Published in final edited form as:

DNA Repair (Amst). 2011 May 5; 10(5): 454–465. doi:10.1016/j.dnarep.2011.01.010.

MODELING OF THE DNA-BINDING SITE OF YEAST PMS1 BY MASS SPECTROMETRY

Allison N. Schorzman¹, Lalith Perera¹, Jenny M. Cutalo-Patterson^{1,#}, Lars C. Pedersen¹, Lee G. Pedersen¹, Thomas A. Kunkel^{1,2}, and Kenneth B. Tomer^{1,*}

¹ Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709

² Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709

Abstract

Mismatch repair (MMR) corrects replication errors that would otherwise lead to mutations and, potentially, various forms of cancer. Among several proteins required for eukaryotic MMR, MutL α is a heterodimer comprised of Mlh1 and Pms1. The two proteins dimerize along their C-terminal domains (CTDs), and the CTD of Pms1 houses a latent endonuclease that is required for MMR. The highly-conserved N-terminal domains (NTDs) independently bind DNA and possess ATPase active sites. Here we use two protein footprinting techniques, limited proteolysis and oxidative surface mapping, coupled with mass spectrometry to identify amino acids involved along the DNA-binding surface of the Pms1-NTD. Limited proteolysis experiments elucidated several basic residues that were protected in the presence of DNA, while oxidative surface mapping revealed one residue that is uniquely protected from oxidation. Furthermore, additional amino acids distributed throughout the Pms1-NTD were protected from oxidation either in the presence of a non-hydrolyzable analog of ATP or DNA, indicating that each ligand stabilizes the protein in a similar conformation. Based on the recently published X-ray crystal structure of yeast Pms1-NTD, a model of the Pms1-NTD/DNA complex was generated using the mass spectrometric data as constraints. The proposed model defines the DNA-binding interface along a positively-charged groove of the Pms1-NTD and complements prior mutagenesis studies of *E. coli* and eukaryotic MutL.

Keywords

Mismatch repair; MutL α ; Pms1; DNA binding; mass spectrometry

1. Introduction

Mismatch repair (MMR) is a complex process that corrects base-base, insertion and deletion mismatches that result from DNA replication and recombination [1–3]. Defective MMR

Address correspondence to: Allison N. Schorzman, Ph.D., NIEHS, PO Box 12233 MD F0-04, Research Triangle Park, NC 27709; schorzmana@niehs.nih.gov or: Kenneth B. Tomer, Ph.D., NIEHS, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709. Fax: 919-541-0220; tomer@niehs.nih.gov.

[#]Currently at the Federal Bureau of Investigation, Washington, D.C.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

proteins compromise genomic stability and increase susceptibility to various forms of cancer in mice and humans [4,5]. Having been reproduced *in vitro* [6], MMR is well-understood in *E. coli* [1,2,7] and requires the cooperative action of the proteins MutS, MutL and MutH. MutS recognizes and binds to the mismatch while the methylation-specific endonuclease MutH nicks the newly replicated, unmethylated strand at the hemi-methylated d(GATC) site of the DNA duplex. MutL binds to MutS and, in the presence of ATP, coordinates excision of the error from the nascent strand to allow correct resynthesis of DNA [1,2]. The eukaryotic MMR process appears to be different and is not as well understood [1,3,7]. For example, eukaryotes do not possess a MutH homolog and the signal used for strand discrimination has yet to be identified. Moreover, eukaryotes encode multiple homologs of MutS and MutL, which form heterodimers to accomplish MMR.

The yeast homolog, MutL α , is comprised of Mlh1 and Pms1, each containing an N-terminal domain (NTD) and C-terminal domain (CTD) separated by a flexible linker. Dimerization occurs between the CTDs [8–13] and the CTD of Pms1 houses a strand-specific endonuclease that is necessary for MMR [14]. The MutL homologs have highly conserved NTDs that harbor ATPase active sites common to the GHL family of ATPases [15]. The NTDs of Mlh1 and Pms1 independently bind and hydrolyze ATP, and they independently bind to DNA [16]. Atomic force microscopy images have illustrated significant asymmetrical structural changes that MutL α undergoes when the nucleotide is bound [17]. It has been hypothesized that these structural changes may expose different interaction sites of MutL α with potential ligands to help coordinate the MMR machinery with the site of DNA mismatch and that MutL α itself acts as an ATP-activated clamp that can tightly bind DNA during MMR. While binding of the Mlh1-Pms1 heterodimer to DNA is essential for the endonuclease activity of Pms1, identification of the DNA-binding surface has yet to be elucidated.

The X-ray crystal structures have been solved for the NTDs of yeast Pms1 (Pms1-NTD) [18], human PMS2 [19] and *E. coli* MutL (LN40) [15]. *E. coli* LN40 dimerizes in the presence of ATP, which generates a positively-charged groove that is capable of accommodating single-stranded DNA [15,20]. Furthermore, an R266E mutation, residing in the positively-charged groove of the LN40 dimer, reduces DNA binding. The Pms1-NTD, as well as human PMS2, does not self-associate in the presence of ATP [16,19]. However, the Pms1-NTD possesses a positively-charged groove on the opposite face of the protein from the ATP-binding motif that may accommodate double-stranded DNA. Mutagenesis of the Pms1-NTD identified two residues (K197 and R198) within that groove that, upon charge reversal, yield mutator phenotypes in yeast and reduce DNA binding *in vitro*, suggesting that they may contribute to a potential DNA-binding surface [18]. Other changes on the surface of Pms1, *e.g.* mutants K218E, R243E/K244E and K328E also yield mutator effects but do not affect DNA binding [16,18]. Binding of the Pms1-NTD to DNA is highly cooperative and sensitive to salt concentration, implicating the involvement of electrostatic interactions [16,21]. Additionally, the presence of DNA does not affect the ATPase activity of PMS2 nor does the presence of ATP affect DNA binding [19].

The large size of the Pms1-NTD/DNA complex and the likelihood that Pms1 moves along DNA strands [2] precludes structural analysis of the complex by NMR or X-ray crystallography. An alternative approach to generating a model employs protein footprinting techniques coupled with mass spectrometry and molecular modeling. Protein footprinting refers to the characterization of protein-ligand interactions and subsequent conformational changes through determination of the accessibility of the peptide backbone and amino acid side chains to enzymatic or chemical modification [22]. Experimental data from protein footprinting techniques coupled with mass spectrometry indicates the DNA-binding surface and can be applied as constraints to generate a computational model for the protein/DNA

interaction. Advances in molecular modeling have facilitated the predictions of protein-ligand complexes for those proteins in which a crystal structure exists [23].

One such protein footprinting technique is limited proteolysis, which uses enzymes to cleave a limited number of peptide bonds in a protein over a monitored time range under conditions that favor incomplete proteolysis (*e.g.*, short reaction time, high substrate-to-enzyme ratio). When bound to a ligand, the protein's conformation becomes stabilized and the ligand may block some cleavage sites, both creating steric barriers to enzymatic attack. The pattern of preferred cleavage sites observed in these experiments show the regions in the protein that are readily accessible to the protease. Mass spectrometry provides single amino acid resolution in the determination of cleavage sites and has been used in the characterization of the tertiary structures of proteins and complexes, including protein/DNA complexes [24–26].

An alternative protein footprinting approach is oxidative surface mapping, which utilizes the hydroxyl radical as a probe [22,27–29]. The hydroxyl radical oxidizes reactive and accessible side chains, resulting in a stable modification with typical mass increases of 16 or 32 Da. The use of hydroxyl radicals as probes for surface mapping provides several advantages in that they are small, with Van der Waals radii and solvent properties similar to that of water [22], allowing them to access amino acid side chains; they are easily and safely generated in high yield with commonly available equipment; and most importantly, they are highly reactive with little specificity so that they provide better sequence coverage than do site-specific modifications (*i.e.* lysine acetylation, Arg-C proteolysis). The hydroxyl radical can be generated using several techniques including Fenton chemistry [30–40], photolysis of hydrogen peroxide [41–48], and radiolysis of water using either X-rays [49–54] or γ -rays [55–61]. Although the hydroxyl radical is generally non-specific, it is selective in the bonds with which it reacts [29]. Reaction occurs most readily for sulfur-containing, aromatic and aliphatic side chains, and typical oxidation products include methionine sulfoxide, N-formylkynurenine, and various hydroxides.

In this study, limited proteolysis and oxidative surface mapping were performed in the presence and absence of DNA and/or ATP analog in order to identify residues by mass spectrometry that are involved in the DNA binding domain of the Pms1-NTD. Both wild-type and mutant proteins were evaluated and the experimental results were applied as constraints to generate a computational model for the Pms1-NTD/DNA complex from the solution structure, which is based on the crystal structure of the Pms1-NTD [18]. The proposed model is consistent with previously reported mutagenesis data and with the electrostatic properties of the Pms1-NTD.

2. Materials and Methods

2.1. Proteins

Expression and purification of the GST-tagged wild-type N-terminal domain of yeast Pms1 (AA 32–396) and the K197E/R198E/K229E mutant have been described by Arana *et al.* [18]. The expression and purification of the His-tagged R311E mutant has been described by Hall *et al.* [62]. The sequence of the wild-type protein, annotated with proteolytic cleavage sites and relevant structural features, is shown in Figure 1.

2.2. dsDNA preparation

The preparation of ~7,250 bp M13mp18 circular DNA has been described [63]. DNA purification was accomplished using the Qiagen Plasmid Mega kit (Qiagen, Valencia, CA) according to the protocol that was provided.

2.3. Limited Proteolysis

The Pms1-NTD wild-type and mutant proteins (4 ng/ μ L) were digested in the absence or presence of either excess double-stranded DNA in a ratio of 725:1 DNA nucleotides:protein or excess single-stranded Poly(dA) DNA (Sigma, St. Louis, MO) in a ratio of 174:1 DNA nucleotides:protein. The proteins were digested for two hours with 0.2 ng/ μ L of Arg-C (Promega, Madison, WI) or Lys-C (Promega) at 37°C in buffer containing 25 mM Hepes, pH 8.0, 10% glycerol, 25 mM NaCl, 4 mM MgCl₂, and 1 mM DTT. Limited Lys-C proteolysis was extended to three hours for the mutant proteins in order to achieve detectable peptides for analysis by mass spectrometry. An aliquot was transferred to a fresh tube every 15 minutes (with the exception of the 120-minute time point), acidified with 0.1% formic acid and desalted with a C18 ZipTip (Millipore, Billerica, MA). The aliquots were analyzed by SDS-PAGE (15 μ L) and mass spectrometry (5 μ L).

2.4. γ -Irradiation

The Pms1-NTD in PBS was diluted to 250 ng/ μ L in water, and catalase (Sigma) was added to a final concentration of 10 nM to quench the H₂O₂ formed upon radiolysis of water. For the 'ATP-bound' Pms1-NTD, the non-hydrolyzable analog of ATP, AMP-PNP (Sigma), was added to a final concentration of 5 mM, whereas the K_m for ATP binding is 1.5 mM [62]. For the DNA-bound Pms1-NTD, dsDNA was added at a concentration of 2.7 μ g/mL correlating to a ratio of 725:1 nucleotide:protein. Samples were placed on a rotating platform and exposed to γ irradiation for 30 minutes at a radiation dosage of 856 rad/min using both sources of the dual-source ¹³⁷Cs irradiator. Irradiated samples were immediately heat denatured and digested overnight using trypsin (Promega) at a ratio of 1:20 at 37°C in 25 mM Tris, pH 7.6, and 0.5 mM CaCl₂. Sample complexity was reduced prior to MS analysis by digesting the dsDNA with 400 gel units of the non-specific endo-exonuclease, micrococcal nuclease (New England Biolabs, Ipswich, MA) per 1 μ g DNA for 2 hours at 37°C with mixing at 500 RPM in 1X Buffer and 20 ng/ μ L BSA. Peptides were purified using a C18 ZipTip prior to analysis by mass spectrometry.

2.5. Mass Spectrometry

All MALDI mass spectra were obtained on an Applied Biosystems Voyager-SUPER DE STR (Framingham, MA). MALDI/MS analyses of limited proteolysis experiments were performed in the reflector positive mode using an accelerating voltage of 20 kV, a grid voltage of 61.5%, and a delay time of 185 ns. All analyses were performed using a saturated solution of α -cyano-4-hydroxycinnamic acid in 50:49.9:0.1 acetonitrile:water:formic acid (vol:vol:vol) as the matrix in a 1:1 ratio with the sample.

LC-ESI quadrupole ion trap data were collected using an Agilent 1100 nano-LC coupled to an Agilent HPLC-Chip/Trap XCT Ultra mass spectrometer system (Agilent, Santa Clara, CA). Peptides were loaded onto an Agilent HPLC-Chip comprised of a 40 nL enrichment column and 75 μ m $\times$ 43 mm separation column packed with Zorbax 300SB-C18 5 μ m material. After a 10 minute wash with 3% ACN/0.1% formic acid, peptides were separated using a 40 minute linear gradient ramped to 50% ACN/0.1% formic acid. Ions were generated using electrospray in the positive-ion mode with a capillary voltage of 2.2 kV, maximum accumulation time of 200 ms, and an m/z range of 300–2000 in full-scan MS and CID MS/MS modes. Data-dependent MS/MS acquisitions were performed for the 6 most abundant ions from the MS mode using threshold abundance of 10,000 and a fragmentation amplitude of 1.0 V. Spectrum Mill MS Proteomics software (Agilent) was used to identify oxidized peptides and sites of oxidation.

Flow-injection ESI mass spectra were collected using a Q-ToF Ultima™ Global mass spectrometer (Waters, Milford, MA). The electrospray solution consisted of 50%

acetonitrile/0.1% formic acid. Ions were generated by positive-ion electrospray ionization using the nanoflow z-spray source with a capillary voltage of 3.0 kV and a cone voltage of 65 V. The full-scan spectra were acquired for 5 minutes in continuum mode using an m/z range of 200–2000.

2.6. Molecular Modeling

The modeling described herein is based on the X-ray crystal structure of the Pms1-NTD (pdb entry: 3H4L, molecule A) in which residues 110–118 and 275–286 were disordered [18]. In addition, side chains of some residues were missing in the crystal structure. The missing segments of residues were added using the loop search algorithm of Sybyl 8.1 (Tripos, Inc, St. Louis, MO). The missing side chains were introduced using the xleap module of Amber.10. The resulting structure was solvated in a bath of water with water molecules extending up to 15 Å away from any protein atom normal to the surface. Prior to equilibration, the system was subjected to energy minimizations at various levels followed by low temperature, constant pressure molecular dynamics to assure a reasonable starting density. Step-wise heating at constant volume to bring the temperature to 300K, followed by a 2 ns constant volume molecular dynamics simulation, completed the equilibration. Final unconstrained trajectories (15 ns) were calculated at 300 K under constant pressure of 1 atm (1 fs time step) using PMEMD (Amber.10) to accommodate long range interactions. The parameters were taken from the FF99SB force field.

A 22-bp DNA segment obtained from the 146-bp DNA palindrome (pdb entry: 1EQZ) was manually docked to the final structure of the simulated Pms1-NTD using Sybyl 8.1 (Tripos, Inc). The docking was guided by the assumption that residues Lys197 and Arg198 have a direct role in the DNA interaction. Also, the positions of residues Arg188, Lys190, Tyr323 and Lys364 were required to be in close contact with DNA in the selection of docked structure. The DNA-bound Pms1-NTD was solvated in a bath of water and following a procedure similar to the one described above, an equilibrated structure of the complex was obtained.

3. Results

3.1. Limited Proteolysis using Arg-C and Lys-C

Limited proteolysis was performed using Arg-C and Lys-C in order to identify basic residues that are potentially involved in the DNA-binding surface of the Pms1-NTD. The extent of proteolysis as a function of time was compared for the Pms1-NTD in the absence and presence of dsDNA and ssDNA. SDS-PAGE analysis of the Arg-C and Lys-C time courses show that the Pms1-NTD is digested rapidly in the absence of DNA with no protein being detected after 15 minutes (Figure 2). However, the Pms1-NTD bound to DNA resists proteolysis and is still detectable up to 120 minutes. Control experiments performed with a non-DNA-binding protein indicate that complete digestion is achieved in the presence of dsDNA within 120 minutes at a similar rate to that of the unbound protein, even at early time points.

The peptides for each time point were analyzed by MALDI-ToF mass spectrometry. Figure 3A shows MALDI-ToF mass spectra for limited Arg-C proteolysis of the Pms1-NTD in the absence and presence of dsDNA and ssDNA. The 45-minute time point was chosen because proteolysis was complete for the protein in the absence of DNA while progressing at a slower rate for the bound protein (Figure 2). Consequently, ion counts for spectra of the Pms1-NTD bound to DNA, particularly for dsDNA, are consistently less than those for the unbound protein. When the Pms1-NTD is bound to dsDNA, several peptides are no longer observed (*i.e.*, peptides R5, R9, R10) while others are present at reduced relative abundances

(i.e., R1, R16, R4). In the presence of ssDNA, peptides R4, R5, R9, R10 and R16 are all observed at reduced relative abundances. Furthermore, the abundance of peptide R15 increases in the presence of DNA.

Figure 3B shows MALDI-ToF mass spectra for the limited Lys-C proteolysis of the Pms1-NTD in the absence of DNA, in the presence of dsDNA and the presence of ssDNA. Limited Lys-C proteolysis appeared to be slower than Arg-C proteolysis, thus the 60-minute time point was chosen for comparison. Again, the total ion counts illustrate the reduced yield of peptides for the protein in the presence of DNA. When the Pms1-NTD is bound to dsDNA, peptide K30 is no longer observed while K7, K12 and K20 appear at reduced relative abundances. In the presence of ssDNA, peptides K7, K12, K20 and K30 are all observed at reduced relative abundances.

In addition to the wild-type protein, limited proteolysis was performed with the Pms1-NTD mutants R311E and K197E/R198E/K229E in the absence and presence of DNA. These mutants were created to interrogate arginines and lysines that are suspected to be involved in DNA binding because they are conserved among species and are observed on the surface of the structure without appearing to play a structural role. In limited Arg-C proteolysis experiments, results for the R311E mutant bound to dsDNA were consistent with those seen for the wild-type Pms1-NTD in that proteolysis occurred slower for R4 and was not observed for R5 (Figure 4A). In addition, upon limited Lys-C proteolysis in the presence of dsDNA, cleavage was slower for K12 while peptides K20, K21 and K30 were not observed (Figure 4B). In the presence of DNA, the K197E/R198E/K229E mutant was more susceptible to Arg-C proteolysis (Figure 5A) than the DNA-bound wild-type Pms1-NTD (Figure 3A, middle panel), which may be due to a reduced affinity for DNA. Only a modest reduction in the formation of peptide R4 was observed and, as this mutant lacks an Arg-C cleavage site at R198, there is no information for peptide R5. Consistent with limited Lys-C proteolysis of the wild-type Pms1-NTD bound to DNA, proteolysis for K12 and K20 was slower in the presence of ssDNA while these peptides were not observed in the presence of dsDNA. Neither peptides K21 nor K30 were observed in the presence of DNA (Figure 5B).

To quantify the extent of proteolysis for the wild-type protein in the absence and presence of DNA, the relative percent abundance of each peptide in the mass spectrum was calculated using the most abundant isotope in the distribution. The area of each peptide peak was divided by the sum of the areas of all peptide peaks present in the spectrum for each time point. Data were measured from three biological replicates and the results subjected to the *t*-test in order to evaluate the significance of their variances. The relative percent abundances of Pms1-NTD peptides in the absence and presence of dsDNA are summarized in Table 1 for the 45-minute time point for Arg-C and the 60-minute time point for Lys-C. Each peptide was classified as protected, not protected or more exposed depending on whether proteolysis was statistically reduced, unchanged or increased, respectively, as determined from the *t*-test. The peptide abundances in the presence of ssDNA have been summarized in Table 2. Notable differences in cleavage between dsDNA and ssDNA were that R1 and K30 were not protected, R13 became more exposed, and R8 and K31 were no longer more exposed. These results indicate some differences in the binding of Pms1-NTD with dsDNA as compared to ssDNA.

The cleavage sites required to form each peptide are listed in Table 1. The cleavage sites associated with the non-protected peptides were eliminated as residues involved in the DNA-binding site, while those that were more exposed were attributed to conformational changes upon DNA binding. All residues protected upon binding to dsDNA were mapped to the Pms1-NTD model and are shown in Figure 6. Of the protected peptides, several cleavage sites could also be eliminated as potential sites of DNA-binding. Arg45 of peptide R1 is

positioned on the N-terminal portion of the truncated Pms1-NTD and, therefore, does not represent its native structure. It likely moves freely in the unbound form readily allowing proteolysis while the DNA causes the structure to adopt a more rigid conformation. Peptide K30 is formed from Lys364 and Lys380. Seeing that K31, formed by cleavage at Lys380, is observed more exposed in the presence of dsDNA, then Lys364 is the protected residue. Furthermore, cleavage sites were eliminated if they were in close proximity to the ATP binding site (Arg170, Lys172), appeared to play a structural role by residing within a set of parallel β sheets (Lys230) or if they were not conserved among homologs (Lys244, Lys271). Arg188, Lys190, Arg198 and Lys364, labeled in bold and italics in Table 1, potentially play a potential role in the DNA binding site of the Pms1-NTD and were utilized in construction of the Pms1-NTD:DNA model.

3.2. Oxidative Surface Mapping

An oxidation time course was performed by γ -irradiation on the Pms1-NTD in the absence of ligand with aliquots of oxidized protein obtained every 15 minutes over a 2 hour period. Upon analysis of the tryptic peptides by flow-injection Q-Tof-MS, it was observed that the rate of oxidation did not change over the first 60 minutes of the experiment. Because the rate of oxidation does not increase, the protein was not unfolding during this time. In all subsequent experiments, proteins were irradiated for 30 minutes to maximize oxidation while minimizing exposure of the protein solution to γ -rays. Tryptic peptides generated from the γ -irradiated Pms1-NTD were initially analyzed by LC/ESI-QIT-MS/MS to identify oxidized peptides and sites of oxidation. While the data were searched for all primary oxidation products as outlined by Xu and Chance [29], only those supported by MS/MS data were utilized considered for these analyses. Table 3 lists the sixteen peptides that were observed oxidized (*i.e.*, increase in mass of +16 or +32 Da) and represents 45% of the Pms1-NTD protein sequence. MS/MS of the singly-oxidized peptide often resulted in a heterogeneous mixture of fragment ions due to the fact that the site of oxidation varied among the accessible, reactive side chains of the peptide; the oxidized amino acids have been highlighted in red in Table 3. As oxidized peptides other than those containing methionine and tryptophan were present in low abundance (<20% oxidized), there may be oxidized side chains that were below our limits of detection.

The oxidized peptides were quantified by ESI-Q-Tof-MS using three experimental replicates in order to obtain statistically relevant results. The percentage of oxidized peptide was calculated by dividing the abundance of the oxidized peptide ions by the sum of the abundances of the unoxidized and oxidized peptide ions, using the most abundant isotope of the envelope. The column labeled *Pms1* in Table 3 represents the percentage and standard deviations of each oxidized peptide in the absence of DNA. Of the peptides listed, methionine residues are the most reactive side chains with percentages of oxidation averaging 76%. The benzene ring and pyrrole moiety of tryptophan are both reactive with the hydroxyl radical yielding 63.5% oxidation of peptide T22. The side chains of phenylalanine, tyrosine and the aliphatic side chains were observed to be less reactive, typically resulting in less than 20% of oxidized peptide.

Because the Pms1-NTD binds ATP during MMR and the crystal structure was solved in the presence of nucleotide [18], the residues that are protected from oxidation in the presence of the non-hydrolyzable form of ATP, AMP-PNP, were interrogated. All peptides that have exhibited a statistically significant decrease in oxidation in the presence of AMP-PNP have been highlighted in gray (column +AMP-PNP in Table 3). The peptides protected from oxidation by AMP-PNP include T2, T15, T22, T28, T29, T38, and T46. It should be noted that Leu124 (peptide T8) is the only amino acid within the ATP-binding motif (Figure 6) that was observed oxidized in the absence of nucleotide. When AMP-PNP is bound, the oxidized peptide T8 was no longer observed.

The oxidative surface mapping experiment was repeated in the presence of dsDNA and the percentage of oxidation for each peptide is shown in the +DNA column of Table 3. All peptides that have exhibited a statistically significant decrease in oxidation in the presence of DNA have been highlighted in gray. It should be noted that there was a global reduction in oxidation when DNA was present because the deoxyribonucleotides, present in solution at four times the abundance of amino acids side chains, are also substrates for oxidation and, therefore, quench hydroxyl radicals that may otherwise oxidize the protein. As a result, all peptide abundances were normalized to peptide T1, which did not exhibit protection. Statistics showed that all amino acid side chains protected from oxidation by AMP-PNP binding, except for T38, were the same as those identified for protection from DNA. In addition, T39 was completely protected from oxidation in the presence of DNA, although it exhibited no significant protection from oxidation by AMP-PNP.

In order to interrogate how each ligand structurally affects Pms1 in the presence of the other, the Pms1-NTD was irradiated in the presence of DNA and AMP-PNP using two sets of conditions: 1) AMP-PNP was introduced to the protein followed by DNA after a 10 minute incubation; and 2) DNA was introduced to the protein followed by AMP-PNP after a 10 minute incubation. The peptides that exhibited statistically significant protection from oxidation are the same as those shared by DNA and AMP-PNP as lone ligands. These results are shown in Supplemental Table S-1.

The protected side chains are distributed throughout the protein and have been mapped to the Pms1-NTD model (Figure 6). With the exception of peptides T38 and T39, the same distribution of protection is observed whether the ligand is AMP-PNP or dsDNA, which is attributed to two reasons. First, DNA and ATP have separate binding domains on the Pms1-NTD as each can bind without affecting the other. In the absence of DNA, we expect that ATP would also interact with the basic DNA binding site through electrostatics. For this reason, it was expected that similar protections from oxidation may be observed for both ligands. Secondly, oxidation of the protein by radiolysis of water for 30 minutes may result in some local unfolding of the protein at oxidized sites. These data suggest that binding of AMP-PNP and/or DNA stabilizes the protein so that it resists unfolding through the duration of γ -irradiation. Peptide T39 was the only peptide that was uniquely protected from oxidation in the presence of DNA, thus Tyr323 is considered to be involved in the DNA-binding site of the Pms1-NTD and was used as a constraint for molecular modeling.

3.3. Modeling of the Pms1-NTD

Prior to creating a model for the Pms1-NTD/DNA complex, a reasonable complete starting structure for the Pms1-NTD was generated based on the X-ray crystal structure. Because missing loops and side chains in the X-ray crystal structure were modeled, equilibration of the solution structure during the course of simulation time (over 15 ns) is essential. The root mean square deviations (RMSDs) calculated for the backbone atoms during the last 6 ns of the simulations with respect to the initial optimized structure (Supplemental Figure S-1) support the conclusion that a relatively stable solution structure of the Pms1-NTD is achieved. Large fluctuations in atomic positions for the loop regions, specifically the two modeled loop structures (segments bracketing residues 110-118 and 275-286), increase the magnitude of the total RMSD (Supplemental Figure S-2). The final structure of the simulation represents a reasonable starting structure for DNA docking.

3.4. Construction of a model for the Pms1-NTD/DNA complex

The limited proteolysis results in this study indicate that Arg188, Lys190, Arg198, and Lys364 are part of the DNA-binding surface while oxidative surface mapping implicates Tyr323. A model was created for the Pms1-NTD/DNA complex that satisfies these mass

spectrometric findings. Only the conformations that allowed DNA to interact with the amino acids that resulted in complete protection from proteolysis (Arg198 and Lys364) and oxidation (Tyr323) in the presence of DNA were considered. Furthermore, direct interaction of DNA with Lys197 [18] was required as well as preference for interaction with Arg188 and Lys190. These residues, labeled yellow and represented in the CPK form (Figure 7A) along with the electrostatic surface potential (Figure 7B) indicate a well-suited surface for DNA binding. After several rounds of optimizations, the solution structure of the complex was evaluated using a constraint-free molecular dynamics simulation. The final complex solution structure is displayed in Figure 7C in which the residues implicated in DNA binding are represented in yellow while the DNA is depicted in tan. This model defines the DNA binding site as the basic groove of the Pms1-NTD on the opposite face of the ATP-binding site. This orientation of the two ligands would make it possible for each to bind to the N-terminal domain simultaneously. Furthermore, the proposed DNA-binding site and the ATP-binding site share several α helices so that the binding of one ligand could influence the binding site of the other.

4. Discussion

The protein footprinting data achieved by mass spectrometry and the subsequent model of the Pms1-NTD/DNA complex correlates well with what is already known about their interaction. The limited proteolysis data has shown that the Pms1-NTD bound to DNA is slower than in the unbound form (Figure 3). While this may be partially attributed to non-specific interactions between the enzyme and DNA precluding proteolysis, it suggests that the Pms1-NTD is in a conformation more stable to proteolysis when in the DNA-bound form. Additionally, proteolysis is slower for protein bound to dsDNA than protein bound to ssDNA indicating that the Pms1-NTD has a higher affinity for dsDNA than ssDNA, which is consistent with previous studies [64]. Several residues exhibited increased proteolysis. While these residues are not considered potential DNA binding sites, they may become more exposed as a result of the DNA-bound conformation to allow potential interactions with other MMR proteins.

Residues Arg188 and Lys190 both exhibited protection from proteolysis in the presence of DNA. They are conserved in LN40, and in PMS2 from human and mouse, implying biological relevance. Along with Lys197 and Arg198, they reside on a bent α helix (Figure 6), which is reminiscent of α helices involved in the structures of other DNA binding proteins with the ability to fit well within the major groove of DNA [65]. It should be noted that mutation rates measured for the K190E mutant did not exhibit a mutator phenotype [18]; therefore, Lys190, which is clearly protected from proteolysis in the data presented here, is not required for DNA binding. For these reasons, Arg188 and Lys190 were considered important in generating a model for the interaction of the Pms1-NTD with DNA.

Two peptides exhibited complete protection when the Pms1-NTD was bound to dsDNA: R5 and K30. The R5 peptide is generated by cleavage at Arg188 and Arg198. Because cleavage occurs at Arg188 in the presence of DNA, albeit significantly reduced as observed for R4 (Figure 3A and Table 1), Arg198 must be directly involved in DNA binding or is highly protected from proteolysis when DNA is bound. Arg198 is located on an α helix in the X-ray structure (Figure 1) and is conserved in the PMS2 human and mouse homologs indicating its significance to protein function. The K197E/R198E/K229E mutant exhibited increased susceptibility to proteases in the presence of DNA suggesting a reduced affinity as a result of one or more of these mutations. Arana *et al.* [18] have shown that Lys197 and Arg198 are both important for DNA binding in experiments where K197E and R198E mutants abolished DNA binding *in vitro* while the K229E mutant does not display a strong mutator phenotype. Strong mutator phenotypes were observed for Lys218, Arg243 and

Lys244 mutants without corresponding reductions in DNA affinity. According to the model presented here, Lys218, Lys229, Arg243, Lys244 and Arg311 are all distant from the DNA-binding site and would not be expected to affect DNA binding. Arana *et al.* also identified a surface consisting of positively-charged residues spanning from Lys197/Arg198 to Lys244 that may potentially interact with DNA. Our model agrees with the basic region containing Lys197 and Arg198, but instead places the DNA-binding region along an adjacent positively charged groove, implicating the aforementioned groove as a potential binding surface for other MMR proteins.

The K30 peptide is generated by cleavage at Lys364 and Lys380. Cleavage occurs at Lys380 because K31 is readily formed, which implies that Lys364 is bound to the DNA or shielded from cleavage due to a subsequent conformational change. Lys364 is conserved in human and mouse Pms2 as well as human PMS2 and *E. coli* LN40, and is located in a loop (Figure 1). Because loops are flexible, it may move allowing Lys364 to interact directly with DNA. Lys364 forms a salt bridge with Glu61 and Asp64 in the modeled structure; therefore, it is also likely that DNA binding disrupts the salt bridge and that the presence of DNA prevents access of the protease to the backbone. Lys364 is taken to be involved in DNA binding, whether through direct contact or protection by the DNA itself, and was used to guide DNA docking in generating the model for the DNA-binding site of the Pms1-NTD.

Oxidative surface mapping data demonstrated similar protection from oxidation with either DNA or ATP as a ligand. This is attributed to electrostatic interactions of AMP-PNP with the DNA-binding site and vice versa, as well as a similar stabilization of the Pms1-NTD to local unfolding during γ -irradiation. Perhaps in the absence of DNA, ATP binding generates an allosteric effect that better presents the binding site to DNA during MMR. Peptide T38 was uniquely protected upon Pms1-NTD binding to AMP-PNP. Phe313 is only ~10% accessible while Tyr315 is completely buried; therefore, in the absence of ligand, local unfolding of the structure may allow access by the hydroxyl radical. Furthermore, the base of the ATP binding site involves the α -helix comprised of residues 55-69 (Figure 1) that constricts the motion of the α helix comprised of residues 195-214 when nucleotide is bound, which in turn protects the structure proximal to Phe313 and Tyr315 from local unfolding. Protection of T38 is attributed to stability of the structure upon AMP-PNP binding.

In the presence of dsDNA, only T39 is uniquely protected from oxidation with statistical significance. There is no significant protection in the presence of AMP-PNP, nor is there significant protection in the presence of both ligands, suggesting that the conformation differs somewhat in the presence of both ligands. Protection from oxidation may result from interaction with DNA instead of, or in addition to, stabilization of the structure. Tyr323, the site of oxidation on T39, is mapped on the crystal structure at the base of the positively-charged groove that also encompasses Arg188, Lys190, Arg198 and Lys364. The diameter of this groove in the crystal structure is approximately 20 Å, making it suitable for accommodating dsDNA.

The electrostatic surface potential diagram of the Pms1-NTD model (Figure 7B) shows the positively charged groove where DNA binding occurs according to our model. Likewise, homodimerization of *E. coli* LN40 in the presence of ADP-PNP forms a highly positive-charged surface potential where ssDNA may bind [20] (Figure 8A). An overlay of the Pms1-NTD/DNA model with the model of the LN40/ADP-PNP complex shows the double-stranded DNA passing through the same groove thought to be the binding site for ssDNA in *E. coli* MutL (Figure 8B). An R266E mutation in this putative DNA binding groove of LN40 (Figure 8A) abolishes DNA binding, providing further evidence for a DNA-binding site [20]. When a homologous mutation, K328E, was made in the Pms1-NTD, there was

little reduction in DNA binding whereas its MutL α binding partner, Mlh1, lost its ability to bind DNA when the homologous R274E mutation was introduced [16]. Our model suggests that DNA binding via a positively-charged groove has been conserved between *E. coli* and eukaryotic homologs of MutL although the basic residues of contact may differ.

Understanding the role of Pms1 is important to human health because mutations in the homologous human *PMS2* gene are associated with hereditary non-polyposis colon cancer (HNPCC), which results in a propensity toward colon cancer, but is also associated with endometrial and ovarian tumors [4]. Most of the *PMS2* mutations that are associated with cancer are deletion, frameshift and nonsense mutations that result in truncation of the protein and prevent ATP binding, DNA binding and heterodimer formation with Mlh1. Four nucleotide substitutions have been reported that correlate with colon or endometrial cancer [66]. Two of the substitutions, which generate the protein variants I18V and R20Q, are located 5' to the sequence that defines the ATP-binding motif and probably disrupt nucleotide binding. The third substitution results in a silent mutation in the protein product, yet is associated with cancer indicating that the mutation may affect transcription. The fourth substitution translates into an A182T mutant in *PMS2* and is associated with endometrial cancer. The homologous version of this mutation on the Pms1-NTD would occur on the same α -helix (Figure 1) immediately downstream of Lys197 and Arg198, and, therefore, within the proposed DNA-binding site. This implies that mutations in the DNA binding site of Pms1 or its homologs could affect its ability to perform MMR with serious consequences to the organism.

The model presented here defines the DNA binding interface along a positively-charged groove of the Pms1-NTD. Our data, as well as previous functional studies, indicate that Lys197 and Arg198 are directly involved while Arg188, Lys190, Tyr323 and Lys 364 define the DNA-binding surface without directly participating in DNA binding. In addition, this study has shown how the complementary techniques of limited proteolysis and oxidative surface mapping coupled with mass spectrometry can be used to obtain structural information from a large and dynamic complex. This model of the Pms1-NTD/DNA complex will provide further insight into the structure-function relationship of Pms1 in mismatch repair.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gratefully acknowledge Dr. Leesa Deterding and Dr. Mercedes Arana for thoughtful review of this manuscript. We thank the Protein Microcharacterization Core Facility for use of their instrumentation. This research was supported by Project Z01 ES0050127 (ANS, JMC, KBT), Project Z01 ES065089 (TAK), Project Z01 ES43010-23 (LP, LGP) and Project Z01 ES102645-01 (LCP), all from the Division of Intramural Research of the National Institute for Environmental Health Sciences/National Institutes of Health.

References

1. Kunkel TA, Erie DA. DNA mismatch repair. *Annu Rev Biochem.* 2005; 74:681–710. [PubMed: 15952900]
2. Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA mismatch repair: Functions and mechanisms. *Chem Rev.* 2006; 106:302–323. [PubMed: 16464007]
3. Modrich P. Mechanisms in eukaryotic mismatch repair. *J Biol Chem.* 2006; 281:30305–30309. [PubMed: 16905530]

4. Fishel R, Kolodner RD. Identification of Mismatch Repair Genes and Their Role in the Development of Cancer. *Current Opinion in Genetics & Development*. 1995; 5:382–395. [PubMed: 7549435]
5. Peltomaki P, Vasen HFA, Bisgaard ML, Buerstedde JM, Friedl W, Grandjouan S, Hutter P, KohonenCorish M, Kolodner R, Kurzawski G, Lindblom A, Lynch HT, Piepoli A, deLeon MP, Radice P, Thibodeau S, Weber W, West S, Wijnen J. Mutations predisposing to hereditary nonpolyposis colorectal cancer: Database and results of a collaborative study. *Gastroenterology*. 1997; 113:1146–1158. [PubMed: 9322509]
6. Lahue RS, Au KG, Modrich P. DNA mismatch correction is a defined system. *Science*. 1989; 245:160–164. [PubMed: 2665076]
7. Schofield MJ, Hsieh P. DNA mismatch repair: Molecular mechanisms and biological function. *Annu Rev Microbiol*. 2003; 57:579–608. [PubMed: 14527292]
8. Pang QS, Prolla TA, Liskay RM. Functional domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations. *Mol Cell Biol*. 1997; 17:4465–4473. [PubMed: 9234704]
9. Guerrette S, Acharya S, Fishel R. The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. *J Biol Chem*. 1999; 274:6336–6341. [PubMed: 10037723]
10. Kondo E, Horii A, Fukushima S. The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res*. 2001; 29:1695–1702. [PubMed: 11292842]
11. Guarne A, Ramon-Maiques S, Wolff EM, Ghirlando R, Hu XJ, Miller JH, Yang W. Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J*. 2004; 23:4134–4145. [PubMed: 15470502]
12. Kosinski J, Steindorf I, Bujnicki JM, Giron-Monzon L, Friedhoff P. Analysis of the quaternary structure of the MutL C-terminal domain. *J Mol Biol*. 2005; 351:895–909. [PubMed: 16024043]
13. Cutalo JM, Darden TA, Kunkel TA, Tomer KB. Mapping the dimer interface in the C-terminal domains of the yeast MLH1-PMS1 heterodimer. *Biochemistry*. 2006; 45:15458–15467. [PubMed: 17176067]
14. Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA, Modrich P. *Saccharomyces cerevisiae* MutLa is a mismatch repair endonuclease. *J Biol Chem*. 2007; 282:37181–37190. [PubMed: 17951253]
15. Ban C, Yang W. Crystal structure and ATPase activity of MutL: Implications for DNA repair and mutagenesis. *Cell*. 1998; 95:541–552. [PubMed: 9827806]
16. Hall MC, Shcherbakova PV, Fortune JM, Borchers CH, Dial JM, Tomer KB, Kunkel TA. DNA binding by yeast Mlh1 and Pms1: implications for DNA mismatch repair. *Nucleic Acids Res*. 2003; 31:2025–2034. [PubMed: 12682353]
17. Sacho EJ, Kadyrov FA, Modrich P, Kunkel TA, Erie DA. Direct visualization of asymmetric adenine nucleotide-induced conformational changes in MutL alpha. *Mol Cell*. 2008; 29:112–121. [PubMed: 18206974]
18. Arana ME, Holmes SF, Fortune JM, Moon AF, Pedersen LC, Kunkel TA. Functional residues on the surface of the N-terminal domain of yeast Pms1. *DNA Repair*. 2010
19. Guarne A, Junop MS, Yang W. Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHF ATPase. *EMBO J*. 2001; 20:5521–5531. [PubMed: 11574484]
20. Ban C, Junop M, Yang W. Transformation of MutL by ATP binding and hydrolysis: A switch in DNA mismatch repair. *Cell*. 1999; 97:85–97. [PubMed: 10199405]
21. Hall MC, Wang H, Erie DA, Kunkel TA. High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer. *J Mol Biol*. 2001; 312:637–647. [PubMed: 11575920]
22. Takamoto K, Chance MR. Radiolytic protein footprinting with mass Spectrometry to probe the structure of macromolecular complexes. *Annu Rev Bioph Biom*. 2006; 35:251–276.
23. Ritchie DW. Recent Progress and Future Directions in Protein-Protein Docking. *Current Protein and Peptide Science*. 2008; 9:1–15. [PubMed: 18336319]
24. Deterding LJ, Kast J, Przybylski M, Tomer KB. Molecular characterization of a tetramolecular complex between dsDNA and a DNA-binding leucine zipper peptide dimer by mass spectrometry. *Bioconjugate Chem*. 2000; 11:335–344.

25. Chen J, Smith DL, Griep MA. The Role of the 6 Lysines and the Terminal Amine of Escherichia Coli Single-Strand Binding Protein in its Binding of Single-Stranded DNA. *Protein Sci.* 1998; 7:1781–1788. [PubMed: 10082375]
26. Gale DC, Smith RD. Characterization of Noncovalent Complexes Formed between Minor Groove Binding Molecules and Duplex DNA by Electrospray Ionization-Mass Spectrometry. *J Am Soc Mass Spectr.* 1995; 6:1154–1164.
27. Maleknia SD, Downard K. Radical Approaches to Probe Protein Structure, Folding and Interactions by Mass Spectrometry. *Mass Spectrom Rev.* 2001; 20:388–401. [PubMed: 11997945]
28. Guan JQ, Takamoto K, Almo SC, Reisler E, Chance MR. Structure and dynamics of the actin filament. *Biochemistry.* 2005; 44:3166–3175. [PubMed: 15736927]
29. Xu G, Chance MR. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem Rev.* 2007; 107:3514–3543. [PubMed: 17683160]
30. Tullius TD, Dombroski BA. Hydroxyl Radical Footprinting - High-Resolution Information About DNA Protein Contacts and Application to Lambda-Repressor and Cro Protein. *P Natl Acad Sci USA.* 1986; 83:5469–5473.
31. Pogozelski WK, Mcneese TJ, Tullius TD. What Species Is Responsible for Strand Scission in the Reaction of [Fe(II)Edta](2-) and H₂O₂ with DNA. *J Am Chem Soc.* 1995; 117:6428–6433.
32. Bridgewater JD, Vachet RW. Using microwave-assisted metal-catalyzed oxidation reactions and mass spectrometry to increase the rate at which the copper-binding sites of a protein are determined. *Anal Chem.* 2005; 77:4649–4653. [PubMed: 16013884]
33. Bridgewater JD, Vachet RW. Metal-catalyzed oxidation reactions and mass spectrometry: The roles of ascorbate and different oxidizing agents in determining Cu-protein-binding sites. *Anal Biochem.* 2005; 341:122–130. [PubMed: 15866536]
34. Lim J, Vachet RW. Development of a methodology based on metal-catalyzed oxidation reactions and mass spectrometry to determine the metal binding sites in copper metalloproteins. *Anal Chem.* 2003; 75:1164–1172. [PubMed: 12641237]
35. Lim J, Vachet RW. Using mass spectrometry to study copper-protein binding under native and non-native conditions: beta-2-microglobulin. *Anal Chem.* 2004; 76:3498–3504. [PubMed: 15228316]
36. Sharp JS, Becker JM, Hettich RL. Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry. *Anal Biochem.* 2003; 313:216–225. [PubMed: 12605858]
37. Zhu Y, Guo TN, Park JE, Li X, Meng W, Datta A, Bern M, Lim SK, Sze SK. Elucidating in Vivo Structural Dynamics in Integral Membrane Protein by Hydroxyl Radical Footprinting. *Mol Cell Proteomics.* 2009; 8:1999–2010. [PubMed: 19473960]
38. Rana TM, Meares CF. Specific Cleavage of a Protein by an Attached Iron Chelate. *J Am Chem Soc.* 1990; 112:2457–2458.
39. Rana TM, Meares CF. Transfer of Oxygen from an Artificial Protease to Peptide Carbon during Proteolysis. *P Natl Acad Sci USA.* 1991; 88:10578–10582.
40. Cheal SM, Ng M, Barrios B, Miao Z, Kalani AK, Meares CF. Mapping Protein-Protein Interactions by Localized Oxidation: Consequences of the Reach of Hydroxyl Radical. *Biochemistry.* 2009; 48:4577–4586. [PubMed: 19354299]
41. Dainton FS. The Primary Quantum Yield in the Photolysis of Hydrogen Peroxide at 3130 Å and the Primary Radical Yield in the X- and γ-Radiolysis of Water. *J Am Chem Soc.* 1956; 78:1278–1279.
42. Weeks JL, Matheson MS. The Primary Quantum Yield of Hydrogen Peroxide Decomposition. *J Am Chem Soc.* 1956; 78:1273–1278.
43. Volman DH, Chen JC. The Photochemical Decomposition of Hydrogen Peroxide in Aqueous Solutions of Allyl Alcohol at 2537-Å. *J Am Chem Soc.* 1959; 81:4141–4144.
44. Aye TT, Low TY, Sze SK. Nanosecond laser-induced photochemical oxidation method for protein surface mapping with mass spectrometry. *Anal Chem.* 2005; 77:5814–5822. [PubMed: 16159110]
45. Sharp JS, Becker JM, Hettich RL. Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry. *Anal Chem.* 2004; 76:672–683. [PubMed: 14750862]

46. Sharp JS, Guo JT, Uchiki T, Xu Y, Dealwis C, Hettich RL. Photochemical surface mapping of C14S-Sml1p for constrained computational modeling of protein structure. *Anal Biochem*. 2005; 340:201–212. [PubMed: 15840492]
47. Hambly DM, Gross ML. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J Am Soc Mass Spectr*. 2005; 16:2057–2063.
48. Charvatova O, Foley BL, Bern MW, Sharp JS, Orlando R, Woods RJ. Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1. *J Am Soc Mass Spectr*. 2008; 19:1692–1705.
49. Maleknia SD, Brenowitz M, Chance MR. Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. *Anal Chem*. 1999; 71:3965–3973. [PubMed: 10500483]
50. Gupta S, Sullivan M, Toomey J, Kiselar J, Chance MR. The beamline X28C of the Center for Synchrotron Biosciences: a national resource for biomolecular structure and dynamics experiments using synchrotron footprinting. *J Synchrotron Radiat*. 2007; 14:233–243. [PubMed: 17435298]
51. Zheng XJ, Wintrode PL, Chance MR. Complementary structural mass spectrometry techniques reveal local dynamics in functionally important regions of a metastable serpin. *Structure*. 2008; 16:38–51. [PubMed: 18184582]
52. Kamal JKA, Chance MR. Modeling of protein binary complexes using structural mass spectrometry data. *Protein Sci*. 2008; 17:79–94. [PubMed: 18042684]
53. Liu RT, Guan JQ, Zak O, Aisen P, Chance MR. Structural reorganization of the transferrin C-lobe and transferrin receptor upon complex formation: The C-lobe binds to the receptor helical domain. *Biochemistry*. 2003; 42:12447–12454. [PubMed: 14580189]
54. Kiselar JG, Janmey PA, Almo SC, Chance MR. Structural analysis of gelsolin using synchrotron protein footprinting. *Mol Cell Proteomics*. 2003; 2:1120–1132. [PubMed: 12966145]
55. Ottinger LM, Tullius TD. High-resolution in vivo footprinting of a protein-DNA complex using gamma-radiation. *J Am Chem Soc*. 2000; 122:5901–5902.
56. Franchetbeuzit J, Spothemaurizot M, Sabattier R, Blazybaudras B, Charlier M. Radiolytic Footprinting - Beta-Rays, Gamma-Photons, and Fast-Neutrons Probe DNA-Protein Interactions. *Biochemistry*. 1993; 32:2104–2110. [PubMed: 8383534]
57. Hayes JJ, Kam L, Tullius TD. Footprinting Protein-DNA Complexes with Gamma-Rays. *Method Enzymol*. 1990; 186:545–549.
58. Sharp JS, Tomer KB. Effects of anion proximity in peptide primary sequence on the rate and mechanism of leucine oxidation. *Anal Chem*. 2006; 78:4885–4893. [PubMed: 16841907]
59. Xu GZ, Chance MR. Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting. *Anal Chem*. 2005; 77:4549–4555. [PubMed: 16013872]
60. Sharp JS, Tomer KB. Analysis of the oxidative damage-induced conformational changes of apo- and holocalmodulin by dose-dependent protein oxidative surface mapping. *Biophys J*. 2007; 92:1682–1692. [PubMed: 17158574]
61. Tong X, Wren JC, Konermann L. gamma-ray-mediated oxidative labeling for detecting protein conformational changes by electrospray mass spectrometry. *Anal Chem*. 2008; 80:2222–2231. [PubMed: 18260674]
62. Hall MC, Shcherbakova PV, Kunkel TA. Differential ATP binding and intrinsic ATP hydrolysis by amino-terminal domains of the yeast Mlh1 and Pms1 proteins. *J Biol Chem*. 2002; 277:3673–3679. [PubMed: 11717305]
63. Bebenek K, Kunkel TA. Analyzing fidelity of DNA polymerases. *Methods Enzymol*. 1995; 262:217–232. [PubMed: 8594349]
64. Drotschmann K, Hall MC, Shcherbakova PV, Wang H, Erie DA, Brownnewell FR, Kool ET, Kunkel TA. DNA binding properties of the yeast Msh2-Msh6 and Mlh1-Pms1 heterodimers. *Biol Chem*. 2002; 383:969–975. [PubMed: 12222686]
65. Beamer LJ, Pabo CO. Refined 1.8 Å crystal structure of the lambda repressor-operator complex. *J Mol Biol*. 1992; 227:177–196. [PubMed: 1387915]
66. Fokkema IFAC, den Dunnen JT, Taschner PEM. LOVD: Easy creation of a locus-specific sequence variation database using an “LSDB-in-a-Box” approach. *Hum Mutat*. 2005; 26:63–68. [PubMed: 15977173]

Figure 1. Primary and secondary structures of the Pms1-NTD

Secondary structural elements derived from the X-ray structure of Pms1-NTD are shown above the sequence (red = α helix; purple = 3_{10} helix; blue = β sheets). The sequences that were unstructured in the crystal structure and represent the modeled-in residues (110-118 and 275-286) are shown highlighted in yellow. The modeled-in side chains that were missing in the crystal structure are shown highlighted in green. The residues underlines in black make up the ATP-binding motif. The peptides generated through limited Arg-C (R) or Lys-C (K) proteolysis and observed by mass spectrometry are defined beneath the sequence, and are numbered so that peptide 1 begins at the N-terminus.

Figure 2. Pms1-NTD resists proteolysis in the presence of DNA

Aliquots of the Pms1-NTD digested by Arg-C or Lys-C in the absence and presence of dsDNA or ssDNA were taken at 15-minute intervals and analyzed by SDS-PAGE. The presence of intact protein at the 120-minute time point is attributed to protein that adhered to the tube throughout the experiment that was resolubilized by the addition of loading buffer prior to SDS-PAGE.

Figure 3. MALDI-ToF-MS identifies peptides with reduced abundance from limited proteolysis in the presence of dsDNA and ssDNA

Pms1-NTD peptides digested by (A) Arg-C for 45 minutes and (B) Lys-C for 60 minutes in the absence of DNA (top panel), in the presence of dsDNA (middle panel) and in the presence of ssDNA (bottom panel). The peptide notation refers to the protease used for limited proteolysis ('R' for Arg-C and 'K' for Lys-C) followed by the number of the peptide assuming that peptide 1 begins at the N-terminus. The peptides observed with a statistically significant change in abundance are denoted with arrows (↓ = protected; ↑ = more exposed). The * indicates a background ion.

Figure 4A

Figure 4. MALDI-ToF-MS analysis of peptides from limited proteolysis of the Pms1-NTD mutant R311E in the absence and presence of dsDNA

(A) The R311E mutant was digested by Arg-C for 45 minutes in the absence of DNA (top panel) and in the presence of dsDNA (bottom panel). The N-terminal tags, which are not part of the Pms1-NTD sequence, are the most abundant peptides in the spectra, particularly in the top panel. Therefore, the regions of the spectrum that flank N-terminal tag (AA1-4) have been expanded. (B) The R311E mutant was digested by Lys-C in the absence of DNA for 120 minutes (top panel) and in the presence of dsDNA for 180 minutes (bottom panel). The peptides observed with a change in abundance that correspond with that observed for the wild-type protein are denoted with arrows (↓ = protected; ↑ = more exposed). The * indicates a background ion and the # indicates a peak that corresponds in mass to an aberrant cleavage of peptide K21 (*i.e.*, cleavage C-terminal to Asp262 rather than Lys271).

Figure 5. MALDI-Tof-MS analysis of peptides from limited proteolysis of the Pms1-NTD mutant K197E/R198E/K229E in the absence and presence of dsDNA and ssDNA

(A) The K197E/R198E/K229E mutant was digested by Arg-C in the absence of DNA for 60 minutes (top panel), in the presence of dsDNA for 120 minutes (middle panel), and in the presence of ssDNA for 120 minutes (bottom panel). Because the arginine at position 198 has been mutated to a glutamic acid, this mutant does not generate an R5 peptide upon Arg-C proteolysis. (B) The K197E/R198E/K229E mutant was digested by Lys-C in the absence of DNA for 120 minutes (top panel), in the presence of dsDNA for 180 minutes (middle panel), and in the presence of ssDNA for 180 minutes (bottom panel). The peptides observed with a change in abundance that correspond with that observed for the wild-type protein are denoted with arrows (↓ = protected; ↑ = more exposed). The * indicates a background ion.

Figure 6. Protected residues mapped to the Pms1-NTD model

Amino acids that were protected from limited proteolysis in the presence of dsDNA, but not implicated in DNA binding are represented as cyan spheres. White spheres indicate oxidized amino acids that exhibited no change in the presence of dsDNA. Those residues that showed a statistically significant protection from oxidation in the presence of dsDNA, but were not implicated in DNA binding are represented in dark blue. The residues protected from limited proteolysis or oxidation in the presence of dsDNA, and used to guide DNA docking in generation of the Pms1-NTD/DNA model, are represented as yellow spheres. The regions that are highlighted in orange designate the ATP-binding motif.

Figure 7A

Figure 7B

Figure 7C

Figure 7. Model of the Pms1-NTD and its interaction with dsDNA

(A) Ribbon diagram of the Pms1-NTD, highlighting (in yellow) the residues involved in DNA binding as determined by experimental mass spectrometry data. (B) Electrostatic surface potential diagram of the modeled Pms1-NTD. The regions with an overall positive charge are shown in red, while the regions with an overall negative charge are shown in blue. These basic surfaces are best suited for interaction with the highly anionic DNA backbone. The basic groove of the Pms1-NTS has been indicated by blue arrows. The orientation of the protein is the same as that observed for the structure shown in (A). (C) Space filled model of the Pms1-NTD bound to DNA (tan) with the electrostatic surface potential included. The residues highlighted in (A) are also shown here in yellow to indicate the interactions with DNA.

Figure 8A

Figure 8B

Figure 8. Comparison of models of the LN40/ADP-PNP and the Pms1-NTD/DNA complexes
 (A) Model of the LN40/ADP-PNP complex (based on pdb entry: 1NHI). The positively-charged putative DNA binding groove has been indicated and Arg266 is shown as sticks.
 (B) Overlay of the Pms1-NTD/DNA model (orange) onto the LN40/ADP-PNP model. The DNA is shown as sticks.

Table 1

Relative abundances of peptides generated by limited proteolysis with Arg-C (45 minutes) and Lys-C (60 minutes) for the Pms1-NTD in the presence of double-stranded DNA. (n = 3)

Peptide	Cleavage sites		No DNA (%)	+dsDNA (%)
Protected				
R1	N _i	45	15.3 ± 1.5	8.9 ± 1.5
R4	170	188	17.7 ± 7.4	2.0 ± 0.9
R5	188	198	14.2 ± 6.6	0 ± 0
K12	172	190	10.9 ± 0.3	1.9 ± 1.8
K20	230	244	11.8 ± 1.1	0.3 ± 0.5
K21*	244	271	N/A	N/A
K30	364	380	2.1 ± 0.6	0 ± 0
Not Protected				
R9	256	273	5.2 ± 6.1	1.4 ± 2.0
R10	273	292	4.5 ± 3.6	0 ± 0
R11	292	306	18.9 ± 6.0	22.6 ± 3.1
R13	311	319	2.4 ± 1.1	4.2 ± 0.7
R16	373	389	6.5 ± 4.5	3.7 ± 2.1
K7	120	139	5.7 ± 0.5	2.6 ± 2.3
K25	294	309	16.2 ± 1.4	14.0 ± 3.3
More Exposed				
R8	243	256	8.4 ± 4.4	18.7 ± 2.5
R15	365	389	6.7 ± 2.6	38.6 ± 1.8
K26	309	318	21.5 ± 2.9	30.0 ± 1.3
K27	318	328	21.5 ± 1.0	35.4 ± 6.3
K31	380	C _t	6.4 ± 0.8	15.8 ± 1.0

All numbers are given in percent. The relative percent abundance for each peptide was calculated by dividing the peak area for the peptide by the sum of the areas for all peptides in the mass spectrum, whereas 0% indicates that the peptide is not observed in the spectrum.

Peptides are denoted with an 'R' or 'K' indicating whether they were generated from proteolysis with Arg-C or Lys-C, respectively.

Two-sample t -statistics were applied assuming equal variances where a change in oxidation was considered significant (highlighted in green) if the t -stat exceeded the two-tail t -critical value. Peptides were classified as 'Protected' or 'More Exposed' if the t -test indicated a significant decrease or increase, respectively, of the relative percent abundance in the presence of DNA.

Cleavage sites shown in italicized bold are those used to guide DNA docking in creation of the Pms1-NTD/DNA model.

* K197E/R198E/K229E mutant only

Table 2

Relative abundances of peptides generated by limited proteolysis with Arg-C (45 minutes) and Lys-C (60 minutes) for the Pms1-NTD in the presence of single-stranded DNA. (n = 3)

Peptide	Cleavage sites	No DNA (%)	+ssDNA (%)
Protected			
R4	170 188	17.7 ± 7.4	4.2 ± 2.9
R5	188 198	14.2 ± 6.6	1.8 ± 1.1
K12	172 190	10.9 ± 0.3	5.1 ± 2.7
K20	230 244	11.8 ± 1.1	2.0 ± 1.8
K21*	244 271	N/A	N/A
Not Protected			
R1	N _i 45	15.3 ± 1.5	12.8 ± 6.9
R8	243 256	8.4 ± 4.4	15.7 ± 2.9
R9	256 273	5.2 ± 6.1	1.6 ± 1.5
R10	273 292	4.5 ± 3.6	0.9 ± 0.8
R11	292 306	18.9 ± 6.0	16.2 ± 1.0
R16	373 389	6.5 ± 4.5	2.8 ± 2.0
K7	120 139	5.7 ± 0.5	2.0 ± 2.4
K25	294 309	16.2 ± 1.4	13.3 ± 2.9
K30	364 380	2.1 ± 0.6	0.6 ± 1.1
K31	380 C _t	6.4 ± 0.8	9.6 ± 2.0
More Exposed			
R13	311 319	2.4 ± 1.1	8.7 ± 2.2
R15	365 389	6.7 ± 2.6	35.3 ± 10.8
K26	309 318	21.5 ± 2.9	38.5 ± 6.8
K27	318 328	21.5 ± 1.0	28.9 ± 3.3

All numbers are given in percent. The relative percent abundance for each peptide was calculated by dividing the peak area for the peptide by the sum of the areas for all peptides in the mass spectrum, whereas 0% indicates that the peptide is not observed in the spectrum.

Peptides are denoted with an 'R' or 'K' indicating whether they were generated from proteolysis with Arg-C or Lys-C, respectively.

Two-sample t -statistics were applied assuming equal variances where a change in oxidation was considered significant (highlighted in green) if the t -stat exceeded the two-tail t -critical value. Peptides were classified as 'Protected' or 'More Exposed' if the t -test indicated a significant decrease or increase, respectively, of the relative percent abundance in the presence of DNA.

Cleavage sites shown in italicized bold are those used to guide DNA docking in creation of the Pms1-NTD/DNA model.

* K197E/R198E/K229E mutant only

Table 3

Percentages of oxidized peptides derived from Pms1-NTD in the absence and presence of AMP-PNP or DNA.
(n = 3)

Peptide	Sequence	Percent Oxidized	Percent Oxidized	Percent Oxidized
T1	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T2	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T3	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T4	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T5	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T6	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T7	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T8	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T9	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T10	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T11	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T12	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T13	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T14	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T15	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T16	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T17	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T18	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T19	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
T20	QVQVQVQVQVQVQVQVQV	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

Tryptic peptides are denoted with a 'T'.

Oxidized amino acids are highlighted in red.

Tryptic peptide T12 was observed alone and as a missed-cleavage peptide with T11; therefore, the data for both have been indicated.

All numbers are given in percent. Percent of oxidized peptide was calculated by dividing the peptide peak abundance for the oxidized peptide by the sum of the unoxidized and oxidized peptide peak abundances. 0% indicates that the oxidized peptide was not observed.

Two-sample *t*-statistics were applied assuming equal variances where a change in oxidation was considered significant (highlighted in green) if the *t*-stat exceeded the two-tail *t*-critical value. The percentages of oxidized peptide that exhibited significant protection when Pms1-NTD was in the presence of ligand are highlighted in gray.