and 20 or 27 °C.

RESULTS

Delimitation of a Globular Domain in the NTR of Metazoan MSH6. Two different fragments of human MSH6 NTR, from residue 68 to either residue 201 or 240, were successfully expressed and purified. A shorter construct (residues 68 to 163) proved to be largely unfolded. All these fragments are found between the PCNA binding site and the nonspecific DNA binding region (Figure 1A; (*11*)). They are rich in hydrophobic amino acids and are highly conserved from fish to humans (Figure 1B). An ^{1}H,^{15}N HSQC spectrum was recorded on the fragment 68–201, and a TROSY-based ^{1}H,^{15}N HSQC spectrum was obtained on the fragment 68–240 (data not shown). These spectra showed a similar localization of the well-dispersed peaks, consistent with the presence of a folded core. However, NMR signals corresponding to the region 202–240 cluster at positions commonly assigned to unstructured residues. The domain (68–201), further named HsMSH6 [68–201], was thus chosen for solution structure determination but was subject to rapid aggregation. Aggregation could be substantially slowed down by optimizing the buffer conditions (*23*) and by adding the strong reducing agent TCEP. However, in order to avoid cysteine oxidation, a mutant HsMSH6 [68–201, Cys → Ser] was produced, in which the three cysteines were changed into serines. ^{1}H,^{15}N HSQC spectra of native and mutated HsMSH6 [68–201] fragments are highly similar, suggesting that these two fragments adopt the same 3D structure (Supporting Information, Figure 1). Further study of the MSH6 region 68–201 was carried out on the mutant HsMSH6 [68–201, Cys → Ser].

Description of the 3D Structure of HsMSH6 [68–201, Cys → Ser]. The HsMSH6 [68–201, Cys → Ser] fragment was analyzed by ^{1}H, ^{15}N, and ^{13}C NMR. Resonance frequencies were assigned using conventional heteronuclear methods (*16*). Analysis of heteronuclear ^{1}H → ^{15}N NOE experiments showed that HsMSH6 [68–201, Cys → Ser] is structured between residues 89 and 194 (data not shown). Calculation of the 3D structure of this fragment was carried out using the software INCA (*17*), on the basis of 2589 interproton distances, 174 Φ,Ψ angle values, and 24 hydrogen bonds. Structural statistics of the 20 energetically best structures are shown in Table 1. The HsMSH6 [68–201, Cys → Ser] fragment is composed of a central five-stranded β-barrel domain $\beta1-\beta5$ ($\beta1$, 95–98; $\beta2$, 106–110; $\beta3$, 126–133; $\beta4$, 139–144; $\beta5$, 148–150) (Figure 2). An additional strand $\beta2'$ (120–123) forms a sheet with the strand $\beta3$ of the β-barrel. Helix 3_{10} α1 (145–147) allows a sharp turn between $\beta4$ and $\beta5$. Helix 3_{10} α2 (157–159) is located in a mostly extended linker (152–169) connecting the central β-barrel to α-helices α3 (170–185) and α4 (187–193) (Figure 2). These helices are packed onto the β-barrel region through hydrophobic interactions with the β-sheet $\beta1\beta2$.

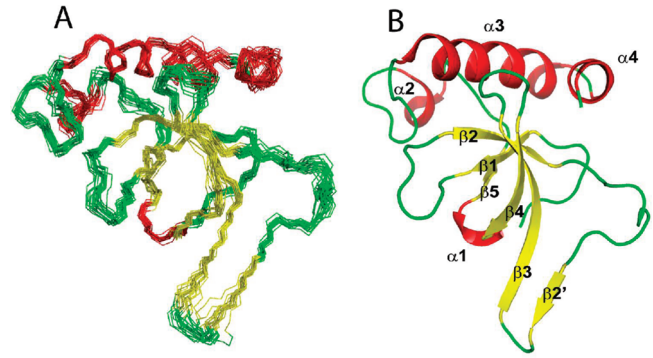

FIGURE 2: Solution NMR structure of HsMSH6 [68–201, Cys → Ser] PWWP domain. (A) Backbone superimposition of the 20 lowest-energy refined structures (red, helices; yellow, β sheets). Only the folded region (89–194) is shown for clarity. (B) Ribbon representation of the best energy structure.

The 3D structure of fragment (89–194) was submitted to the DALI server (*24*). As predicted by domain recognition softwares, this structure resembles known 3D structures of PWWP domains (Figure 3). In particular, structures referenced in the Protein Data Base as 1N27 (from *M. musculus* HDGF), 1H3Z (from *S. pombe* protein SPBC215.07c), 2DAQ (from *H. sapiens* protein WhscIII), and 1KHC (from *M. musculus* Dnmt3b) are structurally aligned by DALI on more than 82 amino acids of HsMSH6 [68–201, Cys → Ser], with a Z-score higher than 6 and a backbone root-mean-square deviation (rmsd) lower than 3.9 Å. PWWP domains all share a 5-stranded β-barrel and a large α-helix (here α3) of variable length packed onto the β-barrel. HsMSH6 [68–201, Cys → Ser] also presents three unique structural regions: strand $\beta2'$, which forms a β-sheet with $\beta3$ of the β-barrel, the large linker (152–169), and helix α4.

Human MSH6 PWWP Domain Binds DNA. As several PWWP domains have been reported to bind DNA (*25, 26*), we have tested the interaction of human MSH6 PWWP domain with DNA by EMSA (electro mobility shift assay) and NMR. First, the *in vitro* interaction of wild type or mutant HsMSH6 [68–201] domain with a 146 bp DNA fragment was followed by electrophoresis in a 5% polyacrylamide gel. DNA forms several complexes with HsMSH6 [68–201, Cys → Ser], as demonstrated by delayed migration relative to free DNA (Figure 4). Saturation of a 146 bp DNA by HsMSH6 [68–201, Cys → Ser] was complete using 4 M excess of protein, and an apparent K_d of 5.6 ± 2.4 nM was deduced. Similar results were obtained with the PWWP wild type domain, even if some aggregation was observed at high protein concentrations (data not shown).

In order to define the structural determinants of the PWWP−DNA interaction, we designed five short DNA fragments of 35 nucleotides containing DNA structures encountered during mismatch repair: double-stranded, nicked,

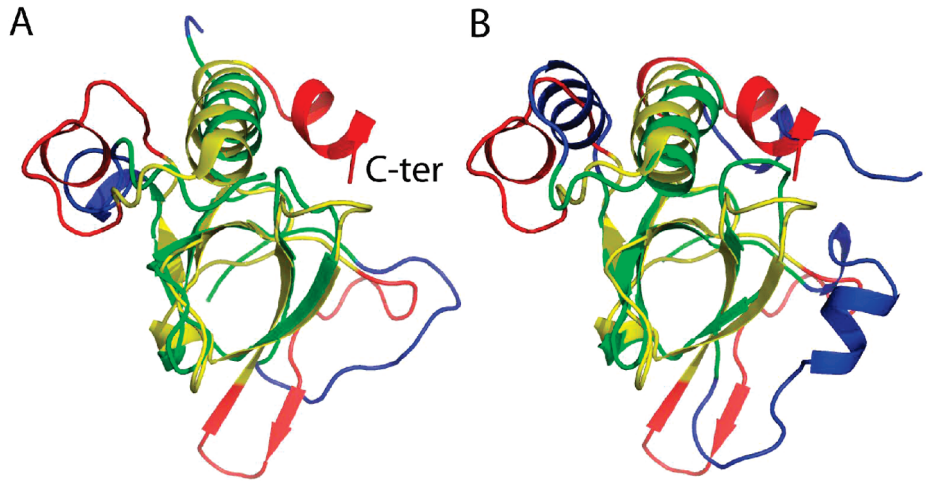

FIGURE 3: Superimposition of HsMSH6 [68–201, Cys → Ser] PWWP domain (89–194) with the two closest PWWP structures (DALI). Structures have been superimposed on the common PWWP motif (β-barrel + long α-helix) with hMSH6 in yellow and the compared structure in green. Additional elements to the canonical PWWP fold are represented in red for hMSH6 and blue for the compared structures. (A) Superimposition with *M. musculus* HDGF (8–96) (PDB 1N27). rmsd = 1.89Å on 67 Cα (β-barrel + long α-helix). (B) Superimposition with *S. pombe* protein SPBC215.07c (123–225) (PDN 1H3Z). rmsd = 2.04Å on 64 Cα (β-barrel + long α-helix).

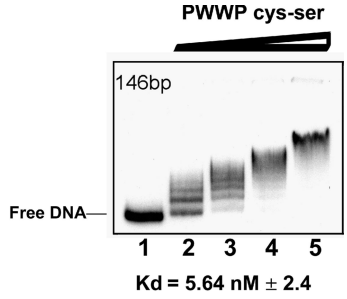

FIGURE 4: PWWP domain binds DNA. The 146 bp DNA fragment at a concentration of 26 nM was incubated in the presence of increasing concentrations of HsMSH6 [68–201, Cys → Ser] PWWP domain: 26 nM (lane 2), 52 nM (lane 3), 104 nM (lane 4), and 208 nM (lane 5). Peptide−DNA complexes with decreasing mobilities were observed. Lane 1 indicates the mobility of free DNA. The dissociation constant K_d was calculated using the equation $K_d = ([D] - [DP])([P] - [DP])/[DP]$, where [D] is the total DNA concentration, [P] the total peptide concentration, and [DP] the peptide−DNA complex concentration. The mean value obtained from three independent experiments is indicated with the standard deviation.

containing a G/T mismatch, 5′ recessed, and single-stranded DNAs. We first tested the interaction of the PWWP domain with double-stranded (ds) and single-stranded (ss) DNA fragments. Figure 5A, left panel shows that incubation of HsMSH6 [68–201, Cys → Ser] with the ds DNA fragment generates two types of complexes, C1 and C2. The slower migrating complex C2 is progressively enriched with increased protein concentrations. The estimated apparent K_d value for these complexes is 134 ± 5 nM. By contrast with the ds DNA, the ss DNA forms only one complex with the PWWP mutant with a reduced affinity ($K_d = 2.5 \pm 0.4$ μM) (Figure 5A, right panel).

Interaction of HsMSH6 [68–201, Cys → Ser] with G/T mismatched, nicked, and 5′ recessed DNA yields two complexes C1 and C2, as observed for homoduplex ds DNA. Furthermore, human MSH6 PWWP mutant possesses a similar apparent affinity toward ds, ds G/T, or ds nicked DNA (Figure 5B). A reduced affinity is observed for 5′ recessed 35 bp DNA that contains 15 unpaired nucleotides ($K_d = 260 \pm 60$ nM; Figure 5B right panel). This probably reflects the weaker binding of the domain to ss versus ds DNA. Thus, HsMSH6 [68–201, Cys → Ser] preferentially binds to homoduplex DNA with or without a G/T mismatch or a nick.

In order to identify the residues of the PWWP domain involved in DNA binding, the 35 bp ds DNA used in the previous biochemical experiment was progressively added to an NMR sample of ^{15}N-labeled HsMSH6 [68–201, Cys → Ser], and the NMR ^{1}H,^{15}N-HSQC fingerprinting of the protein fragment was followed throughout the titration (Figure 6A−C). Upon the addition of DNA, several protein HN resonances experienced significant chemical shift variations suggesting a rapid exchange regime. Chemical shift changes higher than 0.011 ppm (i.e., mean variation plus one standard deviation) all correspond to residues localized on one face of the protein (Figure 6B and C). These residues are found on a line centered on loop $\beta1\beta2$ and the PWWP motif (W97, A98, M100, E101, G102, W105 (Hε), and W106 (Hε)), and going from $\beta2'$ (R121), the loop $\beta2\beta2'$ (F115 and D116), and $\beta3$-strand (G141, W142(Hε), and S144) to the $\beta5$-α3 loop (Q160, G162, and Y166) (Figure 6A and B).

Calculation of the electrostatic potential at the surface of the PWWP domain shows a large positive surface (Figure 6D) formed by several basic conserved residues K99 ($\beta1$), R121 ($\beta2'$), K123 ($\beta2'$), K125 (turn $\beta2'-\beta3$), R128 ($\beta3$), R140 ($\beta4$), and R146 ($\beta3$). The interaction surface found in the NMR titration experiment overlaps this very basic surface. Nevertheless, as the amide proton frequencies of residues 122–128 are not assigned, the role of these residues in DNA binding could not be assessed.

Evaluation of the Impact of the Cancer-Related S144I Mutation of Human MSH6 PWWP Domain. In order to understand the consequences of the S144I pathological mutation for the function of the PWWP domain, the mutation was introduced in HsMSH6 [68–201], as well as in HsMSH6 [68–201, Cys → Ser]. The PWWP S144I mutants were successfully expressed and purified, and their stability was investigated by temperature-dependent unfolding using circular dichroism. These experiments demonstrated a 7 °C decrease in melting temperature for S144I mutants compared to HsMSH6 [68–201] or HsMSH6 [68–201, Cys → Ser] (47

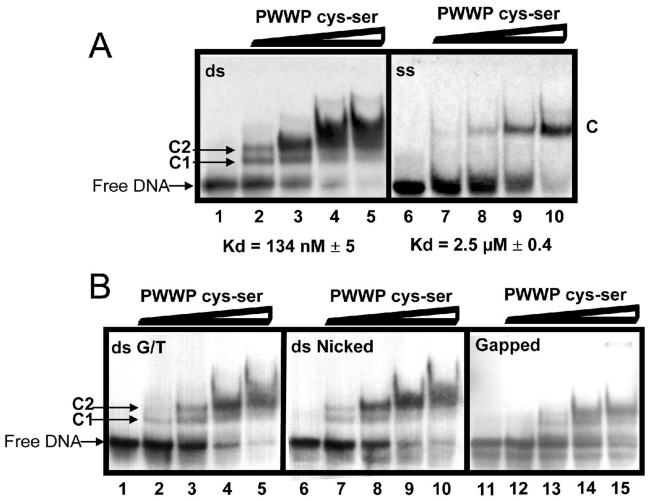

FIGURE 5: PWWP domain binds preferentially to double stranded DNA. (A) Interaction of the HsMSH6 [68–201, Cys → Ser] PWWP domain with double or single stranded DNAs. In the left panel (lanes 1 to 5), the 35 bp DNA fragment (ds) at a concentration of 26 nM was incubated in the presence of increasing concentrations of PWWP Cys-Ser mutant: 26 nM (lane 2), 78 nM (lane 3), 234 nM (lane 4), and 780 nM (lane 5). Two peptide−DNA complexes, C1 and C2, were observed. Lane 1 indicates the mobility of the free DNA. In the right panel (lanes 6−10), the single stranded DNA fragment (ss) at a concentration of 26 nM was incubated in the presence of increasing concentrations of PWWP Cys-Ser mutant: 260 nM (lane 7), 780 nM (lane 8), 2.34 μM (lane 9), and 7.8 μM (lane 10). Only one peptide−DNA complex (C) was formed with the single stranded DNA. Lane 6 indicates the mobility of free DNA. The dissociation constants K_d were calculated as mentioned in the legend for Figure 4. The mean values were obtained from three independent experiments. (B) Interaction of HsMSH6 [68–201, Cys → Ser] PWWP domain with DNA fragments containing other structural elements. The 35 bp DNA fragments containing a G/T mismatch (ds G/T, lanes 1−5), a nick (ds Nicked, lanes 6−10), or a 5′ recessed gap (Gapped, lanes 11−15) at a concentration of 26 nM were incubated with increasing concentrations of PWWP Cys-Ser mutant: 26 nM (lanes 2, 7, and 12), 78 nM (lanes 3, 8, and 13), 234 nM (lanes 4, 9, and 14), and 780 nM (lanes 5, 10, and 15). Formation of the peptide−DNA complexes C1 and C2 was observed with the three DNA fragments. Lanes 1, 6, and 11 indicate the mobility of free DNA.

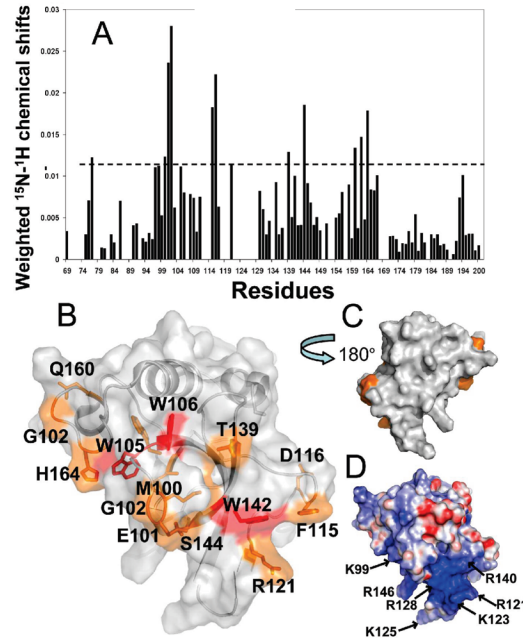

FIGURE 6: Interaction of HsMSH6 [68–201, Cys → Ser] PWWP domain with 35 bp ds DNA monitored by NMR. DNA (35 bp ds) was progressively added to the PWWP domain, and ^{15}N-HSQC were recorded. (A) Weighted amide proton chemical shifts at 1:1 protein/DNA molar ratio. (B) Visualization of DNA interaction on hMSH6 PWWP surface (89–194). Residues with HN (orange), or Trp Hϵ protons (red) experiencing chemical shift variations greater than 0.011 are colored and form a line on the PWWP domain. (C) 180° view of (B) showing that only one face of the PWWP domain is affected by DNA binding. (D) Electrostatic surface (calculated with APBS (*42*)) of hMSH6 PWWP domain (89–194)) in the same orientation as that in (B).

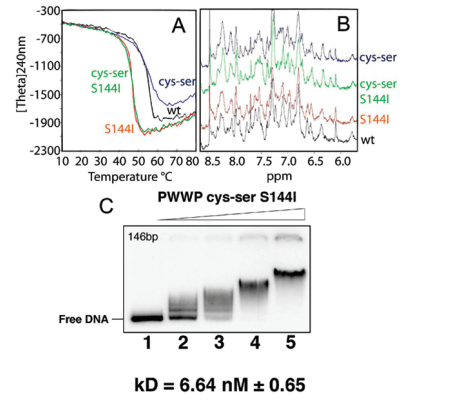

FIGURE 7: Analysis of native and Cys-Ser mutant hMSH6 PWWP domains with or without the S144I lethal mutation. (A) Temperature dependent unfolding hMSH6 native and mutants monitored by circular dichroism at 240 nm. (B) One-dimensional NMR spectra. (C) DNA binding of of HsMSH6 [68–201, Cys → Ser, S144I]. The 146 bp DNA fragment at a concentration of 26 nM was incubated in the presence of increasing concentrations of S144I mutated HsMSH6 [68–201, Cys → Ser] PWWP domain: 26 nM (lane 2), 52 nM (lane 3), 104 nM (lane 4), and 208 nM (lane 5). Lane 1 indicates the mobility of the free DNA. The dissociation constant K_d and the standard deviation were calculated as indicated in the legend for Figure 4.

vs 55 °C, Figure 7A). Nevertheless, all the native and mutated domains exhibit similar folds at 27 °C, as shown by CD and 1D ^{1}H NMR (Figure 7B).

Then, DNA binding of HsMSH6 [68–201, Cys → Ser, S144I] was analyzed by EMSA (Figure 7C). Interaction of this mutant domain with the 146 bp DNA is comparable to the interaction observed for HsMSH6 [68–201, Cys → Ser]: the apparent K_d for DNA yields 6.6 nM ± 0.6 in the case of the mutant versus 5.6 nM ± 2.4 in the case of the wild type domain. Similar results were obtained when observing the interaction of HsMSH6 [68–201, Cys → Ser, S144I] with the shorter double-stranded 35 bp DNA fragment (data not shown). Thus, the S144I mutation has an impact on protein stability but does not affect protein structure and DNA binding.

DISCUSSION

PWWP as DNA Binding Domains. PWWP domains are abundant in eukaryotic multidomain proteins involved in cell division, growth, and differentiation (*27*). PWWP domain proteins for which data are available appear to be nuclear, often DNA-binding proteins that function as transcription factors regulating developmental processes (*27*). However, direct *in vitro* PWWP−DNA interaction was reported only in the case of Dnmt3b (*26*), HDGF (*25*), and LEDGF (*28*).

We have solved the 3D structure of the PWWP domain (89–194) of human MSH6 protein. This structure exhibits a

5-stranded β-barrel stacked onto a large α-helix, as found for other PWWP domains, but also presents unique additional structural elements. A new β-sheet is found between β2 and β3, an extended loop containing a 3_{10} helix connects β5 to α2, and an additional helix α4 is observed after α3 (Figure 3). Three-hundred nineteen PWWP domains are referenced in the SMART database (*29*), and five structures of PWWP domains have been determined to date (*16, 25, 26, 30, 31*). These PWWP domains possess unique combinations of insertion/deletions, and it is reasonable to suggest that the diverse activities assumed by PWWP domains are at least partially carried by these unique elements (*31*).

We report that human MSH6 PWWP domain binds to double-stranded DNA with an apparent affinity that varies with DNA size ($K_{d146bp} = 5$ nM, $K_{d35bp} = 134$ nM). Multiple bands are observed by EMSA when the PWWP domain is in the presence of large DNAs (357 and 211 bp, data not shown; 146 bp; see Figure 4), suggesting that different ratios of peptide/DNA are found within the complexes. It is difficult to assess a possible cooperativity between the multiple DNA binding events. Nevertheless, the number of bands observed by EMSA decreases when DNA size is reduced, indicating that the number of PWWP binding sites is directly linked to DNA size. Only two complexes are observed when the DNA is 35 bp long. This suggests that the 35 bp DNA can bind to one or two PWWP peptides (the measured apparent affinity takes into account both complexes because of the proximity of the two corresponding bands on the EMSA gels) and that a PWWP binding site is approximately 17 bp long. No DNA sequence specificity is observed, as no sequence identity could be found on more than 4 nucleotides between the 35 and 146 bp DNA fragments. Nonspecific DNA binding was reported for Dnmt3b and HDGF PWWP domains (*13, 25*). Only the LEDGF PWWP domain seems to participate in sequence specific binding to Guanine/Adenine repeats (*28*). MSH6 PWWP domain also binds to single-stranded DNA, but with a 20 times lower apparent affinity. Consistently, it binds to the 5′ recessed DNA with an intermediate affinity, twice lower than the affinity for double-stranded DNA but 10 times higher than the affinity for single-stranded DNA.

We have mapped the DNA binding site on the PWWP domain of MSH6. This binding site involves one face of the protein. (Figure 6). It is reminiscent of the DNA binding site mapped by NMR on the PWWP domain of HDGF (*25*). Indeed, both DNA binding surfaces comprise loop β1β2, the β4-strand, loop β4β5, and the helical turns located between β5 and α3. Interestingly, the MSH6 PWWP DNA binding site also involves specific structural elements, namely, loop β2β2′, the β2′-strand, and loop β2′β3. The β-hairpin β2′-β3 is highly positively charged, which makes it a putative DNA binding surface; however, as this hairpin is partially unassigned, its participation to DNA binding was difficult to evaluate.

Structural and Functional Relationships between Tudor and PWWP Domains. The PWWP domain belongs to the royal structural family also containing Tudor, MBT, and chromo domains (*32*). These domains, found in proteins associated with chromatin, are reported to interact with methylated lysines or arginines through a cluster of aromatic residues (*33–36*). The MSH6 PWWP domain also possesses an aromatic cluster (Y103, W106, and F133) at an equivalent position (*35*). Nevertheless, to date, no PWWP domain was reported to bind proteins carrying methylations (*33*). We show that the human MSH6 PWWP aromatic cluster is involved in DNA binding. This could suggest that, whereas PWWP and other royal domains share their β-barrel topology and aromatic cluster, there is no functional relationship between them. Alternatively, the same surface of the PWWP domain might be involved in interactions with both DNA and a yet unidentified methylated partner.

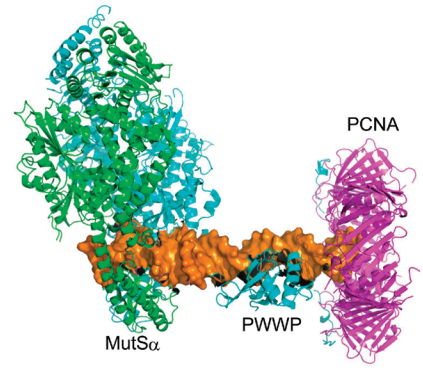

FIGURE 8: Model of the interaction of hMSH6 (cyan) with DNA (orange). The PWWP domain was placed between MutSα (PDB: 2O8B) and PCNA (PDB: 1AXC). MSH2 is colored in green and PCNA in magenta.

Interestingly, if yeast and *C. elegans* MSH6 do not exhibit an N-terminal PWWP domain, *A. thaliana* MSH6 presents an N-terminal Tudor domain (*29*) followed by predicted α-helices (*37*). This Tudor domain probably shares functional properties with the PWWP domain of human MSH6. Thus, domain prediction softwares cannot always distinguish between royal domains presenting functional properties related to Tudor or PWWP domains.

NTR of Human MSH6 Is a Partially Structured Tether to PCNA. Recently, Clark et al. (*11*) reported that both yeast and human MSH6 NTR bind to homoduplex DNA, heteroduplex DNA containing a G/T mismatch, and also single-stranded DNA, although with a lower affinity. They identified a fragment of yeast MSH6, from residue 231 to residue 299, which strongly binds DNA. Deletion of this fragment provokes elevated mutation rates in yeast. Furthermore, the homologous region 360–404 in human MSH6 corresponds to the NTR fragment observed in the X-ray structure of the human MSH6−MSH2 complex (*38*). It adopts an extended conformation with few secondary structure elements and a large number of positively charged residues. Deletion of this fragment from the MSH6 construct studied by X-ray crystallography strongly reduces its binding to cellulose derivatized with single-stranded DNA, consistent with a role of this fragment in nonspecific DNA binding.

We show here that in human, the NTR fragment (68–201) of MSH6 exhibits DNA binding properties similar to those reported by Clark et al. (*11*) for the whole yeast and human NTR, that is, strong binding to homo and heteroduplex DNA, weaker binding to single-stranded DNA. This fragment is specific to metazoan MSH6. It may be responsible for the increased DNA binding observed for hMSH6 NTR compared to that for yMSH6. It is located between the PCNA binding motif of MSH6 and the DNA binding motif described by Clark et al. (*11*) (Figure 8). In yeast, the 234 residues between the PCNA and DNA binding motifs are unstructured. In the metazoan, most of the corresponding 350 residues probably

form an unstructured tether to PCNA, as predicted using a disorder analysis software (*39*), but 106 of these residues fold into a globular domain, the PWWP domain, that binds DNA.

In conclusion, we have discovered a nonspecific DNA binding PWWP domain in the N-terminal region of MSH6 from higher eukaryotes. Mutations in human proteins involved in MMR (Mlh1, Msh2, Msh6, and Exo1) cause inherited somatic defects in MMR resulting in increased development of a cancer predisposition syndrome called HNPCC (hereditary nonpolyposis colorectal cancer). The N-terminal domain of human MSH6 contains one of these mutations, S144I, that has been found in four, nonrelated families, in a particularly well-conserved codon (*2*). This mutation causes endometrial or colorectal cancers at an early age (*14*). The N-terminal domain also contains the mutation R128L, which was identified in a patient suffering from endometrial cancer (*14*). Interestingly, S144 and R128 are in contact. Moreover, they cluster on a positively charged surface involved in DNA binding. Structural analysis of the N-terminal domain mutated at S144 reveals that the fold of this mutant is conserved, while its stability is slightly diminished compared to that of the native domain. Interaction with DNA is not perturbed by the mutation S144I. Interestingly, several phosphorylation sites have been proposed for human MSH6 NTR, including two in the PWWP domain at S137 and S144 (*40, 41*). Thus, the S144I mutation has an impact on protein stability but could also affect an important phosphorylation site involved in the regulation of interactions involving the N-terminal region of MSH6.

ACKNOWLEDGMENT

We thank C. J. Craven for the provision of Felix Macros for spectra analysis, Adrien Favier for help in the setup of NOESY experiments at the IBS in Grenoble. We gratefully acknowledge J. Jiricny for kindly providing hMSH6 cDNA, for fruitful discussions, and for the critical reading of the manuscript. We would also like to thank F. Strauss and C. Gaillard for insightful discussions on DNA binding experiments.

SUPPORTING INFORMATION AVAILABLE

The ^{15}N-HSQC spectra of HsMSH6 [68–201] and HsMSH6 [68–201 Cys-Ser]. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Shell, S. S., Putnam, C. D., and Kolodner, R. D. (2007) The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA. *Mol. Cell 26*, 565–578.
2. Kariola, R., Raevaara, T. E., Lonnqvist, K. E., and Nystrom-Lahti, M. (2002) Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. *Hum. Mol. Genet. 11*, 1303–1310.
3. Stojic, L., Brun, R., and Jiricny, J. (2004) Mismatch repair and DNA damage signalling. *DNA Repair (Amsterdam) 3*, 1091–1101.
4. Jiricny, J. (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol. 7*, 335–346.
5. Kunkel, T. A., and Erie, D. A. (2005) DNA mismatch repair. *Annu. Rev. Biochem. 74*, 681–710.
6. Modrich, P. (2006) Mechanisms in eukaryotic mismatch repair. *J. Biol. Chem. 281*, 30305–30309.
7. Genschel, J., and Modrich, P. (2003) Mechanism of 5′-directed excision in human mismatch repair. *Mol. Cell 12*, 1077–1086.
8. Dzantiev, L., Constantin, N., Genschel, J., Iyer, R. R., Burgers, P. M., and Modrich, P. (2004) A defined human system that supports bidirectional mismatch-provoked excision. *Mol. Cell 15*, 31–41.
9. Zhang, Y., Yuan, F., Presnell, S. R., Tian, K., Gao, Y., Tomkinson, A. E., Gu, L., and Li, G. M. (2005) Reconstitution of 5′-directed human mismatch repair in a purified system. *Cell 122*, 693–705.
10. Constantin, N., Dzantiev, L., Kadyrov, F. A., and Modrich, P. (2005) Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J. Biol. Chem. 280*, 39752–39761.
11. Clark, A. B., Deterding, L., Tomer, K. B., and Kunkel, T. A. (2007) Multiple functions for the N-terminal region of Msh6. *Nucleic Acids Res. 35*, 4114–4123.
12. Ge, Y. Z., Pu, M. T., Gowher, H., Wu, H. P., Ding, J. P., Jeltsch, A., and Xu, G. L. (2004) Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J. Biol. Chem. 279*, 25447–25454.
13. Chen, T., Tsujimoto, N., and Li, E. (2004) The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell. Biol. 24*, 9048–9058.
14. Kariola, R., Hampel, H., Frankel, W. L., Raevaara, T. E., de la Chapelle, A., and Nystrom-Lahti, M. (2004) MSH6 missense mutations are often associated with no or low cancer suceptibility. *Br. J. Cancer 91*, 1287–1292.
15. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.
16. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93–158.
17. Savarin, P., Zinn-Justin, S., and Gilquin, B. (2001) Variability in automated assignment of NOESY spectra and three-dimensional structure determination: a test case on three small disulfide-bonded proteins. *J. Biomol. NMR 19*, 49–62.
18. 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) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D 54*, 905–921.
19. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR 13*, 289–302.
20. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477–486.
21. O'Donohue, M. F., Duband-Goulet, I., Hamiche, A., and Prunell, A. (1994) Octamer displacement and redistribution in transcription of single nucleosomes. *Nucleic Acids Res. 22*, 937–945.
22. Sambrook, J. F., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.
23. Lepre, C. A., and Moore, J. M. (1998) Microdrop screening: a rapid method to optimize solvent conditions for NMR spectroscopy of proteins. *J, Biomol, NMR 12*, 493–499.
24. Holm, L., and Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem. Sci. 20*, 478–480.
25. Lukasik, S. M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., and Bushweller, J. H. (2006) High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. *Protein Sci. 15*, 314–323.
26. Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat. Struct. Biol. 9*, 217–224.
27. Stec, I., Nagl, S. B., van Ommen, G. J., and den Dunnen, J. T. (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett. 473*, 1–5.
28. Singh, D. P., Kubo, E., Takamura, Y., Shinohara, T., Kumar, A., Chylack, L. T., Jr., and Fatma, N. (2006) DNA binding domains and nuclear localization signal of LEDGF: contribution of two helix-turn-helix (HTH)-like domains and a stretch of 58 amino acids of the N-terminal to the trans-activation potential of LEDGF. *J. Mol. Biol. 355*, 379–394.

29. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A. 95*, 5857–5864.
30. Nameki, N., Tochio, N., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Fujikura, Y., Saito, M., Ikari, M., Watanabe, M., Terada, T., Shirouzu, M., Yoshida, M., Hirota, H., Tanaka, A., Hayashizaki, Y., Guntert, P., Kigawa, T., and Yokoyama, S. (2005) Solution structure of the PWWP domain of the hepatoma-derived growth factor family. *Protein Sci. 14*, 756–764.
31. Sue, S. C., Chen, J. Y., Lee, S. C., Wu, W. G., and Huang, T. H. (2004) Solution structure and heparin interaction of human hepatoma-derived growth factor. *J. Mol. Biol. 343*, 1365–1377.
32. Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C. P. (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem. Sci. 28*, 69–74.
33. Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M. T. (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep. 7*, 397–403.
34. Cote, J., and Richard, S. (2005) Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem. 280*, 28476–28483.
35. Sprangers, R., Groves, M. R., Sinning, I., and Sattler, M. (2003) High-resolution X-ray and NMR structures of the SMN Tudor domain: conformational variation in the binding site for symmetrically dimethylated arginine residues. *J. Mol. Biol. 327*, 507–520.
36. 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) Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature 432*, 406–411.
37. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics 14*, 892–893.
38. Warren, J. J., Pohlhaus, T. J., Changela, A., Iyer, R. R., Modrich, P. L., and Beese, L. S. (2007) Structure of the human MutSalpha DNA lesion recognition complex. *Mol. Cell 26*, 579–592.
39. Suyama, M., and Ohara, O. (2003) DomCut: prediction of inter-domain linker regions in amino acid sequences. *Bioinformatics 19*, 673–674.
40. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell 127*, 635–648.
41. Nousiainen, M., Sillje, H. H., Sauer, G., Nigg, E. A., and Korner, R. (2006) Phosphoproteome analysis of the human mitotic spindle. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5391–5396.
42. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.

BI7024639