Biophysical Chemistry 22 (1985) 37–51
Elsevier



BPC 00922

# BASE-PAIRING PROPERTIES OF *O*-METHYLATED BASES OF NUCLEIC ACIDS

## ENERGETIC AND STERIC CONSIDERATIONS

Andrew POHORILLE [a] and Gilda H. LOEW [b]

[a] *Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720 and* [b] *SRI International, Life Science Division, Menlo Park, CA 94025, U.S.A.*

Received 19th May 1984
Revised manuscript received 12th October 1984
Accepted 25th October 1984

*Key words: Methylated nucleic-acid base; Base-pairing; Interaction energy; Base misincorporation*

Base-pairing properties of *O*-methylated nucleic-acid bases have been systematically investigated using both semi-empirical quantum-mechanical methods and a second-order perturbation formalism. The energetic, steric and electronic properties of (a) the individual methylated bases, (b) possible base-pairs formed between *O*-methylated and normal bases, and (c) mini-helices incorporating *O*-methylated bases were calculated. Two types of base-paired complexes were obtained: Those involving classical linear hydrogen bonds, and those involving bifurcated hydrogen-donor–hydrogen-acceptor interactions. In most complexes the presence of mispairs in the helical structure of nucleic acids is expected to create a local perturbation in the structure of the helix. Even though the most stable planar configurations of the mispairs may deviate markedly from those in the regular double helix, the induced deformations in the structure of the backbone are relatively small. Internal energies and geometries of mispairs are strongly affected by the conformation of the exocyclic group of the methylated bases. Another important contribution to the stability of various base-pairing schemes comes from stacking interactions.

## 1. Introduction

Simple alkylating agents are known to be potent carcinogens and mutagens. It has been postulated that mutagenic and carcinogenic activity of these compounds is due to miscoding caused by alkylation. Alkylating agents can convert nucleic-acid bases to derivatives with modified Watson-Crick hydrogen-bonding sites. This, in turn, leads to a change in the hydrogen-bonded scheme of base-pairing. In effect, alkylated bases may direct incorporation of 'incorrect' nucleotides into DNA. Such misincorporations, if unrepaired, may constitute a mutagenic event.

The mutagenic and carcinogenic activity of methylated nucleic acid bases has been the subject of extensive studies [1–8] which established that persistence of $O^6$-alkylguanines in DNA can be correlated with carcinogenic effects in living organisms. The effect of alkylation of the bases on their pairing properties was investigated by in vitro transcription of polymers containing modified nucleotides [9–13]. Possible hydrogen-bonding schemes involving methylated bases were also studied at the level of associations of isolated bases [14,15].

In this work, we have undertaken a systematic study of the steric requirements and energetics of mispairing caused by *O*-methylated bases using the methods of quantum chemistry and computer-

Abbreviations: $O^6$-MeG, $O^6$-methylguanine; $O^2$-MeT, $O^2$-methylthymine; $O^4$-MeT, $O^4$-methylthymine; $N^4$-OHC, $N^4$-hydroxycytosine; $N^4$-OMeC, $N^4$-methoxycytosine; $C^5$-Me-$N^4$-OMeC, 5-methyl-$N^4$-methoxycytosine; A, adenine; C, cytosine; G, guanine; T, thymine; U, uracil.

aided model building. Similar types of mispairs involving unmodified bases have recently been a subject of extensive experimental and theoretical studies [16–19]. Specifically, we are interested in the following questions: (a) what are the stabilities of various possible pairing schemes involving methylated bases, (b) to what extent do hydrogen bonds participate in the formation of the planar complexes, and (c) can the modified base-pair be accommodated in the double helix without a local perturbation of its structure?

In this study, we have investigated the base-pairing properties of $O^6$-MeG, $O^2$-MeT and $O^4$-MeT. $O^2$-MeC was not included since it is very labile in polymers and its incorporation into polynucleotides usually leads to depyrimidation rather than mispairing. On the other hand, our calculations were extended to $N^4$-OMeC and its 5-methyl derivative $C^5$-Me-$N^4$-OMeC. Although these molecules, obtained by reacting cytosine and 5-methylcytosine with methoxyamine, do not qualify as alkylated promutagens, they are important models for studying misincorporation. $N^4$-OMeC and its unmethylated analog, $N^4$-OHC, can exist as a mixture of the imino and amino tautomeric forms, although the former predominates under various experimental conditions [20,21]. This tautomeric equilibrium may explain the observed dual specificity of $N^4$-OMeC and $N^4$-OHC which can substitute for U and C [22–28].

The molecules considered in this study are schematically represented in fig. 1. All of them have exocyclic substituents which can potentially rotate. The exocyclic groups can adopt either a *syn* or *anti* conformation relative to N(1) in the case of $O^6$-MeG or relative to N(3) in the case of pyrimidine derivatives. Thus, in the first step of our study, we performed detailed energy calculations of the rotations of these groups. In the next step, the energies of interaction between free methylated derivatives and normal nucleic-acid bases were evaluated and the most stable geometry of each base-pair was obtained. Finally, energy calculations were performed for various mini-helices built from two base-pairs, one of which included a methylated derivative. This step allows a more realistic representation of the steric constraints present in the polymer.

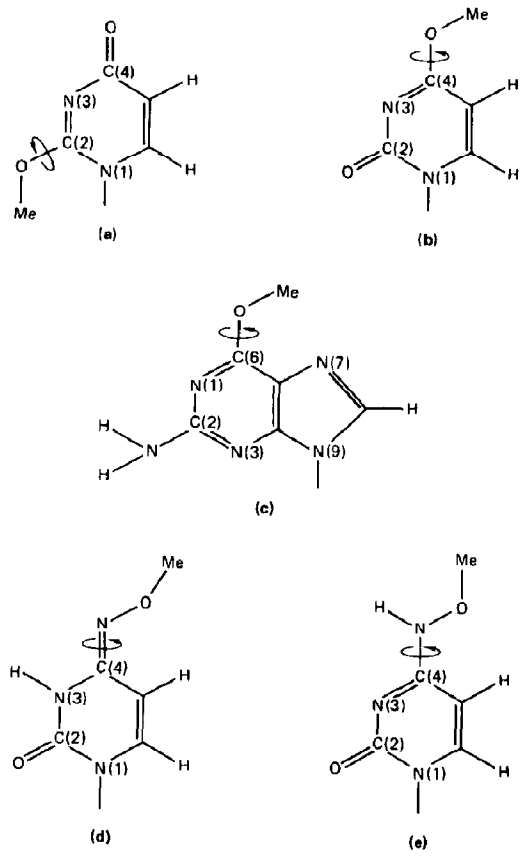


Fig. 1. Schematic representation of (a) $O^2$-MeT, (b) $O^4$-MeT, (c) $O^6$-MeG, (d) the imino form of $N^4$-OMeC, and (e) the amino form of $N^4$-OMeC. Arrows indicate rotating exocyclic groups. All bases are shown in the *anti* conformation of the exocyclic group.

Some of the pairing schemes studied in this work were previously investigated theoretically by Klopman and Ray [29] using the CNDO/2 supermolecule approach. In their work the configuration of each planar complex was kept fixed in the geometry corresponding to that in the double helix. Rotation of the exocyclic group was not allowed and steric constraints from the rest of the helix were not checked. The energies of interaction for various configurations of $O^6$-MeG–U and $O^6$-MeG–C base-pairs have also been reported by Psoda et al. [14]. That study also included preliminary calculations of the rotational barrier for

the *O*-methyl group in $O^6$-MeG. Recently, the conformation of $N^4$-OHC and its C(5)-methylated analog was studied quantum-mechanically using MNDO, MINDO/3 and PCILO methods [30].

## 2. Methods of calculation

The internal rotation of the exocyclic groups in $O^2$-MeT, $O^4$-MeT, $O^6$-MeG, $N^4$-OMeC and $C^5$-Me-$N^4$-OMeC was investigated with the aid of the semi-empirical MNDO method [31]. This method, parametrized to reproduce a variety of structural and energetical properties of many different molecules, is known to be well suited for studying conformational problems of biological molecules. For every angle of rotation around the C-O or C-N bond, complete optimization of all remaining geometrical parameters was performed. Additionally, calculations using the ab initio method with an STO-3G basis set [32] and complete geometry optimization were performed for *syn* and *anti* forms of $C^5$-Me-$N^4$-OMeC.

The energies of interaction between the bases were calculated using a perturbation method. In this approach, the energy of interaction ($E_{int}$) is represented as a sum of electrostatic ($E_{es}$), polarization ($E_{pol}$), dispersion ($E_{disp}$) and repulsion ($E_{rep}$) components

$$E_{int} = E_{es} + E_{pol} + E_{disp} + E_{rep} \tag{1}$$

The electrostatic energy was expressed in terms of interactions between multipoles placed on atoms of interacting molecules. The expansion was truncated after the third term. This leads to the atomic dipole approximation, in which $E_{es}$ is approximated by the sum of charge-charge, charge-dipole and dipole-dipole energies

$$E_{es} = E_{qq} + E_{q\mu} + E_{\mu\mu} \tag{2}$$

Although the atomic dipole approximation tends to underestimate $E_{es}$ by about 20–30%, it can usually satisfactorily predict relative stabilities of various configurations of interacting molecules [33]. Furthermore, calculated higher multipoles are markedly less accurate than atomic dipoles. In our study net atomic charges and atomic dipoles were obtained by means of the CNDO/2 method. The complete formula for the expansion of $E_{es}$ has been given elsewhere [34].

Polarization energy was calculated in the bond polarization approximation using experimental bond polarizabilities [35]. Dispersion and repulsion energies were evaluated as a sum of atom-atom terms according to the semi-empirical formula proposed by Kitaigorodskii [35] and improved by Caillet and Claverie [36].

$$E_{disp} + E_{rep} = \sum_{i<j}\sum E_{ij} \tag{3}$$

where the atom-atom contribution $E_{ij}$ is given in the form

$$E_{ij} = k_i^* k_j^* \left[ -A/z^6 + C^* \left(1 - q_i/N_i^{val}\right)^* \times \left(1 - q^j/N_j^{val}\right)^* \exp(-\alpha z) \right] \tag{4}$$

$A$, $C$ and $\alpha$ are constants independent of the atomic species and $k_i$, $k_j$ are parameters which depend on the type of atom. $N_i^{val}$ is the number of valence electrons of the atom $i$ and $q_i$ is its net charge. The value of $z$ is evaluated as

$$z = R_{ij}/R_{ij}^0 \tag{5}$$

where $R_{ij}$ is the distance between atoms $i$ and $j$ and $R_{ij}^0$ is obtained from the van der Waals radii $R_i^w$ and $R_j^w$ of atoms $i$ and $j$

$$R_{ij}^0 = \sqrt{(2^*R_i^w)^*(2^*R_j^w)} \tag{6}$$

In the cases when hydrogen bonds were formed between interacting molecules, the constants $A$, $C$ and $\alpha$ in the repulsion term were reduced. This reduction, which depends on the distance between the hydrogen and the heavy atom involved in the hydrogen bonding, takes into account the charge-transfer contribution which is attractive and also varies exponentially with the distance. The values of the constants and the formula for the correction are given by Caillet and Claverie [36].

The method employed in this study is very similar to previous approaches used extensively to evaluate the energy of interaction between nucleic-acid bases and their analogs [14,33,34,37–42]. The perturbative treatment proved to be suc-

cessful in predicting relative stabilities of various planar complexes of the bases [34,37–40], treating stacked associations [34,38–41] and analyzing pairing specificity of various rare heterocyclic bases [14,33,42].

The structures of mini-helices were obtained by optimizing the glycosidic angles and all torsional angles in the backbone. Bond lengths, bond angles and geometries of the furanose rings were kept constant. The energy of the system was calculated as a sum of all base-base stacking ($E_{stack}$) and planar ($E_{plan}$) energies of interaction, and torsional contribution from the backbone ($E_{tors}$). The barriers to rotation around various types of bonds, required to calculate torsional energy, were obtained from the works of Millner and Andersen [43] and Govil [44]. Atom-atom contributions involving atoms of the sugar-phosphate backbone were taken into account only in the form of hard sphere interactions. This approximation was necessary since a reliable value of the parameter $k_i$ for phosphorus is not available. It is also quite difficult to represent correctly the electrostatic energy for polyionic molecules.

It should be noted that the energies of mini-helices usually have many local minima which often correspond to conformations considerably different from the standard double-helical structures. Thus, in the case of helices with mismatched pairs, it was of a great importance to obtain a reasonable starting point for minimization. Such an initial configuration was required to be similar to the regular DNA or RNA structure and, at the same time, not to involve intra- or interstrand steric hindrances. This was achieved with the aid of model building and the computer program AIMS [45] coupled with the Evans and Sutherland computer graphics system.

## 3. Results

### *3.1. Rotation of exocyclic groups in the isolated bases*

Rotation of exocyclic groups in methylated nucleic-acid bases can influence the base-pairing scheme in two ways. (1) If the exocyclic group can participate in hydrogen bonding. Such interactions are permitted only if the substituent is oriented *anti* to the Watson-Crick side. (2) In other instances, rotation of the exocyclic group can influence base-pairing by creating steric hindrance between interacting molecules.

#### *3.1.1. $O^2$-MeT and $O^4$-MeT*

The MNDO calculations show that for $O^2$-MeT and $O^4$-MeT the *syn* form, in which the O-Me group is directed towards N(3), is more stable than the *anti* form by 8.8 and 11.1 kcal/mol, respectively. The *syn* form was also found in NMR [46] and crystal structure studies [47] of $O^4$-MeT. The dependence of energy on the angle of rotation is shown in fig. 2. It is worth noting that the energy curves possess only one minimum, for the *syn* conformation, while the *anti* form corresponds to the maximum energy.

#### *3.1.2. $O^6$-MeG*

As shown in fig. 3, the energy as a function of rotational angle C(6)-O exhibits two minima corresponding to the conformers *syn* and *anti* relative to the N(1) of the ring. The *syn* form is favored by 2.7 kcal/mol with the relatively low barrier to rotation equal to 5.6 kcal/mol. These results are in qualitative agreement with previous PCILO calculations [14], in which only the planar angle C(6)-O-$CH_3$ was optimized. The *syn* form was found to be

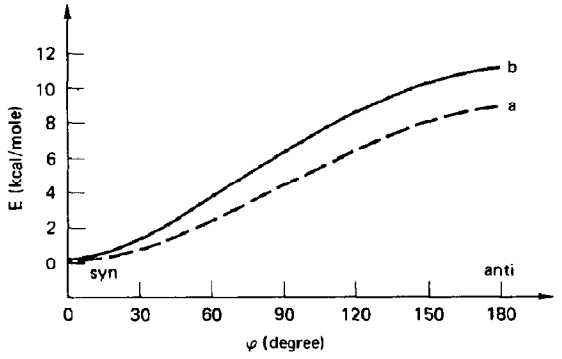


Fig. 2. Energy of rotation of the methyl group around the C(2)-O bond in $O^2$-MeT (curve a) and around the C(4)-O bond in $O^4$-MeT (curve b) as a function of the rotational angle $\varphi$. Minimum energy taken as zero.

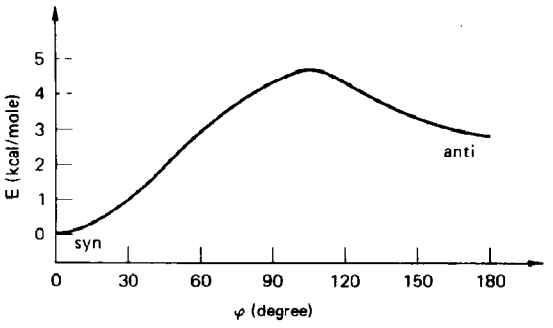


Fig. 3. Energy of rotation of the methyl group around the C(6)-O bond in $O^6$-MeG as a function of the rotational angle φ. Minimum energy taken as zero.

preferred by 1.8 kcal/mol and the barrier to internal rotation was about 2.5 kcal/mol.

### 3.1.3. $N^4$-OMeC

In both the imino and amino tautomers the *syn* and *anti* conformers are very close in energy. As shown in fig. 4, the *syn* conformer is favored by 0.8 kcal/mol in the imino tautomer and the *anti* conformer by 0.9 kcal/mol in the amino form.

In the amino form the barrier to rotation is only 2.0 kcal/mol. In the imino form, however, when rotation around the double C(4) = N bond is required, the barrier is 40 kcal/mol. Such a high barrier obviously precludes facile rotational isomerization. The *syn-anti* interconversion in the imino form could, however, occur by a more probable three-step process: tautomerization to the amino form followed by rotation of the exocyclic group in this form and tautomerization back to the imino form. Our results indicate that the second step is energetically easily accessible and the third step should occur spontaneously. In addition, since 10% of 1-methyl-$N^4$-hydroxycytosine molecules were found to exist in the amino form in aqueous medium [20], it can be anticipated that the energy difference between the imino and amino forms is not larger than a few kcal/mol. Thus, the first step should be also relatively probable. Other mechanisms of the *syn-anti* transition involving deformation of the geometry of the exocyclic group and the change of hybridization on $N^4$ are also possible.

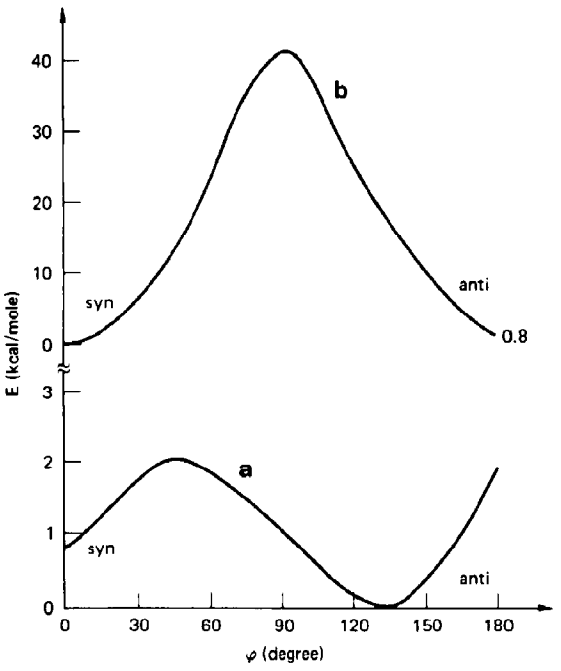


Fig. 4. Energy of rotation of the *O*-methyl group around the C(4)-N bond in the amino (curve a) and the imino (curve b) form of $N^4$-OMeC as a function of the rotational angle φ. Minimum energy taken as zero.

The crystal structures of 1-methyl-$N^4$-hydroxycytosine hydrochloride [48] and 1,5-dimethyl-$N^4$-hydroxycytosine [21] have been found to exist in the imino form with the exocyclic group in the *syn* conformation. Furthermore, it was shown that in the latter molecule simple rotation around the C(4)-N bond by 180° creates significant interference between the exocyclic hydroxyl group and the 5-methyl substituent. On this basis, it was inferred that this molecule cannot exist in the *anti* form [21]. The significance of the *syn* conformation was further stressed by the recent experimental findings [15] that in aqueous solution 1-methyl-$N^4$-methoxycytosine interacts with 2′,3′,5′-tri-*O*-methyladenosine forming a reverse Watson-Crick pair, in which the $O^2$ rather than the $N^4$ is engaged in hydrogen bonding. Such a scheme facilitates participation of the *syn* form of $N^4$-OMeC in base-pair formation. A similar association constant was found for the interactions between 9-

substituted adenine and 1,5-dimethyl-$N^4$-methoxycytosine, suggesting an identical base-pairing scheme with the exocyclic group of $C^5$-Me-$N^4$-OMeC *syn* to the ring N(3). $N^4$-OMeC and its 5-methyl substituent also exhibit similar behavior in polymers. Recently, similar mispairing was found in in vitro transcription experiments for matrices containing residues of $N^4$-OMeC and $C^5$-Me-$N^4$-OMeC [49]. Consequently, if the assumption is made that the 5-methyl substituent constrains the exocyclic group to the *syn* form [21], it would seem to follow that both bases would form the reverse Watson-Crick pairs in polymers. This pairing scheme, however, would be energetically highly unfavorable, because it induces strained conformation of the polynucleotide backbone.

To explain this apparent contradiction, we performed calculations for the *syn-anti* interconversion in the imino form of $C^5$-Me-$N^4$-OMeC allowing geometry relaxation. The results showed that the transition to the *anti* form is accompanied by such a deformation of the planar angle C(5)-C(4)-N(C4) that steric hindrance is avoided. In MNDO calculations the difference between *syn* and *anti* conformers was only 3.2 kcal/mol and the planar angle in the *anti* form was 131°. A similar picture emerges from STO-3G calculations, where the energy difference is 5.0 kcal/mol with a planar angle of 128°. The energy differences between the *syn* and *anti* conformers of the imino tautomer of 5-methyl-$N^4$-hydroxycytosine obtained from MINDO/3 and PCILO calculations with the partial geometry optimization [30] are also of the same order. Thus, the *anti* form cannot be precluded from participation in base-pair formation in polymers.

### 3.2. Interactions of methylated derivatives with the nucleic-acid bases

The energetics of planar interactions between T, A, G, or C and each methylated derivative under consideration was studied in two steps. In the first step single point calculations in which interacting bases were placed in the geometrical arrangement corresponding to that in the double helix were performed. This step provides information about the energetics of planar base-base interactions when practically no changes in the conformation of the helical backbone are involved. In the second step, starting from the double-helical geometry, the nearest local minimum in the energy of interaction was found. This step allows for the deviations from the double helical geometry.

#### 3.2.1. $O^2$-MeT and $O^4$-MeT

Energies of interaction were calculated for base-pairing between $O^2$-MeT and $O^4$-MeT in the *syn* and *anti* form with the four bases T, A, G and C. It should be noted that the *O*-methylated thymines have the unique property of possessing three hydrogen-bond acceptors ($O^2$, N(3), and $O^4$) and no hydrogen-bond donors. Consequently, they are capable of forming two hydrogen bonds only with G, since all other nucleic-acid bases have only one hydrogen-bond donor.

When the base-pairs are in the geometry corresponding to the regular double-helix, significant negative energy of interaction was obtained only for the $O^4$-MeT(*anti*)–G pair. Some stabilization was also found for the $O^2$-MeT(*syn*)–C and $O^2$-MeT(*anti*)–G pairs (see table 1). However, in the latter case, the presence of the pure *anti* conformer of $O^2$-MeT in DNA or RNA is not possible because of the short contact between the $O^2$-methyl group and the H(1′) of the ribose ring. This short contact can be removed without significant defor-

Table 1

Energies of interaction between modified bases and normal nucleic-acid bases in the double-helical configuration (in kcal/mol) [a]

| | A | T | G | C |
|---|---|---|---|---|
| $O^2$-MeT(*anti*) | 1.8 | 1.8 | −3.4 | −0.5 |
| $O^2$-MeT(*syn*) | 37.7 | 1.5 | rep [b] | −3.8 |
| $O^4$-MeT(*anti*) | 10.2 | −0.3 | −13.0 | 3.4 |
| $O^4$-MeT(*syn*) | rep | 3.2 | rep | rep |
| $O^6$-MeG(*anti*) | rep | −5.0 | | 8.6 |
| $N^4$-OMeC(amino,*anti*) | | | −20.6 | |
| $N^4$-OMeC(imino,*anti*) | −3.8 | | | |
| A | rep | −7.3 | rep | rep |
| G | rep | rep | rep | −19.5 |

[a] Energies of interaction between normal purine and pyrimidine bases are given for comparison.

[b] Energies of interaction are larger than 100 kcal/mol.

mation of the double helix only by rotation of the $O^2$-methyl group by at least 60°. Other base-pairs are either sterically prohibited or do not lead to stable configurations because the gain in energy due to the formation of the hydrogen bond is offset by the charge-charge and/or dipole-dipole repulsion. For example, $O^2$-MeT and $O^4$-MeT complexes with T are not stable because of unfavorable $O^2$-$O^2$ or $O^4$-$O^4$ interactions.

When the geometrical constraints of the double helix were relaxed, negative energies of interaction were obtained for all pairing schemes, as can be seen in table 2. The resulting geometrical arrangements, however, differ markedly from those in the double helix (see table 3). As expected, the most stable complexes are formed with G. These base-pairs are stabilized by two nonlinear hydrogen bonds and have energies of interaction in the range of −16.2 to −13.6 kcal/mol. As shown in fig. 5, to avoid steric interference with the bulky *O*-methyl group of $O^2$-MeT in the *syn* form, the guanine molecule shifts away from the $O^2$-Me group, and its two hydrogen-donor groups form a bifurcated hydrogen bond with $O^4$. For the complexes with C and A the orientation of the hydrogen-bond donor of the nucleic-acid base is such that instead of forming a linear hydrogen bond, it interacts simultaneously with two acceptors of the *O*-methylthymines. An example of such interactions with C is shown in fig. 6. In some instances these complexes have energies of interaction comparable to those for the A–T base-pair. For the base-pairs with T, the minimum energy configurations correspond to the formation of a single, nearly linear hydrogen bond. In $O^4$-MeT–T complexes this bond is, however, different for the *syn*

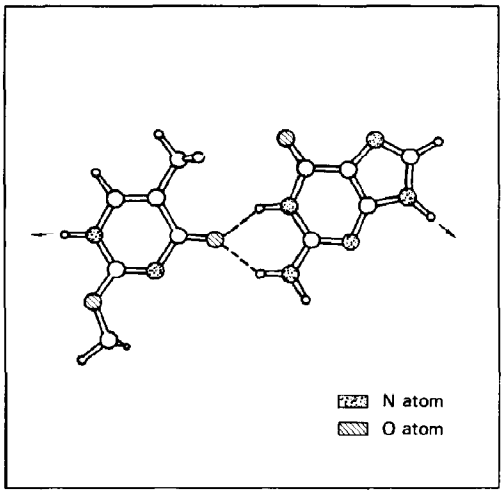


Fig. 5. The minimum energy configuration for the $O^2$-MeT(*syn*)–G base-pair.

Table 2

Energies of interaction between modified bases and normal nucleic-acid bases in the minimum-energy configuration (in kcal/mol) [a]

| | A | T | G | C |
|---|---|---|---|---|
| $O^2$-MeT(*anti*) | −4.1 | −7.9 | −16.2 | −7.6 |
| $O^2$-MeT(*syn*) | −6.5 | −8.1 | −15.8 | −10.6 |
| $O^4$-MeT(*anti*) | −4.0 | −2.2 | −15.1 | −7.9 |
| $O^4$-MeT(*syn*) | −4.2 | −5.6 [b] | −13.6 | −9.5 [b] |
| $O^6$-MeG(*anti*) | −6.6 | −7.9 | | −9.2 |
| $N^4$-OMeC(amino,*anti*) | | | −21.9 | |
| $N^4$-OMeC(imino,*anti*) | −6.0 | | | |
| A | | −8.4 | | −3.4 |
| G | | −14.7 | | −20.8 |

[a] Energies of interaction between normal purine and pyrimidine bases are given for comparison.

[b] Minimum energy was not found; minimization procedure stopped at the predefined limits of geometrical parameters.

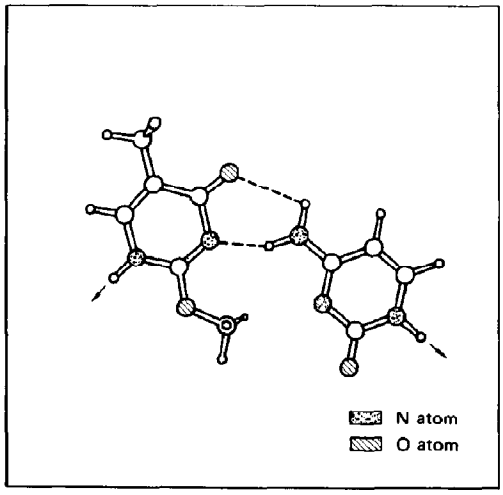


Fig. 6. The minimum energy configuration for the $O^2$-MeT(*syn*)–C base-pair.

Table 3

Geometrical parameters [a] $\varphi_1$, $\varphi_2$ (°) and $l$ (Å) for minimum-energy configurations between modified bases and normal nucleic-acid bases [b]

| | A | | | T | | | G | | | C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\varphi_1$ | $\varphi_2$ | $l$ | $\varphi_1$ | $\varphi_2$ | $l$ | $\varphi_1$ | $\varphi_2$ | $l$ | $\varphi_1$ | $\varphi_2$ | $l$ |
| $O^2$-MeT(*anti*) | 126 | 154 | 9.5 | 108 | 120 | 7.4 | 190 | 132 | 10.5 | 122 | 158 | 8.6 |
| $O^2$-MeT(*syn*) | 134 | 156 | 9.7 | 176 | 138 | 9.0 | 187 | 132 | 10.5 | 122 | 152 | 8.5 |
| $O^4$-MeT(*anti*) | 155 | 180 | 9.7 | 127 | 118 | 7.6 | 136 | 113 | 9.5 | 141 | 196 | 8.7 |
| $O^4$-MeT(*syn*) | 136 | 183 | 8.8 | 90 | 135 | 7.3 | 105 | 130 | 8.9 | 95 | 210 | 8.1 |
| $O^6$-MeG(*anti*) | 126 | 151 | 11.3 | 132 | 122 | 9.0 | | | | 116 | 144 | 9.8 |
| $N^4$-OMeC(amino*anti*) | | | | | | | 125 | 130 | 8.9 | | | |
| $N^4$-OMeC(imino*anti*) | 124 | 127 | 8.9 | | | | | | | | | |
| A | | | | 125 | 127 | 8.8 | | | | 110 | 140 | 9.5 |
| G | | | | 150 | 110 | 8.8 | | | | 132 | 125 | 8.9 |

[a] $\varphi_1$ is a planar angle between glycosidic bond in the modified base and the line joining glycosidic nitrogens, $\varphi_2$ is a planar angle between glycosidic bond in the normal base and the line joining glycosidic nitrogens, and $l$ is a distance between glycosidic nitrogens.

[b] For regular DNA double helix $\varphi_1 = \varphi_2 = 129°$, $l = 9.04$ Å.

and *anti* forms of $O^4$-MeT (see fig. 7a, b). It is worth noting (table 2) that *O*-methylthymines prefer to form pyrimidine-pyrimidine base-pairs with T or C than to interact with A, which is the natural counterpart of the unmodified base.

### 3.2.2. $O^6$-MeG

Energies of interaction were calculated for base-pairing between $O^6$-MeG with the exocyclic group *anti* to the N(1), and C, T or A. The $O^6$-MeG–G base-pair was omitted because it is clearly repulsive in orientations corresponding to the Watson-Crick base-pairing. No significant stabilization is obtained if one of these bases is rotated to the *syn* conformation around the glycosidic bond.

The *syn* form around the C(6)-O(6) bond was not taken into account since the steric hindrance of the *O*-methyl group does not allow Watson-Crick base-pairing. Involvement of the *anti* form

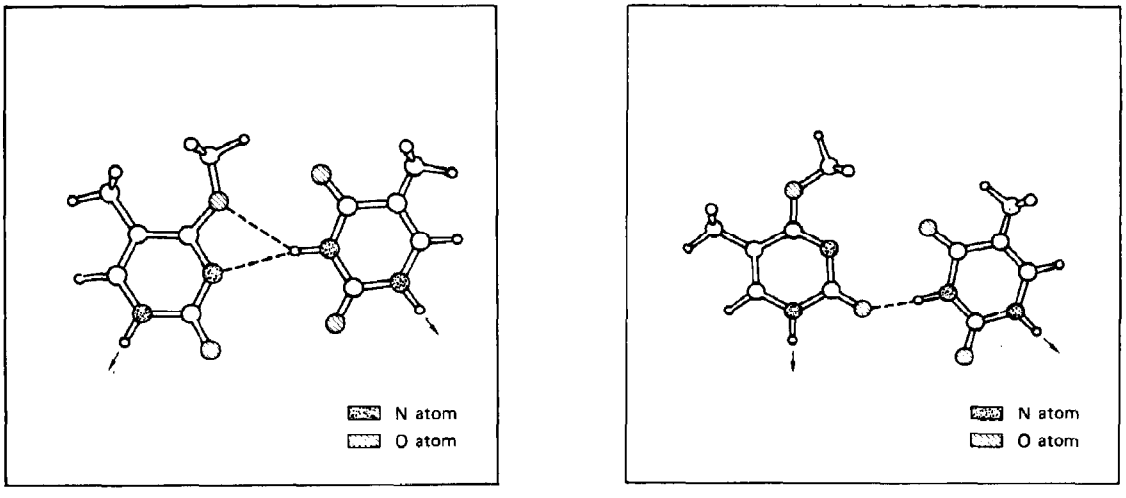


Fig. 7. The minimum energy configuration for the (a) $O^4$-MeT(*syn*)–T and (b) $O^4$-MeT(*anti*)–T base-pair.

in base-pairing is a reasonable assumption, since this form was found to be a local minimum with an energy only somewhat greater than the *syn* conformation (1.8 kcal/mol). In the double-helical configuration, $O^6$-MeG–T interactions are attractive by −5.0 kcal/mol, while $O^6$-MeG–C interactions are repulsive. This steric repulsion is removed in the minimum energy configuration, by increasing the glycosidic N(9)-N(1) distance by 0.8 Å (see table 3). As shown in fig. 8b, the orientation of the interacting bases changes in such a way that two N...H-N hydrogen bonds are formed and the electrostatic repulsion between N(1) of $O^6$-MeG and N(3) of C (fig. 8c) is eliminated. As seen from table 2, the energies of interaction of $O^6$-MeG with C and T in the minimum energy configurations are equal to −9.2 and −7.9 kcal/mol, respectively. Thus, methylation of $O^6$ of G reduces its minimal energy of interaction with C by 11.6 kcal/mol and with T by 6.8 kcal/mol. This reduction is primarily due to the repulsion between negatively charged oxygens or nitrogens of interacting molecules and the loss of one hydrogen bond in the pair with C.

The $O^6$-MeG–A base-pair is sterically not allowed in the regular double-helical configuration (table 1). When this constraint is relaxed (table 2), $O^6$-MeG and A form hydrogen bonds which are analogous to those in the $O^6$-MeG–C base-pair. The corresponding energy of interaction is, however, smaller and the separation between glycosidic nitrogens increases to 11.3 Å (table 3). In this case additional calculations were performed for configurations corresponding to adenine being in the *syn* conformation around the glycosidic bond. It was found that rotation of adenine to the *syn* conformation only partially removes the steric hindrance in the regular double-helical geometry and this geometry remains repulsive by 4.2 kcal/mol. The corresponding unconstrained lowest energy configuration is attractive by 6.9 kcal/mol with the distance between glycosidic nitrogens equal to 9.8 Å.

### 3.2.3. $N^4$-OMeC

Energies of interaction were calculated for base-pairing between $N^4$-OMeC (amino, *anti*) and G, and between $N^4$-OMeC (imino, *anti*) and A. As seen from table 3, these base-pairs fit very well into the regular double-helical structure. Their optimized geometries, shown in fig. 9, appear to be very close to the regular double-helical geometry. Corresponding energies of interaction are given in tables 1 and 2, and compared with those for C-G and T-A, respectively.

## 3.3. Conformations and energies of mini-helices containing modified bases

To evaluate the stability of base-pairs formed with the glycosidic bonds oriented differently than in a regular double helix, we have taken into account the constraints imposed by the sugar-phosphate backbone. A modified pair should join

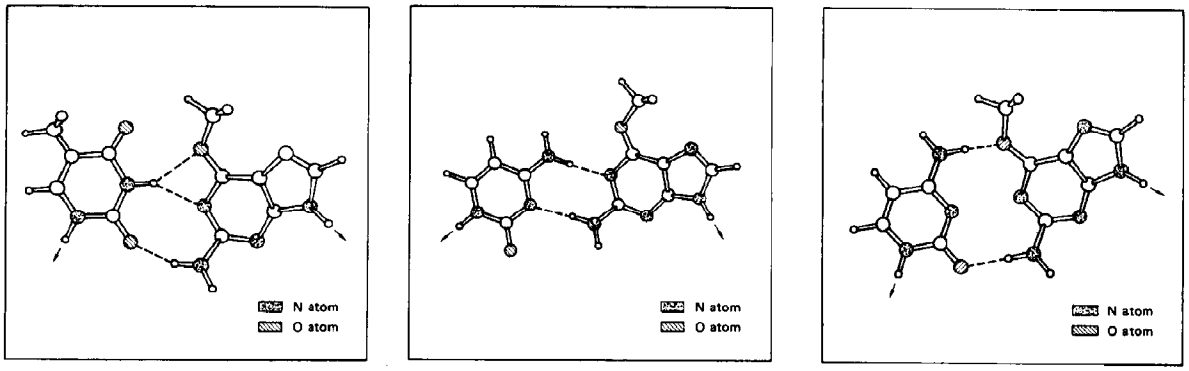


Fig. 8. The minimum energy configuration for the (a) $O^6$-MeG(*anti*)–T base-pair, (b) the minimum energy configuration for the $O^6$-MeG(*anti*)–C base-pair, and (c) the double-helical configuration of the $O^6$-MeG(*anti*)–C base-pair.

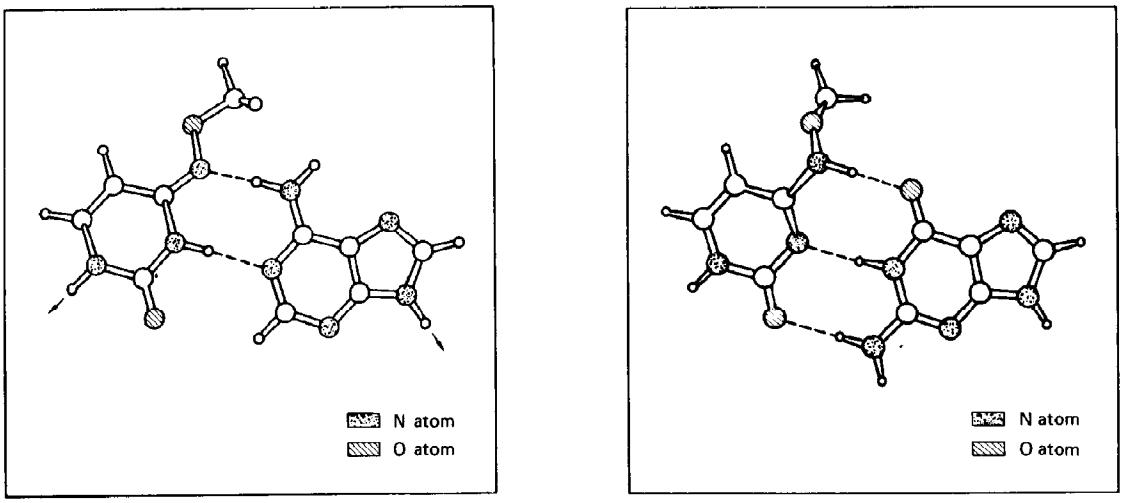


Fig. 9. The minimum energy configuration for (a) the $N^4$-OMeC(imino,*anti*)-A, and (b) $N^4$-OMeC(amino,*anti*)-G base-pairs.

neighboring pairs without creating steric hindrances or large strains in the backbone. This problem is particularly relevant for templates which contain *O*-methylthymines, since these molecules prefer to interact with unmodified bases in configurations which do not fit in the regular double helix. Calculations were therefore carried out for mini-helices, i.e., fragments of DNA containing two base-pairs. One of them was a typical Watson-Crick base-pair. In the second base-pair, *O*-methylthymine was at the 5'-end of one strand and an unmodified base was at the 3'-end of the other strand. This simple model simulates the situation in which an unmodified base would be incorporated opposite to a thymine derivative. The normal base-pairs were chosen in such a way that T (or U), C or A were, each in turn, placed in the same strand as the modified base. Such a choice of position makes the results relevant to the stacking patterns arising during transcription of templates of poly(U), poly(C) and poly(A) containing methylated thymines.

The optimized energies of mini-helices containing $O^2$-MeT, built of deoxyribonucleotides are listed in table 4. The corresponding torsional angles in the backbone are given in table 5. Similar calculations were also performed for mini-helices built of ribonucleotides and containing $O^2$-methyluracil. In both cases planar geometries of the base-pairs after optimization differ slightly from those obtained without constraints but these modifications are not associated with significant losses of planar energy of interaction.

The presence of the modified base also influences stacking interactions. Substitution of *O*-methylthymine for the normal base in the polynucleotide chain can change the stacking energy by 1–6 kcal/mol. The magnitude of this change depends on the base incorporated opposite to the thymine derivative and on the neighboring pair.

Table 4

Energies (in kcal/mol) of mini-helices built of deoxyribonucleotides and containing $O^2$-MeT [a] and T-A base-pair

Energy of T-A/T-A mini-helix is given for comparison.

| | $E_{tot}$ | $E_{plan}$ [b] | $E_{stack}$ | $E_{tors}$ [c] |
|---|---|---|---|---|
| $\downarrow {}^{T\text{-}A}_{O^2\text{-MeT-A}} \uparrow$ | −27.4 | −13.3 | −14.7 | 0.6 |
| $\downarrow {}^{T\text{-}A}_{O^2\text{-MeT-T}} \uparrow$ | −32.1 | −14.1 | −18.8 | 0.8 |
| $\downarrow {}^{T\text{-}A}_{O^2\text{-MeT-G}} \uparrow$ | −35.8 | −21.8 | −14.3 | 0.3 |
| $\downarrow {}^{T\text{-}A}_{O^2\text{-MeT-C}} \uparrow$ | −34.5 | −17.1 | −18.6 | 1.2 |
| $\downarrow {}^{T\text{-}A}_{T\text{-}A} \uparrow$ | −29.4 | −14.6 | −14.8 | 0. |

[a] The exocyclic $O^2$-methyl group of $O^2$-MeT is in the minimum energy (*syn*) conformation.

[b] Planar energy of the T-A base pair equals −7.3 kcal/mol.

[c] Torsional energy of the T-A/T-A mini-helix taken as zero.

Table 5

Backbone and glycosidic torsion angles [a] (°) in mini-helices built of deoxyribonucleotides and containing $O^2$-MeT [b]

Values of corresponding angles in B-DNA are given for comparison.

| | $\downarrow O^2\text{-MeT}\overset{\text{T-A}}{\text{-A}}\uparrow$ | $\downarrow O^2\text{-MeT}\overset{\text{T-A}}{\text{-T}}\uparrow$ | $\downarrow O^2\text{-MeT}\overset{\text{T-A}}{\text{-G}}\uparrow$ | $\downarrow O^2\text{-MeT}\overset{\text{T-A}}{\text{-C}}\uparrow$ | B-DNA [c] |
|---|---|---|---|---|---|
| $\Phi$C4′-C5′ | | | | | |
| strand | 35 | 38 | 32 | 46 | 27–58 |
| strand | 46 | 35 | 80 | 45 | |
| $\Phi$C5′-05′ | | | | | |
| strand | 203 | 183 | 194 | 150 | 160–214 |
| strand | 161 | 170 | 191 | 170 | |
| $\Phi$05′-P | | | | | |
| strand | 332 | 340 | 313 | 323 | 281–335 |
| strand | 311 | 328 | 268 | 314 | |
| $\Phi$P-03′ | | | | | |
| strand | 267 | 260 | 270 | 252 | 224–281 |
| strand | 260 | 261 | 279 | 259 | |
| $\Phi$03′-C3′ | | | | | |
| strand | 160 | 171 | 203 | 206 | 147–194 |
| strand | 180 | 172 | 180 | 173 | |
| $\chi_u^d$ | | | | | |
| strand | 85 | 78 | 82 | 82 | 62–72 |
| strand | 88 | 72 | 57 | 80 | |
| $\chi_l^d$ | | | | | |
| strand | 88 | 52 | 40 | 62 | 62–72 |
| strand | 72 | 61 | 85 | 70 | |

[a] For definition of torsion angles see ref. 51.

[b] The exocyclic $O^2$-methyl group of $O^2$-MeT is in the minimum energy (*syn* conformation).

[c] Compiled from refs. 51–55.

[d] $\chi_u$ are glycosidic torsional angles in the T-A base-pair, and $\chi_l$ are glycosidic angles in the base-pair containing $O^2$-MeT.

## 4. Discussion

The results of our calculations lead to a description of the nature of the hydrogen bonds formed between normal nucleic-acid bases and their alkylated derivatives, and allow the estimation of the relative energetics and steric requirements of mispairs involving modified bases.

Modified nucleic-acid bases studied in this work can be divided into two groups according to the nature of their base-pairing interactions with normal bases. The bases from the first group ($O^6$-MeG and $N^4$-OMeC) are characterized by their ability to form at least two well-defined linear hydrogen bonds with modified bases. Any significant change in the geometry of such hydrogen-bonded base-pairs results in a clear decrease in the energy of interaction. By contrast, $O^2$-MeT and $O^4$-MeT belong to the second group in which the interactions with unmodified bases have greater orientational flexibility, i.e., energy minima for base-base interactions are broader. This flexibility is due to the fact that in the optimal configurations two hydrogen-bond acceptors of the methylated thymines share one donor of the unmodified base, thereby forming bifurcated rather than linear hydrogen bonds. In addition, these configurations allow the distance between negatively charged hydrogen-bond acceptors of interacting molecules to be increased. This, in turn, greatly reduces electrostatic repulsion and further stabilizes the complex. Thus, interactions between the thymine derivatives

and the unmodified bases are essentially of the same (electrostatic) nature as interactions between bases forming linear hydrogen bonds. These results demonstrate that the existence of 'perfect' hydrogen bonds is not a prerequisite for preferential stabilization of certain base-pairs.

Of bases from the first group ($O^6$-MeG and $N^4$-OMeC), $O^6$-MeG forms linear hydrogen-bond complexes with C, T and A; and $N^4$-OMeC, depending on its tautomeric form, pairs with A or G. The most stable geometries of all pyrimidine-purine base-pairs, except for $O^6$-MeG–C, are very close to those in regular nucleic-acid structures. Thus, the conformation of the sugar-phosphate backbone and stacking pattern with neighboring pairs remains essentially unperturbed.

Transcription experiments on poly(dC-dG) containing residues of $O^6$-MeG [12] demonstrated that there is a competitive incorporation of dTMP and dCMP opposite to $O^6$-MeG. When the 5'-triphosphates are present in equal amounts, normal incorporation and miscoding occur with comparable probabilities. Similar competitive incorporation may also occur in vivo [50]. These findings are in accord with our calculations, which give similar energies of interaction between $O^6$-MeG and T or C.

Our values for the minimum energies of interaction of free bases are in a good agreement with those from an earlier study by Psoda et al. [14]. The results from the two studies give, however, different energies of interactions in the geometry exactly corresponding to that in B-DNA. In the earlier work, repulsive contribution to the energy of interaction was not explicitly taken into account. As a consequence, the $O^6$-MeG–C pair in the B-DNA geometry was found to be attractive.

The possibility of formation of $O^6$-MeG–A pairs has been demonstrated in transcription of the copolymer poly(C,$O^6$-MeG) by RNA polymerase of *M. luteus* which leads to incorporation of AMP [9]. Computer-generated models show that the distance between glycosidic nitrogens in $O^6$-MeG–A pairs is increased when adenine is either *anti* or *syn* around the glycosidic bond. Thus, the existence of this base-pair must lead to a local deformation of the double helix.

The ability of $O^6$-MeG to form specific, hydrogen-bonded complexes with potentially complementary bases was put in question by Mehta and Ludlum [10]. On the basis of mixing experiments, which showed that poly($O^6$-methylguanosine monophosphate) does not form helical structures either with poly(C) or with poly(U), they concluded that incorporation of unmodified bases opposite to $O^6$-MeG is nonspecific. It was suggested that alkylation of the $O^6$ position of G disrupts any base-pairing relationship. However, in the same study the poly($O^6$-methylguanosine monophosphate) was shown to possess considerable secondary structure. Thus, we believe that the observed lack of ability of these homopolynucleotides to interact with poly(C) and poly(U) is more likely due to the preference of $O^6$-MeG to autoassociate rather than heteroassociate. A similar result was also observed at the monomer level [14]. It should also be noted that the homopolymer study discussed above is not a good model for nucleic acids at equilibrium structures or in nucleic acid synthesis.

$N^4$-OMeC is the other example studied of a derivative capable of forming well-defined, linear hydrogen bonds with unmodified bases. The imino tautomer forms two hydrogen bonds with A and the amino tautomer pairs with G through three hydrogen bonds (fig. 9). Corresponding energies of interactions were calculated to be very close to those in A-U and G-C pairs, showing that the *anti* conformers of the imino and amino forms of $N^4$-OMeC are perfectly capable if mimicking T (or U) and C, respectively, in polynucleotides. Thus, it is expected that the extent of incorporation of A and G opposite to $N^4$-OMeC will depend on a proportion of the imino and amino forms in a given environment. Such a dual specificity of $N^4$-OMeC and its unmethylated analog $N^4$-OHC was observed under a variety of experimental conditions [22–27]. In transcription with DNA-dependent RNA polymerase in the presence of $Mn^{2+}$ [28] $N^4$-OMeC acts only as U, which can be interpreted as evidence for a strong preference of the imino form.

As was first pointed out by Shugar et al. [21], the ability of $N^4$-OMeC and $N^4$-OHC to base-pair in nucleic acids depends on the conformation of the exocyclic group about the C(4)-N bond. The

$N^4$-OMe or $N^4$-OH group must be oriented *anti* to the cytosine ring N(3) in both the imino and amino form. Otherwise, the Watson-Crick type of base-pairing is no longer possible. On the other hand, the *syn* configuration was found to be intrinsically more stable in experimental studies [21] as well as in our theoretical calculations. However, if total geometry optimization is included, the calculated energy difference between the *syn* and *anti* form is relatively small, even in the case of $C^5$-Me-$N^4$-OMeC. Thus, on balance, the *anti* configuration should be energetically preferred in polynucleotides since the gain in the intermolecular energy, due to formation of a hydrogen-bonded structure, easily compensates for energy loss in going from the favored *syn* rotamer to the *anti* form. This conclusion, of course, does not apply to complexes of free bases, where no geometrical constraints are present.

Our results can also be used to interpret the loss in stability of complexes of poly(U) containing $N^4$-OMe-C residues with poly(A), as manifested in a lower $T_m$. Recently, it was suggested [49] from studying computer models that this loss of stability arises from steric interference between the $O(N^4)$-methyl hydrogens of both the *syn* and *anti* forms of the methoxy derivative and the neighboring base-pairs. This interference, in turn, was thought to cause a distortion of the helical structure. The results of our study, however, do not support this interpretation. We found that with the proper rotation of the methyl group no steric hindrances are present. Rather, our results suggest that the destabilization may be caused by rotation of $N^4$-OMeC to the less favored *anti* form, a mechanism previously postulated for $N^4$-methylcytosine [56], but discounted in the case of $N^4$-OMeC [57].

$O^2$-MeT and $O^4$-MeT belong to the second group of modified bases which form planar complexes having bifurcated hydrogen bonds. The corresponding energy minima are broad, hence the bases have much more orientational flexibility.

An interesting consequence of the orientational flexibility, discussed above, is the possibility of rotating the *O*-methyl groups of methylated thymines to the *syn* conformation. Normally, it is assumed that the presence of *O*-alkyl derivatives in the double helix would require that the alkyl groups be *anti* to the base-pairing side [13]. This restriction, however, does not seem to be justified for *O*-methylated pyrimidines. In fact, base-pairs with the *syn* forms of *O*-methylthymines are not only sterically allowed, but also their lowest interaction energies are very similar to those with the *anti* forms. Thus, it can be expected that the former complexes would be more stable by the *syn-anti* interconversion energy. Calculated stabilities of various mispairing schemes (table 4) were found to coincide with transcription results [57], which show that in the presence of DNA-dependent RNA-polymerase and $Mn^{2+}$ (*O*-alkyluridines direct incorporation of C and G.

Even though the most stable base-paired geometries involving *O*-methylthymines are significantly different from those in the double helices, the corresponding changes in backbone torsional angles are relatively small. In fact, backbone angles found in our minimization procedure are within the range observed for various forms of DNA or RNA (see table 5). This result is in agreement with findings that one can build reasonable mismatches of unmodified base-pairs and cause rather little perturbation of the double helix [18,19].

In spite of similarities in torsional angles, calculated differences in the stacking energy of the mini-helices with different nearest neighbors are different, lending further support to previous results which show that the stability of mispairs vary with their environment in the polynucleotide [19,58–60]. Our calculations, however, do not allow for a detailed analysis of stacking patterns and nearest-neighbor effects. The minimization procedure and geometry search we applied do not guarantee that the resulting structures represent the most stable mini-helices. Thus, it is possible that some structures correspond only to a local energy minimum. Furthermore, the method used for the energy calculation of mini-helices involves several approximations, which could have an influence on the final results.

Our results show that only a few of the mispairs considered in this study are exactly compatible with the structure of the regular Watson-Crick helix. If it is assumed that large constraints imposed by multiprotein replicative machinery do

not allow for deviations from this structure [61], only such mispairs may be formed in vivo and lead to substitution mutations.

Alternative mechanisms have been proposed to explain enhanced discrimination during replication and transcription (for a review see ref. 62). These mechanisms would allow for the possibility that base-pairing with modified bases is accompanied by backbone deformation. Base-pairings of this type were found to be formed during in vitro transcription [9,12]. For example, base-pairs containing *O*-methylthymines do not fit into the regular double-helix. Similarly, $O^6$-MeG is not capable of forming attractive complexes with C and A in the regular double-helical geometry. It should, however, be borne in mind that transcription experiments are usually conducted under conditions which are very different from normal replication. To enhance the level of misincorporation a template is transcribed by using enzymes (usually DNA-dependent RNA polymerase in the presence of $Mn^{2+}$) which are several orders of magnitude less accurate than replication enzymes. Furthermore, the system does not possess any repair capabilities. Thus, the results from transcription experiments cannot be treated as being directly relevant to the situation in vivo.

While still unresolved, there is some evidence that template modifications which cause perturbation of the helical structure may also lead to misincorporation at the in vivo level. For example, when RNA of tobacco mosaic virus was treated with nitrosoguanidine, a correlation between the 3-methylcytidine content of modified RNA and mutagenic frequency was found [63]. As can be shown from model building, the *N*(3)-methyl group of cytosine cannot be accommodated in the regular double-helical geometry.

The results presented here show that there is a good qualitative correlation between calculated internal stabilities of various modified base-pairs and the results of in vitro transcription using DNA-dependent RNA polymerase in the presence of $Mn^{2+}$. It should, however, be borne in mind that interactions with polymerase, which were not taken into account in these calculations could modulate both the difference in free energy between correct and incorrect base-pairings and the extent to which backbone distortion can be accommodated during this process.

## Acknowledgments

This work was supported by NCI contract no. ICP 15730. We are indebted to Dr. S.K. Burt for helpful discussions and assistance in applying computer graphics.

## References

1 A. Loveless, Nature 223 (1969) 206.
2 R. Goth and M.F. Rajewski, Proc. Natl. Acad. Sci. U.S.A. 71 (1974) 639.
3 P.G. Margison and P. Kleihues, Biochem. J. 148 (1975) 521.
4 J.W. Nicoll, P.F. Swann and A.E. Pegg, Nature 254 (1975) 261.
5 P.G. Margison, J.M. Margison and R. Montesano, Biochem. J. 157 (1976) 627.
6 G.P. Margison, N.J. Curtin, K. Snell and A.W. Craig, J. Cancer 40 (1979) 809.
7 B. Singer, J. Natl. Cancer Inst. 61 (1979) 1329.
8 P.F. Swann and R. Mace, Chem.-Biol. Interact. 31 (1980) 239.
9 L.L. Gerchman and D.B. Ludlum, Biochim. Biophys. Acta 308 (1973) 310.
10 J.R. Mehta and D.B. Ludlum, Biochemistry 15 (1976) 4329.
11 J.R. Mehta and D.B. Ludlum, Biochim. Biophys. Acta 521 (1978) 770.
12 P.J. Abott and R. Saffhill, Biochim. Biophys. Acta 562 (1979) 51.
13 B. Singer, in: Molecular and cellular mechanisms of mutagenesis, eds. J.F. Lemontt and W.M. Generoso (Plenum Press, New York, 1982) p. 1.
14 A. Psoda, B. Kierdaszuk, A. Pohorille, M. Geller, T.J. Kusmierek and D. Shugar, Int. J. Quantum Chem. 20 (1981) 543.
15 B. Kierdaszuk, R. Stolarski and D. Shugar, Eur. J. Biochem. 130 (1983) 559.
16 D.J. Patel, S.A. Kozlowski, L.A. Marky, C. Broka, J.A. Rice, K. Itakura and K.J. Breslauer, Biochemistry 21 (1982) 428.
17 D.J. Patel, S.A. Kozlowski, L.A. Marky, C. Broka, J.A. Rice, K. Itakura and K.J. Breslauer, Biochemistry 21 (1982) 445.
18 V.P. Chupina and V.I. Poltiev, Nucleic Acids Res. 11 (1983) 5205.
19 J.W. Keepers, P. Schmidt, T.L. James and P. Kollman, Biopolymers (1984) in the press.
20 D.M. Brown, M.J.E. Hewlins and P. Schell, J. Chem. Soc. C (1968) 1925.

21 D. Shugar, C.P. Huber and G.I. Birnbaum, Biochim. Biophys. Acta 447 (1976) 274.
22 E.I. Budowsky, E.D. Sverdlov and N.T. Spasokuktskaya, FEBS Lett 17 (1971) 336.
23 E.I. Budowsky, E.D. Sverdlov and N.T. Spasokuktskaya, Biochim. Biophys. Acta 287 (1972) 195.
24 R.A. Flavell, D.L. Sabo, E.F. Bandle and C. Weissmann, J. Mol. Biol. 89 (1974) 255.
25 D.L. Sabo, E. Domingo, E.F. Bandle, R.A. Flavell and C. Weissmann, J. Mol. Biol. 112 (1977) 235.
26 T. Taniguchi and C. Weissmann, J. Mol. Biol. 118 (1977) 533.
27 W. Muller, H. Weber, F. Meyer and C. Weissmann, J. Mol. Biol. 124 (1978) 343.
28 B. Singer and S. Spengler, Biochemistry 20 (1981) 1127.
29 G. Klopman and A. Ray, Cancar Biochem. Biophys. 6 (1982) 31.
30 J.S. Kwiatkowski, B. Lesyng, H.M. Palmer and W. Saenger, Z. Naturforch. 37c (1982) 937.
31 M.J.S. Dewar and W. Thiel, J. Am. Chem. Soc. 99 (1977) 4899.
32 W.J. Hehre, R.F. Stewart and J.A. Pople, J. Chem. Phys. 51 (1969) 2657.
33 M. Geller, A. Jaworski and A. Pohorille, Int. J. Quantum Chem. 15 (1979) 369.
34 R. Rein, P. Claverie and M. Pollak, Int. J. Quantum Chem. 2 (1968) 129.
35 A.I. Kitaigorodskii, Tetrahedron 14 (1961) 230.
36 J. Caillet and P. Claverie, Acta Crystallogr. A31 (1975) 448.
37 A. Nash and D.F. Bradley, Biopolymers 3 (1965) 261.
38 P. Claverie, B. Pullman and J. Caillet, J. Theor. Biol. 12 (1966) 419.
39 Z.G. Kudritskaya and V.I. Danilov, J. Theor. Biol. 59 (1976) 303.
40 J. Langlet, P. Claverie, F. Caron and J.C. Boeuve, Int. J. Quantum Chem. 19 (1981) 299.
41 J. Caillet and P. Claverie, Biopolymers 13 (1974) 601.
42 M. Geller, B. Lesyng and A. Pohorille, Int. J. Quantum Chem. 17 (1980) 293.
43 O.E. Millner and J.A. Andersen, Biopolymers 14 (1975) 2159.
44 G. Govil, Biopolymers 15 (1975) 2303.
45 J.T. Egan, J. Hart, S.K. Burt and R.D. MacElroy, Comput. Graphics 6 (1982) 177.
46 F.E. Hruska and W.J.P. Blonski, Can. J. Chem. 60 (1982) 3026.
47 R.G. Brennan, G.G. Prive, W.J.P. Blonski, F.E. Hruska and M. Sundaralingam, J. Am. Chem. Soc. 105 (1983) 7737.
48 G.I. Birnbaum, T. Kulikowski and D. Shugar, Can. J. Biochem. 57 (1979) 308.
49 S. Spengler and B. Singer, Biochemistry 20 (1981) 7290.
50 A.R. Peterson, J.R. Landolph, H. Peterson and C. Heidelberger, Nature 276 (1978) 508.
51 B. Pullman and A. Saran, Progress Nucleic Acid Res. Mol. Biol. 18 (1976) 215.
52 L. Langridge, D.A. Marvin, W.E. Seeds, H.R. Wilson, C.W. Hooper, M.H.F. Wilkins and L.D. Hamilton, J. Mol. Biol 2 (1960) 38.
53 S. Arnott, S.D. Dover and A.J. Wonacott, Acta Crystallogr. B25 (1969) 2192.
54 S. Arnott and D.W.L. Hukins, Biochem. Biophys. Res. Commun. 47 (1972) 1504.
55 S. Arnott and D.W.L. Hukins, J. Mol. Biol. 81 (1973) 93.
56 D.J. Engel and P.H. von Hippel, Biochemistry 13 (1974) 4143.
57 B. Singer, H. Fraenkel-Conrat and T.J. Kusmierek, Proc. Natl. Acad. Sci. U.S.A. 75 (1978) 1722.
58 J.R. Fresco, S. Broitman and A.E. Lane, in: Mechanistic studies of DNA replication and genetic recombination, eds. B.M. Alberts and C.F. Fox, ICN-UCLA Symp. Mol. Cell. Biol. vol. 19 (Academic Press, New York, 1980) p. 753.
59 M.D. Topal and N.K. Sinha, J. Biol. Chem. 258 (1983) 12274.
60 P.N. Borer, B. Dengler, I. Tinoco and O.C. Uhlenbeck, J. Mol. Biol. 86 (1974) 843.
61 M.D. Topal and J.R. Fresco, Nature 263 (1976) 285.
62 L.A. Loeb and T.A. Kunkel, Annu. Rev. Biochem. 51 (1982) 429.
63 B. Singer and H. Fraenkel-Conrat, Biochemistry 8 (1969) 3266.