N^6-METHOXYADENINE-PYRIMIDINE BASE PAIRS AS SUBSTRATES FOR THE MISMATCH REPAIR SYSTEM OF *Escherichia coli*

Josef JIRICNY

Institute of Medical Radiobiology, August Forel-Strasse 7, CH-8008 Zurich, Switzerland; e-mail: jiricny@imr.unizh.ch

Received March 30, 2001
Accepted June 20, 2001

The availability of nucleoside analogues with ambiguous base-pairing properties would be of considerable value in molecular biology. We have incorporated deoxynebularine [9-(2-deoxy-β-D-ribofuranosyl)purine, P], deoxyinosine [9-(2-deoxy-β-D-ribofuranosyl)-6-hydroxypurine, I) and [9-(2-deoxy-β-D-ribofuranosyl)-6-methoxyaminopurine, ^{MeO}A] into hexadecamer oligodeoxyribonucleotides and tested their behaviour in DNA•DNA hybridisations *in vitro*, as well as in oligonucleotide-directed mutagenesis experiments *in vivo*. The results showed that P behaved as an adenine analogue in all assays. Oligonucleotide duplexes containing I/C or I/T base pairs displayed similar thermal stabilities in DNA•DNA hybridisation experiments, however, during DNA synthesis *in vitro* and *in vivo*, hypoxanthine behaved strictly as a guanine analogue. Only ^{MeO}A was truly ambiguous in all assays. The ^{1}H NMR spectrum of the nucleoside demonstrated the existence of two distinct tautomeric forms in a ratio of *ca* 8 : 2, implying that the base might pair with both C and T. Indeed, within the context of synthetic hexadecamer duplexes, ^{MeO}A/C and ^{MeO}A/T pairs brought about a similar thermal destabilisation, with the former base pair being only marginally less favoured. When used as hybridisation probes on single-stranded M13 DNA, the ^{MeO}A-containing hexadecamer oligonucleotides were shown to bind with similar efficiencies to target sequences containing either C or T opposite the analogue. Interestingly, when ^{MeO}A is in the template strand during DNA replication, the polymerase III holoenzyme of *E. coli* reads it predominantly as a G, which indicates that ^{MeO}A exists in B-DNA mostly as the *anti*-imino tautomer.

Keywords: Nucleoside analogues; Ambiguous base-pairing; N^6-Methoxyadenine; Tautomerism; *E. coli* mismatch repair system; Nucleosides; Nucleotides; DNA.

The presence of unfavoured tautomers in cellular nucleotide pools has been proposed to be one of the major factors affecting the fidelity of DNA polymerases during replication[1,2]. All four DNA bases can exist in different tautomeric forms and it is possible to draw A/C and G/T structures that closely resemble the A/T and G/C Watson–Crick base pairs. Although the tautomeric equilibria favour in all cases the amino forms of adenine and cytosine and the keto forms of guanine and thymine by more than four or-

ders of magnitude, incorporation of the less-favoured tautomers into DNA during replication could still lead to significant levels of mutagenesis when one considers the fact that the genomes being replicated have 10^6–10^9 nucleotides. However, due to the ready conversion of the unfavoured tautomers to the favoured ones, the newly-formed A^{imino}/C, A/C^{imino}, G^{enol}/T and G/T^{enol} "pairs" would revert to A/C or G/T mismatches in duplex DNA. This would result in the recognition and correction of these errors either by the proofreading 3′-5′ exonuclease activity of the respective DNA polymerase or by the postreplicative mismatch repair system. The high mutagenic potential of the unfavoured DNA base tautomers, which might predict mutation rates during DNA replication as high as 0.01% per nucleotide added, could thus be reduced to the experimentally-found levels of 10^{-9}–10^{-11} (ref.[1]).

Under normal growth conditions, it is in the interest of most organisms to try and keep the rates of spontaneous mutagenesis as low as possible. They have achieved this goal through efficient DNA polymerases, with error rates of 1 in 10^4 to 1 in 10^5 nucleotides incorporated[3], as well as through efficient proofreading mechanisms, which are associated with all replicative DNA polymerases characterised to date, and which increase the fidelity of the replication process by a further two orders of magnitude or so. In addition, polymerase errors that escape the proofreading function are addressed by postreplicative mismatch repair systems, which make a further 1 000-fold improvement to the precision with which DNA is replicated[4,5]. The mismatch repair systems are known to act also in the correction of recombination-associated mispairs, by either aborting events between sequences that are too divergent, or by correcting DNA synthesis errors that occur in the repair tracts during double-strand break repair[6,7]. Thus, although base/base mismatches bring about only very small distortions in the structure of B-DNA (ref.[8]), most organisms have evolved sophisticated machineries that remove these structures from DNA with high efficiency.

However, there are occasions where it would be advantageous to have high mutation rates. On an evolutionary time scale, it might help organisms escape selective pressures imposed on them by, for example, changes in the environment[9]. On a laboratory time scale, it might significantly facilitate a number of experiments, such as screening of DNA libraries with degenerate probes or random mutagenesis of genes. To achieve the former objective, several laboratories made use of the ambiguous properties of hypoxanthine during DNA•DNA hybridisation (ref.[10] and references therein). Unfortunately, although hypoxanthine appears to form base pairs with both cytosine and thymine, as well as, albeit to a lesser extent, with ade-

nine and guanine, it is unsuitable in mutagenesis, as in template DNA it is read by polymerases as a guanine and its monophosphate is incorporated exclusively opposite cytidine[10]. For facile random mutagenesis of genes, the ideal reagent should fulfil the following criteria: (i) it should have ambiguous base-pairing properties, both in template DNA as well as a DNA polymerase substrate, (ii) it should escape detection by the proofreading exonucleases of DNA polymerases and (iii) it should escape detection by the mismatch repair machinery. In this work we wanted to test whether the latter criteria were satisfied by N^6-methoxyadenine, which is known to exist in two distinct tautomeric forms[11] and to have miscoding properties during transcription[12].

EXPERIMENTAL

All enzymes were purchased from New England Biolabs, and were used according to the manufacturer's instructions. 2′-Deoxy-N^6-methoxydenosine, 3′,5′-di-*O*-(4-methylbenzoyl)-2′-deoxy-N^6-methoxydenosine and the 5′-*O*-(4,4-dimethoxytrityl)-3′-*O*-[*N*,*N*-diisopropylamino(methyl)phosphonyl] nucleotide building blocks of deoxynebularine [9-(2-deoxy-β-D-ribofuranosyl)purine, P], deoxyinosine [9-(2-deoxy-β-D-ribofuranosyl)-6-hydroxypurine, I] and [9-(2-deoxy-β-D-ribofuranosyl)-6-methoxyaminopurine, ^{MeO}A], were synthesised by Dr S. G. Wood using previously published procedures[13,14]. The ^{1}H NMR spectra (Fig. 1b) were recorded on a Brucker 360FT spectrometer, using $CDCl_3$, DMSO-d_6 or D_2O (Aldrich). The hexadecamer oligonucleotides were synthesised on an Applied Biosystems 380A synthesiser. The oligonucleotides were purified as described previously and, where necessary, labelled with ^{32}P (ref.[16]).

Thermal Dissociation of Oligonucleotide Duplexes

The two complementary oligonucleotide strands (0.2 A_{260} units each) in 300 μl of TES buffer (10 mM Tris-HCl pH 7.6, 6 mM $MgCl_2$, 80 mM NaCl, 0.1 mM EDTA) were annealed by heating the mixture to 80 °C (5 min) followed by cooling to 20 °C (10 min) in a microcuvette placed in the Peltier element of a Gilford 2600 spectrophotometer. The heating rate was controlled by a Gilford model 2527 thermoprogrammer. The annealing reaction was considered complete when the absorbance of the mixture (monitored at 260 nm) ceased falling. The melting of the duplexes was measured at the same wavelength, the heating rate being 1 °C/min. The T_m values shown in Fig. 2a were determined from the maxima of the first derivative profiles of the melting curves (not shown).

Dot-Blot Hybridisations

The single-stranded DNAs M13mp9 and M13mp9/op7 (1 μg each)[13,15] were applied onto the nitrocellulose membranes (Schleicher & Schuell, BA85, 10cm) in an alternating pattern shown in Fig. 3. The membranes were then baked at 80 °C *in vacuo* for 2 h, blocked with skimmed milk, washed and hybridised with the ^{32}P-labelled oligonucleotides (10 pmol) in 3 ml of 6 × SSC buffer at room temperature for 4 h, essentially as described previously[16]. Following hybridisation, the membranes were washed for 2 min in 6 × SSC (100 ml) in a plastic

Petri dish (12 cm). The temperature of the wash solution indicated in Fig. 3 was measured directly in the dish. Autoradiography was carried out at –80 °C for 20 min with one intensifying screen, using Kodak XAR-5 film.

Electrophoretic Mobility Shift Assays (EMSA) and Site-Directed Mutagenesis

These experiments were carried out as published previously[13,17], except that the hexadecamer oligonucleotide ^{MeO}A was used in addition to G, A, P and I. As no binding of the purified recombinant MutS protein from *E. coli* to the ^{MeO}A/C and ^{MeO}A/T duplexes could be detected, the EMSA experiments are not shown in the results section. The competition experiments shown in Fig. 4 were carried out as described previously[17]. In the site-directed mutagenesis, the primer extension protocol[13] was used without modifications and the mutant marker yields cited in Table I are yields after reinfection.

RESULTS AND DISCUSSION

During the course of our earlier studies, we examined the thermodynamic and biochemical characteristics of hexadecamer oligonucleotides containing deoxynebularine and deoxyinosine[13,17,18]. Because deoxynebularine was shown to behave as a deoxyadenosine analogue and deoxyinosine was shown to have ambiguous base-pairing properties in DNA•DNA hybridisations, we included these probes in the present study for comparison purposes, together with the G and A controls.

Tautomerisation of N^6-Methoxyadenosine

Due to the electron-withdrawing potential of the methoxy group, N^6-methoxyadenosine can exist in three distinct tautomeric conformations (Fig. 1a). UV, IR and ^{1}H NMR spectroscopy[11,19], as well as crystallographic studies[20] showed that the *syn*-imino form of the nucleoside predominates in aqueous solution approximately four-fold and appears to be the only isomer in the crystal. In agreement with these reports, the ^{1}H NMR spectrum of the 3′,5′-di-*O*-(4-methylbenzoyl)-N^6-methoxy-2′-deoxyadenosine shows the presence of two distinct forms of the nucleoside, as judged by the two double-doublet signals due to the anomeric (C-1′) proton (Fig. 1b). The amino/imino tautomeric ratio was found to be approximately 6 : 4 in deuteriochloroform and 2 : 8 in DMSO-d_6 solution[14]. Under similar experimental conditions only one tautomer of deoxyinosine was detected (results not shown). Deoxynebularine lacks external substituents, and can thus form no analogous tautomeric species.

Thermal Stability of Oligonucleotide Duplexes Containing Purine Analogues

We next wanted to see whether a similar situation existed also in DNA. Following the conversion of the above deoxynucleoside analogues into diisopropylamino(methyl)phosphonyl building blocks, a series of complementary hexadecamer oligodeoxyribonucleotides $^{5'}$GATCCGTC**R**ACCTCGA$^{3'}$ and $^{5'}$TCGAGGT**Y**GACGGATC$^{3'}$ [**R** = guanine (G), adenine (A), purine (P), hypoxanthine (I) and N^6-methoxyadenine (^{MeO}A); **Y** = thymine (T) or cyto-

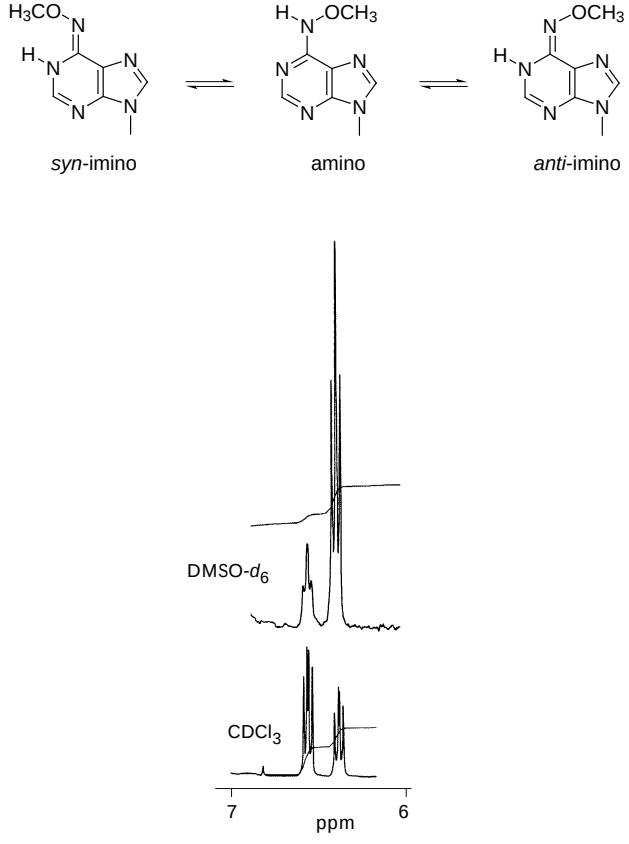

FIG. 1

Tautomeric forms of N^6-methoxyadenine (a). Relative abundance of the imino and amino tautomers as judged from the integration of the double-doublet signals due to the anomeric protons (C1′) in the ^{1}H NMR spectra of 3′,5′-di-*O*-(4-methylbenzoyl)-2′-deoxy-N^6-methoxyadenosine in DMSO-d_6 and $CDCl_3$ (b). The chemical shifts refer to the internal TMS standard

sine (C)] were synthesised and annealed[18]. The effect of the purine base analogues on the stability of the resulting homo- or heteroduplexes was first studied by thermal denaturation. Because these duplexes differ from one another solely by the composition of the R/Y base pair in their centre, the observed differences in their melting temperatures (T_m) are due predominantly to the changes in base stacking and hydrogen bonding at this site. As reported previously[18], the greatest differences in the thermal stability of the duplexes were observed between the Watson–Crick pairs G/C and A/T, and the mispair combinations G/T and A/C. In both these cases, mismatch formation affects stacking interactions and reduces the number of hydrogen bonds by one[8,21,22], which lead to a destabilisation of 7.5 and 7 °C, respectively (Fig. 2a). A similar difference (5.5 °C) was observed between the P/T and P/C structures. The former structure is stabilised by a single hydrogen bond between the N^1 of nebularine and the N^3-H of thymine (Fig. 2b), while the P/C mispair could assume two conformations: one stabilised by a single hydrogen bond between the N^1 of the nebularine and the exocyclic amino group of the cytosine, which would require rotational movement of both bases and thus result in some loss of stacking interactions, and one stabilised solely by stacking. The observed decrease in thermal stability is quite substantial and is probably commensurate with the loss of a hydrogen bond rather than just a change in stacking. The P/C mispair is thus most likely accommodated within the helix solely by stacking interactions[18] (Fig. 2b), even though the possibility that the two bases are extrahelical should also be considered[22]. In contrast, the T_m difference between the I/C and I/T

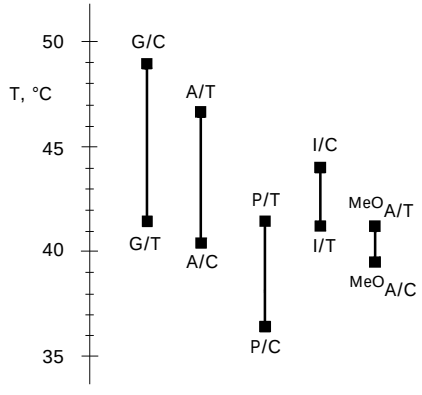

FIG. 2a
Melting temperatures of the hexadecamer oligonucleotide duplexes containing the various base pairs

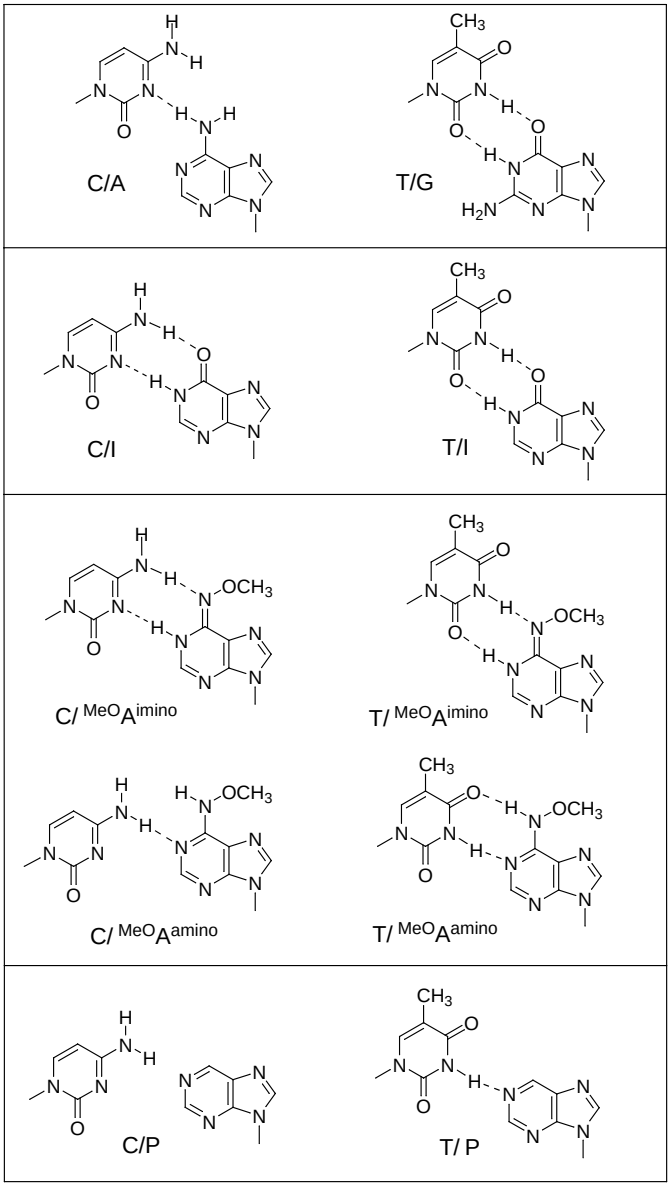

FIG. 2b
Alternative structures of the base pairs of the purine analogues with both pyrimidines. The G/T and A/C structures determined by NMR studies and from crystal structures are also shown

mispairs was substantially smaller (2.5 °C, Fig. 2a), which suggests that both these structures are stabilised by an equal member of hydrogen bonds, but that the I/T mispair is less stable due to the disruption of stacking interactions associated with the rotational movement of the thymine towards the major groove, and of the hypoxanthine into the minor groove (Fig. 2b). It is interesting to note that the T_m of the G/T and I/T duplexes were very similar (Fig. 2a). This implies that both structures are stabilised by two hydrogen bonds and equivalent stacking interactions (Fig. 2b). The exocyclic amino group of the guanine makes apparently no contribution towards the thermal stability of the G/T duplex, unlike in the G/C pair, where its ability to form a third hydrogen bond, with the O^2 of cytosine, is responsible for the 4.5 °C increase in T_m compared to I/C, assuming that the stacking interactions in G/C and I/C are equal.

The T_m difference between the ^{MeO}A/T and ^{MeO}A/C duplexes was 2 °C. Due to the existence of the stable tautomers of ^{MeO}A (Fig. 1a), we expected the amino tautomer to form a Watson–Crick-like "pair" with T, and the *anti*-imino form to form a similar "pair" with C, both stabilised by two hydrogen bonds and stacked normally within the B-DNA helix (Fig. 2b). This supposition was recently substantiated by X-ray crystallographic studies of the oligonucleotide duplexes containing these base combinations[23,24]. However, as the thermal stability of the ^{MeO}A/T and ^{MeO}A/C duplexes was lower than that of A/T or I/C, both of which are also well stacked and stabilised by two hydrogen bonds, our results imply that the hydrogen bond between the methoxyamino, respectively methoxyimino group of the purine and the exocyclic group at the N^4-position of the pyrimidine must be substantially weakened by the electron-withdrawing effect of the methoxy moiety. The pairing is probably also weakened by steric interference brought about by the presence of a substituent on the exocyclic amino group of the adenine. The possibility that the destabilisation is linked with the ^{MeO}A base being in the *syn*-imino form should also not be disregarded, but this tautomer has not been observed in DNA to date[14,23,24] and is unlikely to form base pairs with pyrimidines due to steric interference of the methoxy group (Fig. 1a).

Dot-Blot Hybridisations

The experimental evidence obtained in the above-described thermal dissociation studies was substantiated by a series of dot-blot hybridisations. The polylinker regions of the filamentous bacteriophage Ml3mp9 and Ml3mp9/op7 (ref.[13]) contain the sequences of the C and T oligonucleotides,

respectively (Fig. 5). These single-stranded DNAs were spotted onto nitrocellulose membranes to form an alternating nine-dot pattern, where the "C" DNA was spotted in the four corners and in the centre, and the "T" DNA was in the centre of each side of the square (Fig. 3). The membranes were then hybridised with the 5′-^{32}P-labelled oligonucleotides G, A, I, P and ^{MeO}A at room temperature, under which conditions all nine spots could be visualised (Fig. 3, 25 °C). However, when the filters were washed at elevated temperatures, the G, A and P probes dissociated selectively from the non-complementary DNAs, such that at 56 °C the A and P oligonucleotides remained bound solely to the four "T" DNAs, while, at 62 °C, the G oligonucleotide remained hybridised exclusively to the five "C" DNAs. In contrast, the I and ^{MeO}A oligonucleotides displayed substantially less selectivity, such that at 56 °C they still remained bound to all nine DNA spots. In agreement with the T_m data (Fig. 2a), the ^{MeO}A oligonucleotide was beginning to dissociate from the "C" DNAs at this temperature, while the I probe still displayed significant amount of binding to both phage DNAs even at 62 °C (Fig. 3). The results of this assay show that I and ^{MeO}A can form base pairs of similar thermal stability with both pyrimidines and thus that they are suitable for DNA•DNA hybridisation and PCR experiments, where the template DNA sequence is derived from a peptide sequence and therefore contains ambiguities at the wobble positions of codons. Indeed, it might prove possible to use ^{MeO}A and I for hybridisations with all four nucleotides, as Nishio *et al.* were able to show that ^{MeO}A can form relatively stable base pairs also with both purines[14], while probes containing hypoxanthine at all redundant codons were also successfully used in cloning[25].

Recognition of the Various Base Pairs by the E. coli MutS Protein in vitro

When oligonucleotides containing the ^{MeO}A base were used as templates for DNA polymerases *in vitro*, the modified purine could be shown to code as both G and A (ref.[14]). This suggested either that the base existed in two tautomeric forms also in DNA, or that the proximity of one or the other pyrimidine deoxynucleoside triphosphate within the polymerase active site could induce it to switch from one tautomer to the other. These data implied that once incorporated into DNA, the ^{MeO}A base would be mutagenic also *in vivo*, due to its ability to form both ^{MeO}A/C and ^{MeO}A/T pairs. However, its mutagenic potential *in vivo* does not depend solely on the frequency with which a polymerase reads ^{MeO}A as G or A. It depends also on the recognition of the resulting ^{MeO}A/C or ^{MeO}A/T pairs by the mismatch repair system. We decided to address this question by first testing whether

the latter base pairs were bound by the mismatch recognition factor MutS. The ^{32}P-labelled duplexes ^{MeO}A/C and ^{MeO}A/T were incubated with the purified recombinant MutS protein of *E. coli* as described previously[17], and the formation of protein/DNA complexes was examined in electrophoretic mobility shift assays. Unlike the G/T and I/T duplexes, the A/C, P/C, ^{MeO}A/C and ^{MeO}A/T substrates failed to form detectable protein/DNA complexes with the purified recombinant MutS protein in direct binding assays (ref.[19] and data not shown). In competition assays, where a 40-fold excess of the unlabelled duplexes was used to displace the radioactively-labelled G/T duplex from its complex with MutS, solely the G/T heteroduplex behaved as an efficient competitor. The A/C and I/T duplexes could also be shown to

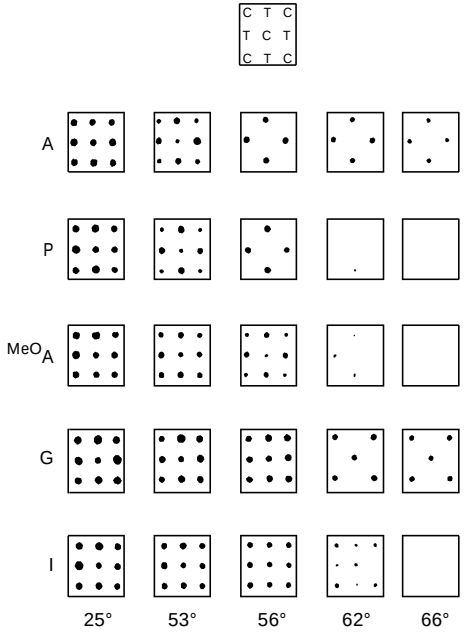

FIG. 3

Dot-blot hybridisations. The single-stranded M13mp9 (**C**) or M13mp9/op7 (**T**) DNAs were spotted on the nitrocellulose membranes in an alternating nine-pin pattern shown above the panel. The membranes were hybridised with the ^{32}P-labelled hexadecamer oligonucleotides A, P, ^{MeO}A, G, or I, and were then washed at increasing temperatures. The A and P oligonucleotides can be seen to hybridise selectively to the T phage DNA, while the G oligonucleotide formed stable hybrids solely with the C phage. The ^{MeO}A and I oligonucleotides displayed ambiguous hybridisation properties, forming hybrids of similar stability with both C and T phage DNAs. See text for details

compete, albeit much less effectively (see also ref.[19]), while only very small decreases in G/T binding were observed in the presence of a 40-fold excess of the P/C, ^{MeO}A/C and ^{MeO}A/T probes (Fig. 4). These data show that only the G/T mispair is a good substrate for the MutS protein *in vitro*, at least as measured by the EMSA assay.

Correction Efficiency of the Base-Pair Combinations in E. coli in vivo

As discussed above, the efficiency of a mutagenesis experiment depends on several criteria. (i) The ambiguous base must be able to form a stable pair with the template base. This requirement is satisfied in the case of both ^{MeO}A and I, as shown by the results of the hybridisation experiments (Figs 2a and 3). (ii) The ambiguous nucleotide must be incorporated into DNA opposite more than one partner, or it must be read ambiguously when present in template DNA. This criterion might be satisfied by ^{MeO}A, as both T and C are incorporated opposite ^{MeO}A by DNA polymerases *in vitro* (ref.[14]). In contrast, I in the form of dITP, behaves as a dGTP analogue and is read by DNA polymerases exclusively as G. (iii) The resulting base pair should es-

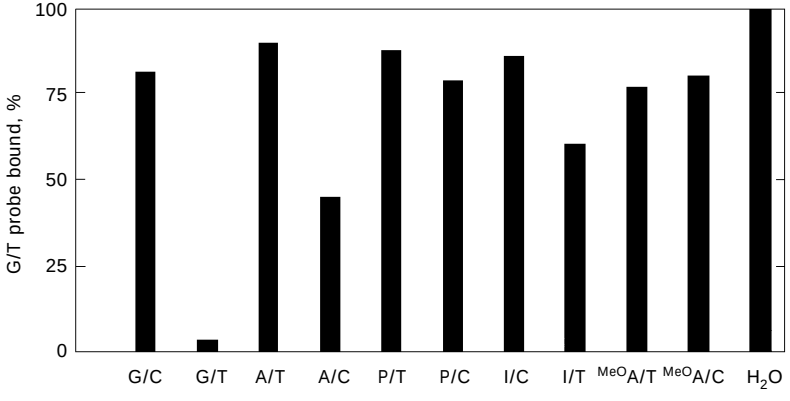

FIG. 4
Electrophoretic mobility shift competition assays. The ^{32}P-labelled oligonucleotide duplex G/T (0.6 pmol) was incubated with the purified recombinant *E. coli* MutS protein (5 pmol) in the presence of a 40-fold excess of the unlabelled duplexes **R**/**Y**. The free probe and the protein/DNA complexes were separated by electrophoresis on a 6% native polyacrylamide gel and the relative intensities of the bound and free bands were quantified using a PhosphoImager. The amount of the labelled G/T probe bound when water was used instead of the competing oligonucleotides was taken as 100%. The results are an average of three independent experiments

cape detection by the mismatch repair machinery. The results of the binding and competition EMSA assays implied that neither ^{MeO}A/C nor the ^{MeO}A/T pair is bound by the *E. coli* MutS protein, and it might therefore be assumed that they will not be addressed by the postreplicative mismatch repair system *in vivo*. However, our previous experiments have shown that the efficiency of binding of different base/base combinations by MutS does not always correlate with the efficiency of their repair *in vivo* (refs[13,17]). We therefore decided to test the repair efficiency of the ^{MeO}A/T and ^{MeO}A/C pairs in mismatch repair-proficient or -deficient *E. coli* strains JM101 or BMH71-18 *mutS*, respectively. The hexadecamer oligonucleotides anneal to the single-stranded M13mp9 DNA such that the purine analogue is positioned opposite the first base of codon 7 of the LacZ α-peptide encoded by this phage (Fig. 5). The wild type M13mp9 contains the sequence CGA (Arg) at this position, and the progeny plaques carrying this sequence are blue on suitable indicator plates. A C→T transition mutation at the first position of this codon gives rise to an opal stop codon TGA and the mutant progeny plaques, named M13mp9/op7, are thus colourless. Conversely, starting with the colourless phenotype, T→C transition mutations give rise to blue mutants[13]. The efficiency of repair can thus be scored by counting the number of mutant progeny plaques. We annealed the ^{MeO}A oligonucleotides to either M13mp9 (C), or M13mp9/op7 (T) single-stranded template DNAs and carried out primer extension reactions to generate the double-stranded heteroduplex molecules containing the desired ^{MeO}A/C and ^{MeO}A/T pairs, respectively (see also ref.[14]). In these constructs, the viral strand containing the C or T was isolated from a *dam*$^+$ bacterial strain and was therefore methylated at the adenines of GATC sequences, while the complementary, *in vitro* synthesised strand, was unmethylated. Upon entry into a repair-proficient bacterium, the binding of the MutS protein to the

	1	2	3	4	5	6	7	8	9	10	11	12
	Pro	Ser	Leu	Ala	Ala	Gly	Arg	Arg	Ile	Pro	Gly	Asn
M13mp9	5′CCA	AGC	TTG	GCT	GCA	GGT	CGA	CGG	ATC	CCC	GGG	AAT3′
13mp9/op7							TGA					
Oligonucleotides						3′A	CGT	CCA	RCT	GCC	TAG5′	

FIG. 5

Polylinker sequence of the filamentous bacteriophage M13mp9 and of its opal mutant M13mp9/op7. The position of the indicator codon 7 is shown in bold. The sequence of the hexadecamer oligonucleotides is also shown (**R** = G, A, P, I, ^{MeO}A)

mispair should initiate the assembly of the repair complex, where the MutH protein should incise specifically the unmethylated strand and direct the repair process such that the mispair would be corrected in favour of the methylated, viral, pyrimidine-containing strand. Should the ^{MeO}A/C or ^{MeO}A/T pairs escape detection by the mismatch repair system, or in a mismatch repair-deficient strain, no correction should take place and the two strands should give rise to similar amounts of progeny DNA during replication.

As shown in Table I, the control G/T mispair was corrected quite efficiently in JM101, giving rise to only 9% mutants, while the A/C mispair yielded 15% of mutants. In the mismatch repair-deficient strain, both markers were left uncorrected and thus gave rise to similar yields of the mu-

TABLE I
Mutant marker yields obtained following transfections/reinfections of mismatch repair-proficient (JM101) and -deficient (BMH71-18 *mutS*) *E. coli* strains with heteroduplexes containing the different base pair combinations

Base pair	Mutant phenotype	Yield of mutant phenotype, %	
		JM101	BMH71-18 *mutS*
G/C	colourless	0.5[a]	5[a]
A/C	colourless	15[b]	50[b]
P/C	colourless	44[b]	48[b]
I/C	colourless	1[a]	4[a]
^{MeO}A/C	colourless	3	11
G/T	blue	9[b]	64[b]
A/T	blue	0[a]	0[a]
P/T	blue	0[a]	0[a]
I/T	blue	3[b,c]	18[b,c]
^{MeO}A/T	blue	9	54

[a] These colourless mutants did not carry the desired CGA→TGA transitions, but CCC→CCCC frameshift mutations in codon 10 (Pro) of the LacZ α-peptide. It must be assumed that a similar proportion of these mutants is present also in the other assays in the CGA→TGA (blue to colourless) series. Frameshift mutants are not detectable in the TGA→CGA (colourless to blue) reversion assay. [b] See also ref.[14]. [c] Hypoxanthine is removed *in vivo* by 3-methyladenine DNA-glycosylase, which converts both I/C and I/T base pairs to G/C and thus reduces mutant marker yields.

tant marker. As discussed earlier, nebularine behaves as an adenine analogue and thus gives rise to no mutants when paired with T. In contrast, the P/C base pair is not detected by the mismatch repair factor MutS *in vitro* (Fig. 4 and ref.[19]) and the finding that it gave rise to similar mutant marker yields in both strains (Table I) implies that it escapes detection also *in vivo*. Hypoxanthine behaved as a guanine, but the I/T mispair produced only low amounts of the mutant M13 phage even in the mismatch repair-deficient strain, as hypoxanthine is efficiently removed from DNA by a DNA-glycosylase, which initiates a base excision repair process that converts the I/C and I/T pairs to G/C and A/T, respectively[26].

Neither the ^{MeO}A/T nor the ^{MeO}A/C base pair was noticeably bound by the *E. coli* MutS protein (Fig. 4) and it was anticipated that both would be left unrepaired also *in vivo*. Interestingly, in the mismatch repair-proficient strain, both heteroduplexes gave rise to mutant marker yields that were similar to those obtained with the efficiently corrected G/T and A/C substrates (Table I). The ^{MeO}A/T heteroduplex gave rise to only 9% of the mutant marker C (Table I), which implies that despite the fact that the hexadecamer heteroduplex containing the ^{MeO}A/T pair was not noticeably bound by the MutS protein *in vitro* (Fig. 4), the ^{MeO}A/T pair was efficiently addressed by the mismatch repair system in *E. coli in vivo*. This would not be unprecedented; the A/C hexadecamer was also only a poor substrate for MutS *in vitro* (see Fig. 4 and ref.[19]), yet repair of the A/C mispair in the same sequence context *in vivo* was quite efficient (see Table I and ref.[14]). However, the mutagenic potential of the ^{MeO}A/T pair *in vivo* is interesting. As shown in Fig. 2b, this pair could exist in two forms: as $^{MeO}A^{amino}$/T or as $^{MeO}A^{anti\text{-}imino}$/T. In the former structure, the base pair would be expected to be stabilised by two hydrogen bonds and stacked similarly to the A/T pair. In this conformation, it might be predicted to resemble a Watson–Crick base pair and thus escape detection by the mismatch repair system. The recently described crystal structure of an oligonucleotide duplex containing the ^{MeO}A/T pair showed that this structure was indeed in the $^{MeO}A^{amino}$/T form[23,24], which would also appear to agree with the results of our competition EMSA assay (Fig. 4), where the ^{MeO}A/T duplex was not appreciably bound by the MutS protein. In contrast, the $^{MeO}A^{anti\text{-}imino}$/T form should resemble a G/T mispair, where the thymine is slightly rotated towards the major groove and the purine is partially displaced into the minor groove of the double helix. It might be anticipated that this latter structure would be addressed by the mismatch repair system, as the methoxy group of the modified adenine would have to move further into the helix, where it might disrupt stacking. How can we reconcile the weak recognition *in vitro*

and the Watson–Crick like structure in the crystal with the apparently efficient repair *in vivo*? Let us assume that the $^{MeO}A^{amino}$/T form is indeed not recognised by the mismatch repair system, while the $^{MeO}A^{anti\text{-}imino}$/T form may be a good substrate for MutS *in vivo* (Fig. 2b). It is possible that both these forms coexist within the flexible B-DNA helix, because the base prefers to be in the imino tautomer, as indicated by the ^{1}H NMR study (Fig. 1b), and because DNA "breathes" at the temperature of the assay and is thus often single stranded, which should facilitate the tautomeric switch. In this scenario, each time the pair assumes the $^{MeO}A^{anti\text{-}imino}$/T form, it could be addressed by the mismatch repair system, and because the repair process degrades the ^{MeO}A-containing strand, resynthesis will give rise to an A/T pair. Where replication precedes mismatch correction, or in the mismatch repair-deficient strain, the progeny will originate from both strands. The T strand can give rise solely to A/T pairs and thus to colourless plaques. However, when the ^{MeO}A strand serves as a template for the replicating DNA polymerase, it could produce either ^{MeO}A/T or ^{MeO}A/C progeny. The finding that the yield of the blue (C) marker in the mismatch repair deficient BMH71-18 cells following transfection/reinfection with the ^{MeO}A/T heteroduplex was lower (54%) that that seen with the control G/T mispair (64%, see Table I), indicates that about 20% of the ^{MeO}A strand were replicated to form new ^{MeO}A/T pairs. Interestingly, this correlates well with the estimated abundance of the amino tautomer in the solution of the toluoylated 2′-deoxy-N^6-methoxyadenosine (Fig. 1b) and suggests that this tautomer is read as an adenine in the template strand by the polymerase III holoenzyme during DNA synthesis in *E. coli in vivo*.

The ^{MeO}A/C pair gave rise to 3% of mutant progeny in the mismatch repair-proficient strain. This result could be interpreted to mean that it was efficiently corrected to G/C. This may indeed be the case, however, the ^{MeO}A/C substrate was not bound by the MutS protein *in vitro* and the crystal structure of the ^{MeO}A/C pair showed it to be in a Watson–Crick conformation within B-DNA (refs[23,24]), which would predict poor recognition by the mismatch repair system *in vivo*. Thus, an alternative explanation for the low mutant marker yield could be that the ^{MeO}A/C pair, in which the purine exists predominantly in the *anti*-imino tautomer (Fig. 2b), is unrecognised by the mismatch repair machinery, but that during DNA replication, the ^{MeO}A in the template strand behaves mostly like G. The mutant progeny would thus arise solely in the rare cases where the replicating DNA polymerase read the ^{MeO}A as an A, which would presumably require a tautomeric switch from the *anti*-imino tautomer to the amino form of ^{MeO}A. That the latter scenario contributes to the mutagenicity of ^{MeO}A is wit-

nessed by the fact that the yield of the mutant marker following transfection/reinfection into the mismatch repair-deficient strain BMH71-18 was approximately 80% lower than that obtained with A/C, and more than two-fold higher than that observed for G/C. Thus, ^{MeO}A in the template strand does indeed have ambiguous base-pairing properties, even though it is read predominantly as a G due to the predominance of the imino tautomer.

Our results agree well with the findings of Brown and colleagues[27], who investigated the mutagenic properties of N^6-methoxy-2,6-diaminopurine and N^6-methoxyaminopurine 2′-deoxyribonucleosides and their triphosphates *in vitro*, as well as in a series of reporter *E. coli* strains. In PCR reactions, both triphosphates were found to produce AT→GC transition mutations at least ten-fold more frequently than GC→AT, suggesting that the analogues were incorporated preferentially opposite T and that they were read by the *Taq*-polymerase as G when in the template DNA.

CONCLUSIONS

The results presented in this study show that both hypoxanthine and N^6-methoxyadenine have ambiguous base-pairing properties during DNA•DNA hybridisation. Thus, oligonucleotides containing either I or ^{MeO}A can be used with similar efficiency as degenerate probes in the screening of cDNA libraries. In contrast, only the latter base is capable of pairing with either pyrimidine during DNA synthesis, both *in vitro* and *in vivo*. Due to the fact that N^6-methoxyadenine codes for a G approximately four times more often than for an A, it should be possible to develop protocols that would facilitate random mutagenesis of open reading frames. The availability of such a procedure would be extremely useful in studies of structure/function relationships of proteins and of protein/protein interactions.

I would like to thank Dr S. G. Wood for synthesising the N^6-methoxyadenine nucleosides, as well as all the oligonucleotides. I am also grateful to Ms D. Martin for excellent technical assistance and to Dr O. Schärer for critical reading of the manuscript.

REFERENCES

1. Fersht A. R., Knill-Jones J. W.: *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4251.
2. Topal M. D., Fresco J. R.: *Nature* **1976**, *263*, 285.
3. Kunkel T. A.: *J. Biol. Chem.* **1992**, *267*, 18251.
4. Modrich P.: *Annu. Rev. Biochem.* **1987**, *56*, 435.
5. Modrich P.: *Annu. Rev. Genet.* **1991**, *25*, 229.

6. Radman M.: *Genome* **1989**, *31*, 68.
7. Harfe B. D., Jinks-Robertson S.: *Annu. Rev. Genet.* **2000**, *34*, 359.
8. Brown T., Kennard O., Kneale G., Rabinovich D.: *Nature* **1985**, *315*, 604.
9. Taddei F., Radman M., Maynard-Smith J., Toupance B., Gouyon P. H., Godelle B.: *Nature* **1997**, *387*, 700.
10. Kamiya H., Sakaguchi T., Murata N., Fujimuro M., Miura H., Ishikawa H., Shimizu M., Inoue H., Nishimura S., Matsukage A.: *Chem. Pharm. Bull.* **1992**, *40*, 2792.
11. Stolarski R., Kierdaszuk B., Hagberg C. E., Shugar D.: *Biochemistry* **1984**, *23*, 2906.
12. Singer B., Spengler S.: *Biochemistry* **1981**, *20*, 1127.
13. Wood S. G., Ubasawa A., Martin D., Jiricny J.: *Nucleic Acids Res.* **1986**, *14*, 6591.
14. Nishio H., Ono A., Matsuda A., Ueda T.: *Nucleic Acids Res.* **1992**, *20*, 777.
15. Bruni R., Martin D., Jiricny J.: *Nucleic Acids Res.* **1988**, *16*, 4875.
16. Sambrook J., Fritsch E. F., Maniatis T.: *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York 1989.
17. Jiricny J., Su S. S., Wood S. G., Modrich P.: *Nucleic Acids Res.* **1988**, *16*, 7843.
18. Jiricny J., Wood S. G., Martin D., Ubasawa A.: *Nucleic Acids Res.* **1986**, *14*, 6579.
19. Stolarski R., Kierdaszuk B., Hagberg C. E., Shugar D.: *Biochemistry* **1987**, *26*, 4332.
20. Birnbaum G. I., Kierdaszuk B., Shugar D.: *Nucleic Acids Res.* **1984**, *12*, 2447.
21. Quignard E., Fazakerley G. V., van der Marel G., van Boom J. H., Guschlbauer W.: *Nucleic Acids Res.* **1987**, *15*, 3397.
22. Boulard Y., Cognet J. A., Gabarro-Alpa J., Le Bret M., Sowers L. C., Fazakerley G. V.: *Nucleic Acids Res.* **1992**, *20*, 1933.
23. Chatake T., Ono A., Ueno Y., Matsuda A., Takenaka A.: *J. Mol. Biol.* **1999**, *294*, 1215.
24. Chatake T., Hikima T., Ono A., Ueno Y., Matsuda A., Takenaka A.: *J. Mol. Biol.* **1999**, *294*, 1223.
25. Miura N., *et al.*: *Gene* **1985**, *38*, 271.
26. Karran P., Lindahl T.: *J. Biol. Chem.* **1978**, *253*, 5877.
27. Hill F., Williams D. E., Loakes D., Brown D. M.: *Nucleic Acids Res.* **1998**, *26*, 1144.

I was born in Prague, on the 30th April 1951. In the summer of 1969, I finished my secondary education and moved to England, where I continued with my studies. In 1972, I obtained the degree of bachelor of science (BSc) in chemistry from the University of Aston in Birmingham and moved to London, where I obtained my PhD in chemistry from the University in 1977. I then worked as a postdoctoral research assistant in the laboratory of Prof. Colin B. Reese, where I investigated diverse aspects of nucleoside synthesis. In 1979, I moved to the Imperial Cancer Research Laboratory in Lincoln's Inn Fields, London. During the following three and a half years, I made a transition from chemistry to molecular biology. My research activities focused on the synthesis of oligonucleotides containing modified or damaged bases and on the study of these substrates in diverse biological systems. This work was continued following my move, in 1983, to the Friedrich Miescher Institute in Basel, where I stayed as a Senior Group Leader until 1990. It was at the FMI where I became interested in spontaneous damage of DNA and in the mechanism of its repair. The efforts of my group concentrated primarily on the study of the repair of the hydrolytic deamination of 5-methylcytosine and of

postreplicative mismatch correction. In 1990 I moved to Rome, where I was invited to help start up a new institute of molecular biology. This institute, IRBM, is financed by the pharmaceutical industry and its focus lies principally in applied research, in the areas of viral hepatitis and gene therapy. In 1993, the link between mismatch repair defects and hereditary colon cancer was discovered. As I was unable to carry on competitive research in this field at IRBM, I decided to return to academic research. In 1996, I was elected full professor at the Medical Faculty of the University of Zürich and Director of the Institute of Medical Radiobiology. Last month I was elected full professor at the Faculty of Mathematics and Natural Sciences of the University of Zürich.

My collaborators and I are currently working on two long-term projects. (i) Investigation of the molecular mechanism of postreplicative mismatch correction in human cells and the link of mismatch repair defects to hereditary colon cancer. This study attempts to identify the missing members of the mismatch repair system in humans and to reconstitute the repairosome from the individual purified components. We are also trying to understand how mutations in mismatch repair genes cause cancer in hereditary nonpolyposis colon cancer (HNPCC) families and which signal transduction pathways are involved in signalling the presence of DNA damage to the apoptotic machinery in mismatch repair proficient and deficient cells. (ii) Identification and characterisation of novel DNA repair enzymes. This project focuses on DNA repair activities in the hyperthermophilic archaeon Pyrobaculum aerophilum. This organism lives at 100 °C and we are interested in learning how DNA damage is recognised and corrected at these high temperatures.

A more detailed description of my research activities and the list of my publications can be found on the web site of my institute (http://www.imr.unizh.ch).