

# Influence of Local Duplex Stability and $N^6$ -Methyladenine on Uracil Recognition by Mismatch-Specific Uracil-DNA Glycosylase (Mug)

Victoria Valinluck,<sup>†</sup> Pingfang Liu,<sup>†,‡</sup> Artur Burdzy,<sup>†</sup> Junichi Ryu,<sup>†</sup> and Lawrence C. Sowers<sup>\*,†,‡</sup>

Department of Biochemistry and Microbiology, School of Medicine, Loma Linda University, Loma Linda, California 92350, and Graduate School of Biological Sciences, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, California 91010

Received July 2, 2002

To maintain genomic integrity, DNA repair enzymes continually remove damaged bases and lesions resulting from endogenous and exogenous processes. These repair enzymes must distinguish damaged bases from normal bases to prevent the inadvertent removal of normal bases, which would promote genomic instability. The mechanisms by which this high level of specificity is accomplished are as yet unresolved. One member of the uracil–DNA glycosylase family of repair enzymes, *Escherichia coli* mismatch-specific uracil–DNA glycosylase (Mug), is reported to distinguish U:G mispairs from U:A base pairs based upon specific contacts with the mispaired guanine after flipping the target uracil out of the duplex. However, recent studies suggest other mechanisms for base selection, including local duplex stability. In this study, we used the modified base  $N^6$ -methyladenine to probe the effect of local helix perturbation on Mug recognition of uracil.  $N^6$ -Methyladenine is found in *E. coli* as part of both the mismatch repair and restriction-modification systems. In its cis isomer,  $N^6$ -methyladenine destabilizes hydrogen bonding by interfering with pseudo-Watson–Crick base pairing. It is observed that the selection of uracil by Mug is sequence dependent and that uracil residues in sequences of reduced thermostability are preferentially removed. The replacement of adenine by  $N^6$ -methyladenine increases the frequency of removal of the uracil residue paired opposite the modified adenine. These results are in accord with suggestions that local helix stability is an important determinant of base recognition by some DNA repair enzymes and provide a potential strategy for identifying the sequence location of modified bases in DNA.

## Introduction

The human genome is under constant insult from both exogenous and endogenous DNA damaging agents. The number of lesions resulting from endogenous DNA damage sources such as water, reactive oxygen species, and alkylating agents is estimated to be on the order of 10 000/cell/day (1–3). To maintain genomic integrity and stability, DNA repair pathways are implemented. The major mechanisms that exist for the repair of these lesions include base excision repair, mismatch repair, nucleotide excision repair, and recombination repair, although base excision repair is most commonly employed in the repair of endogenous lesions (4–9). Base excision repair glycosylases must distinguish the estimated  $10^4$  damaged bases in need of repair from the  $10^9$  normal bases also present. The selectivity of these glycosylases must be greater than 1 in  $10^5$  to prevent the excision of undamaged bases, thereby preventing inadvertent damage to the genome that would contribute to genomic instability. The mechanisms by which glycosylases find a damaged base and distinguish that base from normal bases are under investigation.

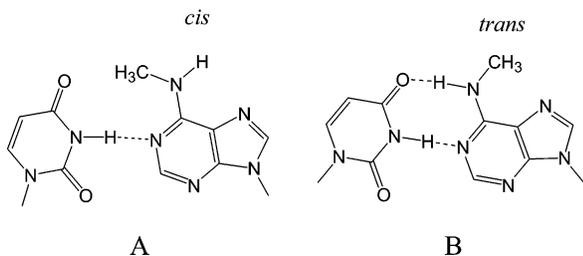
*Escherichia coli* mismatch-specific uracil–DNA glycosylase (Mug)<sup>1</sup> is a member of the uracil glycosylase family and only cleaves its substrate in the context of a DNA duplex, explaining its initial name, double-strand specific uracil-DNA glycosylase (10, 11). Later, it was reported that Mug recognizes U:G mispairs, but not U:A base pairs (10–12), and that Mug specifically recognizes the “widowed” guanine opposite the target uracil via three specific hydrogen bonds (12, 13), thus the designation mismatch-specific uracil–DNA glycosylase (Mug) (12). However, other investigators refer to this base excision repair enzyme as ethenoC-DNA glycosylase based on its high excision activity against ethenocytosine in duplex DNA (14). Recent studies suggest that thermal stability rather than recognition of the “widowed” guanine is the primary determinant of Mug substrate selectivity (15). Mug has been reported, by several groups, to excise uracil from pseudo-Watson–Crick U:A base pairs as well as wobble U:G mispairs, though at a lower frequency than the U:G mispairs (14–16). In this study, we investigate whether thermal destabilization ( $\Delta T_m$ ) may be a key factor in the mechanism for Mug substrate selectivity. We do so by

\* To whom correspondence should be addressed. Telephone: (909) 558-4480. Fax: (909) 558-4035. E-mail: lsowers@som.llu.edu.

<sup>†</sup> Loma Linda University.

<sup>‡</sup> City of Hope National Medical Center.

<sup>1</sup> Abbreviations: Mug, mismatch-specific uracil–DNA glycosylase; APE, human apurinic/apyrimidinic endonuclease; m6A,  $N^6$ -methyladenine; tms, trimethylsilyl;  $T_m$ , oligonucleotide melting temperature; dU, 2'-deoxyuridine; GC/MS, gas chromatography–mass spectrometry; DMT, dimethoxytrityl.

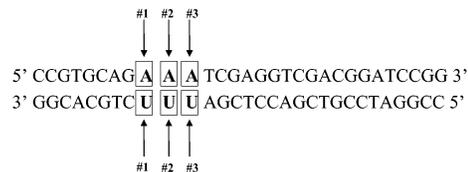


**Figure 1.** (A) Base pairing between uracil and N1-cis rotational isomer of *N*<sup>6</sup>-methyladenine, and (B) pseudo-Watson–Crick base pairing between uracil and N1-trans rotational isomer of *N*<sup>6</sup>-methyladenine. Dashed lines represent hydrogen-bonding interactions.

studying the effect of modest base pair destabilization, induced by m6A base paired with uracil in duplex oligonucleotide DNA, on Mug substrate recognition.

*N*<sup>6</sup>-Methyladenine (m6A) is involved in mismatch repair in *E. coli*. After DNA replication, the newly synthesized daughter strand is transiently unmethylated at adenine in GATC sequences while the parent strand retains the adenine methylation pattern (17). This adenine hemimethylation allows for strand discrimination and directs the mismatch repair in *E. coli* (17–19). For the free nucleoside, there is a 20:1 preference for the m6A cis rotational isomer over the trans rotational isomer (20). However, in DNA, the methyl group of m6A exists as either cis or trans to N1 with a preference for the trans isomer (20) (Figure 1). In both isomers, the methyl group is coplanar with the heterocyclic rings of adenine. When in the trans configuration, the methyl group of the m6A protrudes into the major groove of the helix. This positioning of the methyl group in the trans orientation does not disturb hydrogen bonding between the m6A and the opposite pyrimidine, allowing for pseudo-Watson–Crick base pairing. However, when in the cis configuration, the methyl group of m6A is located on the N1 side of the base, interfering with pseudo-Watson–Crick base pairing (20, 21). It has been shown that the presence of m6A in DNA destabilizes base pairing and helix stability (22). In this study, we use this base-pairing destabilization property of m6A to create selected sites of instability within a duplex to study the substrate recognition mechanism of the base excision repair enzyme, mismatch-specific uracil–DNA glycosylase (Mug).

We prepared an oligonucleotide containing a run of three adenines base-paired with three uracils, as well as a series of oligonucleotide duplexes containing uracil paired with m6A at selected positions. These oligonucleotides were then subjected to uracil excision by Mug. Examination of Mug-induced uracil excision opposite a run of three adenine residues revealed that the uracil excision efficiency is consistent with the local stability of the DNA duplex. We also find that, within the same sequence context, uracil opposite m6A is rendered more susceptible to Mug-mediated uracil excision as compared to uracil opposite adenine. These results, in conjunction with *T<sub>m</sub>* data obtained from corresponding 11-mer duplexes, confirm the importance of slight differences in thermal stability for Mug substrate recognition. Furthermore, the recognition of modest base pair destabilization by Mug can potentially be exploited as a strategy for determining the location of m6A in a DNA duplex, applicable in the characterization of restriction-modification systems.



**Figure 2.** Sequence of 30-mer duplexes used in Mug-mediated uracil excision assays. The duplex without m6A is **oligo A**. Introduction of m6A into the duplex at position 1 gives **oligo A1**, at position 2 gives **oligo A2**, and at position 3 gives **oligo A3**.

## Materials and Methods

**m6A Phosphoramidite Preparation (23–25).** *N*<sup>6</sup>-Methyl-2'-deoxyadenosine (Aldrich) was coevaporated with anhydrous pyridine (Aldrich). The 1.13 mmol of dry product was tritylated with 1.22 mmol DMT-Cl (Aldrich) in 7 mL of anhydrous pyridine overnight at room temperature. Completion of the reaction was determined via thin-layer chromatography on general purpose silica gel glass plates with UV indicator (Aldrich) using 5% (v/v) MeOH in dichloromethane (Aldrich) as the solvent. Extraction of the tritylation reaction was completed using dichloromethane and saturated sodium bicarbonate (Aldrich). The organic phase was retained and evaporated. The dried product was dissolved in dichloromethane with 0.5% (v/v) triethylamine (Aldrich). A Silica H gel column (Aldrich) with dichloromethane with 0.5% (v/v) triethylamine as the solvent was used to purify the *N*<sup>6</sup>-methyl-2'-deoxyadenosine-DMT, with a total yield of 75%. *N*<sup>6</sup>-methyl-2'-deoxyadenosine-DMT (0.85 mol or 0.483 g) was phosphitylated with 300  $\mu$ L of 2-cyanoethyl tetraisopropylphosphoramidite 97% (Aldrich), in the presence of 75 mg of *N,N*-diisopropylammonium tetrazolide (Aldrich) and 2 mL of acetonitrile overnight, under argon at room temperature. Completion of the reaction was determined via thin-layer chromatography using 5% (v/v) MeOH in dichloromethane as the solvent. Extraction of the phosphitylation reaction was completed with 20% (v/v) sodium chloride (Aldrich) in dH<sub>2</sub>O with equal volume of triethylamine and ethyl acetate. The organic phase was retained and evaporated. The dried product was dissolved in ethyl acetate. The phosphoramidite purification was completed using a Silica H gel column with ethyl acetate and hexane at varying concentrations with 0.5% (v/v) triethylamine as the solvent. The final product was dried under reduced pressure.

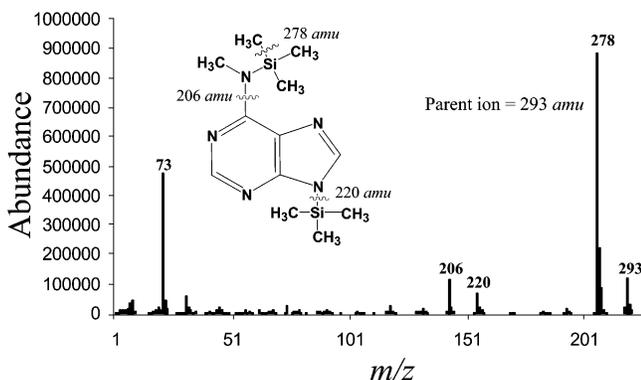
**dU Phosphoramidite Preparation.** 2'-Deoxyuridine was converted to its corresponding phosphoramidite using the same protocol as described previously for the *N*<sup>6</sup>-methyl-2'-deoxyadenosine phosphoramidite (24, 25), except on the gram scale rather than the milligram scale.

**Oligonucleotide Synthesis and Purification.** Oligonucleotide 30-mers (Figure 2) were prepared by standard solid-phase synthesis using either the Gene Assembler Plus (Pharmacia) or Expedite Nucleic Acid Synthesis System (Applied Biosystems) automated DNA synthesizers (24, 25). Phosphoramidites other than those for m6A and dU, were obtained from Glen Research. The oligonucleotides were removed from the solid support and deprotected in aqueous ammonia (Aldrich) at 60 °C overnight, after which they were purified with Poly-Pak II cartridges (Glen Research). The composition of the oligonucleotides was verified by gas chromatography–mass spectrometry (GC/MS) after hydrolysis in 200  $\mu$ L of 88% (v/v) formic acid (Aldrich) at 140 °C for 30 min, and silylation in 100  $\mu$ L of dry acetonitrile (Aldrich) and 100  $\mu$ L of bis-(trimethylsilyl) trifluoroacetamide containing 1% (v/v) trimethylchlorosilane (Aldrich) (26, 27). The GC/MS analysis was conducted with a Hewlett-Packard 5970 GC interfaced with a 5890 mass selective detector. The *m/z* of predominant ions from the mass spectrum of *N*<sup>6</sup>-methyladenine is shown in Figure 3. Oligonucleotides for the melting temperature studies and for the oligonucleotide size markers used in the Mug-mediated uracil excision assays were synthesized in the same manner as described for the 30-mer oligonucleotides.

**Table 1. Comparison of  $T_m$  ( $^{\circ}$ C) for 11-mer Duplexes Containing Adenine or  $N^6$ -Methyladenine at Different Positions Paired with either Uracil or Thymine<sup>a</sup>**

	GCAGA <sub>1</sub> A <sub>2</sub> A <sub>3</sub> TCGA (A)	GCAGMA <sub>2</sub> A <sub>3</sub> TCGA (m6A1)	GCAGA <sub>1</sub> MA <sub>3</sub> TCGA (m6A2)	GCAGA <sub>1</sub> A <sub>2</sub> MTCGA (m6A3)
TCGAUUUCTGC (U)	44.0 ± 0.6	41.5 ± 1.1	41.4 ± 0.4	42.0 ± 0.5
TCGATTTCTGC (T)	46.6 ± 0.6	42.6 ± 0.3	43.1 ± 0.4	44.4 ± 0.9

<sup>a</sup> The measurements were carried out in 0.1 M NaCl, 0.01 M sodium phosphate, and 0.1 mM EDTA (pH 7). The  $T_m$  values shown were determined at a total strand concentration of 25  $\mu$ M. The values shown are the averages of five to six separate determinations. **M** denotes  $N^6$ -methyladenine. The names of the single-stranded 11-mers are in parentheses, italicized and bolded.

**Figure 3.** Mass spectrum of doubly silylated  $N^6$ -methyladenine.

The sequences of the 11-mer oligonucleotides used in the melting temperature measurements are seen in Table 1. The sequences of the oligonucleotides used as the oligonucleotide size markers in the Mug-mediated uracil excision assays are as follows:

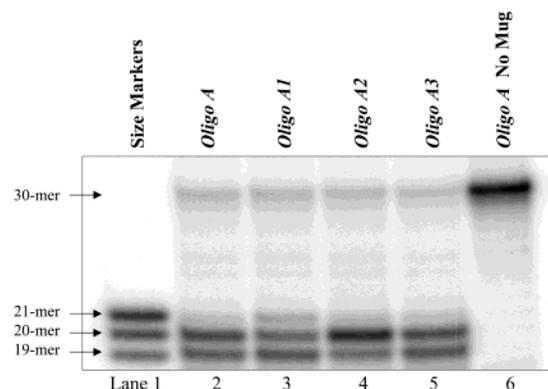
19-mer: 5' CCGGATCCGTCGACCTCGA 3'

20-mer: 5' CCGGATCCGTCGACCTCGAT 3'

21-mer: 5' CCGGATCCGTCGACCTCGATT 3'

**Oligonucleotide Labeling and Annealing.** The 30-mer oligonucleotide with the run of three uracil residues (Figure 2, bottom strand) was 5'-<sup>32</sup>P-end labeled by T4 polynucleotide kinase (New England Biolabs) with [ $\gamma$ -<sup>32</sup>P]ATP (ICN Life Sciences) under conditions recommended by the enzyme supplier. The labeled oligonucleotides were purified using G50 Sephadex columns (Boehringer Mannheim). The labeled oligonucleotides were annealed with a complementary strand containing m6A at one of three different positions, resulting in three different duplexes with m6A replacing adenine at positions 1, 2, or 3 (Figure 2). The annealing was carried out in 20 mM Tris-HCl, pH 8.0, 0.1 mg/mL BSA, 1 mM EDTA, 1 mM EGTA, and 1 mM DTT, with a 2-fold molar excess of the unlabeled complementary strand. The annealing mixture was heated to 95  $^{\circ}$ C for 5 min and then allowed to slowly cool to room temperature. To confirm duplex formation, the duplexes were digested with *Sa*II (New England Biolabs) under conditions recommended by the enzyme supplier, and the products were sized on denaturing 20% (v/v) polyacrylamide gels (not shown). The 19-, 20-, and 21-mers used for the oligonucleotide size markers (Figure 4, lane 1) were also 5'-<sup>32</sup>P-end labeled and purified as described above.

**Mug-mediated Uracil Excision Assays.** *E. coli* Mug, Human AP Endonuclease (APE), and their respective reaction buffers were obtained from Trevigen Inc. Each Mug-mediated uracil excision reaction was conducted with approximately 2 pmol of double-stranded oligonucleotide substrate (Figure 2) and 1  $\mu$ L of Mug (1 unit/ $\mu$ L) in 20 mM Tris-HCl, pH 8.0, 0.1 mg/mL BSA, 1 mM EDTA, 1 mM EGTA, and 1 mM DTT in a 10  $\mu$ L reaction volume for 5, 10, 15, 20 or 25 min at 37  $^{\circ}$ C. Trevigen defines 1 unit of Mug as the amount that cleaves 1 pmol of a <sup>32</sup>P-labeled oligonucleotide probe containing 3, $N^6$ -ethenocytosine within an oligonucleotide duplex in 1 h at 37  $^{\circ}$ C. The apyrimi-

**Figure 4.** Twenty minute Mug-mediated uracil excision of 30-mer duplexes with uracil paired with either adenine or  $N^6$ -methyladenine at different positions. Lane 1: 19-, 20-, and 21-mer size markers. Lanes 2–5: 20 min Mug-mediated uracil excision of *oligo A*, *oligo A1*, *oligo A2*, and *oligo A3*, respectively. Lane 6: negative control, *oligo A* incubated in 20 mM Tris-HCl, pH 8.0, 0.1 mg/mL BSA, 1 mM EDTA, 1 mM EGTA, and 1 mM DTT for 20 min without the addition of Mug. This reaction was conducted in a 10  $\mu$ L reaction volume, and was treated with human APE as described for all other Mug-mediated uracil excision reactions.

dic sites were cleaved by the addition of 1  $\mu$ L of human APE (1 unit/ $\mu$ L) under 10 mM HEPES-KOH, pH 6.5, 100 mM KCl, 10 mM MgCl<sub>2</sub> buffer conditions in a volume of 15  $\mu$ L and then incubated for 1 h at 37  $^{\circ}$ C. Trevigen defines 1 unit of human APE as the amount that cleaves 1 pmol of a <sup>32</sup>P-labeled apurinic/aprimidinic site oligonucleotide in 1 h at 37  $^{\circ}$ C. The reactions were stopped by the addition of equal volumes of Maxam-Gilbert loading buffer (98% formamide, 0.01 M EDTA, 1 mg/mL xylene cyanol, and 1 mg/mL bromophenol blue). The quenched reactions were electrophoresed on 20% (v/v) denaturing polyacrylamide gels, and the bands corresponding to the Mug-mediated uracil excision products were visualized and quantified using a phosphorimager (Molecular Dynamics) and the ImageQuant 5.0 software (Molecular Dynamics).

**Determination of Melting Temperatures of Oligonucleotides.** Samples of the single-stranded 11-mer oligonucleotides, previously described, were prepared by combining aliquots of the appropriate complementary single strands (Table 1) into cuvettes for a 12.5  $\mu$ M final concentration of each strand with 100 or 300  $\mu$ L (depending upon cuvette size) of the melting buffer (0.1 M NaCl, 0.01 M sodium phosphate, and 0.1 mM EDTA, pH 7.0) (28). Oligonucleotide melting temperatures ( $T_m$ ) were determined using a Varian Cary 100 Bio UV-Visible spectrophotometer. Four temperature ramps were performed on each sample per run: (1) 15 to 98  $^{\circ}$ C at a rate of 1  $^{\circ}$ C/min, (2) 98 to 15  $^{\circ}$ C at a rate of 0.5  $^{\circ}$ C/min, (3) 15 to 98  $^{\circ}$ C at a rate of 0.5  $^{\circ}$ C/min, and (4) 98 to 15  $^{\circ}$ C at a rate of 0.5  $^{\circ}$ C/min. Data were collected at 0.5  $^{\circ}$ C intervals while monitoring the temperature with a probe inserted into a cuvette containing only buffer. The oligonucleotides were allowed to anneal in ramp 2. Only ramps 2–4 were used for data analysis. Using the Cary WinUV Thermal software (Varian), the  $T_m$  of each duplex was determined using a total of five to six independent  $T_m$  measurements (Table 1) (29).

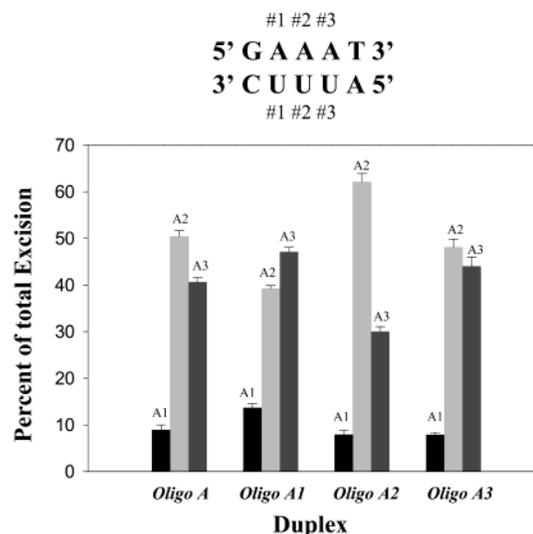
## Results and Discussion

### Oligonucleotide Synthesis and Characterization.

The sequence of the oligonucleotide duplexes used in this study is shown in Figure 2. The 30-mer was chosen based upon studies done by Taylor et al. (30) on the substrate recognition and selectivity of *EcoR124I* methyltransferase. The sequence originally used by Taylor et al. (30) was slightly modified. Adenine instead of thymine was included as the eleventh base from the 5' end (Figure 2, upper strand) as to allow for a run of three adenine residues. This sequence includes the recognition sequence for the type I restriction-modification system *EcoR124I* (31). The three tandem adenine residues are paired with three uracil residues in the duplexes used for the Mug-mediated uracil excision assays. Introduction of m6A at position A2 in the duplex (Figure 2) corresponds to the natural methylation site of *EcoR124I* methyltransferase in the 5' portion of its recognition sequence (31–32). Of the four 30-mer oligonucleotide duplexes synthesized, one duplex, **oligo A**, does not contain modifications in the run of three adenines in the upper strand (Figure 2), whereas the other three duplexes have m6A at either position A1, A2, or A3, and will be referred to as **oligo A1**, **oligo A2**, and **oligo A3**, respectively (Figure 2).

The composition of the oligonucleotides was confirmed via GC/MS. The mass spectrum for the doubly silylated m6A is shown in Figure 3. The doubly silylated m6A, with a retention time of 11.71 min, is the major silylated form observed. The singly silylated m6A has a retention time of 10.48 min and is the minor form. We observe the parent ion (*m*) for the doubly silylated m6A at 293 amu. Also observed are a major peak at 278 amu (*m*-15), and minor peaks at 73 amu (*tms*), 206 amu (*m*-N-*tms*), and 220 amu (*m*-*tms*). The observation of these peaks indicates the correct incorporation of the m6A into the oligonucleotides.

**Melting Temperature Studies.** Table 1 shows the average melting temperatures for the 11-mer duplexes. The relative melting temperatures obtained for the four 11-mer duplexes are representative of the relative melting temperatures of the corresponding 30-mer duplexes used in the Mug-induced uracil excision assays. The shorter 11-mer duplexes were used in the melting temperature studies because the effect of destabilization of one base pair on the melting temperature is more readily apparent in a shorter duplex. Melting temperatures determined for the *U/m6A1*, *U/m6A2*, and *U/m6A3* duplexes are the same within experimental error (Table 1), but the average of the three is 2.4 °C lower than the  $T_m$  determined for the *U/A* duplex, which does not contain m6A. The same trend is true for the *T/m6A1*, *T/m6A2*, and *T/m6A3* duplexes. The melting temperatures of the three duplexes containing m6A are experimentally indistinguishable (Table 1), but they are on average 3.2 °C lower than the  $T_m$  determined for the *T/A* duplex without m6A. As expected, the melting temperatures obtained for the duplexes containing three uracil residues are on average 2.6 °C lower than the corresponding *T* containing duplexes. The decreased base-stacking interactions of uracil in the duplex, as compared with thymine, decreases the duplex melting temperature (33). These melting studies show that the addition of a single m6A base to the 11-mer duplex lowers the  $T_m$  of that duplex, whether the opposite strand is the *U* or the *T* oligo



**Figure 5.** Quantification of Mug-mediated excision of uracil opposite adenine and  $N^6$ -methyladenine in 30-mer duplexes. The data shown is an average obtained from four sets of 20 min Mug-mediated uracil excision reactions on the four duplexes tested. The sequence shown includes only the central five base pairs in the 30-mer duplexes. **Oligo A** does not contain m6A, and the excision pattern seen,  $2 > 3 > 1$ , is reproducible and significant with a *p*-value of 0.018, determined using a two-tailed Friedman test. Systematic replacement of adenine with m6A in the duplex at each of the three positions significantly increases the Mug excision of uracil opposite the m6A. In **oligo A1**, **oligo A2**, and **oligo A3**, adenine was replaced with m6A at positions 1, 2, and 3, respectively. The *p*-value for the increase in uracil excision at positions 1, 2, and 3 after introduction of m6A were determined to be 0.0105, 0.0105, and 0.0415, respectively, using a one-tailed Mann–Whitney test.

(Table 1). Though we can detect the thermal instability induced by the introduction of m6A into the duplex, the decrease in  $T_m$  of the duplex is the same with m6A at each of the three positions examined. Therefore, we cannot distinguish the position of the m6A within the duplex based upon the  $T_m$  data alone.

**Effect of  $N^6$ -Methyladenine-Induced Thermal Instability on Uracil Excision by Mug.** Mug-mediated uracil excision reactions were completed on the set of four 30-mer duplexes (Figure 2). Mug-mediated uracil excision was allowed to proceed for varying lengths of time: 5, 10, 15, 20, and 25 min. Multiple data sets revealing essentially identical excision patterns for all time points were reproducibly obtained. Though the excision pattern is the same at all time points, the percent of product formation is proportional to the length of time the substrates were incubated with Mug. These data indicate that Mug will only excise one of the three possible target uracils from each 30-mer duplex. Figure 4 shows a denaturing polyacrylamide gel of 20 min Mug-mediated uracil excision reactions completed with the four duplexes. The Mug uracil excision pattern seen in this gel is representative of that seen at all experimental time points tested. Cleavage of uracil at positions 1, 2, and 3 (Figure 2) generates 21-, 20-, or 19-mer bands, respectively (Figure 4). Figure 5 shows the quantitation of four sets of data obtained from 20 min Mug-mediated uracil excision of the four duplexes.

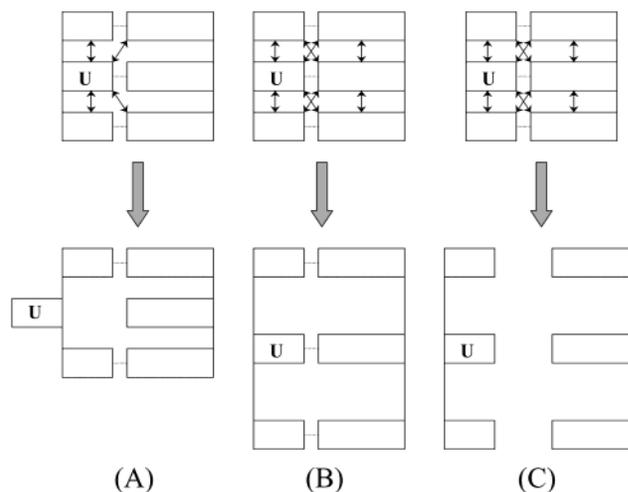
The pattern seen with Mug-mediated uracil excision of **oligo A** (no m6A) provides the baseline to which excision patterns of **oligo A1**, **oligo A2**, and **oligo A3**, with m6A at positions 1, 2, and 3, respectively, are

compared. Mug-mediated uracil excision of *oligo A* produces the reproducible pattern where the observed excision frequency is uracil at position 2 > 3 > 1. The addition of one m6A residue opposite uracil at any of the three positions alters this pattern in a reproducible manner. The new pattern resulting from substitution of one m6A residue is indicative of the position of the substitution. Without variation, the uracil opposite the m6A at any of the three positions tested is rendered more susceptible to Mug excision as compared to uracil opposite adenine within the same sequence context (Figures 4 and 5). The methyl group of m6A in its N1-cis isomer is known to disrupt pseudo-Watson–Crick base pairing (Figure 1) (20–22). Disruption of the interstrand hydrogen bonding between uracil and m6A due to the methyl group destabilizes the base pair, thereby decreasing the thermal stability of that base pair (Table 1) and increasing its susceptibility to Mug-mediated uracil excision. Although the position of the m6A within the substituted oligonucleotides cannot be distinguished from the melting studies, the Mug-mediated uracil excision pattern of *oligo A* is specifically altered by the m6A substitution at each of the three target sites. Mug excision of the uracil residue opposite m6A is selectively enhanced relative to the corresponding U:A base pair in *oligo A* (Figure 5).

**Examination of Models to Explain the Sequence-Dependence of Uracil Excision by Mug.** The mismatch-specific uracil–DNA glycosylase, Mug, was named for its reported ability to remove uracil mispaired with guanine, but not uracil paired with adenine (10, 12). The apparent mechanism for this selectivity was suggested to result from the formation of specific hydrogen-bonding interactions between the “widowed” guanine opposite the target uracil and the glycosylase (12, 13). However, the results reported here, in accord with those of Liu et al. (15) and Sung et al. (16), demonstrate that Mug cleaves uracil opposite adenine and opposite m6A. A recent study suggests that selection of the target base by Mug is more likely influenced by local helix stability rather than the specific recognition of the base opposite the target uracil residue, reflecting the energy cost of flipping the target base from the helix (15).

In the experimental system examined here, Mug can only excise one of the three uracil residues within a target duplex. The data presented in Figure 5 show that the Mug substrate selection in the duplex containing three uracil residues paired with adenine is specific, resulting in a pattern of uracil excision that is reproducible, with substantial and highly significant differences in excision frequency of the three uracils. The uracil in the central position, U2, is cleaved more frequently than U3, while U1 is the least frequently excised. It is known that Mug, like several other DNA glycosylases and DNA methylases, flips the target base from the helix (12, 34–36). It is plausible that the observed Mug-mediated uracil excision pattern (Figure 5) is related to the relative energy required to extrude a given uracil residue from the helix prior to enzymatic cleavage of the glycosidic linkage.

We considered several models to explain the observed Mug sequence-dependent uracil excision preferences as diagrammed in Figure 6. The simplest model involves disrupting the inter- and intrastrand base-stacking interactions between the target uracil residue and the adjacent base pairs, and the hydrogen-bonding interaction of the target U:A base pair as shown in Figure 6A.



**Figure 6.** Models for observed Mug sequence-dependent uracil excision preferences. Lines with double-headed arrows represent base-stacking interactions, and dotted lines represent hydrogen-bonding interactions. (A) Model of Mug-mediated uracil excision based upon base stacking and base pairing of the target uracil. In this model, disruption of the base stacking and base pairing of only the target uracil is considered in Mug uracil excision preference. According to this model, the predicted pattern of uracil excision of *oligo A* would be position 3 > 2 > 1, which is inconsistent with the observed relative excision frequency 2 > 3 > 1. (B) Model of Mug-mediated uracil excision based upon base-stacking interactions of the target base pair. This model takes into consideration the energy required to disrupt the base-stacking interactions of the target base-pair rather than simply the target uracil. According to this model, the predicted pattern of uracil excision of *oligo A* would be position 2 > 1 > 3, which is not the pattern experimentally observed. (C) Model of Mug-mediated uracil excision based upon base-stacking and base-pairing interactions of the target base and adjacent base pairs. In this model, the energy required to disrupt the hydrogen bonding as well as the energy that is required to disrupt base-stacking interactions of the target uracil and the adjacent base pairs are considered. This model predicts the relative uracil excision frequency of *oligo A* to be 2 > 3 > 1 which coincides with the observed frequency.

In this model, the magnitude of the base-stacking interactions would depend on the identity of the adjacent bases, while the disruption of the base-pairing hydrogen-bonding interaction with adenine would be sequence independent. Previously, Kollman et al. (37) calculated inter- and intrastrand base-stacking interactions for all possible nearest neighbors in DNA. Their calculations were performed with thymine rather than uracil. Although the replacement of thymine with uracil in a duplex decreases the oligonucleotide melting temperature due to comparatively decreased stacking interactions of the uracil (33), we assume that the differences in stacking interactions between uracil and thymine would be similar for the three sequence positions examined. Upon the basis of the data of Kollman et al. (37), the inter- and intrastrand base-stacking interactions for a thymine residue at positions 1, 2, and 3 (Figure 5), corresponding to the triplets 3' CT<sub>1</sub>T, TT<sub>2</sub>T, and TT<sub>3</sub>A in the sequence 3' CT<sub>1</sub>T<sub>2</sub>T<sub>3</sub>A, would be −17.0, −16.2, and −15.1 kcal/mol, respectively. The T residue demonstrating the least stacking interactions, and therefore the most easily excluded from the helix according to the model shown in Figure 6A, would be T3, followed by T2 and T1.

The observed Mug excision preference for *oligo A* is position 2 > 3 > 1 (Figure 4, lane 2), whereas the relative excision frequency predicted based upon disruption of only base stacking with adjacent bases is position 3 > 2

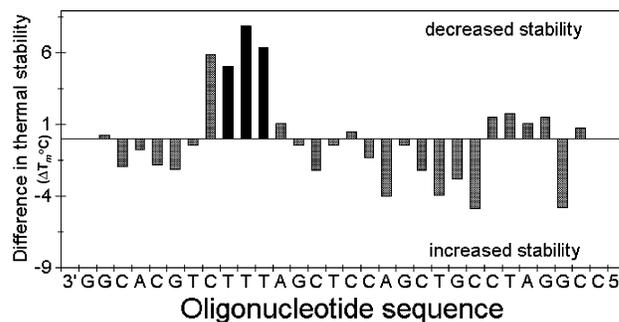
> 1. We therefore considered a model in which the stacking interactions of the target base pair, rather than simply the target base, is disrupted by the glycosylase prior to base flipping (Figure 6B). The corresponding base pair stacking interaction energies based upon the data reported by Kollman et al. (37) would be  $-39.2$ ,  $-38.2$ , and  $-40.0$  kcal/mol for positions 1, 2, and 3, respectively. The predicted relative excision frequency based upon this model, position 2 > 1 > 3, does not correspond with the observed Mug excision frequency for **oligo A**.

We then considered a third model (Figure 6C) in which both base-stacking and base-pairing interactions are disrupted for the target and adjacent base pairs. The corresponding base-stacking and hydrogen-bonding interaction energies would be  $-66.9$ ,  $-58.9$ , and  $-60.9$  kcal/mol for positions 1, 2, and 3, respectively, predicting a relative excision frequency of 2 > 3 > 1 (37). This trend coincides with the observed relative Mug-mediated uracil excision frequency of **oligo A**, suggesting that the preference of Mug for excision of a target uracil residue is inversely related to the energy required to disrupt both hydrogen-bonding and base-stacking interactions of the target and adjacent base pairs.

The variation of oligonucleotide melting temperature with sequence has been attributed to both sequence-dependent base-stacking and base-pair dependent hydrogen-bonding interactions (38–40). Several experimental data sets have been reported which have been used to interpret the observed sequence-dependent differences in thermostability. Doktycz et al. (41) measured the melting temperatures of a series of octanucleotides in which the terminal and central three base pairs were systematically altered. It was demonstrated that this data set could be used to predict the melting temperatures of octanucleotides of any sequence (41). The melting temperatures for a series of octanucleotides containing the central trimers 3'CTT, TTT, and TTA, whose central residues correspond to positions 1, 2, and 3, were reported to be 36.5, 33.7, and 35.2 °C, respectively. Upon the basis of the disruption of the trimer surrounding the target base as diagramed in Figure 6C, this relative thermostability predicts the relative Mug-mediated uracil excision frequency of **oligo A** to be position 2 > 3 > 1. This trend is consistent with that obtained from the theoretical studies discussed above for the third model.

Whereas the melting of short oligonucleotides is cooperative and fits well to a two-step denaturation model, the melting behavior of longer oligonucleotides is more complex. Local regions of lower thermostability melt with greater ease, creating denatured bubbles that then propagate throughout the sequence. We used the data of Doktycz et al. to predict the relative thermostability along the sequence of the oligonucleotide sequence utilized in this study and shown in Figure 2. The melting temperature of each overlapping duplex trimer (Table 1 in ref 41) was compared with the average duplex trimer melting temperature for the 30-base oligonucleotide duplex. Relative thermostability at each nucleotide position was calculated by taking the difference between the average and "local" thermostability. Differences in thermostability across the 30-base oligonucleotide are plotted in Figure 7. Positive values correspond to regions of lower thermostability.

The region of the oligonucleotide in which the three target uracils are located has the lowest relative thermostability within the oligonucleotide. Among the three



**Figure 7.** Thermal stability of each overlapping triplet duplex relative to the average thermal stability of the triplets within the 30-mer sequence corresponding to **oligo A**. The run of uracil residues in **oligo A** is replaced with thymine residues in the sequence seen above. The region in which the three target uracils would be located is the least thermally stable region of the oligonucleotide. Among the three target bases, indicated as solid black bars, the relative order of decreased thermal stability corresponds with the observed relative Mug excision frequency of **oligo A**, position 2 > 3 > 1.

bases in the target region, the order of relative local decreased thermostability is position 2 > 3 > 1 (Figure 7), which corresponds to the observed relative excision efficiency demonstrated by Mug at each of the three positions (Figure 5). Similar results were obtained with the experimental oligonucleotide melting temperature set of Gotoh et al. (42). While we do not wish to overinterpret our data, the observed relative cleavage of the uracil residues in the target oligonucleotide by Mug is best explained by differences in local thermostability. These data are in accord with our studies, discussed above, in which the introduction of an  $N^6$ -methyladenine residue decreases the stability of the substituent base pair, rendering that uracil residue more susceptible to Mug excision.

The model presented here is consistent with the pinch-pull-push mechanism described by Mosbaugh and co-workers for uracil-DNA glycosylase (43). Our model is similar to thumbing through a book in search of a page that will fall open due to the presence of a bookmark. In this analogy, the thickness of the bookmark would represent the degree of local instability of the target base. Mug would thumb through the helix, inducing stress to disrupt areas of the duplex with weaker base-stacking and base-pairing interactions, causing that region of the duplex to open up (Figure 6C, bottom). This mechanism of recognition would provide a conceivably simple and quick way for Mug to scan a duplex in search of its target lesion.

The significance of thermal stability as a parameter in substrate recognition for repair enzymes has been discussed previously (15, 28, 44–46). Our results confirm the importance of thermal stability as a recognition mechanism for Mug. We provide a model where Mug recognition is inversely correlated with local duplex thermostability due to the base stacking of adjacent base pairs as well as interstrand hydrogen bonding. Furthermore, the ability of Mug to target and preferentially excise uracil paired with m6A over uracil paired with adenine can be exploited as a method for locating m6A in duplexes.

**Acknowledgment.** This study was supported by NIH Grants GM50351 and CA84487.

**Note Added after ASAP.** This article posted to the Web as an ASAP article November 20, 2002. Subsequently, two corrections in the text were made. The article reposted to the Web December 16, 2002.

## References

- Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993) Oxidants, antioxidants, and the degenerative diseases of aging. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7915–7922.
- Mullaart, E., Lohman, P. H. M., Berends, F., and Vijg, J. (1990) DNA damage metabolism and aging. *Mutat. Res.* **237**, 189–210.
- Loeb, L. A. (1989) Endogenous carcinogenesis: Molecular oncology into the twenty-first century – Presidential address. *Cancer Res.* **49**, 5489–5496.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715.
- Lindahl, T., and Wood, R. D. (1999) Quality control by DNA repair. *Science* **286**, 1897–1905.
- Cleaver, J. E., Karplus, K., Kashani-Sabet, M., and Limoli, C. L. (2001) Nucleotide excision repair “a legacy of creativity”. *Mutat. Res.* **485**, 23–36.
- Hoeijmakers, J. H. J. (2001) Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374.
- Thompson, L. H., and Schild, D. (2001) Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutat. Res.* **477**, 131–153.
- Seeborg, E., Eide, L., and Bjoras, M. (1995) The base excision repair pathway. *Trends Biochem. Sci.* **20**, 391–397.
- Gallinari, P., and Jiricny, J. (1996) A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* **383**, 735–738.
- Pearl, L. H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.* **460**, 165–181.
- Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L. H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell* **92**, 117–129.
- Barrett, T. E., Scharer, O. D., Savva, R., Brown, T., Jiricny, J., Verdine, G. L., and Pearl, L. H. (1999) Crystal structure of a thwarted mismatch glycosylase DNA repair complex. *EMBO J.* **18**, 6599–6609.
- Saparbav, M., and Laval, J. (1998) 3,N<sup>4</sup>-ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine DNA-glycosylase. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8508–8513.
- Liu, P., Burdzy, A., and Sowers, L. C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG and TDG. *Chem. Res. Toxicol.* **15**, 1001–1009.
- Sung, J. S., and Mosbaugh, D. W. (2000) *Escherichia coli* double-strand uracil DNA glycosylase: involvement in uracil-mediated DNA base excision repair and stimulation of activity by endonuclease IV. *Biochemistry* **39**, 10224–10235.
- Marinus, M. G., and Morris, N. R. (1973) Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. *J. Bacteriol.* **114**, 1143–1150.
- Pukkila, P. J., Peterson, J., Herman, G., Modrich, P., and Meselson, M. (1983) Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *Escherichia coli*. *Genetics* **104**, 571–582.
- Radman, M., and Wagner, R. (1984) Effects of DNA methylation on mismatch repair, mutagenesis, and recombination in *Escherichia coli*. *Curr. Top. Microbiol. Immunol.* **108**, 23–28.
- Engel, J. D., and von Hippel, P. H. (1978) Effects of methylation on the stability of nucleic acid conformations. Studies at the polymer level. *J. Biol. Chem.* **253**, 927–934.
- Sternglanz, H., and Bugg, C. E. (1973) Conformations of N<sup>6</sup>-monosubstituted adenine derivatives. Crystal structure of N<sup>6</sup>-methyladenine. *Biochim. Biophys. Acta* **308**, 1–8.
- Quignard, E., Fazakerley, G. V., Teoule, R., Guy, A., and Guschlbauer, W. (1985) Consequences of methylation on the amino group of adenine. A proton two-dimensional NMR study of d(GGATATCC) and d(GGm<sup>6</sup>ATATCC). *Eur. J. Biochem.* **152**, 99–105.
- Ono, A., and Ueda, T. (1987) Synthesis of decaoxynucleotides containing N<sup>6</sup>-methyladenine, N<sup>4</sup>-methylcytosine, and 5-methylcytosine: recognition and cleavage by restriction endonucleases (nucleosides and nucleotides part 74). *Nucleic Acids Res.* **15**, 219–232.
- Nielsen, J., Taagaard, M., Marugg, J. E., van Boom, J. H., and Dahl, O. (1986) Application of 2-cyanoethyl-N,N,N',N'-tetraiso-propylphosphorodiamidite for *in situ* preparation of deoxyribonucleoside phosphoramidites and their use in polymer-supported synthesis of oligodeoxynucleotides. *Nucleic Acids Res.* **14**, 7391–7403.
- Gait, M. J. (1984) *Oligonucleotide Synthesis, A Practical Approach*, IRL Press, Oxford, U.K.
- Dizdaroglu, M. (1985) Application of capillary gas chromatography–mass spectrometry to chemical characterization of radiation-induced base damage of DNA: Implications for assessing DNA repair processes. *Anal. Biochem.* **144**, 593–603.
- Fujimoto, J., Tran, L., and Sowers, L. C. (1997) Synthesis and cleavage of oligodeoxynucleotides containing a 5-hydroxyuracil residue at a defined site. *Chem. Res. Toxicol.* **10**, 1254–1258.
- Hang, B., Sagi, J., and Singer, B. (1998) Correlation between sequence-dependent glycosylase repair and the thermal stability of oligonucleotide duplexes containing 1,N<sup>6</sup>-ethenoadenine. *J. Biol. Chem.* **273**, 33406–33413.
- Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601–1620.
- Taylor, I., Watts, D., and Kneale, G. (1993) Substrate recognition and selectivity in the type IC DNA modification methylase M.EcoR124I. *Nucleic Acids Res.* **21**, 4929–4935.
- Kan, N. C., Lautenberger, J. A., Edgell, M. H., and Hutchison, C. A., III (1979) The nucleotide sequence recognized by the *Escherichia coli* K12 restriction and modification enzymes. *J. Mol. Biol.* **130**, 191–209.
- Bickle, T. A. (1982) The ATP-dependent restriction endonucleases. *Nucleases*. Cold Spring Harbor Laboratory Press, Plainville, NY.
- Sowers, L. C., Shaw, B. R., and Sedwick, W. D. (1987) Base stacking and molecular polarizability: Effect of a methyl group in the 5-position of pyrimidines. *Biochem. Biophys. Res. Commun.* **148**, 790–794.
- Roberts, R. J., and Cheng, X. (1998) Base flipping. *Annu. Rev. Biochem.* **67**, 181–198.
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* **384**, 87–92.
- Stivers, J. T., Pankiewicz, K. W., and Watanabe, K. A. (1999) Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **38**, 952–963.
- Kollman, P. A., Weiner, P. K. and Dearing, A. (1981) Studies of nucleotide conformations and interactions. The relative stabilities of double-helical B-DNA sequence isomers. *Biopolymers* **20**, 2583–2621.
- Devoe, H., and Tinoco, I., Jr. (1962) The stability of helical polynucleotides: Base contributions. *J. Mol. Biol.* **4**, 500–517.
- Cröthers, D. M., and Zimm, B. H. (1964) Theory of the melting transition of synthetic polynucleotides: Evaluation of stacking free energy. *J. Mol. Biol.* **9**, 1–9.
- Petruska, J., Sowers, L. C., and Goodman, M. F. (1986) Comparison of nucleotide interactions in water, proteins, and vacuum: Model of DNA polymerase fidelity. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1559–1562.
- Doktycz, M. J., Morris, M. D., Dormady, S. J., Beattie, K. L., and Jacobson, K. B. (1995) Optical melting of 128 octamer DNA duplexes. Effects of base pair location and nearest neighbors on thermal stability. *J. Biol. Chem.* **270**, 8439–8445.
- Gotoh, O., and Tagashira, Y. (1981) Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers* **20**, 1033–1042.
- Wong, I., Lundquist, A. J., Bernards, A. S., and Mosbaugh, D. W. (2002) Presteady-state analysis of a single catalytic turnover by *Escherichia coli* uracil-DNA glycosylases reveals a “Pinch-Pull-Push” mechanism. *J. Biol. Chem.* **277**, 19424–19432.
- Sagi, J., Hang, B., and Singer, B. (1999) Sequence-dependent repair of synthetic AP sites in 15-mer and 35-mer oligonucleotides: role of thermodynamic stability imposed by neighbor bases. *Chem. Res. Toxicol.* **10**, 917–932.
- Sagi, J., Perry, A., Hang, B., and Singer, B. (2000) Differential destabilization of the DNA oligonucleotide double helix by a T:G mismatch, 3,N<sup>4</sup>-ethenocytosine, 3,N<sup>4</sup>-ethanocytosine, or an 8-(hydroxymethyl)-3,N<sup>4</sup>-ethenocytosine adduct incorporated into the same sequence contexts. *Chem. Res. Toxicol.* **13**, 839–845.
- Biswas, T., Clos, L. J., II, SantaLucia, J., Jr., Mitra, S., and Roy, R. (2002) Binding of specific DNA base-pair mismatches by N-methylpurine-DNA glycosylase and its implication in initial damage recognition. *J. Mol. Biol.* **320**, 503–513.