Intrinsic Conformational Energetics Associated with the Glycosyl Torsion in DNA: A Quantum Mechanical Study

Nicolas Foloppe,* Brigitte Hartmann,† Lennart Nilsson,* and Alexander D. MacKerell, Jr.‡

*Center for Structural Biology, Department of Bioscience, Karolinska Institutet, S-141 57, Huddinge, Sweden; †Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France; and ‡Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, USA

ABSTRACT The glycosyl torsion (χ) in nucleic acids has long been recognized to be a major determinant of their conformational properties. χ torsional energetics were systematically mapped in deoxyribonucleosides using high-level quantum mechanical methods, for north and south sugar puckers and with γ in the g^+ and *trans* conformations. In all cases, the *syn* conformation is found higher in energy than the *anti*. When γ is changed from g^+ to *trans*, the *anti* orientation of the base is strongly destabilized, and the energy difference and barrier between *anti* and *syn* are significantly decreased. The barrier between *anti* and *syn* in deoxyribonucleosides is found to be less than 10 kcal/mol and tends to be lower with purines than with pyrimidines. With $\gamma = g^+/\chi = anti$, a south sugar yields a significantly broader energy well than a north sugar with no energy barrier between χ values typical of A or B DNA. Contrary to the prevailing view, the *syn* orientation is not more stable with south puckers than with north puckers. The *syn* conformation is significantly more energetically accessible with guanine than with adenine in 5-nucleotides but not in nucleosides. Analysis of nucleic acid crystal structures shows that $\gamma = trans/\chi = anti$ is a minor but not negligible conformation. Overall, χ appears to be a very malleable structural parameter with the experimental χ distributions reflecting, to a large extent, the associated intrinsic torsional energetics.

INTRODUCTION

DNA is a very flexible molecule, and this flexibility is central to its many biological functions. The flexibility is apparent, for instance, in the polymorphism of the overall structure of DNA (Franklin and Gosling, 1953; Leslie et al., 1980; Hartmann and Lavery, 1996), the wrapping of DNA around the histone core (Lüger et al., 1997), the structure of the Holliday junction (Ortiz-Lombardia et al., 1999), bending and kinking of DNA upon binding to a variety of proteins (Dickerson, 1998; Jones et al., 1999), DNA deformation upon drug binding (Takahara et al., 1995; Coste et al., 1999), and the growing number of unusual DNA conformations (Lebrun and Lavery, 1997). This flexibility largely reflects the presence of several rotatable bonds in each of the polymer building blocks (Saenger, 1984). The intrinsic conformational energetics associated with rotation about these torsions is expected to contribute significantly to the overall structure and dynamics of DNA. Accordingly, this contribution to DNA energetics is of fundamental importance. The overall DNA conformation, which is probed experimentally, however, integrates other energetic contributions, such as the influence of environmental factors (e.g., water activity, crystal packing effects, interactions with proteins or drugs) or noncovalent interactions between different parts of the DNA itself. It is therefore very difficult, if not impossible, to dissect the contribution of the intrinsic conformational energetics of the polymer solely based on the analysis of experimental data. That is why considerable efforts have been dedicated to analyzing the conformational energetics of nucleic acid components in vacuo using theoretical methods.

Early theoretical approaches for the conformational analysis of nucleic acid building blocks included hard-sphere models (Wilson and Rahman, 1971; Haschmeyer and Rich, 1967; Sundaralingam, 1969; Lakshminarayanan and Sasisekharan, 1970; Olson and Flory, 1972a), molecular mechanics potentials (Lakshminarayanan and Sasisekharan, 1969; Olson and Flory, 1972b; Olson, 1973, 1982; Yathindra and Sundaralingam, 1973; Gabb and Harvey, 1993) and semiempirical quantum mechanical calculations (Jordan and Pullman, 1968; Berthod and Pullman, 1971a,b, 1973; Saran et al., 1972, 1973). Although these studies have provided valuable insights into the gross conformational properties of DNA constituents, such coarse physical models cannot yield accurate intrinsic conformational properties. More recently, however, it has become possible to analyze these conformational properties using ab initio quantum-mechanical calculations, which are a priori much more accurate than the earlier models. This has led to a renewed interest in the conformational analysis of the nucleic acid building blocks (Foloppe and MacKerell, 1998, 1999a,b; Florián et al., 1998; Brameld and Goddard, 1999; Hocquet et al., 2000; Hocquet and Ghomi, 2000; Leulliot et al., 1999a; Shishkin et al., 2000; Hocquet, 2001; Banavali and MacKerell, 2001). These studies have provided new insights into the conformational properties of the deoxyribose (Foloppe and MacKerell, 1998, 1999a; Brameld and

Submitted July 18, 2001, and accepted for publication November 20, 2001.

Address reprint requests to Alexander D. MacKerell, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201. Tel.: 410-707-7442; Fax: 410-706-0346; E-mail: alex@outerbanks.umaryland.edu.

Nicolas Foloppe's present address: RiboTargets Ltd., Granta Park, Abington, Cambridge CB1 6GB, United Kingdom.

0006-3495/02/03/1554/16 $2.00

Goddard, 1999), and the intrinsic torsional energetics about other DNA rotatable bonds (Florián et al., 1998; Foloppe and MacKerell, 1999b; Hocquet et al., 2000; Banavali and MacKerell, 2001).

The glycosyl torsion, χ, about the bond which links the base to the sugar, is of special interest because it is related to the overall architecture of the DNA molecule (Dickerson et al., 1982; Hartmann and Lavery, 1996). It is experimentally well documented that χ preferentially adopts an *anti* conformation in which the base is oriented away from the sugar. This is observed in crystals (Altona and Sundaralingam, 1972; Gelbin et al., 1996) and nuclear magnetic resonance (NMR) (Altona and Sundaralingam, 1973; Davies and Danyluk, 1974; Davies, 1978) structures of nucleosides and nucleotides, and in the A and B forms of DNA (Arnott and Hukins, 1969, 1972; Dickerson et al., 1982). In the alternative *syn* conformation, χ is such that the base is oriented over the furanose ring (Donohue and Trueblood, 1960). Based on simple steric models, it has long been suggested that unfavorable steric hindrance between the base and sugar in the *syn* orientation was responsible for the preference for the *anti* conformation (Haschmeyer and Rich, 1967; Wilson and Rahman, 1971; Saenger, 1984). This is believed to be the reason why the *syn* conformation is observed significantly less with pyrimidines than with purines (Rao and Sundaralingam, 1970; Gelbin et al., 1996), as the six-member ring of pyrimidines is bulkier than the corresponding five-member ring in purines. This is in agreement with data on nucleosides and nucleotides, obtained from NMR studies (Davies, 1978) and statistical analyses of crystal structures (Gelbin et al., 1996). The *syn* conformation was, however, observed with pyrimidines (Saenger and Scheit, 1970; Gelbin et al., 1996), including in Z DNA (Wang et al., 1985; Ho and Mooers, 1997). To be characterized experimentally in the *syn* conformation, the bases were frequently chemically modified so as to destabilize the *anti* orientation (Tavale and Sobell, 1970; Rao and Sundaralingam, 1970), stressing the difficulty to experimentally study the *syn* conformation for the standard bases.

NMR data in solution indicate that guanine has a significant propensity to adopt a *syn* orientation in 5′-nucleotides (Son et al., 1972), and early theoretical investigations have suggested that this propensity is higher with guanine than with adenine (Olson, 1973; Yathindra and Sundaralingam, 1973). This, combined with a number of crystal structures of nucleosides (Haschemeyer and Sobell, 1965; Rao and Sundaralingam, 1970), where the guanine adopts the *syn* orientation, has led to the notion that guanine imparts special properties to nucleosides and nucleotides in terms of the *syn* versus *anti* equilibrium (Guschlbauer, 1972; Saenger, 1984). This view was reenforced by the elucidation of Z DNA (Wang et al., 1979) and Z RNA (Davis et al., 1990) structures, where the *syn* purines are mostly guanines. Z DNA typically occurs in alternating purine-pyrimidine sequences dominated by GC basepairs (Pohl and Jovin, 1972; Wang et al., 1979; Arnott et al., 1980; Drew et al., 1980; Ho and Mooers, 1997), and the ability to assume the Z form becomes increasingly difficult as the GC content decreases (Pohl and Jovin, 1972; Ho et al., 1986; Peticolas et al., 1988; Wang et al., 1984, 1987). Although a role for Z-DNA in vivo has not yet been firmly established, there are several lines of evidence suggesting that Z-DNA may be biologically relevant (Herbert and Rich, 1996; Schwartz et al., 1999; Malinina et al., 1994). That GC base pairs favor Z DNA more than AT base pairs has long been attributed to the higher propensity of guanine to adopt a *syn* conformation, as compared with adenine (Saenger, 1984). The *syn* conformation found in the Z_I and Z_{II} forms of DNA (Wang et al., 1981) may serve as a more general model of the *syn* conformation in other contexts. *Syn* guanine bases were also found in the crystal structure of a dinucleotide bound to the enzyme barnase (Baudet and Janin, 1991), in DNA quadruplexes (Smith and Feigon, 1992; Kang et al., 1992), which may represent the structure of telomeric DNA, in mismatched base pairs (Cognet et al., 1991; Skelly et al., 1993; Roll et al., 1999), in stretched DNA (Lebrun and Lavery, 1996), and in some RNA tetraloops (Cheong et al., 1990). Therefore, interest in the *syn* conformation is not limited to Z DNA.

The intrinsic energetics associated with the *syn* conformation in deoxy-nucleosides and deoxy-nucleotides remains, however, incompletely characterized, despite some recent efforts in this direction (Leulliot et al., 1999b; Hocquet et al., 2000; Shishkin et al., 2000). The present work addresses this question using high level ab initio calculations to investigate the intrinsic energetics associated with χ in nucleosides and 5′-nucleotides, taking into account the coupling between χ and γ. When χ = *anti*, γ is largely g^+, as seen in the A and B forms of DNA (Dickerson et al., 1982), in the pyrimidine nucleotides in Z DNA (Dickerson, 1992; Hartmann and Lavery, 1996), and in crystal structures of nucleosides and nucleotides where χ = *anti* (Gelbin et al., 1996). It has long been proposed, however, that χ = *syn* tends to be associated with γ = *trans* (Olson, 1973; Yathindra and Sundaralingam, 1973), and this is illustrated by the structure of Z DNA (Wang et al., 1979; Dickerson et al., 1982; Hartmann and Lavery, 1996). Much less attention, however, has been devoted to the occurrence of γ = *trans* when χ = *anti*. It was noted that γ = *trans* is frequently observed in nucleosides and nucleotides bound to proteins even when χ = *anti* (Moodie and Thornton, 1993). A study of the γ = *trans*/χ = *anti* conformation in nucleic acids is, to our knowledge, not available. That is why, in addition to the energetics of γ = *trans*/χ = *syn*, $\gamma = g^+/\chi$ = *anti* and $\gamma = g^+/\chi$ = *syn* conformations, the present work also investigates the γ = *trans*/χ = *anti* conformation. The relevance of the γ = *trans*/χ = *anti* conformation is supported by an analysis of crystal structures of DNA and RNA.

The present results provide new, systematic, and more quantitative insights into the conformational properties of DNA constituents and are interpreted in the context of DNA structure and dynamics. They provide guidelines for the structural modeling of nucleic acids, the calibration of empirical force fields, and may also be useful for the development of algorithms designed to predict genomic sequences prone to adopt the Z form of DNA (Ho et al., 1986) and other unusual conformations.

METHODS

The atom names and dihedral angle nomenclature are as in Saenger (1984). Deoxyribose puckering pseudorotation angles (P) and amplitudes (τ) have been determined following Altona and Sundaralingam (1972), using the same reference state for $p = 0.0°$. The pseudorotation space is divided into four equally sized quadrants centered around $p = 0.0°$, $p = 90.0°$, $p = 180.0°$, and $p = 270.0°$ that are referred to as the north, east, south, and west quadrants, respectively. The pseudorotation angles and amplitudes were extracted from the energy-minimized structures. All energy minimizations were performed with the GAUSSIAN 98 program (Frisch et al., 1998) to the default tolerances.

Nucleosides

The *anti* to *syn* potential energy surfaces were calculated by changing the value of χ by 10° increments from the *anti* range to the *syn* range. This was performed starting with both a north and a south pucker for the deoxyribose and with γ in both the g^+ and *trans* ranges. These two conformations of γ were selected because they are observed in the A, B, and Z forms of DNA (Dickerson, 1992; Hartmann and Lavery, 1996). The initial value of χ in *anti* was 200° with north sugars and 230° with south sugars. Then, χ was progressively changed toward the *syn* conformation through the two a priori available paths via $\chi = 0°$ or $\chi = 120°$. These two paths were explored independently but are represented together in Fig. 1. All the *anti* to *syn* energy profiles were calculated to a final value of $\chi = 60°$. During the *anti* to *syn* transition, the C4′-C5′-O5′-H5′ torsion can reorient, driven by the formation of an intramolecular hydrogen bond between the 5′-OH group and the base. Although such hydrogen bonds are of interest for understanding the properties of nucleosides themselves, they are not relevant to the study of full nucleic acids, which is the primary focus of this work. Therefore, β was fixed at 180° when calculating the *anti* to *syn* energy surfaces in nucleosides, consistent with the observation that β only populates the *trans* range in A, B, and Z DNA (Dickerson, 1992; Hartmann and Lavery, 1996; Schneider et al., 1997). This constraint on β was maintained during all calculations with the nucleosides. No other constraint was applied when calculating the *anti* to *syn* energy surfaces, which were computed at the restricted Hartree-Fock (HF) level of theory, with the 6-31G* basis set. Each conformation along a surface was energy minimized in the presence of the above-mentioned constraints. In addition, the energy minima in the *syn* and *anti* range were located by energy optimization without restraints on χ, starting from the lowest energy conformations obtained by discrete sampling. The conformations corresponding to the energy minima at the HF/6-31G* level were then used as initial structures for energy minimization at the second-order Møller-Plesset (MP2) level of theory with the 6-31G* basis set.

The energy barrier between the *syn* and *anti* minima is defined here as the point of highest energy obtained by discrete sampling, along the path of lowest energy between the two minina. No attempt was made to locate the true energy maximum between the two energy minima. The *syn* to *anti* potential energy surfaces were obtained with the four standard DNA bases.

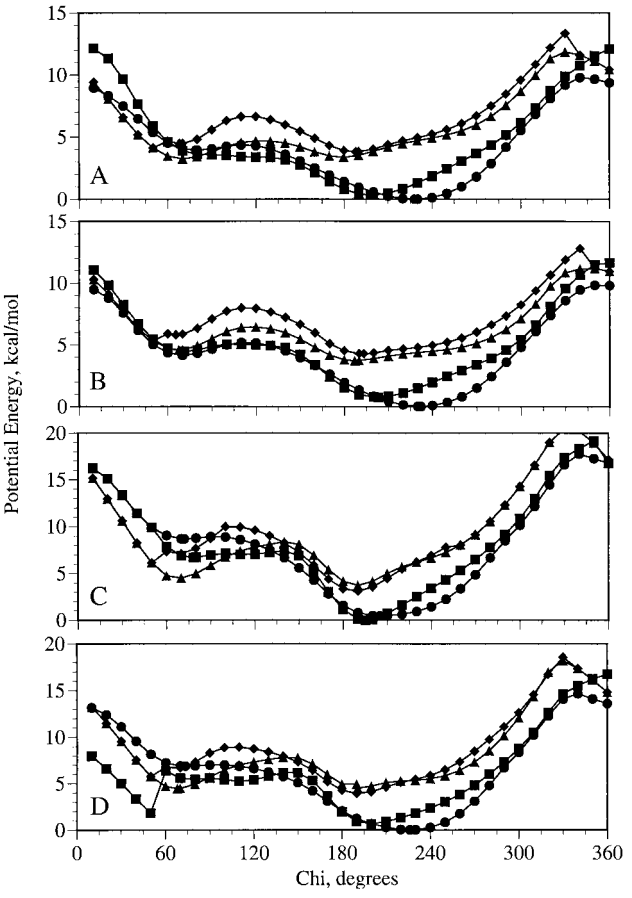

FIGURE 1 Potential energy surfaces (HF/6-31G*) obtained for (*A*) deoxyadenosine, (*B*) deoxyguanosine, (*C*) deoxycytidine, and (*D*) deoxythymidine by varying the glycosyl (χ) torsion in 10° increments from the *anti* to the *syn* range. These energy surfaces were obtained with a north sugar and $\gamma = g^+$ (*squares*), a north sugar and $\gamma =$ *trans* (*triangles*), a south sugar and $\gamma = g^+$ (*circles*), and a south sugar and $\gamma =$ *trans* (*diamonds*). The values of χ corresponding to the energy minima are reported in Table 5. χ was brought from *anti* to *syn* independently through the two available pathways, i.e., through $\chi = 0.0°$ and $\chi = 120.0°$ (for details, see Methods). On the profiles with $\gamma =$ *trans* and initiated with a south sugar, the sugar switches to the north range in the vicinity of $\chi = 340°$ or $\chi = 50°$ for all four nucleosides. In the energy surface for deoxythymidine (*D*) with a north pucker and $\gamma = g^+$, the sharp drop in energy when going from $\chi = 360.0°$ to $\chi = 50.0°$ is due to γ switching from g^+ to g^-, leading to the formation of a hydrogen bond between the 2′-OH and 5′-OH groups.

For a given nucleoside, the potential energy surfaces were offset relative to the lowest energy obtained for this nucleoside.

Nucleotides

The 5′-mononucleotides of adenine (5′-AMP) and guanine (5′-GMP) were studied with a mono-anionic phosphate group to mimic the situation in nucleic acids, with the 5′-phosphoryl moiety being capped with a proton. These nucleotides were energy minimized in a B-DNA like conformation with the base in *anti*, and in the Z_I and Z_{II} DNA-like conformations (Wang et al., 1981) with the base in *syn*. To prevent the formation of intramolecular hydrogen bonds between the 3′ hydroxyl and 5′ phosphoryl groups, the torsions α, β, γ, ϵ, ζ , and χ were constrained to relevant crystallo-

TABLE 1 Occurrence of torsion γ in the *trans* conformation in crystal structures of DNA and RNA

	B DNA*	B DNA/Protein†	A DNA*	A RNA*	A RNA/protein†	A RNA/DNA‡
Number of oligomers	48	45	50	15	3	18
Number of steps						
Total	805	1148	529	479	82	274
γ = *trans*	4	74	66	15	4	16
γ = *trans*/χ = *anti*	4	74	66	13	3	15
Percent of steps						
γ = *trans*	0.5	6.5	12.5	3	0.5	6

*Nucleic acid not bound to a protein.
†Nucleic acid bound to a protein.
‡Mixed duplexes of RNA and DNA.
See Methods for the selection of the analyzed structures.

graphically observed values (see Table 5 in Dickerson, 1992), depending on the DNA form to be modeled. These constraints were $\alpha = 295°$, $\beta = 167°$, $\gamma = 51°$, $\epsilon = 203°$, $\zeta = 240°$, and $\chi = 257°$ in B DNA; $\alpha = 48°$, $\beta = 179°$, $\gamma = 190°$, $\epsilon = 256°$, $\zeta = 80°$, and $\chi = 67°$ for purines in Z_I DNA; $\alpha = 93°$, $\beta = 193°$, $\gamma = 157°$, $\epsilon = 182°$, $\zeta = 74°$, and $\chi = 57°$ for purines in Z_{II} DNA. It is worth noting that in the Z DNA-like conformations the constraints to be applied on ζ at the 5′ termini of the 5′-nucleotides are those formally associated with the pyrimidines in Z DNA. None of the furanose endocyclic torsion angles were constrained during the energy minimizations. These structures were first energy minimized at the HF/6-31+G* level of theory. The resulting structures were subsequently energy minimized at the HF/6-31+G** level or using density functional theory and the B3LYP functional (Becke, 1993) combined with the 6-31+G* basis set. Diffuse functions on heavy atoms were added for the nucleotides due to the presence of the negatively charged phosphoryl group. Some of these energy minimizations were pursued by removing the constraint on torsion χ to allow for the optimization of the interactions between the base and the rest of the nucleotide.

Analysis of crystal structures

An analysis of DNA and RNA crystal structures available in the nucleic acid (Berman et al., 1992) or protein (Bernstein et al., 1977) databanks was carried out to investigate the conformations adopted by torsion γ. The analyzed structures included nucleic acids alone or bound to a protein, the A and B forms of DNA, and sequences with mismatches or unusual bases. End-to-end effects at the junctions between oligonucleotides in crystals frequently distort the terminal base pairs, and these were not included in the statistics. Only structures resolved to 2.0 Å or better were taken into account because some relevant structural features, notably backbone conformation, are unclear below this resolution (Hahn and Heinemann, 1993). The number of oligomers in each category and the number of base steps in each category are given in Table 1.

The C3′···N3 distances in nucleotides with a *syn* base were obtained from Z DNA crystal structures solved at a resolution of 1.5 Å or less. Modified bases, mismatches, and terminal bases were excluded, leaving a set of 28 distances.

RESULTS AND DISCUSSION

One major aim of the present calculations is to investigate the relative potential energies of the *anti* versus *syn* conformation of the glycosyl torsion, the paths between them, and selected aspects of the associated geometries. In the first part of this work (see Conformational analysis at the nucleoside level), these properties were studied with the four standard deoxyribonucleosides. In the second part (see The *syn* conformation in 5′-mononucleotides: adenine versus guanine), we analyze the influence of a 5′-phosphoryl group on the relative energies of the *syn* versus *anti* conformations. This was performed only with the purines, given that the *syn* orientation is mostly observed with these bases (see Introduction). The third part of the work further characterizes the γ = *trans*/χ = *anti* conformation, by analyzing its occurrence in crystal structures of nucleic acids, and presenting relevant ab initio calculations on nucleosides.

The size of the nucleosides, combined with the need to carefully sample the χ torsional energy surface, has dictated the level of theory (HF/6-31G*) at which the energy surfaces presented in Fig. 1 were initially calculated. The general features of the torsional energy profiles obtained for a number of nucleic acid building blocks at the HF/6-31G* level are similar to their MP2/6-31G* counterpart (Foloppe et al., 2001). This is also the experience of other authors when comparing the relative energies of selected conformations of deoxyribonucleosides (Cheatham et al., 1999). This indicates that the HF/6-31G* level should provide a reasonable guide when mapping the overall features of χ energy surfaces. Conformations of special interest (energy minima and energy barriers), however, were also energy minimized at the MP2/6-31G* level. The discussion of the geometries and relative energies of the nucleosides in these conformations is based on the results obtained at the MP2/6-31G* level.

Calculations performed in the present study provide potential energies in vacuo and not free energies in the condensed phase. The merits and limitations of this approach have already been assessed in a similar context (Foloppe and MacKerell, 1998, 1999a,b; Hocquet et al., 2000), showing that this type of calculation can provide useful insights into the intrinsic properties of the nucleosides and their counterparts in nucleic acids.

Because the calculations presented in this work were performed on 2′-deoxy compounds, the results are primarily discussed in relationship to DNA. It is, however, expected

TABLE 2 Energy differences* (kcal/mol) of the nucleosides with χ = *anti* or χ = *syn*, relative to their global energy minimum†

	χ = *anti*		χ = *syn*	
	$\gamma = g^+$	$\gamma = t$	$\gamma = g^+$	$\gamma = t$
North				
dA	0.4‡	3.0	ne†	3.7
dG	0.7‡	3.7	4.2	5.1
dC	0.0‡	4.2	7.1	5.1
dT	0.9‡	5.2	6.4	5.8
South				
dA	0.0‡	3.2	3.3	3.8
dG	0.0‡	3.6	3.6	5.2
dC	0.3‡	2.4	7.7	6.7
dT	0.0‡	3.5	6.7	6.9

At the MP2/6-31G level of theory, with β constrained to 180°.
†Nonexistent (ne): the *north*/$\gamma = g^+/\chi$ = *syn* conformation is not stable with dA, in agreement with the corresponding energy profile obtained at the HF/6-31G* level (see Fig. 1).
‡From reference Foloppe and MacKerell (1999b).

TABLE 3 Pseudorotation angles for the energy minima* obtained for the four deoxynucleosides

	χ = *anti*		χ = *syn*	
	$\gamma = g^+$	$\gamma = t$	$\gamma = g^+$	$\gamma = t$
North				
dA	7.0‡	331.9	ne†	24.8
dG	9.6‡	334.3	41.5	33.5
dC	8.8‡	4.9	40.2	24.8
dT	12.4‡	335.6	35.5	34.7
South				
dA	168.3‡	215.4	155.0	178.5
dG	168.6‡	213.7	151.2	173.1
dC	162.1‡	210.5	155.8	176.4
dT	162.7‡	210.8	153.7	174.1

Obtained at the MP2/6-31G level.
†Nonexistent: the *north*/$\gamma = g^+/\chi$ = *syn* conformation is not stable with adenine, in agreement with the corresponding energy profile obtained at the HF/6-31G* level (see Fig. 1).
‡From Table 2 in reference Foloppe and MacKerell (1999b).

that the present conformational analysis should also be relevant, to a large extent, to ribo compounds and RNA.

Conformational analysis at the nucleoside level

The χ torsional energy profiles going from *anti* to *syn* for deoxyadenosine (dA), deoxycytidine (dC), deoxyguanosine (dG), and deoxythymidine (dT) are shown in Fig. 1, *A* through *D*, respectively. The two conformations for γ (g^+ or *trans*) combined with the two sugar puckers (north or south) led to four *anti* to *syn* potential energy profiles (HF/6-31G*) for each deoxyribonucleoside. The corresponding four energy profiles follow a similar pattern with the four nucleosides. The relative energies of the derived energy minima at the MP2/6-31G* level are listed in Table 2, and the associated pseudorotation angles (*P*), amplitudes of puckering (τ), χ, and γ values are given in Tables 3, 4, 5, and 6, respectively.

Shape of the energy well associated with χ in anti

With all deoxyribonucleosides, the conformation of lowest energy is in $\gamma = g^+/\chi$ = *anti*, with a south (adenine, guanine, thymine) or north (cytosine) sugar, in agreement with previous theoretical studies (Foloppe and MacKerell, 1999a; Hocquet et al., 2000) and the large body of experimental data available for nucleosides and nucleotides (Davies, 1978; Gelbin et al., 1996) as well as DNA (Dickerson, 1992). Detailed discussions of the energy differences between the north and south energy minima in nucleosides with $\gamma = g^+/\chi$ = *anti* were published previously (Cheatham et al., 1999; Foloppe and MacKerell, 1999a; Hocquet et al., 2000) and are not repeated here. By sampling the entire *anti* range, however, the present results complement these previous calculations. They also complement a recent study where χ was sampled systematically (Foloppe and MacKerell, 1999b) but with a larger grid and model compounds where the sugar was represented only by a constrained furanose. The present detailed sampling of the χ energy surface confirms and strengthens the previous observation (Foloppe and MacKerell, 1999b) that the energy well associated with the *anti* orientation is broad with no energy barrier separating the low from the high χ values (in the 200–260° range). With $\gamma = g^+$, the fine sampling of χ emphasizes the much narrower *anti* energy well associated with a north sugar as compared with a south sugar. This unambiguously mirrors the narrower χ distribution observed in crystal structures of A DNA as compared with the broader B DNA distribution (Foloppe and MacKerell, 1999b), and thereby confirms the relevance of these energy

TABLE 4 Puckering amplitudes for the energy minima* obtained for the four deoxynucleosides

	χ = *anti*		χ = *syn*	
	$\gamma = g^+$	$\gamma = t$	$\gamma = g^+$	$\gamma = t$
North				
dA	39.7‡	38.4	ne†	31.6
dG	39.4‡	38.0	36.2	33.0
dC	39.4‡	38.7	36.9	23.0
dT	39.3‡	36.2	35.5	28.4
South				
dA	39.7‡	37.5	41.4	38.4
dG	36.7‡	37.4	42.5	38.5
dC	37.4‡	36.8	42.2	37.4
dT	37.7‡	37.0	42.6	38.2

Obtained at the MP2/6-31G level.
†Nonexistent: the *north*/$\gamma = g^+/\chi$ = *syn* conformation is not stable with adenine, in agreement with the corresponding energy profile obtained at the HF/6-31G* level (see Fig. 1).
‡From Table 2 in Follope and MacKerell (1999b).

TABLE 5 Values of the glycosyl torsion χ (deg.) for the energy minima* obtained for the four deoxynucleosides

	χ = *anti*		χ = *syn*	
	$\gamma = g^+$	$\gamma = t$	$\gamma = g^+$	$\gamma = t$
North				
dA	192.3‡	165.1	ne†	64.4
dG	198.2‡	169.3	63.6	62.4
dC	194.8‡	187.5	75.5	66.7
dT	197.9‡	181.7	76.0	65.4
South				
dA	230.2‡	181.9	77.0	68.8
dG	232.8‡	184.0	68.0	66.3
dC	207.0‡	186.7	68.0	68.0
dT	230.6‡	187.9	69.0	65.9

Obtained at the MP2/6-31G level.
†Nonexistent: the *north*/$\gamma = g^+/\chi$ = *syn* conformation is not stable with adenine, in agreement with the corresponding energy profile obtained at the HF/6-31G* level (see Fig. 1).
‡From Table 2 in reference Foloppe and MacKerell (1999b).

profiles. The present results indicate that this difference simply reflects the intrinsic energetics associated with the glycosyl torsion. To our knowledge, this difference between the two distributions has yet to be explained. The broader well with the south sugar versus the north sugar may entropically favor the B form of DNA over the A form.

Analysis of sugar puckers in crystal structures shows the range of *P* sampled in the north range to be narrower than in the south range (Foloppe and MacKerell, 1998). Accordingly, the variation of *P* with respect to χ was analyzed in the regions of χ sampled in the A and B forms of DNA. However, no clear correlation that could explain the crystal *P* distributions was observed (data not shown).

Of note is the limited influence of sugar pucker on the shape of the *anti* energy wells when γ = *trans* (Fig. 1). It is interesting to note that the pseudorotation angle values are outside the typical C2′endo and C3′endo ranges for the γ = *trans*/χ = *anti* energy minima (Table 3). The north and south values of *P* are in the 330° (except for dC) and 210° ranges, respectively. A few values in these ranges were observed crystallographically (Gelbin et al., 1996). In addition, the energy minima with γ = *trans*/χ = *anti* are systematically shifted to lower values of χ, as compared with the energy minima with $\gamma = g^+/\chi$ = *anti* (Fig. 1 and Table 5). The γ = *trans*/χ = *anti* energy minima with the purines have similar energies for both sugar conformations, although the north energy minima are higher than the south with the pyrimidines (Table 2). The most striking feature regarding the *anti* conformations, however, is the influence of γ on their relative energies.

TABLE 6 Values of the torsion γ (deg.) for the energy minima* obtained for the four deoxynucleosides

	χ = *anti*		χ = *syn*	
	$\gamma = g^+$	$\gamma = t$	$\gamma = g^+$	$\gamma = t$
North				
dA	53.0‡	177.6	ne†	187.9
dG	50.9‡	177.3	48.7	188.5
dC	53.7‡	183.3	52.1	188.4
dT	52.4‡	173.3	53.3	188.2
South				
dA	50.5‡	175.1	45.3	184.8
dG	49.5‡	175.5	43.4	185.3
dC	52.9‡	174.5	38.1	186.3
dT	50.3‡	174.6	42.3	184.7

Obtained at the MP2/6-31G level.
†Nonexistent: the *north*/$\gamma = g^+/\chi$ = *syn* conformation is not stable with adenine, in agreement with the corresponding energy profile obtained at the HF/6-31G* level (see Fig. 1).
‡From Table 2 in Foloppe and MacKerell (1999b).

Fig. 1 and Table 2 show that with all nucleosides the energies with γ = *trans*/χ = *anti* are significantly higher than with $\gamma = g^+/\chi$ = *anti*. These energy differences are all $>$ 2.4 kcal/mol, showing that switching γ from g^+ to *trans* is enough to strongly destabilize the *anti* conformation. This is in agreement with the combination $\gamma = g^+/\chi$ = *anti* being much more frequently observed in DNA and its components than γ = *trans*/χ = *anti* (Table 1) and helps to rationalize this observation. This leads to the prediction that the γ = *trans*/χ = *anti* conformation should be rare in noncanonical DNA conformations, including single stranded DNA. The γ = *trans*/χ = *anti* conformation, however, is not a negligible aspect of nucleic acids structure (see Conformation γ = *trans*/χ = *anti* in nucleic acids structures).

The elevated energy of γ = *trans*/χ = *anti* is consistent with favorable interactions between the base and the 5′-OH group in $\gamma = g^+/\chi$ = *anti*. Analysis of experimental structures in this conformation has long suggested that the interaction between C8—H8 (purines) or C6—H6 (pyrimidines) and the O5′ oxygen includes some hydrogen bond character (Rubin et al., 1972; Sussman and Seeman, 1972; Wahl and Sundaralingam, 1997), and this is consistent with ab initio calculations (Hocquet et al., 2000; Hocquet, 2001). Because a hydrogen bond cannot be formed in γ = *trans*/χ = *anti*, the strong destabilization of this conformation relative to $\gamma = g^+/\chi$ = *anti* supports the view that the interaction between C8—H8/C6—H6 and O5′ involves some hydrogen bond character. It must be kept in mind, however, that the γ = *trans* conformation is intrisically less stable as compared to $\gamma = g^+$ in the absence of the base (Foloppe and MacKerell, 1999b). Consequently, it is difficult to evaluate the strength of the interaction between C8—H8/C6—H6 and O5′ in $\gamma = g^+/\chi$ = *anti* simply by comparing the associated energies to those obtained for γ = *trans*/χ = *anti*. The strong destabilization of the γ = *trans*/χ = *anti* combination observed for the associated energy minima extends to the other values of χ in the *anti* range, including those corresponding to the A-like and B-like DNA conformations. The best-documented occurrence of γ = *trans* is that systematically observed with the purine nucleotides in

Z DNA with an associated *syn* orientation for χ. Therefore, the energetics of γ going to *trans* while χ remains in *anti* is a priori of interest for a better understanding of the possible mechanisms of transitions from A or B DNA to Z DNA. For instance, Fig. 1 shows that when γ is in *trans*, the energy difference and barrier between the *anti* and *syn* minima are significantly decreased, as compared with their counterpart when γ is in g^+. This effect is even more pronounced with north sugars, which are associated with the *syn* orientations of purines in Z DNA.

Syn *versus* anti *orientation of* χ *in nucleosides*

Many of the energetic and structural properties of the deoxyribonucleosides with χ in *syn* are still to be characterized, due to the relatively small number of crystal structures (Gelbin et al., 1996) and high level ab initio calculations available with χ in this orientation. Although less frequently observed with pyrimidines than purines, the *syn* conformation can occasionally be adopted by pyrimidines (Gelbin et al., 1996), including in DNA (Wang et al., 1985). Only a few theoretical studies investigated the properties of the *syn* nucleosides ab initio, using the DFT/B3LYP/6-31G* approach (Hocquet et al., 2000; Shishkin et al., 2000; Leulliot et al., 1999b). The calculations of Hocquet et al. (2000) on deoxyribonucleosides with χ in *syn* were restricted to north sugars and obtained with γ = *trans* for the purines and $\gamma = g^+$ for the pyrimidines. Shishkin et al. (2000) presented calculations for a variety of *syn* conformations, but that study stressed the importance of intramolecular hydrogen bonds between the 5′-OH group and the bases in *syn*, although such interactions cannot occur in DNA. A number of conformations of direct interest to DNA structure, such as γ = *trans*/χ = *syn* for purines with a north sugar, were not analyzed by Shishkin et al. (2000). Besides, the relative potential energies obtained by Shishkin et al. differ from those reported by Hocquet et al. for equivalent conformations. For instance, Shishkin et al. found that the four deoxyribonucleosides with $\gamma = g^+/\chi$ = *anti* are more stable with a north pucker, although this was found to be the case only for dC in the work of Hocquet et al. In addition, none of these previous studies considered the A-like or B-like conformations of χ, and it is therefore difficult to relate the results in the context of DNA structure. In the present work, the four standard deoxyribonucleosides with χ in *syn* were systematically energy minimized at the MP2/6-31G* level with both north and south sugars, and γ in *trans* and g^+. The resulting potential energies, relative to the global energy minimum ($\gamma = g^+/\chi$ = *anti*), are given in Table 2. The corresponding χ and γ values are listed in Tables 5 and 6, respectively.

An energy minimum corresponding to the *syn* conformation was found for all the combinations of sugar puckering and γ studied, except for dA with a north sugar and $\gamma = g^+$. The *syn* energy well is typically narrower than the *anti* energy well, suggesting that the *anti* conformation is entropically favored over the *syn*. The pseudorotation angle values (*P*) when χ = *syn* depend on γ (Table 3). North sugars with χ = *syn* have higher *P* values than with $\gamma = g^+/\chi$ = *anti*, and this is even more pronounced with $\gamma = g^+$ than with γ = *trans*. This trend mirrors the distribution of *P* values in purines in Z DNA (Foloppe and MacKerell, 1998), suggesting that these experimental values reflect an intrinsic property of the corresponding building blocks. In the north/γ = *trans*/χ = *syn* category, the value of *P* is significantly affected by the nature of the base with no clear trend in terms of purine versus pyrimidine (Table 3). The puckering amplitudes are also sensitive to the nature of the base and orientation of the furanose substituents (Table 4). We note the distinct flattening of the north furanose when χ = *syn*, as compared with what is obtained when the furanose is south. Overall, these results suggest that the fine details of the deoxyribose conformation are very sensitive to the details of its local environment.

The structures in *syn* are all higher in energy (Table 2) than the global minimum ($\gamma = g^+/\chi$ = *anti*), in agreement with experimental observations (Gelbin et al., 1996; Davies, 1978). The relative MP2/6-31G* energies of the *syn* conformations are higher than that of the available B3LYP/6-31G* counterparts (Hocquet et al., 2000) by ~1.3 (dA) to ~2.0 kcal/mol (dC). The MP2/6-31G* energies of the *syn* conformations, relative to $\gamma = g^+/\chi$ = *anti*, range from ~3.0 to ~7.0 kcal/mol (Table 2). This argues against the notion (Wang et al., 1985) that purine nucleosides can adopt the *syn* conformation as easily as they can adopt the *anti* conformation. For γ in a given range, the *syn* orientation is almost systematically higher in energy with pyrimidines than with purines, as was qualitatively predicted based on simple steric models (Haschmeyer and Rich, 1967; Olson, 1973; Yathindra and Sundaralingam, 1973; Wilson and Rahman, 1971), and observed experimentally (Gelbin et al., 1996; Davies, 1978). The comparison of dC with dG in north/γ = *trans*/χ = *syn* provides an exception to this trend, where both compounds have the same energy. This is not in contradiction with the *syn* conformation being overwhelmingly assumed by guanines rather than cytosines in Z DNA (Ho and Mooers, 1997), because the 5′-phosphoryl group has to be taken into account to understand the trends observed in Z DNA (see The *syn* conformation in 5′-mononucleotides: adenine versus guanine).

With χ = *syn*, the conformations with a north sugar have lower or equal energies as compared with the conformations with a south sugar, except for dG with $\gamma = g^+$ (Table 2). However, the energy differences between the north and south sugar conformers with χ = *syn* are less than 0.6 kcal/mol, except with the pyrimidines with γ = *trans*. Despite these small energy differences, the present results are useful because they do not support the prevailing view that the *syn* conformation is sterically more destabilized with north sugars than with south sugars (Haschmeyer and

TABLE 7 Selected distances* (Å) between the base† in *syn* and the sugar in the four deoxyribonucleosides

	North				South			
	dA	dG	dC	dT	dA	dG	dC	dT
$\gamma = g^+$								
N3(O2)···O4	ne‡	3.19	3.16	3.25	3.62	3.39	3.03	3.13
N3(O2)···C2′	ne‡	3.15	2.78	2.81	3.14	3.15	2.93	2.95
N3(O2)···C3′	ne‡	3.18	2.95	2.98	4.34	4.28	3.99	4.05
N3(O2)···H3′	ne‡	2.49	2.33	2.34	4.70	4.65	4.41	4.47
N3(O2)···C5′	ne‡	4.11	4.20	4.22	4.44	4.26	3.85	4.00
N3(O2)···O5′	ne‡	3.33	3.51	3.49	3.52	3.46	3.18	3.25
$\gamma = trans$								
N3(O2)···O4′	3.34	3.24	2.95	3.00	3.45	3.37	3.04	3.06
N3(O2)···C2′	3.25	3.23	2.84	2.89	3.17	4.13	2.90	2.94
N3(O2)···C3′	3.16	3.64	2.85	2.93	4.14	3.16	3.80	3.87
N3(O2)···H3′	2.41	2.47	2.29	2.32	2.36	2.36	4.12	4.26
N3(O2)···C5′	3.84	3.90	3.40	3.65	3.55	3.61	3.27	3.45
N3(O2)···H5″	3.06	3.14	2.64	2.93	2.56	2.63	2.32	2.42

At the MP2/6-31G level of theory, with β constrained to 180°.
†The base atoms considered here are either N3 (purines) or O2 (pyrimidines).
‡Nonexistent.

Rich, 1967; Wilson and Rahman, 1971; Saenger, 1984; Gelbin et al., 1996). With dC and dT and $\gamma = trans/\chi = syn$, the north pucker is favored over the south by 1.6 and 1.1 kcal/mol, respectively. This is consistent with the *syn* pyrimidines adopting a north pucker in Z DNA (Wang et al., 1985). Adjustments in the furanose pseudorotation angle (Table 3) and amplitude (Table 4) alleviate a potential steric clash between a north pucker and the base, in contrast with what was obtained in earlier models where the flexibility of the furanose could not be taken into account (Haschmeyer and Rich, 1967).

Another important aspect of understanding the relative stabilities of the *syn* and *anti* orientations is the coupling between the orientations of χ and γ (Olson, 1973; Yathindra and Sundaralingam, 1973). With purines, $\gamma = g^+/\chi = syn$ is more stable than $\gamma = trans/\chi = syn$, except with dA with a north sugar. With pyrimidines, $\gamma = trans/\chi = syn$ is more stable than $\gamma = g^+/\chi = syn$, except with dT with a south sugar. Therefore, no clear trend is observed at the nucleoside level regarding the coupling of χ and γ, even in terms of purines versus pyrimidines. This is expected to be different in nucleic acids, where steric hindrance due to the 5′-phosphoryl group should strongly disfavor the existence of $\gamma = g^+/\chi = syn$. That is why the *syn* bases are associated with $\gamma = trans$ in a given nucleotide in Z DNA, where this combination is mostly observed with guanine nucleotides. Interestingly, the present results indicate that at the nucleoside level, the $\gamma = trans/\chi = syn$ combination is less favored with guanine than adenine for both the north and south sugars. The same observation was made based on B3LYP/6-31G* calculations (Hocquet et al., 2000) (with a north sugar only). Together, these results suggest that deoxyguanosine does not have any particular intrinsic propensity to be in *syn*, contrary to what is sometimes believed (e.g., Saenger (1984), p. 77). The higher propensity of guanine to adopt the *syn* conformation depends on the presence of a 5′-phosphoryl group, which interacts favorably with the guanine 2-amino group (see The *syn* conformation in 5′-mononucleotides: adenine versus guanine). In dG, the distance between the 2-amino group nitrogen and the O5′ atom is 4.0, 4.3, 5.2, and 6.1 Å in north/$\gamma = g^+/\chi = syn$, south/$\gamma = g^+/\chi = syn$, south/$\gamma = trans/\chi = syn$, and north/$\gamma = trans/\chi = syn$, respectively. Therefore, no clear hydrogen bond is formed between this amino group and the O5′ atom in any of these conformations.

Selected distances between the N3 (purines) or O_2 (pyrimidines) atom of the base in *syn* and some sugar atoms are given in Table 7. With some exceptions, these distances are typically shorter with pyrimidines than with purines, as expected. The N3(O2)···O4′ distance is systematically <3.45 Å (except in one case), and probably represents an unfavourable interaction, contributing to the *syn* conformation being systematically higher in energy than *anti*. We also note the particularly short N3(O2)···C2′ and N3(O2)···C3′ distances, especially with a north furanose. Some of the N3(O2)···C5′ distances are also relatively short (<3.5 Å) with $\gamma = trans$. Some of these short distances may correspond to energetically favorable interactions, given the accumulating evidence supporting the existence of C—H···O and C—H···N hydrogen bonds (Wahl and Sundaralingam, 1997; Taylor and Kennard, 1982; Gu et al., 1999; Steiner, 2000; Vargas et al., 2000; Brandl et al., 1999; Sutor, 1962).

If such interactions exist in nucleotides with a *syn* base, they would presumably involve "activated" C—H bonds, where the carbon is linked to an electron withdrawing group. This prompted us to investigate the distribution of the C3′···N3 distance in nucleotides with a *syn* purine in Z DNA crystal structures, because C3′ is bound to an oxygen. The average for this distance is 3.3 Å (SD $\pm$ 0.1 Å). This is

TABLE 8 Torsional energy barrier* (kcal/mol) about χ, between the *anti* and *syn* ranges

	$\gamma = g^+$				$\gamma = trans$			
	χ	P	τ	ΔE	χ	P	τ	ΔE
North								
dA	ne†	ne†	ne†	ne†	120	322.9	39.4	3.7
dG	100	330.3	35.4	3.0	120	317.9	40.5	5.0
dC	140	334.3	43.6	6.2	140	325.4	43.6	7.6
dT	120	326.1	40.0	4.3	140	324.6	43.0	7.4
South								
dA	110	158.9	40.7	4.0	120	152.3	42.7	7.1
dG	110	156.5	41.8	4.8	120	147.0	43.7	8.3
dC	90	152.8	43.3	8.4	100	203.2	38.5	9.5
dT	120	151.3	44.0	6.8	110	145.9	45.5	9.8

At the MP2/6-31G level of theory, with β constrained to 180°. The energy barriers (ΔE) were calculated relative to the conformation with χ = *anti* for the γ conformation of interest. P (deg.) and τ (deg.) are the furanose pseudorotation angle and amplitude, respectively.
†Nonexistent.

compatible with a weakly favorable interaction between C3′ and N3, but it could also represent a repulsive interaction. A shortening of the C—H covalent bond has been observed in some instances when the C—H group acts as a hydrogen bond donor (Gu et al., 1999). Whether the base is *syn* or *anti*, no significant change in the length of the C3′—H3′ bonds was observed in the structures calculated in the present work, including with the 5′-mononucleotides (see The *syn* conformation in 5′-mononucleotides: adenine versus guanine).

Pathways between χ = anti *and* χ = syn

Ab initio calculations provide a powerful tool to probe the energetics of conformational energy barriers. These energy barriers are difficult, if not impossible, to characterize experimentally when many degrees of freedom can vary simultaneously, as is the case with a nucleoside. Given the well-documented importance of the glycosyl torsion on the structure of nucleic acids and their components, the energy barrier between χ = *anti* and χ = *syn* is of particular interest. This barrier remains, however, poorly characterized with relatively little experimental (Rhodes and Schimmel, 1971; Kuramoto et al., 1998; Nishikawa et al., 1999, 2000) or theoretical (Foloppe and MacKerell, 1999b) work addressing this question. There are a priori two pathways for the transition of χ between *anti* and *syn*, that is, through χ = 0° or χ = 120°. Previous work indicates the low energy path to be via χ = 120°, although the model compounds lacked the 3′-OH and 5-OH groups and had the furanose constrained (Foloppe and MacKerell, 1999b). The present study uses full nucleosides without constraints on the furanose conformation. The energy profiles are shown in Fig. 1. The main features of these surfaces are similar to those previously reported for smaller model compounds (Foloppe and MacKerell, 1999b) with minima in the *syn* and *anti* regions and maxima in the regions of χ = 0° and χ = 120°. These results confirm that the lowest energy barrier between *anti* and *syn* is through χ = 120°. We note that in the γ = *trans*/south surfaces of the four nucleosides there is a discontinuity in the vicinity of χ = 45°, where the sugar switched to the north pucker. This, however, occurs only when approaching the *syn* minimum through χ = 0°, and the *syn* minimum with the correct pucker is obtained via χ = 120°. Therefore, the discontinuity does not prevent analysis of the corresponding *syn* minimum.

Detailed analysis of the HF/6-31G* energy surfaces in the χ = 120° region shows their shape and the location of the maxima to depend on the base and on the conformation of γ. For example, with dA, the location of the maximum covers a 20° range, and there is apparently no barrier in the north/$\gamma = g^+$ surface. Another example is the north/$\gamma = g^+$ surface with dT, where there are two small barriers with a local minimum at 110°. In addition, in dT with $\gamma = g^+$ the location of the barriers for the two sugar puckers vary by 50°.

To obtain more accurate estimates of the barrier heights, MP2/6-31G* optimizations were performed with χ constrained to the energy maximum obtained at the HF/6-31G* level. The corresponding energy barriers obtained at the MP2/6-31G* level in the vicinity of χ = 120° are given in Table 8. Overall, the MP2 trends mirror their HF counterpart with the lowest barriers occurring in the $\gamma = g^+$/north surfaces and the highest in the γ = *trans*/south surfaces. However, comparison of the barrier heights in Table 8 with the relative energies of the *syn* minima (Table 2) show that some energy barriers are lower in energy than the *syn* minima (e.g., dC and dT data with $\gamma = g^+$/north). While such results are somewhat disconcerting, they are not unexpected. The often small energy difference between the barrier and the associated *syn* minimum, combined with the various locations of the HF energy barriers, suggests that the exact location of the energy barrier as a function of χ is not the same when calculated at the HF versus the MP2 level. Full determination of the χ torsional energy surface at the MP2/6-31G* level will be required to more rigorously char-

TABLE 9 Selected descriptors of the 5′-nucleotides of adenine and guanine in Z_I and Z_{II} DNA-like conformations

	Adenine				Guanine				$\Delta\Delta E$	
	Z_I		Z_{II}		Z_I		Z_{II}			
	$\chi^\dagger$	$\Delta E_{ZI\text{-}B}$	$\chi^\dagger$	$\Delta E_{ZII\text{-}B}$	$\chi^\dagger$	$\Delta E_{ZI\text{-}B}$	$\chi^\dagger$	$\Delta E_{ZII\text{-}B}$	Z_I	Z_{II}
HF/631+G*, χ cons*		10.5		15.6		6.9		5.1	3.6	10.5
HF/631+G*, χ relax†	73.8	10.4	78.9	14.6	43.5	5.6	52.6	5.1	4.8	9.5
HF/631+G**, χ cons*		10.6		15.7		7.0		5.5	3.6	10.2
HF/631+G**, χ relax†	74.2	10.5	78.9	14.7	43.8	5.8	52.3	5.5	4.7	9.2
B3LYP/631+G*, χ cons*	9.4		15.0			na		na	na	na
B3LYP/631+G*, χ relax†	78.3	9.2	85.7	13.6	na	na	na	na	na	na
B3LYP/631+G**, χ cons*	9.5		15.2			na		na	na	na
B3LYP/631+G**, χ relax†	78.5	9.3	85.5	13.8	na	na	na	na	na	na

$\Delta E_{ZI\text{-}B}$ and $\Delta E_{ZII\text{-}B}$ (kcal/mol) are the energy of the Z-like (Z_I and Z_{II}, respectively) conformation minus that of the B-like conformation, for a given nucleotide. $\Delta\Delta E$ is the value of $\Delta E_{ZI\text{-}B}$ (or $\Delta E_{ZII\text{-}B}$) with adenine minus the value of $\Delta E_{ZI\text{-}B}$ (or $\Delta E_{ZII\text{-}B}$) with guanine.

*After energy minimization with χ constrained (cons) to a Z_I (χ = 67.0°) or a Z_{II} (χ = 57.0°) conformation.

†After energy minimization with no constraints (relax) on χ; the values of χ after energy minimization are given in this table.

na, Not available.

acterize the barrier between *anti* and *syn* in the χ = 120° region. The present MP2/6-31G* results provide a useful estimate of the magnitude of the *syn* to *anti* intrinsic torsional energy barrier in deoxyribonucleosides, which is found to be less than 10 kcal/mol in all cases (Table 8).

The *syn* conformation in 5′-mononucleotides: adenine versus guanine

Theoretical work has long suggested that guanine favors the *syn* conformation more than adenine in nucleotides with a 5′-phosphoryl group (Olson, 1973; Yathindra and Sundaralingam, 1973), which is compatible with experiment (Son et al., 1972; Wang et al., 1979, 1984; Guschlbauer, 1972). This difference has largely been attributed to favorable interactions between the 2-amino group of guanine and the 5′-phosphoryl group, which cannot exist when the base is adenine (Olson, 1973; Yathindra and Sundaralingam, 1973). Indeed, in nucleosides, guanine does not have a higher propensity to adopt the *syn* conformation as compared with adenine (see *Syn* versus *anti* orientation of χ in nucleosides). The potential energy difference favoring the *syn* conformation in 5′-GMP as compared with 5′-AMP, however, remains elusive. To clarify this point, a series of calculations were carried out on the 5′-nucleotides of adenine and guanine, in the B DNA-like and Z_I and Z_{II} DNA-like conformations with χ constrained or allowed to relax, and at various levels of theory (Table 9). The interest in studying the Z_I and Z_{II} conformations goes beyond Z DNA, because these two conformations provide a more general model to analyze the *syn* orientation of purines, which is not limited to Z DNA (see Introduction).

The energies of the 5′-nucleotides in the Z_I and Z_{II} DNA-like conformations were calculated relative to their energy in a B DNA-like conformation (see $\Delta E_{ZI\text{-}B}$ and $\Delta E_{ZII\text{-}B}$ values in Table 9). The energies of the Z DNA-like conformations are always found to be higher than the B DNA-like reference, and this energy difference is systematically higher with adenine than with guanine. The values of $\Delta E_{ZI\text{-}B}$ range from 9.2 to 10.6 kcal/mol with adenine and from 5.6 to 7.0 kcal/mol with guanine. The values of $\Delta E_{ZII\text{-}B}$ range from 13.6 to 15.7 kcal/mol with adenine and from 5.1 to 5.5 kcal/mol with guanine. Therefore, adenine is found to destabilize the Z form of DNA significantly more than guanine, as observed experimentally (Wang et al., 1984). In addition, the present calculations suggest that adenine destabilizes the Z_{II} form significantly more than the Z_I form by approximately 5 kcal/mol (Table 9). The opposite trend is predicted with guanine, but the energy difference between the Z_I and Z_{II} conformations is clearly smaller than with adenine. The energy differences between 5′-AMP and 5′-GMP, for the Z_I or the Z_{II} conformation, and relative to a B DNA conformation, are summarized by $\Delta\Delta E$ values in Table 9. The values of $\Delta\Delta E$ suggest that the intrinsic energetics of the 5′-nucleotides destabilizes the Z_I form of DNA by ~3.5 to 5.0 kcal/mol when replacing guanine by adenine. This destabilization energy is much larger ($\Delta\Delta E \sim$ 10.0 kcal/mol) with the Z_{II} form of DNA.

The qualitative relationship between the propensity of guanine to adopt the *syn* conformation and the decreasing ability of DNA to assume the Z form as the GC content decreases has long been noted (Pohl and Jovin, 1972; Ho et al., 1986; Peticolas et al., 1988; Wang et al., 1984, 1987). The present results show more quantitatively how differences in the intrinsic energetics of the 5′-GMP and 5′-AMP building blocks contribute to GC-rich sequences being more amenable to the Z form than AT-rich sequences. Replacing cytosine with thymine is expected to have a lesser, although not negligible, influence on the equilibrium between the B and Z forms of DNA (Foloppe and MacKerell, 1999a). Therefore, in terms of the intrinsic energetics, the relative energies of 5′-GMP and 5′-AMP in *syn* are the factors that

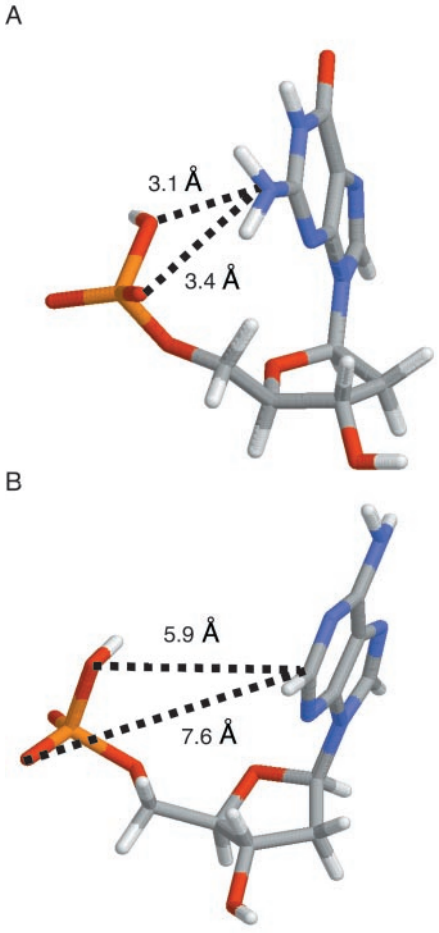

FIGURE 2 Structures of (*A*) 5′-GMP and (*B*) 5′-AMP after energy minimization at the HF/6-31+G** level with χ allowed to relax. The distances between two of the phosphoryl oxygens and the 2-amino nitrogen (5′-GMP) or carbon 2 (5′-AMP) are given along the corresponding broken lines.

dominates the higher propensity of GC-rich sequences to adopt the Z form of DNA.

Fig. 2 illustrates the interaction between the 2-amino group of guanine and the 5′-phosphoryl group, and its absence in adenine, in the structures optimized at the HF/6-31+G** level with χ relaxed. Although this interaction has long been suggested to play a key role with respect to the differences between adenine and guanine, it has not been described as involving some hydrogen bond character (Saenger, 1984; Olson, 1973; Yathindra and Sundaralingam, 1973). One may speculate that the 2-amino functionality and its hydrogens were maintained planar in earlier molecular mechanics studies (Olson, 1973; Yathindra and Sundaralingam, 1973), which may hinder formation of a hydrogen bond. Indeed, visual inspection of these structures shows that the 2-amino group is strongly pyramidalized in both the Z_I and Z_{II} conformations. This pyramidal character brings the 2-amino hydrogens closer to the phosphoryl

TABLE 10 Selected descriptors* of the interaction between the 2-amino group and the 5′-phosphoryl group in 5′-GMP

	Z_I HF/6-31+G**	Z_{II} HF/6-31+G**
Distances (Å)		
$N_2 \cdots O_3'$	3.3	3.1
$N_2 \cdots O_1$	5.3	3.4
$H_{22} \cdots O_3'$	2.4	2.7
$H_{22} \cdots O_1$	4.3	2.5
$H_{21} \cdots O_3'$	3.6	3.0
$H_{21} \cdots O_1$	5.7	3.5
Angles (°)		
$N_2—H_{22} \cdots O_3'$	149.6	98.2
$N_2—H_{21} \cdots O_3'$	63.5	81.5
$N_2—H_{22} \cdots O_1$	179.1	147.7
$N_2—H_{21} \cdots O_1$	57.8	74.0

*In energy minimized structures with χ allowed to relax, at the HF/6-31+G** level of theory.

oxygens, which presumably strengthens the favorable interactions between the two functionalities. Selected distances and angles between the 2-amino group and the 5′-phosphoryl group are listed in Table 10, which show the presence of a hydrogen bond between $N_2—H_{22}$ and O_3' in the Z_I conformation, and between $N_2—H_{22}$ and O_1 in the Z_{II} conformation. In the latter, the $N_2—H_{22} \cdots O_3'$ interaction may also be regarded as having some hydrogen bond character.

To further test the effect of the pyramidal character of the 2-amino group on the interaction between guanine and the 5′-phosphoryl group, 5-GMP was energy minimized (HF/6-31+G**) with its 2-amino group constrained to be planar in the Z_I and Z_{II} conformations with χ allowed to relax. The Z_I and Z_{II} conformations with a planar 2-amino group are 1.2 and 3.2 kcal/mol higher in energy, respectively, than their counterparts where this amino group can assume a pyramidal geometry. In the Z_{II} conformation, constraining the 2-amino group to be planar leads to significant differences in the final structures, with the $N_2 \cdots O_1$ distance increasing from 3.4 Å (pyramidal 2-amino) to 4.6 Å (planar 2-amino). This illustrates how the pyramidal character of a base amino group can play a role in stabilizing a biologically relevant conformation, such as the *syn* conformation of guanine.

It is the presence of the hydrogen bonds between the guanine 2′-amino group and the 5′-phosphoryl that prompted us to perform some calculations with polarization functions on the hydrogens because these functions can have an impact on the computed properties of hydrogen-bonded complexes (Scheiner, 1991). In the present calculations, however, this extension of the basis set has only a minor influence on the results (Table 9). The available relative energies obtained using DFT/B3LYP are also close to their HF counterpart, suggesting that the present HF results are relevant to a large extent, despite the limitations inherent to this approach (Hehre et al., 1986). This is expected to be especially true in the present context where

one only considers energy differences between analogous molecules and where these differences are relatively large.

The value of χ was allowed to relax in a number of calculations on the Z DNA-like conformations to allow for optimization of the interactions between the base and the 5′-phosphoryl group. With adenine, this led to increases in χ of approximately 7° and 22° for the Z_I and Z_{II} conformations, respectively (Table 9). An increase in χ moves the base farther away from the phosphoryl group. With guanine, χ decreases when it is relaxed by ~23° and 5° for the Z_I and Z_{II} conformations, respectively. These opposite trends in the adjustments of the χ values further illustrate the different properties of adenine and guanine with respect to the *syn* conformation in 5′-nucleotides. The largest of these adjustments are associated with changes in energy on the order of ~1.0 kcal/mol.

Conformation γ = *trans*/χ = *anti* in nucleic acid structures

Occurrence in crystal structures

As explained in the Introduction, the presence of the γ = *trans*/χ = *syn* conformation in nucleic acids has long been recognized. The possible occurrence of the γ = *trans*/χ = *anti* conformation, however, has not been clearly documented. Therefore, the available crystal structures of DNA and RNA were analyzed to determine if this conformation is relevant to nucleic acid structure. In the following, free DNA or RNA will refer to nucleic acids not complexed to a protein. The number of selected oligomers and base steps for this analysis are given in Table 1, which also lists the number and percentage of steps with γ = *trans*. Table 1 also gives the percentage of γ = *trans* associated with χ = *anti*.

The γ = *trans* conformation is uncommon, as evidenced by the highest observed percentage being 12.5% in A DNA. In free B DNA, steps with γ = *trans* are nearly nonexistent. But if B DNA is bound to a protein, the percentage of γ = *trans* becomes significant (6.5%). In this category, 35% of the γ = *trans* torsions are associated with an unusual α = *trans* torsion, while β remains in the *tran*s range. Nevertheless, the majority of the steps with γ = *tran*s include $\alpha = g^+$. Such changes lead to a more extended backbone. This is reminiscent of what was reported for the nucleosides and nucleotides bound to proteins. The conformation with γ = *trans* was qualified as "open" versus "closed" with $\gamma = g^+$, conveying the notion that an "open" nucleotide offers more possibilities for favorable interactions with the protein (Moodie and Thornton, 1993).

Although γ = *tran*s is also present in A RNA and hybrid DNA/RNA structures (Table 1), the majority of base steps with this conformation are part of mismatches or distorted structures. This is not the case, however, in A DNA. In A DNA, γ = *trans* is associated with α = *trans* and β = *tran*s. Only a marginal number of γ = *trans* torsions were found in A RNA when bound to a protein, but this may have to be reexamined as more RNA structures become available.

TABLE 11 Values* of χ (deg.) in A and B form helices of nucleic acids,† for γ in g^+ and in *trans*

Helix type	χ for $\gamma = g^+$	χ for γ = *trans*
A form	200 ± 21	189 ± 11
B form	253 ± 19	244 ± 21

*Average values.
†Includes all categories itemized in Table 1.

An interesting finding is that γ = *trans*/χ = *anti* is almost systematically associated with a sugar in the north or east range, even in the B form helices. In protein-DNA complexes, 80% of the nucleotides with γ = *tran*s have a north or east sugar, although this type of puckering is found in only 33% of the nucleotides with $\gamma = g^+$. This observation is consistent with molecular mechanics calculations suggesting that the south/γ = *tran*s combination leads to destacking of the nucleotide, and increased torsional strain in the backbone (Hartmann, unpublished results). In all categories of nucleic acids, the sugars of nucleotides with γ = *trans* cover a wide range of pseudorotation angles in the north range; the average P value is 15°, with extreme north values ($P \sim 330°$) being slightly populated. With the south/γ = *trans* combination (16 cases), the average P value is 150°, in good agreement with the calculations (see below).

Average values of χ in *anti* for γ = *trans*, when found in A or B helices are given in Table 11 and compared with the values of χ in *anti* for $\gamma = g^+$. This shows that γ = *trans* shifts χ to lower values by ~10° for both north and south puckers. This reflects the general shape of the energy profiles shown in Fig. 1, which indicate that higher values of χ in *anti* are more destabilized when γ = *trans* than when $\gamma = g^+$. This illustrates how the intrinsic energetics associated with γ = *trans*, yield χ values intermediate to those of the isolated nucleosides and canonical A or B helices. We note that the shift in χ values when γ = *trans* can be explained without any consideration of the α conformation. Thus, quantum-mechanical calculations allow for dissection of the influence of various degrees of freedom on the properties of the polymer building blocks and allows for separation of constraints due to the double helix from those imposed by the intrinsic properties of the building blocks.

Calculated properties of nucleosides with γ = trans and χ in A DNA-like or B DNA-like conformation

The values of χ associated with the energy minima in γ = *trans*/χ = *anti* (Table 5) are significantly lower than the average values of χ observed in A and B DNA (Schneider et al., 1997). In the context of nucleic acid structure, it is therefore more relevant to use the A-like and B-like values of χ as the reference state for this torsion (Foloppe and MacKerell, 1999a) rather than the energy minima found with the nucleosides. To

TABLE 12 Selected descriptors of the conformations with γ = *trans* for the nucleosides with χ in an A DNA-like* or a B DNA-like† conformation

	A-DNA like χ*			B-DNA like $\chi^{\dagger}$					
				South sugar			North sugar		
	P	γ	$\Delta E^{\ddagger}$	P	γ	$\Delta E^{\ddagger}$	P	γ	$\Delta E^{\ddagger}$
dA	10.9	186.3	4.4	152.4	181.3	6.3	33.1	188.6	6.3
dG	11.3	186.9	4.1	152.6	181.8	6.1	30.4	188.9	5.9
dC	14.4	185.4	4.7	149.7	180.7	6.6	50.1	188.6	6.4
dT	13.2	185.5	4.9	153.4	180.5	6.8	49.3	188.3	6.6

*Sugar in north and χ = 201.1°.
†Sugar in south and χ = 258.1°. See main text for the choice of these values of χ. p is the pseudorotation angle and γ is the backbone torsion.
‡Energy (kcal/mol) of the conformation with γ = *trans* minus that of the conformation with $\gamma = g^+$, for identical values of χ, obtained at the MP2/6-31G* level of theory.

be consistent with a previous theoretical study (Foloppe and MacKerell, 1999a) we use the values χ = 201.1° and χ = 258.1° to represent A and B DNA, respectively. Therefore, the deoxyribonucleosides were energy minimized in the γ = *trans*/χ = 201.1° and γ = *trans*/χ = 258.1° conformations at the MP2/6-31G* level. Given the results of the crystal survey (see Occurrence in crystal structures), the nucleosides with γ = *trans*/χ = 258.1° were energy minimized with both a north and a south pucker. The derived values of P and γ are given in Table 12. These values of P are significantly different as compared with what was obtained for the γ = *trans*/χ = *anti* energy minima (Table 3), by 9 to 39° with a north pucker and 57 to 63° with a south pucker. This again illustrates that the pseudorotation angle depends significantly on the value of χ in the *anti* range. With the DNA-like values of χ, the calculated values of P are significantly closer to their crystallographic counterparts (see Occurrence in crystal structures) than the values obtained when χ is unconstrained.

The energy differences between the γ = *trans* and $\gamma = g^+$ conformations for the A and B DNA-like values of χ are given in Table 12. These energy differences are all >4.1 kcal/mol with A-like conformations and >6.1 kcal/mol with B-like conformations. This confirms that the conformation γ = *trans* is intrinsically unfavorable in an A or B DNA context. This, to our knowledge, had not been clearly shown before based on reliable energy calculations. The lower energy difference between γ = *trans* and $\gamma = g^+$ with A-like χ values, as compared with B-like χ values, is consistent with this conformation being significantly more populated in A DNA than in B DNA (Table 1). This, however, does not explain why γ = *trans*/χ = *anti* is less populated in A RNA than in A DNA. It may be that the 2′-OH in RNA alters the intrinsic energetics of the ribonucleosides relative to that of the deoxyribonucleosides. With χ = 258.1°, the energy difference between north and south puckers is negligible. Therefore, the predominance of north puckers associated with γ = *trans*/χ = *anti* in B helices does not seem to be a consequence of the building block intrinsic energetics.

CONCLUSIONS

The flexibility about the glycosyl bond was investigated in deoxyribonucleosides for both north and south sugars, with γ in g^+ or *trans*. The $\gamma = g^+/\chi$ = *anti* conformation corresponds to a broad energy well. This energy well is broader with a south than with a north pucker, which explains the significantly narrower χ distribution in A DNA as compared with B DNA. Such differences may entropically favor the B form of DNA over the A form. This illustrates the influence of the intrinsic conformational properties of the nucleoside building blocks on nucleic acid structure. When γ changes from *trans* to g^+ the energy minimum associated with a south pucker is shifted to lower values of χ leading to similar shapes of the γ = *trans*/χ = *anti* profiles whether the pucker is north or south. The energy well associated with γ = *trans*/χ = *anti* is also broad, but this conformation is strongly destabilized as compared with $\gamma = g^+/\chi$ = *anti*. Overall, the calculations indicate that, in terms of intrinsic energetics, χ in *anti* arguably represents one of the most flexible torsions in nucleic acids. This is expected to play a role in nucleic acid dynamics and should be of interest when modeling noncanonical nucleic acid structures.

The destabilization of χ = *anti* when γ = *trans* significantly decreases the energy difference and energy barrier between the *anti* and *syn* orientations of the base, consistent with the presence of *syn* bases associated with γ = *trans* in Z DNA. In nucleosides, the path of lowest energy between *anti* and *syn* is through χ = 120° and the corresponding energy barrier was less than 10 kcal/mol. This barrier tends to be lower with purines than with pyrimidines and depends on the conformation of γ. When β in the nucleosides is constrained to *trans*, as in a DNA context, the *syn* conformation is systematically found to be higher in energy than the *anti*.

The energy difference between the *syn* conformation and the global minimum ($\gamma = g^+/\chi$ = *anti*) ranges from ~3 kcal/mol to ~8 kcal/mol, depending on the base, the sugar

pucker, and the conformation of γ. As expected, the relative energy of the *syn* conformation tends to be higher with pyrimidines than with purines, but some significant differences are observed between dA and dG as well as between dC and dT. The sugar pseudorotation angle and amplitude undergo significant adjustments when the base is in *syn*, as compared with the $\gamma = g^+/\chi = anti$ reference. In particular, the puckering amplitudes are significantly lowered in $\gamma = trans/\chi = syn$ with a north sugar. This flexibility in the furanose ring helps to explain why the *syn* conformation does not tend to be higher in energy with a north pucker than with a south pucker, in contrast to a widely accepted view (Haschmeyer and Rich, 1967; Gelbin et al., 1996).

Interestingly, the Z-DNA like $\gamma = trans/\chi = syn$ conformation with a north sugar is less stable with guanine than with adenine at the nucleoside level. With 5′-nucleotides, however, 5′-GMP favors the *syn* conformation more than 5′-AMP by ~3.5 to 5.0 kcal/mol in the Z_I conformation, and by ~10.0 kcal/mol in the Z_{II} conformation. Comparison of these nucleosides to the nucleotides shows that this is due to a favorable interaction between the guanine 2-amino group and the 5′-phosphoryl group, confirming previous conclusions (Olson, 1973; Yathindra and Sundaralingam, 1973). This interaction involves a hydrogen bond, which is significantly strengthened when the 2-amino group is pyramidalized rather than planar, which could not be observed in simpler models. The present calculations are consistent with the relative energies of 5′-GMP and 5′-AMP in *syn* being the dominant factor, which facilitates the transition to the Z form with GC-rich sequences more than with AT-rich sequences. The present results should also be useful to analyze and model the *syn* conformation of guanine when it occurs outside a Z DNA context (see Introduction).

The strong influence of γ on the nucleoside energetics when χ is *anti* led us to investigate the occurrence of the $\gamma = trans/\chi = anti$ conformation in nucleic acid crystal structures. This conformation is significantly more frequent in B DNA bound to a protein than in free DNA, and it is even more frequent in A DNA. It is scarcely observed in A RNA, but this may have to be revised as more RNA structures become available. The $\gamma = trans/\chi = anti$ conformation tends to be associated with north sugars, including in B DNA. The higher frequency of $\gamma = trans/\chi = anti$ in A DNA than in free B DNA is consistent with the intrinsic energetics in nucleosides. These energetics are also in agreement with the shift of χ toward lower *anti* values when $\gamma = trans$, and the overall predominance of the $\gamma = g^+/\chi = anti$ conformation in nucleic acids.

The present results will act as a useful guide for modeling of nucleic acids and will also be valuable toward a more accurate parametrization of nucleic acid force fields.

This work has been financially supported by National Institutes of Health grant GM51501. We also thank the NSF PACI program, DOD ASC Major Shared Resource Computing and High Performance Computing, the Pittsburgh Supercomputing Center, and NCI's Frederick Biomedical Supercomputing Center for providing computational resources. This work was also supported financially by the Swedish Natural Science Research Council, and by the Swedish Research Council for Engineering Sciences. We thank Dr. N. Banavali for useful discussions.

REFERENCES

Altona, C., and M. Sundaralingam. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides: a new description using the concept of pseudorotation. *J. Am. Chem. Soc.* 94:8205–8212.

Altona, C., and M. Sundaralingam. 1973. Conformational analysis of the sugar ring in nucleosides and nucleotides: improved method for the interpretation of proton magnetic resonance coupling constants. *J. Am. Chem. Soc.* 95:2333–2344.

Arnott, S., R. Chandrasekaran, D. L. Birdsall, A. G. W. Leslie, and R. L. Ratcliff. 1980. Left-handed DNA helices. *Nature.* 282:743–745.

Arnott, S., and D. W. L. Hukins. 1969. Conservation of conformation in mono and poly-nucleotides. *Nature.* 224:886–888.

Arnott, S., and D. W. L. Hukins. 1972. Optimized parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun.* 47:1504–1510.

Banavali, N. B., and A. D. MacKerell, Jr. 2001. Reevaluation of stereoelectronic contributions to the conformational properties of the phosphodiester and *N*-3′-phosphoramidate moieties of nucleic acids. *J. Am. Chem. Soc.* 128:6747–6755.

Baudet, S., and J. Janin. 1991. Crystal structure of a barnase-d(GpC) complex at 1.9 Å resolution. *J. Mol. Biol.* 219:123–132.

Becke, A. D. 1993. Density-functional thermochemistry: III. The role of exact exchange. *J. Phys. Chem.* 98:5648–5652.

Berman, H. M., W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.-H. Hsieh, A. R. Srinivasan, and B. Shneider. 1992. The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* 63:751–759.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer-based archival file from macromolecular structures. *J. Mol. Biol.* 122:535–542.

Berthod, H., and B. Pullman. 1971a. Molecular orbital calculations of the conformation of nucleic acids and their constituents: II. Conformational energies of β-nucleosides with C(3′)- and C(2′)-endo sugars. *Biochim. Biophys. Acta.* 232:595–606.

Berthod, H., and B. Pullman. 1971b. Molecular orbital calculations of the conformation of nucleic acids and their constituents: II. Conformational energies of β-nucleosides with C(3′)- and C(2′)-exo sugars. *Biochim. Biophys. Acta.* 246:359–364.

Berthod, H., and B. Pullman. 1973. Nucleotides: rigid or flexible? *FEBS Lett.* 30:231–235.

Brameld, K. A., and W. A. Goddard, III. 1999. Ab initio quantum mechanical study of the structures and energies for the pseudorotation of 5′-dehydroxy analogues of 2′-deoxyribose and ribose sugars. *J. Am. Chem. Soc.* 121:985–993.

Brandl, M., K. Lindauer, M. Meyer, and J. Suhnel. 1999. C-H. . . O and C-H. . . N interactions in RNA structures. *Theor. Chem. Acc.* 101: 103–113.

Cheatham, T. E., I. I. I., P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16:845–862.

Cheong, C., G. Varini, and I. Tinoco. 1990. Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. *Nature.* 346: 680–682.

Cognet, J. A. H., J. Gabarro-Arpa, M. Le Bret, G. A. van der Marel, J. H. van Boom, and G. V. Fazakerley. 1991. Solution conformation of an oligonucleotide containing a G.G mismatch determined by nuclear magnetic resonance and molecular machanics. *Nucleic Acids Res.* 19: 6771–6779.

Coste, F., J.-M. Malinge, L. Serre, W. Shepard, M. Roth, M. Leng, and C. Zelwer. 1999. Crystal structure of a double-stranded DNA containing a

cisplatin interstrand cross-link at 1.63 Å resolution: hydration at the platinated site. *Nucleic Acids Res.* 27:1837–1846.

Davies, D. B. 1978. Conformations of nucleosides and nucleotides. *Prog. Nucl. Magn. Reson. Spectrosc.* 12:135–226.

Davies, D. B., and S. S. Danyluk. 1974. Nuclear magnetic resonance studies of 5′-ribo- and deoxyribonucleotide structure in solution. *Biochemistry.* 13:4417–4434.

Davis, P. W., R. W. Adamiak, and I. J. Tinoco. 1990. Z-RNA: the solution NMR structure of r(CGCGCG). *Biopolymers.* 29:109–122.

Dickerson, R. E. 1992. DNA structure from A to Z. *Methods Enzymol.* 211:67–111.

Dickerson, R. E. 1998. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* 26:1906–1926.

Dickerson, R. E., H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini, and M. L. Kopka. 1982. The anatomy of A-, B-, and Z-DNA. *Science.* 216:475–485.

Donohue, J., and K. N. Trueblood. 1960. Base pairing in DNA. *J. Mol. Biol.* 2:363–371.

Drew, H., T. Takano, S. Tanaka, K. Itakura, and R. E. Dickerson. 1980. High-salt d(CpGpCpG), a left-handed Z'-DNA double helix. *Nature.* 286:567–573.

Florián, J., M. Strajbl, and A. Warshel. 1998. Conformational flexibility of phosphate, phosphonate, and phosphorothioate methyl esters in aqueous solution. *J. Am. Chem. Soc.* 120:7979–7966.

Foloppe, N., and A. D. MacKerell, Jr. 1998. Conformational properties of the deoxyribose and ribose moieties of nucleic acids: a quantum mechanical study. *J. Phys. Chem. B.* 102:6669–6678.

Foloppe, N., and A. D. MacKerell, Jr. 1999a. Intrinsic conformational properties of deoxyribonucleosides: implicated role for cytosine in the equilibrium among the A, B and Z forms of DNA. *Biophys. J.* 76: 3206–3218.

Foloppe, N., and A. D. J. MacKerell. 1999b. Contribution of the phosphodiester backbone and glycosyl linkage intrinsic torsional energetics to DNA structure and dynamics. *J. Phys. Chem. B.* 103:10955–10964.

Foloppe, N., L. Nilsson, and A. D. MacKerell, Jr. 2002. *Ab initio* conformational analysis of nucleic acid components: intrinsic energetic contributions to nucleic acid structure and dynamics. *Biopolymers.* In press.

Franklin, R. E., and R. G. Gosling. 1953. Molecular configuration in sodium thymonucleate. *Nature.* 171:740–741.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. 1998. Gaussian 98. Pittsburgh, PA.

Gabb, H. A., and S. C. Harvey. 1993. Conformational transitions in potential and free energy space for furanoses and 2′-deoxynucleosides. *J. Am. Chem. Soc.* 115:4218–4227.

Gelbin, A., B. Schneider, L. Clowney, S.-H. Hsieh, W. K. Olson, and H. M. Berman. 1996. Geometric parameters in nucleic acids: sugar and phosphate constituents. *J. Am. Chem. Soc.* 118:519–529.

Gu, Y., T. Kar, and S. Scheiner. 1999. Fundamental properties of the C-H···O interaction: is it a true hydrogen bond? *J. Am. Chem. Soc.* 121:9411–9422.

Guschlbauer, W. 1972. Why is guanosine different? *Jerus. Symp. Quant. Chem. Biochem.* 4:297–310.

Hahn, M., and U. Heinemann. 1993. DNA helix structure and refinement algorithm: comparison of models for (CCAGGCm^5CTGG) derived from NUCLSQ, TNT, and X-Plor. *Acta Cryst.* D49:468–477.

Hartmann, B., and R. Lavery. 1996. DNA structural forms. *Q. Rev. Biophys.* 29:309–368.

Haschemeyer, A. E. V., and H. M. Sobell. 1965. The crystal structure of a hydrogen bonded complex of deoxyguanosine and 5-bromodeoxycytidine. *Acta Crystogr.* 19:125–130.

Haschmeyer, A. E. V., and A. Rich. 1967. Nucleoside conformations: an analysis of steric barriers to rotation about the glycosidic bond. *J. Mol. Biol.* 27:369–384.

Hehre, W. J., L. Radom, P. Schleyer, and J. A. Pople. 1986. Ab initio molecular orbital theory. John Wiley and Sons, New York.

Herbert, A., and A. Rich. 1996. The biology of left-handed Z-DNA. *J. Biol. Chem.* 271:11595–11598.

Ho, P. S., M. J. Ellison, G. J. Quigley, and A. Rich. 1986. A computer-aided thermodynamic approach for predicting the formation of Z-DNA in naturally occurring sequences. *EMBO J.* 5:2737–2744.

Ho, P. S., and B. H. M. Mooers. 1997. Z-DNA crystallography. *Biopolymers.* 44:65–90.

Hocquet, A. 2001. Intramolecular hydrogen bonding in 2′-deoxyribonucleosides: an AIM topological study of the electronic density. *Phys. Chem. Chem. Phys.* 2:3192–3199.

Hocquet, A., and M. Ghomi. 2000. The peculiar role of cytosine in nucleoside conformational behaviour: hydrogen bond donor capacity of nucleic bases. *Phys. Chem. Chem. Phys.* 2:5351–5353.

Hocquet, A., N. Leulliot, and M. Ghomi. 2000. Ground-state properties of nucleic acid constituents studied by density functional calculations: 3. Role of sugar puckering and base orientation on the energetics and geometry of 2′-deoxyribonucleosides and ribonucleosides. *J. Phys. Chem. B.* 104:4560–4568.

Jones, S., P. van Heyningen, H. M. Berman, and J. M. Thornton. 1999. Protein-DNA interactions: A structural analysis. *J. Mol. Biol.* 287: 877–896.

Jordan, F., and B. Pullman. 1968. Molecular orbital calculations on the preferred conformation of nucleosides. *Theor. Chim. Acat.* 9:242–252.

Kang, C., X. Zhang, R. Ratliff, R. Moyzis, and A. Rich. 1992. Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature.* 356: 126–131.

Kuramoto, N., S. Nishikawa, and F. Jordan. 1998. Reexamination of ultrasonic relaxation kinetics of aqueous solutions of nucleosides: evidence for fast syn-anti glycosyl isomerization in adenosine 5′-monophosphate and adenosine 5′-diphosphate. *J. Phys. Chem. B.* 102: 9181–9186.

Lakshminarayanan, A. V., and V. Sasisekharan. 1969. Stereochemistry of nucleic acids and polynucleotides: IV. Conformational energy of base-sugar units. *Biopolymers.* 8:475–488.

Lakshminarayanan, A. V., and V. Sasisekharan. 1970. Stereochemistry of nucleic acids and polynucleotides: II. Allowed conformations of the monomer unit for different ribose puckerings. *Biochim. Biophys. Acta.* 204:49–59.

Lebrun, A., and R. Lavery. 1996. Modeling extreme stretching of DNA. *Nucleic Acids Res.* 24:2260–2267.

Lebrun, A., and R. Lavery. 1997. Unusual DNA conformations. *Curr. Opin. Struct. Biol.* 7:348–354.

Leslie, A. G. W., S. Arnott, and R. Chandrasekaran. 1980. Polymorphism of DNA double helices. *J. Mol. Biol.* 143:49–72.

Leulliot, N., M. Ghomi, G. Scalmani, and G. Berthier. 1999a. Ground state properties of the nucleic acid constituents studied by density functional calculations: I. Conformational features of ribose, dimethyl phosphate, uridine, cytidine 5′-methyl phosphate-uridine:and3′-methyl phosphate-uridine. *J. Phys. Chem. A.* 103:8716–8724.

Leulliot, N., M. Ghomi, G. Scalmani, and G. Berthier. 1999b. Ground state properties of the nucleic acids constituents studied by density functional calculations: I. Conformational features of ribose, dimethyl phosphate, uridine, cytidine 5′-methyl phosphate-uridine:and3′-methyl phosphate-uridine. *J. Phys. Chem. A.* 103:8716–8724.

Lüger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 389:251–257.

Malinina, L., L. Urpí, X. Salas, T. Huynh-Dinh, and J. A. Subirana. 1994. Recombination-like structure of d(CCGCGG). *J. Mol. Biol.* 243: 484–493.

Moodie, S. L., and J. M. Thornton. 1993. A study into the effects of protein binding on nucleotide conformation. *Nucleic Acids Res.* 21:1369–1380.

Nishikawa, S., H. Huang, and F. Jordan. 2000. Structural effects of nucleotides on syn-anti glycosyl isomerization kinetics by ultrasonic relaxation methods. *J. Phys. Chem. B.* 104:1391–1394.

Nishikawa, S., H. Kuramoto, H. Huang, and F. Jordan. 1999. Kinetic and thermodynamic study of syn-anti isomerization in aqueous solutions of AMP, ADP and ATP by ultrasonic relaxation methods. *J. Phys. Chem. B.* 103:3754–3757.

Olson, W. K. 1973. Syn-anti effects on the spatial configuration of polynucleotides chains. *Biopolymers.* 12:1787–1814.

Olson, W. K. 1982. How flexible is the furanose ring? 2. An updated potential energy estimate. *J. Am. Chem. Soc.* 104:278–286.

Olson, W. K., and P. J. Flory. 1972a. Spatial configuration of polynucleotide chains: I. Steric interactions in polyribonucleotides: a virtual bond model. *Biopolymers.* 11:1–23.

Olson, W. K., and P. J. Flory. 1972b. Spatial configuration of polynucleotides chains: II. Conformational energies and the average dimensions of polyribonucleotides. *Biopolymers.* 11:25–56.

Ortiz-Lombardia, M., A. González, R. Eritja, J. Aymamí, F. Azorín, and M. Coll. 1999. Crystal structure of a DNA Holliday junction. *Nat. Struct. Biol.* 6:913–917.

Peticolas, W. L., Y. Wang, and G. A. Thomas. 1988. Some rules for predicting the base-sequence dependence of DNA conformation. *Proc. Natl. Acad. Sci. U.S.A.* 85:2579–2583.

Pohl, F. M., and T. M. Jovin. 1972. Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.* 67:375–396.

Rao, S. T., and M. Sundaralingam. 1970. Stereochemistry of nucleic acids and their constituents: XIII. The crystal and molecular structure of 3′-*O*-acetyladenosine:conformational analysis of nucleosides and nucleotides with syn glycosidic torsion angle. *J. Am. Chem. Soc.* 92: 4963–4970.

Rhodes, L. M., and P. R. Schimmel. 1971. Nanosecond relaxation processes in aqueous mononucleoside solutions. *Biochemistry.* 10: 4426–4433.

Roll, C., C. Ketterle, G. V. Fazakerley, and Y. Boulard. 1999. Solution structures of a duplex containing an adenine opposite a gap (absence of one nucleotide). An NMR study and molecular dynamics simulations with explicit water molecules. *Eur. J. Biochem.* 264:120–131.

Rubin, J., T. Brennan, and M. Sundaralingam. 1972. Crystal and molecular structure of a naturally occuring dinucleoside monophsopshate: uridylyl-(3′-5′)-adenosine hemihydrate: conformational "rigidity" of the nucleotide unit and models for polynucleotide chain folding. *Biochemistry.* 11:3112–3128.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York.

Saenger, W., and K. H. Scheit. 1970. A pyrimidine nucleoside in the synconformation: molecular and crystal structure of 4-thiouridine-hydrate. *J. Mol. Biol.* 50:153–169.

Saran, A., D. Perahia, and B. Pullman. 1973. Molecular orbital calculations of the conformation of nucleic acids and their constituents: VII. Conformation of the sugar ring in β−nucleosides: the pseudorotational representation. *Theor. Chim. Acta.* 30:31–44.

Saran, A., B. Pullman, and D. Perahia. 1972. Molecular orbital calculations on the conformation of nucleic acids and their constituents: IV. Conformation about the exocyclic C(4′)-(C5′) bond. *Biochim. Biophys. Acta.* 287:211–231.

Scheiner, S. 1991. Calculating the Properties of Hydrogen Bonds by *ab Initio* Methods: Reviews in Computational Chemistry, vol. 2. K. B. Lipkowitz and D. B. Boyd, editors. VCH Publishers, Inc. 165–210.

Schneider, B., S. Neidle, and H. M. Berman. 1997. Conformation of the sugar-phosphate backbone in helical DNA crystal structures. *Biopolymers.* 42.

Schwartz, T., M. A. Rould, K. Lowenhaupt, A. Herbert, and A. Rich. 1999. Crystal structure of the Zα domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science.* 284:1841–1845.

Shishkin, O. V., A. Pelmenshikov, D. M. Hovorun, and J. Leszczynski. 2000. Molecular structure of free canonical 2′-deoxyribonucleosides: a density functional study. *J. Mol. Struct.* 526:329–341.

Skelly, J. V., K. J. Edwards, T. C. Jenkins, and S. Neidle. 1993. Crystal structure of an oligonucleotide duplex containing G-G base pairs: influence of mispairing on DNA backbone conformation. *Proc. Natl. Acad. Sci. U.S.A.* 90:804–808.

Smith, F. W., and J. Feigon. 1992. Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature.* 356:164–168.

Son, T.-D., W. Guschlbauer, and M. Guéron. 1972. Flexibility and conformations of guanosine monophosphates by the overhauser effect. *J. Am. Chem. Soc.* 94:7903–7911.

Steiner, T. 2000. Influence of C-H. . . O interactions on the conformation of methyl groups quantified from neutron diffraction data. *J. Phys. Chem. A.* 104:433–435.

Sundaralingam, M. 1969. Stereochemistry of nucleic acids and their constituents: IV. Allowed and preferred conformations of nucleosides, nucleosides mono-, di-, tetraphosphates, nucleic acids, and polynucleotides. *Biopolymers.* 7:821–860.

Sussman, J. L., and N. C. Seeman. 1972. Crystal structure of a naturally occurring dinucleoside phosphate: uridylyl 3′-5′-adenosine phosphate model for RNA chain folding. *J. Mol. Biol.* 66:403–421.

Sutor, D. J. 1962. The C-H. . . O hydrogen bonds in crystals. *Nature.* 195:68–69.

Takahara, P. M., A. C. Rosenweig, C. A. Frederick, and S. J. Lippard. 1995. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature.* 377:649–652.

Tavale, S. S., and H. M. Sobell. 1970. Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the *syn* conformation. *J. Mol. Biol.* 48:109–123.

Taylor, R., and O. Kennard. 1982. Crystallographic evidence for the existence of C-H. . . O, C-H. . . N, and C-H. . . Cl hydrogen bonds. *J. Am. Chem. Soc.* 104:5063–5070.

Vargas, R., D. Garza, A. Dixon, and B. P. Hay. 2000. How strong is the Ca-H. . . O = C hydrogen bond? *J. Am. Chem. Soc.* 122:4750–4755.

Wahl, M. C., and M. Sundaralingam. 1997. C-H···O hydrogen bonding in biology. *TIBS.* 22:97–102.

Wang, A. H.-J., R. V. Gessner, G. A. van der Marel, J. H. van Boom, and A. Rich. 1985. Crystal structure of Z-DNA without an alternating purine-pyrimidine sequence. *Proc. Natl. Acad. Sci. U.S.A.* 82: 3611–3615.

Wang, A. H.-J., T. Hakoshima, G. van der Marel, J. H. van Boom, and A. Rich. 1984. AT base pairs are less stable than GC base pairs in Z DNA: the crystal structure of d(m5CGTAm5CG). *Cell.* 37:321–331.

Wang, A. H.-J., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature.* 282:680–686.

Wang, A. H. J., G. J. Quigley, F. J. Kolpak, G. van der Marel, J. H. van Boom, and A. Rich. 1981. Left-handed double helical DNA: variations in the backbone conformation. *Science.* 211:171–176.

Wang, Y., G. A. Thomas, and W. L. Peticolas. 1987. Sequence dependent conformations of oligomeric DNA's in aqueous solutions and in crystals. *J. Biomol. Struct. Dyn.* 5:249–274.

Wilson, H. R., and A. Rahman. 1971. Nucleoside conformation and nonbonded interactions. *J. Mol. Biol.* 56:129–142.

Yathindra, N., and M. Sundaralingam. 1973. Conformational studies on guanosine nucleotides and polynucleotides: the effect of the base on the glycosyl and backbone conformations. *Biopolymers.* 12:2075–2082.