Photocross-linking of the NH₂-terminal Region of Taq MutS Protein to the Major Groove of a Heteroduplex DNA

(Received for publication, March 5, 1997, and in revised form, June 25, 1997)

Vladislav A. Malkov, Indranil Biswas, R. Daniel Camerini-Otero, and Peggy Hsieh‡
From the Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-1810

The MutS DNA mismatch repair protein recognizes heteroduplex DNAs containing mispaired or unpaired bases. To identify regions of MutS protein in close proximity to the heteroduplex DNA, we have utilized the photocross-linking methodology. The relative affinities of the protein for heteroduplex DNA have been determined for two classes of mismatches. The affinities of the protein for heteroduplex DNA containing unpaired or mispaired bases compare with those for DNA containing guanine-thymine mismatches or base pair mismatches and insertion/deletion mutations, mis-

DNA mismatch repair is critical for mutation avoidance in virtually all organisms. In addition to its role in the repair of base pair mismatches and insertion/deletion mutations, mismatch repair serves as a barrier to homologous recombination between evolutionarily divergent sequences (reviewed in Ref. 1). The importance of mismatch repair is highlighted in its role in cancer surveillance; defects in highly conserved components of mismatch repair have been implicated in both sporadic and hereditary tumors (reviewed in Refs. 1, 2, and 5); however, the molecular details of heteroduplex DNA recognition by MutS, including the identity of structural determinants involved in mismatch recognition, remain obscure.

To identify regions of MutS protein in close contact with the heteroduplex DNA, we have carried out photocross-linking of nucleoprotein complexes containing a MutS protein from Thermus aquaticus bound to a derivatized heteroduplex DNA containing 5-iododeoxyuridine (5-IdUrd),1 a photoactivated zero-length cross-linker (6). Taq MutS protein has a higher affinity in vitro for heteroduplexes containing insertion/deletions compared with a G:T mismatch and, on the basis of chemical probe experiments, appears to interact with both the major and minor grooves of a heteroduplex DNA in the immediate vicinity of an unpaired base (7, 8).

Photocross-linking of DNA-protein complexes has been used to identify regions of a protein in close proximity to a bound DNA. The success of the approach hinges on the requirement for close proximity of an amino acid especially in the case of a zero-length cross-linking moiety, the need for the amino acid to assume an appropriate geometry for cross-linking, and the chemical reactivity of a potential target amino acid. For these reasons, not all amino acid residues in close contact with the DNA will be identified by cross-linking. Conversely, a cross-linked region, while close to the DNA, need not constitute part of a DNA-binding domain. Nevertheless, in the absence of a high-resolution structure, cross-linking experiments combined with mutational analysis of a candidate region can provide a starting point for identifying regions of a protein involved in DNA binding.

Photocross-linking of halogenated pyrimidines has been used to identify point contacts in DNA- and RNA-protein complexes (9–11). Specific DNA-protein photocross-linking utilizing the photoactivated 5-IdUrd cross-linking moiety has been reported by several groups. Willis et al. (6) demonstrated selective cross-linking of 5-IdUrd-substituted DNA to the Oxytricha telomere protein α-subunit following long-wavelength UV irradiation at 325 nm. Recently, we (12) and others (13) have utilized photocross-linking of a 5-IdUrd-substituted DNA to identify the DNA-binding domain of RecA protein, thereby confirming predictions based on crystallographic studies of RecA protein.

Photocross-linking with the 5-IdUrd cross-linker has several desirable characteristics. First, 5-IdUrd is nearly isoteric with thymine since the van der Waals radius of iodine is only 8% larger than that of the methyl group it replaces (6). Second, long-wavelength UV irradiation of 5-IdUrd-derivatized substrates results in specific cross-linking to 5-IdUrd while minimizing photodamage to other chromophores in the protein or DNA (6). Third, incidental cross-linking to regions of a protein not involved in DNA binding is minimized because the 5-IdUrd

1 The abbreviations used are: 5-IdUrd, 5-iododeoxyuridine; DTT, dithiothreitol; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]-glycine.

‡ To whom correspondence should be addressed: NIH, Bldg. 10, Rm. 9D04, 10 Center Dr. MSC 1810, Bethesda, MD 20892-1810. Tel.: 301-496-0306; E-mail: hsieh@ncifcrf.gov.
moeity is a zero-length cross-linker. Fourth, several amino acids can form adducts with photoactivated pyrimidines, including cysteine, serine, methionine, lysine, arginine, histidine, tryptophan, phenylalanine, and tyrosine (14).

In this paper, we show that a highly conserved region near the NH₂ terminus of Taq MutS protein is cross-linked to the unpaired base of a heteroduplex DNA. Substitution of Ala for Phe-39, the point of cross-linking, results in a mutant protein that is significantly impaired in its ability to bind heteroduplex DNA, although it retains the ability to oligomerize, is thermostable, and retains functional ATPase activity. These results implicate the region in the vicinity of Phe-39 as being crucial for heteroduplex DNA binding by Taq MutS protein.

MATERIALS AND METHODS

Reagents—Taq MutS protein isolated from an Escherichia coli BL21 overproducing strain was purified to apparent homogeneity on Source 30Q and MonoQ HR 10/10 anion-exchange columns (Pharmacia Biotech Inc.) as described elsewhere (8). The concentration of protein refers to protein monomers and was based on a molar extinction coefficient of ε₂₈₀ = 8.1 × 10⁴ M⁻¹ cm⁻¹ determined by amino acid analysis (8). Oligonucleotides were synthesized by automated β-cyanoethyl phosphoramidite DNA synthesis using 5-IdUrd-cyanoethyl phosphoramidite (Glen Research Corp.) on a Model 308B DNA synthesizer (Applied Biosystems, Inc.). DNA concentrations reflect the number of DNA molecules.

DNA Binding Assays—Gel mobility shift assays were carried out with 15 nM ³²P-labeled DNA substrates and 280 nM Taq MutS protein as described previously (7). For competition studies, ~420 nM ³²P-labeled monosubstituted DNA substrate ΔI-100 was incubated for 15 min at 37 °C with 850 nM Taq MutS protein and the indicated molar excess of unlabeled substrate ΔI over ΔI-100 in 10 μl of 20 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, and 0.1 mM DTT. Reactions were electrophoresed on 5% native polyacrylamide gels containing 7 μM MgCl₂ followed by quantitation using a Fuji BAS 2000 phosphorimager.

Analytical Scale UV Irradiation, Protease Digestion, and Purification of Cross-linked Peptides—Approximately 100 nM ³²P-labeled DNA substrates was preincubated for 15 min at 37 °C with 440 nM Taq MutS protein and the indicated molar excess of unlabeled substrate ΔI over ΔI-100) in 10 μl of 20 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, and 0.1 mM DTT. The sample was eluted in elution buffer 1. The buffer was adjusted to 0.2% SDS, 2 mM DTT, and 50 mM Tris-HCl, pH 8.0, in a reaction volume of 0.2 ml. The tube was heated at 75 °C for 20 min to inactivate the protein and cooled. 1.5 ml of clostripain buffer (250 mM Tris-HCl, pH 7.5, 50 mM CaCl₂, and 2 mM DTT) was added. After 4 h of clostripain digestion with 60 μM of enzyme at 37 °C, the reaction was further diluted by the addition of 1.2 ml of clostripain buffer containing 40 μg of clostripain and incubated overnight.

The sample was subjected to one cycle of clostripain digestion followed by three cycles of trypsin digestion as follows. Fractions containing ³²P-labeled DNA were brought to 0.2% SDS, 3 mM DTT, and 50 mM Tris-HCl, pH 8.0, in a reaction volume of 0.2 ml. The tube was heated at 75 °C for 20 min to inactivate the protein and cooled. 1.4 ml of clostripain buffer (250 mM Tris-HCl, pH 7.5, 50 mM CaCl₂, and 2 mM DTT) was added. After 4 h of clostripain digestion with 60 μM of enzyme at 37 °C, the reaction was further diluted by the addition of 1.2 ml of clostripain buffer containing 40 μg of clostripain and incubated overnight.

The sample was subjected to an oligo(dT)-cellulose column (0.9 g) as described above. The cross-linked material was eluted in elution buffer 2 (5 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 0.2 mM DTT, and 1 mM urea) and concentrated in a SpeedVac. Approximately 100 nM ³²P-labeled DNA sub-

RESULTS

Heteroduplex DNA Binding by Taq MutS Protein—The photocross-linking scheme necessitated modification of heteroduplex DNA substrates containing an unpaired base, usually a thymidine. The modified heteroduplexes differ from the canonical heteroduplex (ΔI) in that the former (Δ1-I-0, Δ1-I-2, and Δ1-I-3) each contain a single photoactivated 5-IdUrd cross-linking moiety at one of three different positions in the heteroduplex (Fig. 1). In addition, the modified heteroduplexes used

Downloaded from http://www.jbc.org/ by guest on July 25, 2018
Cross-linking of Taq MutS Protein to a Heteroduplex DNA

in cross-linking experiments contain an oligo(dA)18 single-strand region at the 5'-end of the strand bearing the 5-IdUrd moiety that facilitates purification of the cross-linked species. Since the van der Waals radius of iodine is only 8% larger than that of the methyl group it replaces (6), it is unlikely that the monosubstitution of thymine by photoactive 5-IdUrd would significantly alter the structure of the MutS nucleoprotein complex.

Here, we show that Taq MutS fails to recognize either the poly(dA)18 single-strand region or a 5-IdUrd-bearing homoduplex. This result is consistent with previous observations showing that Taq MutS does not readily form complexes with single-strand DNA or DNA homoduplexes (7). The ability of Taq MutS protein to bind to each of the heteroduplex DNAs and their corresponding control homoduplexes was assessed in a gel mobility shift assay. The positions of nucleoprotein complexes and unbound DNA are indicated on the right.

FIG. 1. The presence of an unpaired nucleotide is necessary and sufficient for Taq MutS binding. Top, shown are the sequences of all duplex DNAs used in this study. X denotes the photoactive 5-IdUrd moiety. Bottom, binding of Taq MutS protein to 32P-labeled DNA substrates was monitored in a gel mobility shift assay. The positions of nucleoprotein complexes and unbound DNA are indicated on the right.

All DNA binding experiments were conducted at 37 °C, and UV irradiation of nucleoprotein complexes was carried out at 30–40 °C (see “Materials and Methods”). We have previously shown that the extent of DNA binding by Taq MutS and the discrimination between heteroduplex and homoduplex DNAs are unchanged over a wide temperature range from 4 to 70 °C (7).

The mode of heteroduplex binding by Taq MutS protein to the 5-IdUrd-substituted ΔI(0) heteroduplex was assessed in competition experiments. As shown in Fig. 2, an increasing molar excess of the unsubstituted ΔI heteroduplex effectively competes for binding to the 32P-labeled, 5-IdUrd-substituted ΔI(0) heteroduplex. Quantitation of the extent of competition reveals that a 2-fold molar excess of the ΔI heteroduplex reduces binding of Taq MutS to the derivatized ΔI(0) heteroduplex by 50%. From the data shown in Figs. 1 and 2, we conclude that Taq MutS protein binds specifically to heteroduplex DNAs bearing an unpaired base and that this binding is not appreciably altered in heteroduplexes containing the photocross-linking derivative.

Photocross-linking of Heteroduplex-MutS Complexes—Analytical scale photocross-linking experiments were carried out with the ΔI(0) heteroduplex and Taq MutS protein to ascertain that cross-linking was specific for heteroduplex-MutS complexes. DNA binding reactions between Taq MutS protein and either the ΔI(0) heteroduplex or the corresponding AT-I(0) homoduplex (see Fig. 1) were carried out as described above. Samples were then irradiated at 312 nm to effect cross-linking as described under “Materials and Methods,” and the resulting material was visualized after SDS-polyacrylamide gel electrophoresis on 10–20% Tricine gels (Fig. 3). Cross-linking was achieved only when Taq MutS protein was incubated with the ΔI(0) heteroduplex and UV-irradiated and resulted in the appearance of a single cross-linked species as judged by electrophoretic mobility (Fig. 3, lane 4). The yield of cross-linked complexes was routinely 10–13%. Omission of either MutS protein or UV irradiation failed to yield any cross-linked species. As expected, no cross-linked material was detected in the case of the 5-IdUrd-substituted homoduplex (lane 8), consistent with the DNA substrate specificity of Taq MutS protein shown in Fig. 2.

Mapping of Cross-linked Peptides—To determine the optimal placement of the 5-IdUrd crossing-linking group within the heteroduplex DNA and to obtain preliminary information on the nature of any cross-linked peptide, we carried out cross-linking on an analytical scale followed by peptide fingerprinting. Dimethyl sulfate footprinting experiments revealed that
Cross-linking of Taq MutS Protein to a Heteroduplex DNA

![Image](http://www.jbc.org/)

**FIG. 3. Photocross-linking of heteroduplex-MutS complexes.** The $^{32}$P-labeled $\Delta 1$-I(0) heteroduplex (lanes 1–4) or AT-I(0) homoduplex (lanes 5–8) was incubated for 15 min in the presence (lanes 3, 4, 7, and 8) or absence (lanes 1, 2, 5, and 6) of a 2-fold molar excess of Taq MutS protein. Samples were monitored by SDS-polyacrylamide gel electrophoresis directly (lanes 1, 3, 5, and 7) or after irradiation at 312 nm (lanes 2, 4, 6, and 8).

Interactions of Taq MutS protein with the major groove of a heteroduplex are limited to several base pairs on either side of a mispaired or unpaired base (8). Correspondingly, the 5-IdUrd heteroduplexes are limited to several base pairs on either side of interactions of Taq MutS protein with the major groove of a 3-strand oligo(dA) tails were recovered by affinity chromatography over oligo(dT) columns. The use of DNA affinity chromatography for the purification of cross-linked peptides afforded several advantages over other methods that depend on characteristics of the peptide moiety. First, since the chromatography step is based on the recovery of heteroduplex DNAs rather than on any property of the cross-linked peptide, we minimized any bias in the recovery of cross-linked species. Second, the recovery of cross-linked material was compared at each round of purification to the recovery of uncross-linked DNAs, providing a means for monitoring selective as opposed to nonspecific losses of cross-linked peptides. Third, since Taq MutS protein exhibits insignificant binding to single-strand DNA (7), the risk of contaminating free protein and resulting peptides was lowered.

The cross-linked proteins were digested with clostripain, which cleaves at the carboxylic side of Arg, and were visualized by autoradiography after urea-denaturing polyacrylamide gel electrophoresis (Fig. 4). Each of the three heteroduplex DNAs cross-linked to a clostripain-derived peptide (denoted peptide A) was observed to have the same apparent electrophoretic mobility (Fig. 4, lanes 1, 5, and 9). Inspection of Fig. 4 reveals that the efficiency of cross-linking to different heteroduplexes showed significant differences. Whereas 10–13% of the complexes were cross-linked in the case of the $\Delta 1$-I(0) substrate, $<1\%$ of the complexes were cross-linked to the $\Delta 1$-I(−2) or $\Delta 1$-I(−3) heteroduplex.

Peptide fingerprints of the cross-linked peptide obtained by sequential rounds of digestion with two additional proteases strongly suggest that an identical MutS peptide is cross-linked to each of the three heteroduplexes. The clostripain-treated peptides were digested with trypsin, which cleaves at the carboxylic side of both Arg and Lys residues, followed by digestion with S. aureus V8 protease, which cleaves after Glu and Asp residues. Trypsin digestion did not result in the appearance of any new species (Fig. 4, lanes 2, 6, and 10), suggesting that the clostripain peptides cross-linked to DNA do not contain internal lysines and are the shortest tryptic peptides cross-linked to 5-IdUrd (see below). These data also suggest that only one region of Taq MutS protein is being efficiently cross-linked to the 5-IdUrd-substituted heteroduplex DNA.

Subsequent incubation with V8 protease under conditions in which cleavage occurs only at Glu, in NH$_4$HCO$_3$ buffer (Fig. 4, lanes 3, 7, and 11), or under conditions in which cleavage at Glu is favored over Asp, in NaH$_2$PO$_4$ buffer (lanes 4, 8, and 12), resulted in the appearance of a major cross-linked product (denoted peptide A′) as well as several minor species with identical electrophoretic mobilities for all three heteroduplexes. This result indicates that the cross-linked, clostripain-derived peptide has internal Glu residues. We have routinely observed that V8 protease cleaves at Glu residues more efficiently in the presence of the phosphate buffer compared with the carbonate buffer (12). This most likely accounts for the increase in abundance of product A′ in lanes 4, 8, and 12 compared with lanes 3, 7, and 11.

Discrete $^{32}$P-labeled products migrating more rapidly than the unbound oligodeoxynucleotide were also observed. As previously shown, these are UV photocleavage products of heteroduplex DNAs containing 5-IdUrd (16). We surmise that the greater abundance of DNA photocleavage products in the case of the $\Delta 1$-I(−2) and $\Delta 1$-I(−3) substrates compared with the...
MutS protein is relatively resistant to proteolytic cleavage by clostripain. We have also observed that the thermostable Taq labeled species (labeled sequencing was 36%. More important, the yield of cross-linked round of trypsinization revealed the presence of three 32P-gel electrophoresis of the cross-linked material after the final go(dT)-cellulose column and subjected to three additional cy-4, 0.5 where in the peptide fingerprinting experiment shown in Fig. 5, 1 mg of cross-linked protein was digested with 20 m M of clostripain, the addition of SDS and EDTA, and passage over an oligo(dT) column to remove unbound MutS protein (see above), the cross-linked material was subjected to clostripain digestion. A small aliquot was analyzed by denaturing gel electrophoresis, revealing that the clostripain digestion was incomplete, resulting in larger, underdigested products (compare Fig. 4 (lane 1) with Fig. 5 (lane 2)). Such a result was not unexpected since, in the preparative cross-linking experiment in Fig. 5, 1 mg of cross-linked MutS protein was digested with 100 µg of clostripain, whereas in the peptide fingerprinting experiment shown in Fig. 4, 0.5 µg of cross-linked protein was digested with 20 µg of clostripain. We have also observed that the thermostable Taq MutS protein is relatively resistant to proteolytic cleavage by a variety of proteases.

The clostripain-treated material was repurified on an oligo(dT)-cellulose column and subjected to three additional cycles of trypsin digestion, each followed by purification on an oligo(dT)-cellulose column. Urea-denaturing polyacrylamide gel electrophoresis of the cross-linked material after the final round of trypsinization revealed the presence of three 32P-labeled species (labeled A, B, and C) as well as free 32P-labeled oligonucleotide and DNA photolysis products (Fig. 5, lane 3). Material from bands A–C as well as unbound DNA (Oligo) were subjected to amino acid sequencing. After the extensive proteolysis and purification scheme, the final yield of all 32P-labeled material that was subjected to amino acid sequencing was 36%. More important, the yield of cross-linked DNA corresponding to bands A–C after proteolysis and purification was 8% of all 32P-labeled material compared with 10% of all 32P-labeled material prior to proteolysis and purification. Thus, the recoveries of cross-linked and non-cross-linked DNAs were similar throughout the proteolysis and purification procedure. At each cycle of proteolytic digestion and oligo(dT) purification, the recovery of 32P-labeled material was in excess of 70%. These results suggest that our experimental scheme did not result in selective loss of the major cross-linked species.

Amino acid sequencing of peptide A yielded an unambiguous match with a region from the NH₂ terminus of Taq MutS corresponding to amino acid residues 25–49 (Fig. 5). After 25 cycles of sequencing terminating in an Arg residue, no additional amino acid sequence was detected. As expected from the pattern of clostripain and trypsin digestion seen in Figs. 4 and 5, no Lys residues were identified in peptide A, and the deduced peptide sequence of the intact Taq MutS protein predicts that peptide A is immediately preceded by Arg-24. No amino acid was identified in cycle 15, which corresponds to Phe-39. Since Phe-34 and Phe-43 were detected in cycles 10 and 19, respectively, this strongly suggests that Phe-39 is the point of cross-linking.

The cross-linked material corresponding to band A migrated with the identical mobility as the clostripain-derived peptide obtained from analytical cross-linking experiments shown in Fig. 4 (band A) and represents the limit digest with trypsin. The recovery of the same cross-linked tryptic peptide after the abbreviated purification scheme in the analytical experiment and after extensive rounds of purification and proteolytic digestion in the preparative case strongly suggests that, under our cross-linking conditions, only one region of MutS protein is efficiently cross-linked to the heteroduplex DNA.

Based on electrophoretic mobilities, proteolytic precursor-product relationships, and amino acid sequencing, the peptides corresponding to bands B and C were shown to be derived from incomplete proteolysis of peptide A. First, the slower migrating B and C bands were not detected during analytical peptide mapping where the extent of proteolysis was greater (Fig. 4). Second, after the second cycle of preparative trypsin digestion, bands B and C predominated and band A was a minor species (data not shown), whereas after the third round of trypsin digestion, band A predominated at the expense of bands B and C (Fig. 5, lane 3). Third, amino acid sequencing of peptides from bands B and C indicated that they both initiate at Gly-7, consistent with trypsin cleavage at Lys-6 (7). Sequencing was carried out for 19 cycles ending at Asp-25, indicating that peptides B and C start upstream of peptide A and overlap with peptide A. The exact position of cross-linking of peptides B and C was downstream of Asp-25 and could not be determined. As expected, no amino acid sequence was obtained from the material corresponding to free oligonucleotide.

Loss of Heteroduplex Binding in a F39A Mutant Protein—Site-specific mutagenesis of the coding region of Taq MutS protein was carried out, resulting in a single amino acid substitution of Ala for Phe-39, the point of cross-linking. The wild-type and F39A mutant Taq MutS proteins were purified in parallel from E. coli.

The F39A protein exhibited the same anomalous behavior upon gel filtration on Sephacryl S-300 as the wild-type protein (Fig. 6A), yielding apparent sizes of 290 kDa for the F39A mutant compared with 280 kDa for the wild-type protein. Since the wild-type Taq MutS protein exists primarily as a dimer in solution, this result indicated that the substitution of Ala for Phe-39 did not significantly alter the overall structure of the protein and did not prevent oligomerization of the mutant monomers. Circular dichroism revealed that the F39A mutant

---

Fig. 5. Identification of cross-linked MutS peptides. Shown are the results from preparative urea-denaturing polyacrylamide gel electrophoresis of cross-linked peptides. Lane 1, heteroduplex Δ1-I(0) alone; lane 2, cross-linked heteroduplex Δ1-I(0)-Taq MutS nucleoprotein complex treated with clostripain alone (R); lane 3, cross-linked peptides after exhaustive clostripain and trypsin treatment (R+K). Material from bands A–C as well as unbound DNA (Oligo) were subjected to amino acid sequencing. The white F denotes the point of photocross-linking at Phe-39 for peptide A.

Δ1-I(0) substrate is due to the following. First, the frequency of photocleavage is greater when the 5-IdUrd moiety is paired with an adenine and stacked within the DNA helix as opposed to being unpaired. Second, the presence of a bound protein in the vicinity of 5-IdUrd may inhibit DNA photocleavage.

Amino acid Sequencing of Cross-linked Tryptic Peptides—The identity of the photocross-linked MutS peptide was unambiguously determined by amino acid sequencing. Taq MutS protein was incubated with the 32P-labeled Δ1-I(0) heteroduplex on a preparative scale and irradiated at 312 nm. As described for the analytical cross-linking shown in Fig. 3, the yield of cross-linking was ~10%. Due to the low levels of cross-linking of the Δ1-I(−2) and Δ1-I(+3) heteroduplexes described above, peptide sequencing of these complexes was not attempted. Following cross-linking of the Δ1-I(0) heteroduplex, the addition of SDS and EDTA, and passage over an oligo(dT) column to remove unbound MutS protein (see above), the cross-linked material was subjected to clostripain digestion. A small aliquot was analyzed by denaturing gel electrophoresis, revealing that the clostripain digestion was incomplete, resulting in larger, underdigested products (compare Fig. 4 (lane 1) with Fig. 5 (lane 2)). Such a result was not unexpected since, in the preparative cross-linking experiment in Fig. 5, 1 mg of cross-linked MutS protein was digested with 100 µg of clostripain, whereas in the peptide fingerprinting experiment shown in Fig. 4, 0.5 µg of cross-linked protein was digested with 20 µg of clostripain. We have also observed that the thermostable Taq MutS protein is relatively resistant to proteolytic cleavage by a variety of proteases.

The clostripain-treated material was repurified on an oligo(dT)-cellulose column and subjected to three additional cycles of trypsin digestion, each followed by purification on an oligo(dT)-cellulose column. Urea-denaturing polyacrylamide gel electrophoresis of the cross-linked material after the final round of trypsinization revealed the presence of three 32P-labeled species (labeled A, B, and C) as well as free 32P-labeled oligonucleotide and DNA photocleavage products (Fig. 5, lane 3). Material from bands A–C and free DNA (Oligo) were excised from the gel and subjected to amino acid sequencing. After the extensive proteolysis and purification scheme, the final yield of all 32P-labeled material that was subjected to amino acid sequencing was 36%. More important, the yield of cross-linked

---

2 I. Biswas and P. Hsieh, unpublished observations.
ATP is hydrolyzed to ADP and P\(_i\) (7). As shown in Fig. 6B, the ATPase activity of the F39A mutant protein was very similar to that of the wild-type protein at 70 °C. In addition, both the wild-type and mutant MutS proteins retained full ATPase activity after incubation for 20 min at 70 °C. Thus, the substitution at Phe-39 has no significant effect on either the thermostability or the ATPase activity of MutS protein. The identical size as determined by gel filtration chromatography, the similar CD spectra, and the same thermal stability and ATPase activity of the wild-type and F39A mutant proteins are wholly consistent with the absence of gross conformational changes in the mutant protein.

The ability of the F39A mutant protein to bind heteroduplex DNA was assessed in gel mobility shift assays. As shown in Fig. 6C, the wild-type Taq MutS protein readily bound to a heteroduplex DNA containing an unpaired thymidine residue. In contrast, the F39A mutant protein bound very weakly to the same heteroduplex, requiring protein concentrations in excess of 10⁻⁷ M (although the size of the DNA-F39A complex was unchanged as judged by electrophoretic mobility). Examination of the binding data indicates that substitution of Ala for Phe-39 may have lowered the relative affinity for heteroduplex DNA by some 3 orders of magnitude. The best fit of the data in Fig. 6C, assuming that at equilibrium one dimer of MutS binds to a single heteroduplex DNA, yielded apparent dissociation binding constants of 6 x 10⁻¹⁰ and 1 x 10⁻⁸ M for the wild-type and F39A proteins, respectively. While these are only rough approximations of the relative affinities for heteroduplex DNA, the data in Fig. 6C clearly establish that heteroduplex binding is severely impaired in the F39A mutant. In fact, the high protein concentration (>10⁻⁷ M) required for residual heteroduplex binding of the mutant protein is similar to that required for binding to perfect homoduplex DNA by the wild-type Taq MutS protein (data not shown). We were unable to detect any binding to homoduplex DNA by the F39A mutant protein under these conditions.

**DISCUSSION**

Photocross-linking of Taq MutS protein to a derivatized heteroduplex DNA containing a 5-IdUrd cross-linking moiety reveals that a region at the NH₂ terminus of MutS is closely associated with the major groove of the heteroduplex DNA. Peptide sequencing of the limit trypsin digest of the cross-linked peptide indicates that it maps to residues 25–49, with Phe-39 being the point of cross-linking. Substitution of Ala for Phe-39 linked peptide indicates that it maps to residues 25–49, with Phe-39 as being critical for heteroduplex binding by Taq MutS protein. The severe deficiency in DNA binding resulting from a single amino acid change at Phe-39 is not attributable to a gross alteration in the conformation of the mutant protein. The F39A mutant protein is able to dimerize like its wild-type counterpart. In addition, the F39A mutant protein retains thermostability and an ATPase activity of the wild-type and mutant MutS proteins. The identical size at Phe-39 has no significant effect on either the thermostability or the ATPase activity of MutS protein. The identical size as determined by gel filtration chromatography, the similar CD spectra, and the same thermal stability and ATPase activity of the wild-type and F39A mutant proteins are wholly consistent with the absence of gross conformational changes in the mutant protein.

The thermal stability of the mutant protein as well as its ATPase activity were assessed in parallel with the wild-type protein. We have previously shown that the wild-type Taq MutS protein has a thermostable ATPase activity in which protein has to a first approximation retained the native conformation observed for the wild-type protein and ruled out the possibility that the mutant polypeptide is grossly misfolded or denatured (data not shown).

The thermal stability of the mutant protein as well as its ATPase activity were assessed in parallel with the wild-type protein. We have previously shown that the wild-type Taq MutS protein has a thermostable ATPase activity in which...
The NH2-terminal region of MutS is involved in the recognition of intermediates of homologous recombination rather than mismatch repair since the first amino acids of MSH4 and MSH5 align at Taq MutS positions 73 and 101, respectively. Multiple sequence alignments of MutS family members using Pileup (Genetics Computer Group, Inc., Madison, WI) suggest that MSH4 and MSH5 may have lost the putative DNA-binding domain identified in this study since the first amino acids of MSH4 and MSH5 align at the N terminus of MutS involved in DNA binding may be larger or smaller.

While the cross-linked region is conserved in virtually all MutS homologs, there are two notable exceptions. A search of GenBank™ data base sequences (Release 100) using BLAST for homology to the first 50 amino acids of Taq MutS protein identified all known members of the MutS family except Saccharomyces cerevisiae MSH4 and MSH5. Multiple sequence alignments of MutS family members using Pileup (Genetics Computer Group, Inc., Madison, WI) suggest that MSH4 and MSH5 may have lost the putative DNA-binding domain identified in this study since the first amino acids of MSH4 and MSH5 align at the N terminus of MutS involved in DNA binding may be larger or smaller.

The geometry of the unpaired 5-IdUrd base within the nucleoprotein complex containing heteroduplex ΔI-I(0) is unknown, the fact that the identical clostripain peptide was cross-linked to heteroduplexes ΔI-I(−2) and ΔI-I(+3) in which the 5-IdUrd moieties are stacked within the duplex (20) suggests that the NH2-terminal region of Taq MutS protein is in close proximity to the major groove near an unpaired base. These data suggest that the specificity of heteroduplex DNA recognition is achieved by the interaction of MutS protein with the mispaired or unpaired base and, at a minimum, constituents residing in the major groove of flanking residues near the lesion. Interactions involving more distal regions of the heteroduplex detected by DNase I footprinting (8, 21) may serve to increase the affinity of MutS protein for nonspecific DNA contacts.

Photocross-linking of Phe-39 to a heteroduplex and characterization of a mutant protein bearing a substitution of Ala for Phe-39 suggest that Phe-39 is involved in heteroduplex DNA binding by MutS protein. We note, however, that the exact role of Phe-39 in DNA binding has not been established. Phe-39 may make direct contacts with the unpaired base. If so, the low levels of cross-linking to the heteroduplexes substituted at positions −2 and +3 involve residues other than Phe-39, a prediction we have not tested. Alternatively, despite its close proximity to the unpaired base, Phe-39 may have an indirect role in DNA binding, contributing, for example, to the conformational stability of the DNA-binding site. Finally, the identification of other residues critical for DNA binding that reside in the vicinity of Phe-39 as well as elsewhere in the MutS polypeptide awaits further study.

Acknowledgments—We thank the members of the W. M. Keck Foundation Biotechnology Resource Laboratory, especially Drs. Ken Williams and Kathy Stone, for peptide sequencing and helpful discussions. We are grateful to Howard Nash for comments on the manuscript, to George Poy for oligonucleotide syntheses, and to Linda Robinson for assistance.

REFERENCES
1. Modrich, P., and Lahue, R. (1996) Annu. Rev. Biochem. 65, 101–133
2. Kolodner, R. (1996) Genes Dev. 10, 1433–1443
3. Kiegerl, J., Umar, A., Rizzo, J., Berezney, A., Kunkel, T. A., and Barrett, J. C. (1996) Nat. Genet. 14, 102–105
4. LeClerc, J. E., Li, B., Payne, W. L., and Cebula, T. A. (1996) Science 274, 1209–1211
5. Modrich, P. (1991) Annu. Rev. Genet. 25, 229–253
6. Willis, M. C., Hicke, B. J., Uhenbeck, O. C., Cech, T. R., and Koch, T. H. (1993) Science 262, 1255–1257
7. Biswas, I., and Hsieh, P. (1996) J. Biol. Chem. 271, 5040–5048
8. Biswas, I., and Hsieh, P. (1997) J. Biol. Chem. 272, 13355–13364
9. Blatter, E. E., Ehrig, Y. W., and Ehrig, R. H. (1992) Nature 359, 350–352
10. Hicke, B. J., Willis, M. C., Koch, T. H., and Cech, T. R. (1994) Biochemistry 33, 3578–3574
11. Willis, M. C., LeCuyer, A. K., Meisenheimer, K. M., Uhenbeck, O. C., and Koch, T. H. (1994) Nucleic Acids Res. 22, 4947–4952
12. Malkov, V. A., and Camerini-Otero, R. D. (1995) J. Biol. Chem. 270, 30230–30233
13. Wang, Y., and Adzuma, K. (1996) Biochemistry 35, 3563–3571
14. Chodosh, L. A. (1996) in Current Protocols in Molecular Biology (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) Vol. 12.5.1–12.5.8, John Wiley & Sons, Inc., New York
15. Drakeau, G. R. (1976) Methods Enzymol. 45, 469–475
16. Murray, V., and Martin, R. F. (1989) Nucleic Acids Res. 17, 2675–2691
17. Ross-MacDonald, P., and Roeder, G. S. (1994) Cell 79, 1069–1080
18. Hollingsworth, N. M., Pante, L., and Halsey, C. (1995) Genes Dev. 9, 1728–1739
19. Jirinov, J., Hughes, M., Corman, N., and Rudkin, B. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8660–8664
20. Cantor, C. R., and Schimmel, P. R. (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules, pp. 311–341, W. H. Freeman & Co., New York
21. Su, S.-S., Lahue, R. S., Au, K. G., and Modrich, P. (1988) J. Biol. Chem. 263, 6829–6835
Photocross-linking of the NH$_2$-terminal Region of Taq MutS Protein to the Major Groove of a Heteroduplex DNA

Vladislav A. Malkov, Indranil Biswas, R. Daniel Camerini-Otero and Peggy Hsieh

J. Biol. Chem. 1997, 272:23811-23817.
doi: 10.1074/jbc.272.38.23811

Access the most updated version of this article at http://www.jbc.org/content/272/38/23811

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 20 references, 9 of which can be accessed free at http://www.jbc.org/content/272/38/23811.full.html#ref-list-1