Flexibility of the B-DNA backbone: effects of local and neighbouring sequences on pyrimidine–purine steps

H.-O. Bertrand, T. Ha-Duong[1], S. Fermandjian and B. Hartmann[1],*

Département de Biologie et Pharmacologie Structurales, Institut Gustave Roussy, URA CNRS 147, 39 rue Camille Desmoulins, 94805 Villejuif Cedex, France and [1]Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, UPR CNRS 9080, 13 rue Pierre et Marie Curie, 75005 Paris, France

ABSTRACT

The structurally correlated dihedral angles ε and ζ are known for their large variability within the B-DNA backbone. We have used molecular modelling to study both energetic and mechanical features of these variations which can produce BI/BII transitions. Calculations were carried out on DNA oligomers containing either YpR or RpY dinucleotides steps within various sequence environments. The results indicate that CpA and CpG steps favour the BI/BII transition more than TpA or any RpY step. The stacking energy and its intra- and inter-strand components explain these effects. Analysis of neighbouring base pairs reveals that BI/BII transitions of CpG and CpA are easiest within $(Y)_n(R)_n$ sequences. These can also induce a large vibrational amplitude for TpA steps within the BI conformation.

INTRODUCTION

Although B-DNA fine structure has generated much interest in recent years, the link between the B-DNA backbone and base sequence has been largely overlooked. However, several groups, using ^{31}P NMR studies, have claimed that the phosphodiester backbone can play a major role, notably in protein–DNA recognition. It has been demonstrated that sequence-dependent structural variations along the DNA backbone can influence recognition by DNA binding proteins (1–3). The backbone perturbations induced by proteins suggest that both the conformation and the flexibility of the phosphodiester backbone may also be important for DNA–protein recognition (4–6). Studies such as those of Gorenstein *et al.* (1–3) and El Antri *et al.* (7) have highlighted effects of base composition on the conformation of the phosphodiester backbone.

Analysis of crystallized B-DNA structures shows that the backbone dihedral angles ε (C4′-C3′-O3′-P) and ζ (C3′-O3′-P-O5′) are those which exhibit the largest variability (8). The BI conformation (clearly the most common) and the BII conformation are characterized by ε and ζ values corresponding to (t/g-) and (g-/t) respectively. Since changes in ε and ζ are correlated, the conformation of the phosphodiester backbone of a given dinucleotide junction can be characterized by the

difference ($\varepsilon - \zeta$) which is roughly –90° for the BI conformation and roughly +90° for the BII conformation (9). According to crystallographic data (10–13) and molecular modelling (14) BII conformations can decrease the width of the minor groove, decrease the stacking of the bases and may generate kinked structures. These significant structural modifications may have biological implications.

BII conformations have been thoroughly analysed by crystallography. According to certain authors BII conformations are favoured by crystal packing effects (11,15). However, in solution NMR studies provide evidence for BII sites. Examples include the double mismatch G:A (16–18) and the CpG step contained in the CTTCGAAG octamer (19).

Semi-empirical and *ab initio* molecular orbital chemical shift calculations have suggested that ^{31}P chemical shifts should be roughly –4.6 p.p.m. for the BI conformation and roughly –3.00 p.p.m. for BII conformations (8). Experimentally, ^{31}P chemical shifts are found between these extreme values and it is not yet clear whether they characterize a BI/BII equilibrium or reflect the large variability of the ε and ζ dihedral angles within the BI conformation. Thus, in agreement with crystallographic data (20), the value of –4.1 p.p.m. for the TpA step within the GTACGTAC sequence may indicate a BI conformational substate with a high twist value and an ($\varepsilon - \zeta$) of around –30°, rather than the presence of a small amount of the BII conformation (7,21). However, a chemical shift of –3.9 p.p.m. for the central CpG within CTTCGAAG shows all the signs of a BI/BII conformational equilibrium, giving to its large H2′-H6/H8 internucleotide distance (19).

In an attempt to assess the impact of DNA sequence on both the mechanics, the energetics of the BI/BII transition and on flexibility of the phosphodiester backbone we have carried out a systematic energy mapping and normal mode study of several repeating sequences as well as of six octamers. Oligonucleotides were chosen to contain YpR dinucleotide steps which are suspected to show a greater propensity to undergo the BI/BII transition. These steps are also contrasted with RpY steps, which favour the BI form.

MATERIALS AND METHODS

Calculations were performed with the JUMNA 10.00 program (22–23), which uses a combination of helical and internal

---

*To whom correspondence should be addressed. Tel: +33 1 43 25 26 09; Fax: +33 1 43 29 56 45; Email: brigitte@ibpc.fr

variables to treat the flexibility of nucleic acids. Dihedral and valence angles are used to model the flexibility of each nucleotide, while the nucleotides are positioned in space using helical rotations and translations with respect to a reference system. All bond lengths are kept fixed and the junctions between successive 3′-monophosphate nucleotides, as well as the closure of the sugar rings, are ensured using quadratic restraints on the O5′-C5′ and C4′-O4′ distances. This representation leads to the use of roughly 10 times fewer variables than Cartesian coordinate molecular mechanics and, thus, speeds up and facilitates the conformational search. Solvent effects were modelled by a simple distance-dependent dielectric function of sigmoidal form and counterion damping was dealt with by a reduction of the total phosphate charge to –0.5 e. While this is a crude representation of the solvent effects, it nevertheless appears to be well adapted to model DNA oligomers for which previous work has shown good agreement with experimental data (19,21,24,25).

The sequences studied are given in Table 1. They were chosen to contain dinucleotide steps belonging to the YpR family (CpA ≡ TpG; CpG; TpA) or to the RpY family (ApC ≡ GpT; GpC; ApT) in various sequence environments.

Table 1. DNA sequences studied

Repeated dinucleotides	Repeated tetramers	Octamers
ATATATATATAT	ACGTACGTACGT	CTTCGAAG
	AGCTAGCTAGCT	CATCGATG
GCGCGCGCGCGC		CTACGTAG
	TGCATGCATGCA	CTTTAAAG
ACACACACACAC	TCGATCGATCGA	CATTAATG
		CTATATAG

For each sequence the conformational space was explored by scanning the sugar phase angles from 130 to 200°. We have previously demonstrated that this approach enables us to locate all the stable B-family conformations of such sequences (19,26–28). In some cases several distinct substates were found, but they were always separated by <1 kcal/mol. Starting from each of these minima, BI/BII transitions were induced by scanning the dihedral angle difference (ε – ζ) between –130 and +160°, in steps of 10°. BI/BII transitions were generated in both strands, either at the same dinucleotide level or at two different levels. Therefore, in both cases the energies presented refer to two concurrent transitions.

Both model and crystallographic structures were analysed with the CURVES program (29–30), which provides a rigorous way to obtain the locus of an overall helical axis for irregular nucleic acids. All helical parameters were calculated with respect to this helical axis and are thus global parameters. They obey the Cambridge convention for DNA conformation (31). Results from our modelling study were compared with crystallographic data of 13 B-DNA decamers (10–13,15,33–39) which were each resolved to at least 2.0 Å. Below this resolution backbone features become very dependent on the refinement procedure (10,32). Junctions perturbed by mismatches (G:A; 10) and unusual bases (such as oxoguanine; 36) were not taken into account. Our results were also compared with seven structures we have already resolved by NMR (19,21,27,40; Tisné *et al.*, in preparation) and to nine from the Protein Data Bank (41–46).
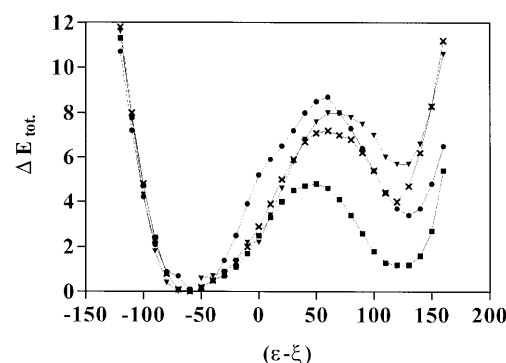

Figure 1. Total energy variation for the BI/BII transition (kcal/mol) as a function of (ε − ζ) values (degrees) for CpG, TpA and ApT dinucleotide steps contained in the respective sequence oligomers (CG)$_n$ and (TA)$_n$. CpG (■) and ApT (●) indicate two BI/BII transitions at the same dinucleotide level; CpG (×) and TpA (◆) indicate two BI/BII transitions at different dinucleotide levels.

The normal mode calculations and their analysis were performed by the methodology described in a previous paper (47). The trajectory of each normal mode is constructed within the limits of thermal agitation (0.6 kcal/mol at 300 K) and the series of structures generated are analysed in terms of helicoidal parameters using the CURVES program (29–30). The modes were then automatically scanned to detect significant backbone angle flexibility and to detect maximum variations of the helicoid parameters for individual base pairs and base pair steps.

RESULTS AND DISCUSSION

Energetics of the BI/BII transition as a function of the dinucleotide step involved

Total transition energy. Whatever the dinucleotide sequence and the type of BI/BII transition considered (i.e. either creating two BII sites facing each other or two staggered BII sites) the BI conformation was always found to be more stable. The variability in (ε – ζ) values are, however, somewhat larger for BI minima and these minima also appeared to be more flexible, since for an energy cost of 2 kcal/mol around the BI minima the (ε – ζ) difference covers a range of 60–80°, whereas around the BII minima this difference covers a range of 40–50°. This corroborates the crystallographic data where the BI steps have (ε – ζ) values of –80° ± 40°, while BII steps have values of +80° ± 30°.

For the oligomers presented in Table 1 it was always easier to generate two BII sites facing each other than two staggered BII sites. This is illustrated in Figure 1, where the total energy variation of some dinucleotide steps is presented as a function of (ε – ζ). As an example, the average total transition energy of CpG is 4.8 kcal/mol when the two BII sites are facing each other and 7.1 kcal/mol for two staggered BII sites.

The results presented in Figure 1 indicate that for both types of BI/BII transition the relative energies of the BI and BII minima and the transition energy barriers reflect clear sequence effects. However, for (ε – ζ) values either below –60° or above 140° no sequence effects are observed and the increase in energy is mainly due to backbone steric clashes. For (ε – ζ) values between –60 and –120° the distance between the H2′ proton of the 5′-sugar and the O5′ oxygen of the phosphate group becomes <3 Å, with an
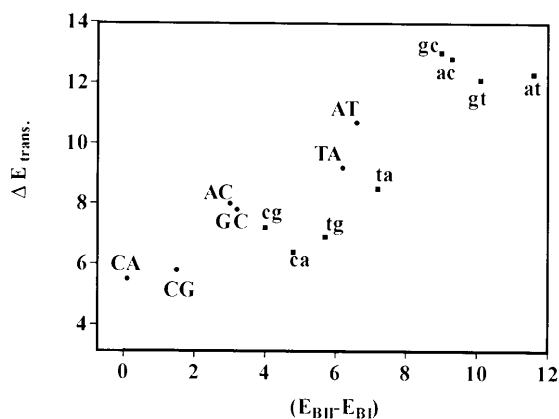
Figure 2. Variation in total transition energy (kcal/mol) as a function of internal energy variation between the BI and BII conformations (kcal/mol). For a given dinucleotide step values were obtained by averaging all the possible transitions within the oligomers studied. Capital letters indicate two BI/BII transitions at the same dinucleotide level, lower case letters indicate two BI/BII transitions at different dinucleotide levels.

Table 2. Analysis of the backbone conformation of the YpR and RpY dinucleotide steps contained in 13 crystallographic B-DNA decamers resolved to at least 2.0 Å (10–13,15,33–39)

Dinucleotide step		Occurrence	Number of BII sites $(\varepsilon - \zeta) > 50°$
YpR	TpA	6	1
	CpG	24	11
	CpA	18	16
	TpG	18	16
RpY	GpC	7	0
	ApT	24	0
	ApC	1	0
	GpT	1	0

associated unfavourable Lennard–Jones energy. For $(\varepsilon - \zeta)$ values above ~135° the distance between the H2′ proton of the 5′-sugar and the O3′ oxygen of the phosphate group becomes very short (~2.5 Å). Keeping that distance constant requires a significant repuckering of the sugar (essentially a large increase in amplitude), which is again energetically unfavourable.

All dinucleotide sequence effects are summarized in Figure 2. For a given dinucleotide junction all the possible BI/BII transitions within the various oligomers studied were carried out. The average values of the total transition energy variation ($\Delta E_{transition}$) are reported versus the energy difference between BI and BII conformations ($E_{BII} - E_{BI}$). The smaller the transition barrier and the energy difference ($E_{BII} - E_{BI}$), the easier is the BI/BII transition. This process shows a clear sequence dependence: CpA and CpG appear to be the most malleable dinucleotide steps. Such effects are also reflected in the crystallographic data. An analysis of backbone parameters for all the YpR and RpY steps available is reported in Table 2. It confirms that CpA/TpG (90% BII) and CpG (50% BII) are most likely to adopt the BII conformation. In contrast, no BII site was observed for either the ApT and GpC step and only one BII site was found for a TpA step.

We note that for CpA and CpG steps the values of the backward transition barriers and those of the forward transition are of the same order of magnitude. In contrast, for more rigid steps barrier values of the backward transition are significantly smaller. These theoretical results are consistent with NMR data which also show that CpA and CpG steps adopt a BI/BII conformational equilibrium (19,49; Tisné *et al.*, in preparation).

In summary, the analysis of the total transition energy provides strong evidence for clear sequence effects; CpA and CpG steps exhibit a different behaviour from the RpY steps and also with respect to TpA steps. Decomposition of the total transition energy helps in our understanding of these effects.

Decomposition of the total transition energy. Let us now consider the variation in the total transition energy during the BI/BII transition. The energy variation of each oligonucleotide studied is essentially localized at the dinucleotide step where the transition occurs. When the two BII sites are generated facing one another the energy of the neighbouring nucleotides stays virtually constant. When the two BII sites are generated at different dinucleotides levels the perturbation is somewhat more delocalized to include the closest neighbours. It also appears that the variation in total transition energy mainly reflects variations in intra-strand energy. Within this intra-strand energy two components vary together.

(i) The torsional energy, which reaches its maximum for $(\varepsilon - \zeta)$ $= +50°$. Taking into account variations in the dihedral angle β, the change in torsional energy lies between 2.5 and 3.7 kcal/mol.

(ii) The intra-strand base–base interaction energy.
Since sequence effects cannot be directly explained by variation in the torsional energy, which is only dependent on dihedral angle, it seems logical to assess the role of stacking energy.

Figure 3 shows the average total energy barrier versus the stacking energy barrier. A relatively good correlation is seen. Decomposition of the total stacking energy into its intra- and inter-strand components shows that the principal change involves a loss of the intra-strand stacking energy. For the TpA, ApT, ApC and GpC dinucleotide steps this component is largely dominant. Calculated average stacking energies of the model BI dinucleotide steps can be compared with the calculated average stacking energies of various BI dinucleotide steps deduced from crystallographic or NMR structures (stacking energies were simply re-calculated using the JUMNA 10.00 program without any energy minimization). Results are given in Table 3. Although the total stacking energy does not fully account for sequence effects, its decomposition into inter- and intra-strand components shows that CpA and CpG steps possess weaker intra-strand stacking compared with any other dinucleotide steps. As shown below, a BI/BII transition involves displacement of two consecutive bases and, hence, the weaker the initial intra-strand stacking, the easier the transition becomes. When the BI/BII transition concerns staggered dinucleotide steps the loss of intra-strand stacking energy is found to be larger and the loss of inter-strand stacking also becomes significant. This is probably due to difficulties in moving facing bases while their
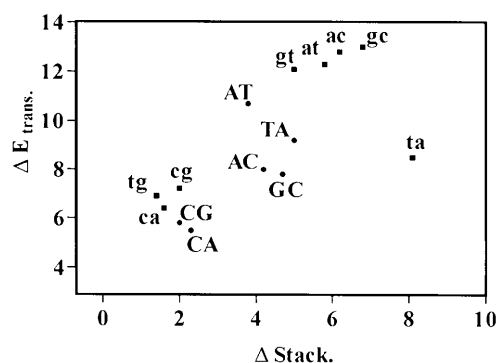
Figure 3. Variation in total transition energy (kcal/mol) as a function of stacking energy variations (kcal/mol). For a given dinucleotide step values were obtained by averaging all the possible transitions within the oligomers studied. Capital letters indicate two BI/BII transitions at the same dinucleotide level, lower case letters indicate two BI/BII transitions at different dinucleotide levels.

backbone remains in the BI conformation. As indicated in Figure 3, this is particularly true for TpA, ApT, ApC and GpC.

Table 3. Total, intra- and inter-strand stacking energies for the dinucleotide steps of interest

Dinucleotide step	Stacking energy (kcal/mol)		
	Total	Intra-strand	Inter-strand
CA–MM	–8.8	–1.3	–7.5
CA–Cryst	–8.0[a]	–1.5[a]	–6.5[a]
CA–NMR	–9.4	–1.6	–7.8
CG–MM	–7.6	–0.2	–7.5
CG–Cryst	–9.2	–3.6	–5.5
CG–NMR	–7.0	–0.1	–6.9
TA–MM	–9.7	–8.8	–0.9
TA–Cryst	–6.4	–5.9	–0.5
TA–NMR	–7.4	–7.6	0.2
AC–MM	–10.4	–7.7	–2.7
AC–Cryst	–10.4	–6.6	–3.8
AC–NMR	–10.4	–7.4	–2.9
AT–MM	–9.9	–9.8	–0.2
AT–Cryst	–8.1	–8.2	0.1
AT–NMR	–5.1	–8.6	3.4
GC–MM	–13.6	–18.2	4.7
GC–Cryst	–9.5	–12.2	2.8
GC–NMR	–10.7	–17.0	6.4

MM, values deduced from modelled structures; Cryst and NMR, values deduced from crystallographic and NMR structures.
[a]All the available CpA steps exhibit BII conformations.

It is worth noting that the total stacking energy has only been considered in a few theoretical studies, with often contradictory results. Usually TpA steps are thought to be as weakly stacked as CpG or CpA steps. Calculations by Hunter (48), using a model of
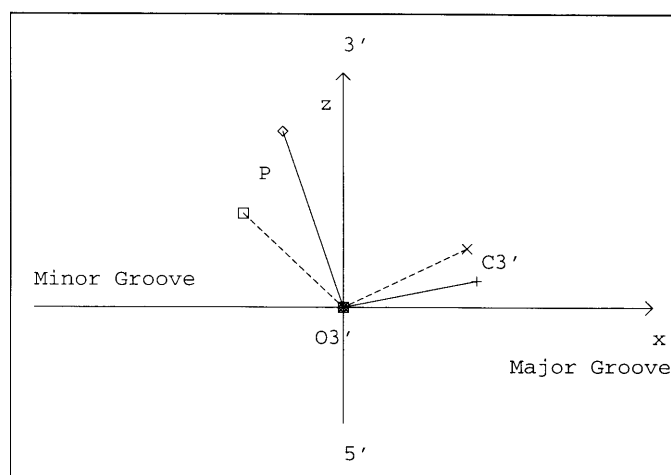

Figure 4. C3′-O3′ and O3′-P vector projections (standard deviation ±5°) onto the *xz* plane (as defined in 31) for the BI minima (solid line) and the BII minima (dashed line).

π–π interactions, predict YpR dinucleotide steps to be more weakly stacked than RpY ones. In contrast, we find TpA steps more strongly stacked than CpA or CpG steps. They exhibit an intermediate behaviour between either CpA or CpG and RpY steps. While our calculations do not use a special model for base stacking, analysis of micro-calorimetry data obtained from large DNA fragments (up to 200 bp) (50) confirms this result. Moreover, micro-calorimetric stacking values demonstrate a large variability for TpA steps. This is in agreement with crystallographic data (20) and with our previous calculations (51).

Thus BI/BII transitions are dependent on the dinucleotide sequence when they occur, an effect which appears to be governed mainly by the intra-strand stacking energies. Since CpG and CpA steps exhibit the weakest intra-strand stacking, they undergo BI/BII transitions most easily.

The destacking mechanism associated with BI/BII transitions

BI/BII transitions involve the variation in helical parameters, particularly twist and slide, and generate both a negative roll and a positive X_{disp} (8,14). From a mechanical point of view this can be geometrically explained by projecting the C3′-O3′ and O3′-P vectors onto the *xz* plane (the *x*-, *y*- and *z*-axes are defined according to the Cambridge convention for DNA conformation; 31). Note that the two bonds chosen correspond to the ε and ζ dihedral angles respectively. Figure 4 shows the average values of the projected C3′-O3′ and O3′-P vectors corresponding to the BI and BII minima (the standard deviation of these orientations is 5°).

In the BI conformations the C3′-O3′ vector projection is roughly parallel to the *x*-axis (i.e. the base pair pseudo-dyad axis) and the O3′-P vector projection is roughly parallel to the *z*-axis (i.e. the helical axis). Changing the dihedral angle ε involves a scissor movement of two consecutive bases almost in the *y* direction (i.e. the long axis of the base pair; see Fig. 4) and generates a slide. Changing the dihedral angle ζ involves the contra-rotation of two consecutive bases almost around the *z*-axis (see Fig. 4) and generates a twist. As long as ε and ζ move in the t/g- domain, which corresponds to the first part of the BI/BII
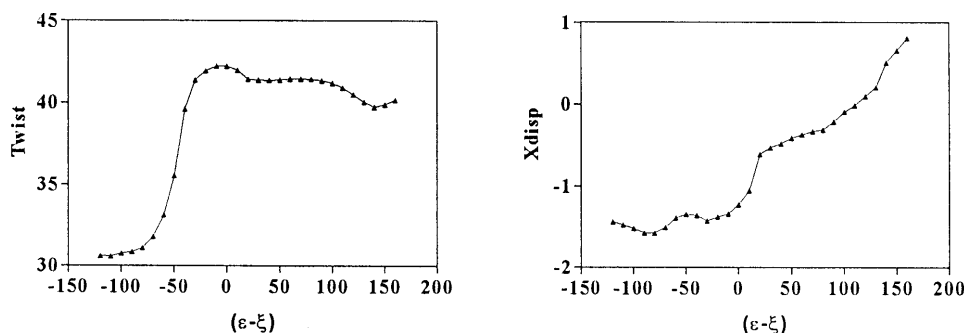
Figure 5. Twist and X_{disp} variations (in degrees and Å respectively) as a function of $(\varepsilon - \zeta)$ values (degrees) for the CpG dinucleotide step.

transition, the slide and the twist increase and intra-strand stacking decreases. This fact provides a simple explanation of the well-known correlation between slide and twist linked to ε/ζ (12,28,51). For BII conformations the projections of both the C3′-O3′ and O3′-P vectors undergo a rotation around the *y*-axis (see Fig. 4), but the O3′-P vector also rotates around the *x*-axis, thus the *x* component decreases and the *y* component becomes significant. Such a spatial re-orientation leads to a negative roll (rotation around the *y*-axis) and a positive X_{disp} (translation along the *x*-axis towards the major groove).

The variation in twist and X_{disp} versus $(\varepsilon - \zeta)$ for the CpG dinucleotide step is shown in Figure 5. As mentioned above, the twist values increase during the first part of the transition and stay constant during the second part, whereas the X_{disp} values remain constant in the first part of the transition and then increase.

Table 4. Total transition energies and relative stabilities (kcal/mol) of the octamers studied

Octamers	$\Delta E_{transition}$	$E_{BII} - E_{BI}$
CTT**CG**AAG	5.6	0.0
CTA**CG**TAG	6.2	3.2
CAT**CG**ATG	7.2	2.8
CTT**TA**AAG	8.9	7.6
CTA**TA**TAG	9.4	7.7
CAT**TA**ATG	9.6	8.1

These values are related to a double BI/BII transition at the same dinucleotide level (i.e. to the central YpR step).

Neighbouring sequence effects

BI/BII transition. Previous NMR data have shown that CpA and CpG dinucleotide steps were very sensitive to the neighbouring sequences (19,49; Tisné *et al.*, in preparation). In the present study the variability in total transition energy is found to be greater for CpA and CpG (±1.8 kcal/mol) compared with the other dinucleotide steps (±0.8 kcal/mol). To learn more about such sequence effects we studied the CpG and TpA dinucleotide steps at the centre of the six octamers shown in Table 1. CpG, which has been extensively studied by NMR spectroscopy (7,14,19,21,40), is compared with TpA, which displays a RpY-like behaviour. The results are summarized in Table 4.

Neighbouring sequence effects within the studied octamers are revealed by the variability in both the transition barriers and the energy differences between the BII and BI conformations. The corresponding energy ranges being: CpG, 1.6 and 3.2 kcal/mol; TpA, 0.7 and 0.5 kcal/mol. According to these values the CpG step appears to be much more sensitive to the flanking sequences than TpA. This is partially explained by the value of the stacking energy of the CpG BI minimum within the CTT**CG**AAG environment, which is 1 kcal/mol weaker than for the other CpG BI minima. Within this octamer it is also found that the energy variation is completely localized at the CpG dinucleotide step. The rigidity of the TpT, TpC and CpT dinucleotide steps has already been reported by other authors (52,53). Here we find that BI/BII transitions are easier when CpG is preceded by a pyrimidine tract and followed by a purine tract. Analysis of NMR and crystallographic data provides evidence for CpA and CpG steps with BII conformations occurring in Y_n(**CA/CG**)R_n sequences. Although TpA belongs to the YpR family, it does not easily undergo BI/BII transitions. The question of how TpA behaves within a $(Y)_n$**TA**$(R)_n$ environment remains open.

Normal modes. TpA steps differ from other YpR steps as they display a low propensity to undergo BI/BII transitions. As shown by several NMR studies (54,55), they also display a particular dynamic behaviour in T_nA_n sequences ($n \geq 2$). To investigate the sequence-dependent flexibility of TpA with respect to CpG we performed an internal coordinate normal mode analysis of the six octamers given in Table 1 in their BI conformation. For each sequence we searched for the maximal vibrational amplitudes of the backbone angles and the helical parameters versus the normal mode frequencies.

Results indicate that ζ is the most flexible backbone angle, as its vibrational amplitude is on average ~15°. Next come the angles β (~8°) and α (~6°), followed by ε (~5°) and γ (~5°). The difference $(\varepsilon - \zeta)$ presents an average vibrational amplitude of 17°. Figure 6a indicates that TpA is the most flexible step within the CTT**TA**AAG sequence, particularly in terms of ε, ζ and β angle vibrations. The properties of the TpA step are particular within the TTTAAA tract and are unlike those of any other sequences studied to date (56). In contrast, the backbone vibrational amplitudes of CpG in the CTT**CG**AAG sequence appear only slightly higher compared with those displayed by this step in the two other sequences studied. Thus $(Y)_n(R)_n$ flanking sequences exert strong effects on the central YpR dinucleotide steps, either in terms of backbone vibrations, for TpA, or in terms
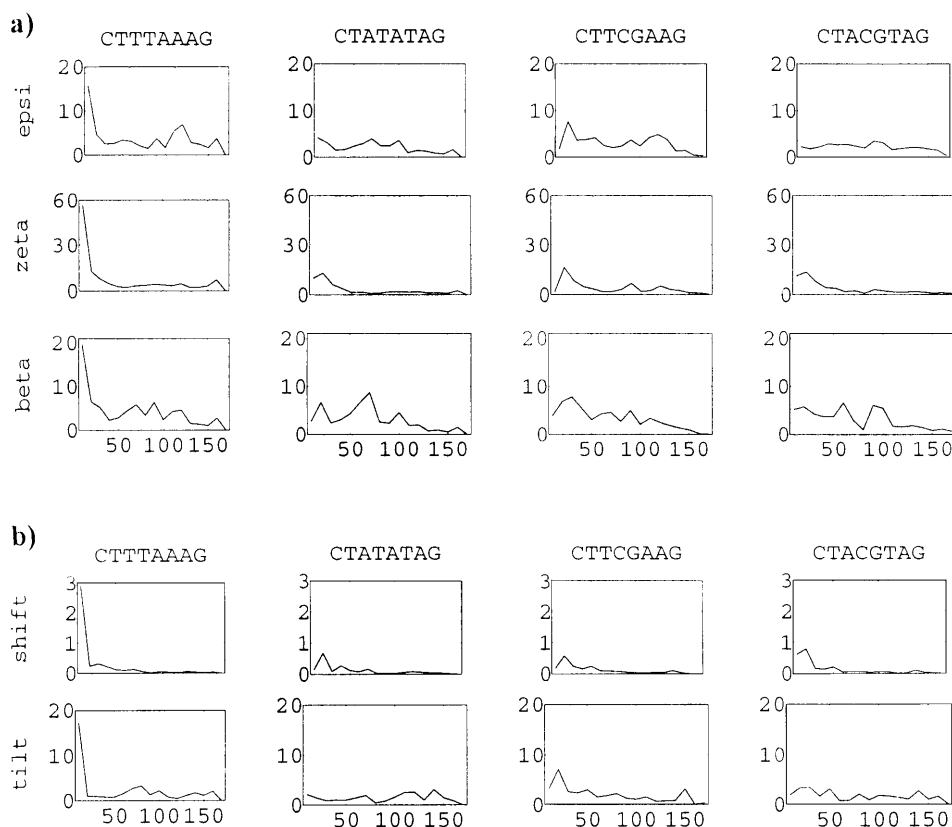
Figure 6. (**a**) Maximum amplitude of vibration of ε, ζ and β angles (degrees) as a function of normal mode frequencies (per cm) for the central step (TpA or CpG) of the octamers studied. (**b**) Maximum amplitude of vibration for shift (Å) and tilt (degrees) as a function of normal mode frequencies (per cm) for the central step (TpA or CpG) of the octamers studied.

of ease of the BI/BII transition, for CpG. Interestingly, the high flexibility of the CTTTAAAG phosphodiester backbone, seen by normal mode analysis, is not restricted to the central TpA step ($\Delta\varepsilon \sim 16°, \Delta\zeta \sim 57°, \Delta\beta \sim 20°$). It also concerns the adjacent T_3pT_4 ($\Delta\varepsilon \sim 9°, \Delta\zeta \sim 35°, \Delta\beta \sim 19°$) step as well as the two A_5pA_6 and A_6pA_7 steps ($\Delta\varepsilon \sim 17°, \Delta\zeta \sim 53°, \Delta\beta \sim 17°$). The particular dynamic properties of the central TpA step in the **CTTTA**AAG segment are also reflected in the large vibrational amplitudes of the shift and tilt parameters at the central step (Fig. 6b), also seen at the adjacent junctions. Whatever the central step, variations in twist are similar in the oligomers we studied (Δtwist $\sim 4°$). Nevertheless, the central CpG step exhibits a larger overall vibrational amplitude in terms of rise ($\sim$0.8 Å) and roll ($\sim 12°$) than a central TpA step (Δrise ~ 0.3 Å, Δroll $\sim 6°$). We note that the CpG step in the CTTCGAAG sequence undergoes weaker vibrations in terms of rise (~ 0.3 Å) and roll ($\sim 7°$) than in the CTACGTAG and CATCGATG sequences (Δrise ~ 1 Å, Δroll $\sim 14°$).

In summary, normal mode analysis of the six oligonucleotides shown in Table 1 reveals that the CTTTAAAG sequence is more flexible in terms of both backbone angles (ε, ζ and β) and of shift and tilt variations than the five other DNA oligomers. The analysis confirms the effect of the flanking sequences on the dynamics of a given dinucleotide. These results are in line with NMR data which shows that adenine proton resonances at the TpA junction of T_3A_3 tracts are broadened due to unusually large motions of this step (54,55). Compared with other sequences very large vibrational amplitudes are observed at the central YpR

junction of TTYRAA tracts when it is occupied by a TpA step, but not with a CpG step. Comparison with the BI/BII transition studies presented above suggest that YpR steps within a Y_nR_n sequence has a particular flexibility characterized either by facilitated BI/BII transition (CpG) or a larger vibrational amplitude of backbone angles within the BI conformation (TpA).

CONCLUSION

Molecular modelling has been used to systematically study the flexibility of YpR and RpY dinucleotide steps, in terms of ε and ζ dihedral angle variations. Calculations show that whatever the flanking sequences, the CpA and CpG steps are more malleable than either TpA or RpY steps in terms of BI/BII transitions. It is also generally easier to create two BII sites facing each other than two BII sites at different dinucleotide levels. The dinucleotide sequence effect on the transition energy barrier can be correctly accounted for by studying the variation in stacking energy. From a mechanical point of view the variations in ε and ζ lead first to twist and slide deformations, which correspond to an important loss of intra-strand stacking energy. This could explain the flexible behaviour of the CpG and CpA backbones.

Analysis of the effects of the neighbouring bases around a given dinucleotide step indicates that the flexibility of a central YpR dinucleotide is enhanced within $(Y)_n(R)_n$ sequences. This flexibility results in either easier BI/BII transitions (CpA or CpG) or induces very large vibrational amplitudes in backbone dihedral

angles (TpA). These theoretical results are in line with available NMR data. The latter both provide evidence for a BI/BII conformational equilibrium within CpA and CpG steps and suggest large backbone motions at the TpA steps when they occur in an $(Y)_n(R)_n$ environment. We have also demonstrated the large flexibility of YpR steps compared with RpY steps and shown that within the YpR family TpA steps have very different backbone flexibility compared with either CpA or CpG steps.

ACKNOWLEDGEMENTS

REFERENCES

1 Shroeder,S.A., Fu,J.M., Jones,C.R. and Gorenstein,D.G. (1987) *Biochemistry*, **26**, 3812–3821.
2 Gorenstein,D.G., Shroeder,S.A., Fu,J.M., Metz,J.T., Roogta,V. and Jones,C.R. (1988) *Biochemistry*, **27**, 7223–7237.
3 Shroeder,S.A., Roogta,V., Fu,J.M., Jones,C.R. and Gorenstein,D.G. (1989) *Biochemistry*, **28**, 8292–8303.
4 Karslake,C., Schroeder,S., Wang,P.L. and Gorenstein,D.G. (1990) *Biochemistry*, **29**, 6578–6584.
5 Karslake,C., Botuyan,M.V. and Gorenstein,D.G. (1992) *Biochemistry*, **31**, 1849–1858.
6 Botuyan,M.V., Keire,D.A., Kroen,C. and Gorenstein,D.G. (1993) *Biochemistry*, **32**, 6863–6874.
7 El antri,S., Bittoun,P., Mauffret,O., Monnot,M., Convert,O., Lescot,E. and Fermandjian,S. (1993) *Biochemistry*, **32**, 7079–7083.
8 Gorenstein,D.G. (1992) *Methods Enzymol.*, **211**, 254–284.
9 Fratini,A.V., Kopka,M.L., Drew,H.R. and Dickerson,R.E. (1982) *J. Biol. Chem.*, **257**, 14686–14707.
10 Privé,G.G., Heinemann,U., Chandrasegaran,S., Kan,L.S., Kopta,M.L. and Dickerson,R.E. (1987) *Science*, **231**, 498–504.
11 Privé,G.G., Yanagi,K. and Dickerson,R.E. (1991) *J. Mol. Biol.*, **217**, 177–199.
12 Grzeskowiak,K., Yanagi,K., Privé,G.G. and Dickerson,R.E. (1991) *J. Biol. Chem.*, **266**, 8861–8883.
13 Heinemann,U. and Hahn,M.J. (1992) *J. Biol. Chem.*, **267**, 7332–7341.
14 Hartmann,B., Piazzola,D. and Lavery,R. (1993) *Nucleic Acids Res.*, **21**, 561–568.
15 Yuan,H., Quintana,J. and Dickerson,R.E. (1992) *Biochemistry*, **31**, 8009–8021.
16 Chou,S.-H., Cheng,J.-W., Fedoroff,O.Y., Chuprina,V.P. and Reid,B.R. (1992) *J. Am. Chem. Soc.*, **114**, 3114–3115.
17 Chou,S.-H., Cheng,J.-W. and Reid,B.R. (1992) *J. Mol. Biol.*, **228**, 138–155.
18 Nikonowicz,E.P. and Gorenstein,D.G. (1992) *J. Am. Chem. Soc.*, **114**, 7494–7503.
19 Lefebvre,A., Mauffret,O., Lescot,E., Hartmann,B. and Fermandjian,S. (1996) *Biochemistry*, **35**, 12560–12569.
20 Leonard,G.A. and Hunter,W.N. (1993) *J. Mol. Biol.*, **234**, 198–208.
21 Lefebvre,A., Mauffret,O., Hartmann,B., Lescot,E. and Fermandjian,S. (1995) *Biochemistry*, **34**, 12019–12028.
22 Lavery,R., Zakrzewska,K. and Sklenar,H. (1995) *Comp. Phys. Commun.*, **91**, 135–158.
23 Lavery,R. (1988) In Olson,W.K., Sarma,R.H., Sarma,M.H. and Sundaralingam,M. (eds), *Structure and Expression*. Adenine Press, New York, Vol. III, pp. 191–211.
24 Lebrun,A. and Lavery,R. (1996) *Nucleic Acids Res.*, **24**, 2260–2267.
25 Sanghani,S.R., Zakrzewska,K., Harvey,S. and Lavery,R. (1996) *Nucleic Acids Res.*, **24**, 1632–1637.
26 Lavery,R. and Hartmann,B. (1994) *Biophys. Chem.*, **50**, 33–45.
27 Sodano,P., Hartmann,B., Rose,T., Wain-Hobson,S. and Delepierre,M. (1995) *Biochemistry*, **34**, 6900–6910.
28 Lefebvre,A., Fermandjian,S. and Hartmann,B. (1997) *Nucleic Acids Res.*, **25**, 3854–3861.
29 Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dynam.*, **6**, 63–91.
30 Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dynam.*, **6**, 655–661.
31 Dickerson,R.E., Bansal,M., Calladine,C.R., Diekmann,S., Hunter,W.N., Kennard,O., Lavery,R., Nelson,H.C.M., Olson,W.K., Saenger,W., Shakked,Z., Sklenar,H., Soumpasis,D.M., Tung,C.S., von Kitzing,E., Wang,A.H.J. and Zhurkin,V.D. (1989) *J. Mol. Biol.*, **205**, 787–791.
32 Hahn,M. and Heinemann,U. (1993) *Acta Crystallogr.*, **D49**, 468–477.
33 Lipanov,A., Kopka,M.L., Kaczor-Grzeskowiak,M., Quintana,J. and Dickerson,R.E. (1993) *Biochemistry*, **32**, 1373–1389.
34 Quintana,J.R., Grzeskowiak,K., Yanagi,K. and Dickerson,R.E. (1992) *J. Mol. Biol.*, **225**, 379–395.
35 Heinemann,U. and Alings,C. (1989) *J. Mol. Biol.*, **210**, 369–381.
36 Lipscomb,L.A., Peek,M.E., Morningstar,M.L., Verghis,S.M., Miller,E.M., Rich,A., Essigmann,J.M. and Williams,L.D. (1995) *Proc. Natl. Acad. Sci. USA*, **3**, 719–723.
37 Grzeskowiak,K., Goodsell,D.S., Kaczor-Grzeskowiak,M., Cascio,D. and Dickerson,R.E. (1993) *Biochemistry*, **32**, 8923–8931.
38 Baikalov,I., Grzeskowiak,K., Yanagi,K., Quintana,J. and Dickerson,R.E. (1993) *J. Mol. Biol.*, **231**, 768–784.
39 Goodsell,D.S., Kopka,M.L., Cascio,D. and Dickerson,R.E. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 2930–2934.
40 Mauffret,O., Hartmann,B., Convert,O., Lavery,R. and Fermandjian,S. (1992) *J. Mol. Biol.*, **227**, 852–862.
41 Baleja,J.D., Germann,M.W., van de Sande,J.H. and Sykes,B.D. (1990) *J. Mol. Biol.*, **215**, 411–428.
42 Schmitz,U., Pearlman,D.A. and James,T.L. (1991) *J. Mol. Biol.*, **221**, 271–292.
43 Kim,S.G. and Reid,B.R. (1992) *Biochemistry*, **31**, 12103–12116.
44 Kim,S.G., Lin,L.J. and Reid,B.R. (1992) *Biochemistry*, **31**, 3564–3574.
45 Schmitz,U., Sethson,I., Egan,W.M. and James,T.L. (1992) *J. Mol. Biol.*, **227**, 510–531.
46 Mujeeb,A., Kerwin,S.M., Kenyon,G.L. and James,T.L. (1993) *Biochemistry*, **32**, 13419–13431.
47 Ha Duong,T. and Zakrzewska,K. (1997) *J. Comp. Chem.*, **18**, 796–811.
48 Hunter,C.A. (1993) *J. Mol. Biol.*, **230**, 1025–1054.
49 Tisné,C., Simenel,C., Hantz,E., Schaeffer,F.and Delepierre,M. (1996) *Magn. Resonance Chem.*, **34**, 115–124.
50 Schaeffer,F., Rimsky,S. and Spassky,A. (1996) *J. Mol. Biol.*, **260**, 523–539.
51 Poncin,M., Hartmann,B. and Lavery,R. (1992) *J. Mol. Biol.*, **226**, 775–794.
52 Suzuki,M., Yagi,N. and Finch,J.T. (1996) *FEBS Lett.*, **379**, 148–152.
53 Hassan,M.A. and Calladine,C.R. (1996) *J. Mol. Biol.*, **259**, 95–103.
54 Kennedy,M.A., Nuutero,S.T., Davis,J.T., Drobny,G.P. and Reid,B.R. (1993) *Biochemistry*, **32**, 8022–8035.
55 Lingbeck,J., Kubinec,M.G., Miller,J., Drobny,G.P., Reid,B.R. and Kennedy,M.A. (1996) *Biochemistry*, **35**, 719–734.
56 Ha Duong,T. and Zakrzewska,K. (1997) *J. Biomol. Struct. Dynam.*, **14**, 691–701.