Conformational Deformability of RNA: A Harmonic Mode Analysis

M. Zacharias* and H. Sklenar†
*Institute for Molecular Biotechnology, 07745 Jena, and †Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

ABSTRACT The harmonic mode analysis method was used to characterize the conformational deformability of regular Watson-Crick paired, mismatch- and bulge-containing RNA. Good agreement between atomic Debye-Waller factors derived from x-ray crystallography of a regular RNA oligonucleotide and calculated atomic fluctuations was obtained. Calculated helical coordinate fluctuations showed a small sequence dependence of up to ~30–50%. A negative correlation between motions at a given base pair step and neighboring steps was found for most helical coordinates. Only very few calculated modes contribute significantly to global motions such as bending, twisting, and stretching of the RNA molecules. With respect to a local helical description of the RNA helix our calculations suggest that RNA bending is mostly due to a periodic change in the base pair step descriptors slide and roll. The presence of single guanine:uridine or guanine:adenine mismatches had little influence on the calculated RNA flexibility. In contrast, for tandem guanine:adenine base pairs the harmonic mode approach predicts a significantly reduced conformational flexibility in the case of a sheared arrangement and slightly enhanced flexibility for a face-to-face (imino proton) pairing relative to regular RNA. The presence of a single extra adenine bulge nucleotide stacked between flanking sequences resulted in an increased local atomic mobility around the bulge site (~40%) and a slightly enhanced global bending flexibility. For an adenine bulge nucleotide in a looped-out conformation a strongly enhanced bulge nucleotide mobility but no increased bending flexibility compared to regular RNA was found.

INTRODUCTION

Many biological processes involve the participation of RNA molecules either in complex with proteins or as the sole functional unit. The number of experimentally determined three-dimensional (3D) RNA structures has grown rapidly in recent years. Besides the detailed atomic resolution structure an understanding of the mobility and conformational deformability of RNA structures is important for interpreting its function. Currently, only some aspects of the conformational mobility of nucleic acids can be measured experimentally. In particular, the relation between conformational deformability at the nucleotide or even the atomic level and global motions of nucleic acids can be difficult to investigate experimentally. Theoretical approaches based on a model energy function describing the atomic interactions could be helpful for gleaning at least some hints about possible motions of RNA molecules.

A useful computational method for studying the dynamics of RNA is the molecular dynamics (MD) method, including explicit solvent and ions (Beveridge et al., 1994; Cheatham et al., 1997; Auffinger and Westhof, 1998). However, rather long simulation times are often necessary for the study of global motions and to relax various conformational degrees of freedom and the solvent and ion atmosphere around a nucleic acid molecule (Feig and Pettitt, 1998).

The method of normal mode (NM) analysis, although limited to the local motion in the vicinity of one stable energy minimum structure, is computationally less demanding and does not suffer from the convergence problems of molecular dynamics simulations (Case, 1994). The NM method can be useful to obtain a first impression of the flexible degrees of freedom of an RNA molecule and how it may change when part of the molecule is replaced with a noncanonical structural motif. Results of a normal mode analysis could also serve as a guideline for studying nucleic acids by more accurate methods such as MD.

The harmonic mobility of DNA has been studied using a description of the nucleic acid molecule in terms of Cartesian coordinates (Tidor et al., 1983; Irikura et al., 1985; Garcia and Soumpasis, 1989; Kottalam and Case, 1990; Zacharias and Sklenar, 1999a) and using torsion angle variables (Ha Duong and Zakrzewska, 1997a,b; Lin et al., 1997; Ha Duong and Zakrzewska, 1999; Matsumoto and Go, 1999). Most related to the present approach are the studies by Ha Duong and Zakrzewska (1997a) and Matsumoto and Go (1999). In both studies the global flexibility of DNA has been analyzed for periodic DNA sequences, $d(CG)_n$ and $d(TA)_n$, respectively, using the normal mode method in torsion angle space (including sugar pucker flexibility).

The conformational mobility of tRNAs has also been investigated using NM analysis in torsion angle space (Nakamura and Doi, 1994; Matsumoto et al., 1999). However, no systematic study of the harmonic flexibility of RNA and the influence of small noncanonical motifs has been published so far.

In the present study harmonic modes in terms of the helical and internal coordinates (including sugar pucker flexibility) used within the nucleic acid modeling program Jumna (Lavery et al., 1995) have been calculated for a

Received for publication 24 September 1999 and in final form 10 February 2000.

Address reprint requests to Dr. Martin Zacharias, Theoretical Biophysics Group, Institute for Molecular Biotechnology, Beutenbergstr. 11, 07745 Jena, Germany. Tel.: 011-49-3641-656491; Fax: 011-49-3641-656495; E-mail: zacharia@imb-jena.de.

0006-3495/00/05/2528/15 $2.00

number of RNA molecules. In contrast to normal modes, only the eigenvectors of the second derivative matrix of the energy function with respect to the coordinates used in Jumna have been considered (termed harmonic modes).

In the first part of the paper calculated atomic mobilities are compared with Debye-Waller factors from x-ray crystallography of a RNA structure determined at high resolution. The purpose of this comparison is to get an impression of the precision and usefulness of the harmonic mode approach. The second part of the paper gives a detailed analysis of helical as well as global fluctuations of several periodic RNA model structures (in A-form geometry) and A-DNA. It is possible to study and interpret the mechanism of RNA bending in terms of helical coordinate changes at the base-pair level. It is also possible to estimate the chain stiffness and twist flexibility of double-stranded RNA within the harmonic approximation. The calculation of the covariation of helical coordinate changes along the sequence allows us to get an idea of the interdependence of base-pair motions in RNA and how fast a small conformational deformation at one position relaxes along the sequence.

Finally, in the last part, the deformability of RNAs that contain single mismatches, tandem GA mismatches, or an extra adenine bulge nucleotide are compared to canonical dsRNA.

MATERIALS AND METHODS

Energy minimization

Energy minimization (EM) and harmonic mode calculations have been carried out using a modified version of the Jumna (junction minimization of nucleic acids) program (Lavery et al., 1995; Flatters et al., 1997) and the Amber4.1 force field (Cornell et al., 1995). In Jumna each nucleic acid strand is considered as a chain of 3′-monophosphate nucleotides that are placed in space using helicoidal coordinates. These are three translational variables (Xdisp, Ydisp and Rise) and three rotational variables (Inclination, Tip and Twist) obeying the Cambridge convention for nucleic acids (Dickerson et al., 1989). In addition, the internal flexibility of each nucleotide is described by dihedral angles around the glycosidic sugar base link and the phosphodiester backbone dihedral angles ϵ (around C3′-O3′ bond) and ζ (around O3′-P bond). The sugar rings are broken at the C4′-O4′ bond, leading to a linearized sugar with five degrees of freedom (the dihedral angles: $\tau 1$: O1′-C1′-C2′-C3′, $\tau 2$: C1′-C2′-C3′-C4′; and three valence angles: O4′-C1′-C2′, C1′-C2′-C3′, and C2′-C3-C4′). Valence angles outside the sugar ring and all bond lengths except for the connection between nucleotides and the C4′-O4′ bond are fixed at their optimum value. The constraints on the sugar ring and internucleotide connections are imposed via harmonic penalty terms (maximum deviation: <0.05 Å and $<1°$ for bonds and angles, respectively, from the optimum geometry; Lavery et al., 1995). The independent variables for each nucleotide are the six helicoidal variables, five dihedral angles, and three valence angles (within the sugar ring). All other variables are dependent and are determined by the closure conditions between nucleotides (one bond and two valence angles) and the sugar ring (Sklenar et al., 1986; Lavery et al., 1995).

Force field

The Amber4.1 force field within Jumna consists of the pairwise additive Coulomb-type electrostatic and 6–12 dependent Lennard-Jones potential terms complemented by harmonic valence angle terms and cosine dihedral barriers (including terms for the dependent dihedral angles):

$$V = \sum_{\text{angles}} \frac{1}{2} K_\theta(\theta - \theta_0)^2 + \sum_{\text{torsions}} K_\phi(1 + \cos(n\phi - \delta)) + \sum_{i<j} \left(\frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon(r_{ij}) r_{ij}} \right) \quad (1)$$

It turned out that to achieve good agreement of energy-minimized structures with available experimental RNA structures, the choice of electrostatic parameters is most critical. A distance-dependent dielectric function, $\epsilon(r_{ij})$, with a sigmoidal distance dependence (Lavery et al., 1995) of the dielectric constant (slope = 0.2, ϵ_o (dielectric constant at zero distance) = 2.0, ϵ_{wat} (dielectric constant of water) = 78.0) and a scaling of the phosphate charge by 0.33, gave good agreement of energy-minimized structures with a standard A-form RNA (for the definition of the standard A-form, see below):

$$\epsilon(r_{ij}) = \epsilon_{wat} - \frac{1}{2}(\epsilon_{wat} - \epsilon_0)\{(\text{slope } r_{ij})^2 + 2(\text{slope } r_{ij}) + 2\} \cdot \exp(-\text{slope } r_{ij}) \quad (2)$$

For example, with the above parameters EM of the high-resolution x-ray structure of an RNA with the sequence $(CCCCGGGG)_2$ (Egli et al., 1996) results in a structure with an all-atom rmsd (Cartesian root mean square deviation) of <0.8 Å from the x-ray start structure. In case of a recently published high-resolution x-ray analysis of a G:U wobble containing RNA (Mueller et al., 1999), energy minimization of either a standard A-form RNA (with the G:U pair) or the x-ray structure leads to the same energy minimum structure with an rmsd of ~0.7 Å (from the x-ray structure).

Generation of RNA model structures

All structures were generated using Jumna (Lavery et al., 1995). An A-form starting structure (Xdisp = −5.28 Å, Ydisp = 0.01 Å, Rise = 2.56 Å, Inclination = 20.7°, Tip = −3.77°, Twist = 32.7°; backbone dihedral angles around C5′-C4′ bond: 59°; C4′-C3′: 78°; C3′-O3′: −155°; O3′-P: −67°; P-O5′: −75°; O5-C5′: −179°) was used for energy minimization of double-stranded RNA, DNA, or RNA with central non-Watson-Crick base pairs. Energy minimization of regular RNA as well as RNA with single mismatches resulted in structures close to the A-form starting geometry (rmsd $<$ 1.5 Å). The periodic Watson-Crick paired RNA model structures consisted of 18 bp with the sequences $(CG)_9$ (RNA and A-DNA), $(UA)_9$, $C(GGCC)_4G$, and $U(AAUU)_4A$.

In the case of a single guanine:uridine (G:U) mismatch (CGCGC**G**-GCGCG/CGCGC**U**GCGCG, mismatch in bold) the hydrogen bonding geometry was similar to the geometry observed in a recent high-resolution x-ray structure (Mueller et al., 1999). For the single G:A mismatch (CGCGC**G**GCGCG/CGCGC**A**GCGCG) a face-to-face (or imino proton) pairing was assumed with a G(anti):A(anti) orientation, as observed in an x-ray structure of single G:A mismatches (Leonhard et al., 1994). In the case of tandem G:A mismatches the backbone geometry as observed in the NMR structures of a sheared (N7-amino, amino-N3 pairing; SantaLucia and Turner, 1993) and a face-to-face G:A pairing (imino hydrogen-bonded; Wu and Turner, 1996), respectively, were used. In the NMR studies a sequence context C**GA**G was found to stabilize the sheared mismatch arrangement (SantaLucia and Turner, 1993), and a G**GA**C context resulted

in a face-to-face pairing (Wu and Turner, 1996). However, to be consistent with the sequence context of other mismatches of the present study, only one sequence context $(CGCGC\mathbf{GA}GCGCG)_2$ for both arrangements was considered.

An RNA structure with an extra adenine bulge base stacked between flanking dsRNA was built using Jumna starting from regular A-form RNA with a central A:U base pair. The uridine nucleotide of the central base pair was removed, and the structure was energy minimized (CGCG-C**A**GCGCG/CGCGCGCGCG). All backbone torsion angles of the resulting structure are within the range observed for regular A-form RNA. The bulge structure with the bulge base in the looped-out form was built based on the torsion angle pattern observed in the x-ray structure of a single adenine bulge in a chimeric RNA/DNA structure in A-form geometry (Portmann et al., 1996).

Harmonic mode calculation

Calculation of the harmonic modes at an energy minimum of a given nucleic acid structure requires the calculation of the second derivative of the conformational energy with respect to the coordinates. This second derivative matrix was calculated numerically following an approach by Levitt et al. (1985):

$$F_{ij} = \frac{\partial^2 E}{\partial q_i \partial q_j} = \frac{1}{\delta}\left(\left(\frac{\partial E}{\partial q_i}\right)_{q_j = q_{jo}+\delta} - \left(\frac{\partial E}{\partial q_i}\right)_{q_j = q_{jo}}\right) \quad (3)$$

The parameter δ was chosen to be sufficiently small that the matrix elements do not dependent on δ. By solving the eigenvalue equation,

$$(F - \lambda I)Y = 0 \quad (4)$$

a set of eigenvectors Y_i (harmonic modes) and corresponding eigenvalues λ_i were calculated. The harmonic modes form a complete set of orthogonal directions to describe any deformation of the given energy minimum structure. For small deformations the energy change is a quadratic function for the deformation in any harmonic mode direction independent of other HM deformations:

$$\Delta E(\Delta Y_i) = \frac{1}{2}\lambda_i(\Delta Y_i)^2 \quad (5)$$

Thermal coupling of the molecule to the environment leads to deformations of the energy minimum structure in each HM direction (due to collisions with other molecules). At equilibrium the form of the energy function (Eq. 5) implies a Gaussian probability distribution for a given HM deformation with variance $\sigma_i = kT/\lambda_i$ (k, Boltzmann constant; T, selected temperature). This variance corresponds to the square of an elongation in the HM direction that results from exciting the HM with an energy of $\frac{1}{2}kT$. Because for small deformations every HM contributes at every time point independently (each HM represents one degree of freedom of the molecule) to the variance of a selected coordinate, one can calculate the mean (or total) variance by summing over all HM contributions (excited by $\frac{1}{2}kT_R$, $T_R = 298$ K). Note that the HM directions can differ from so-called normal-mode coordinates that are the eigenvectors of the eigenvalue equation

$$(F - \lambda MI)Y = 0 \quad (6)$$

This equation results from a solution of the equation of motion for the molecule, and M is the mass matrix (or kinetic energy matrix) of the molecule under study (Brooks and Karplus, 1983; Brooks et al., 1995). In the case of Cartesian coordinates, M is diagonal, and then HMs and NMs have the same directions (only the eigenvalues differ by a constant). In the general case M is nondiagonal, and HM and NM directions might be (slightly) different. Normal-mode analysis provides vibrational coordinates and frequencies and can be used to describe the vacuum dynamics of the molecule. However, in the present study we are only interested in the equilibrium probability for any deformation of a structure around an energy minimum in the harmonic approximation. The assumption is that all motions are overdamped and the eigenvalues of each harmonic mode are only a measure of the deformability (harmonic force constant) of the structure in the corresponding harmonic mode. Any possible difference in the diffusional relaxation of the various modes has not been considered. Cartesian displacements of each atom with respect to the position in the energy minimum structure for harmonic mode deformations were calculated after superposition of the deformed structure on the energy minimum structure (Kabsch, 1976; Levitt et al., 1985) and are given as rmsd (in Å). The mean square fluctuation of atomic Cartesian coordinates, $\langle \Delta r_a^2 \rangle$, is calculated by summing over all harmonic mode contributions (each HM excited by an energy of $\frac{1}{2}kT_R$). The crystallographic atomic Debye-Waller (B) factors were calculated using

$$B_a = \left(\frac{8\pi^2}{3}\right)\langle \Delta r_a^2 \rangle \quad (7)$$

Helical coordinate variations and changes in the global geometry upon deformation in the harmonic modes were calculated using an approach by Ha Duong and Zakrzewska (1997a). Each structure deformed in the kst HM (excited with a thermal energy of $\frac{1}{2}kT_R$) was analyzed using the CURVES program (Lavery and Sklenar, 1988a,b). The CURVES program generalizes the calculation of a global helical axis to irregular nucleic structures. The helical axis segments defining the globally curved helical axis are calculated by minimizing the sum of variations given by the changes in the helical axis and in the associated axis-base parameters (Xdisp, Ydisp, Inclination, Tip) between successive nucleotides (see Lavery and Sklenar, 1988a,b). The program outputs the helical nucleotide and base pair descriptors with respect to the global axis or with respect to a local dinucleotide step description. In the case of the present RNA structures the calculated global axis is almost linear and follows the z axis if the RNA helix is aligned along the z axis. The difference of a given helical coordinate, $R_{i,k}$, between the deformed structure and the energy minimum structure ($\Delta R_{i,k}$) for the given HM was calculated. This is the contribution of the kst HM to a given helical or global coordinate change. The square of these contributions summed over all HMs is the (thermal) variance of the selected coordinate, and in the following the square root of this variance, Rssd (root summed square deviation), is called the (thermal) harmonic fluctuation of the selected coordinate:

$$\mathrm{Rssd}(R_i) = \sqrt{\textstyle\sum_{k=1,N}(\Delta R_{i,k})^2} \quad (8)$$

Some of the helical coordinate fluctuations are correlated. This correlation is characterized by the covariation between two selected coordinates (e.g., R_i, R_j). A contribution from the kst HM is given by the product $\Delta R_{i,k} \cdot \Delta R_{j,k}$. The sum of these products over all HMs gives the covariance of the fluctuation of the two coordinates:

$$\sigma(R_i, R_j) = \sum_{k=1,N} \Delta R_{i,k} \cdot \Delta R_{j,k} \quad (9)$$

RESULTS

Comparison of experimental and calculated atomic B-factors

In addition to atomic coordinates, x-ray crystallography can also provide some information about the thermal fluctuation of each atom of a structure. The Debye-Waller or B-factors are proportional to the mean square position fluctuation of each atom in the crystal structure. Atomic mean square

fluctuations can also be calculated using the calculated harmonic modes for an energy minimum close to the crystal structure. Because the present approach is approximate with respect to the model energy function and assumes a nearly quadratic energy function around the energy minimum, it is interesting to see the degree to which experimental atomic B-factors can be reproduced. It should also be mentioned that the present HM approach does not include bond length and bond angle flexibility (some angles are included; see Materials and Methods), although such flexibility is present in real molecules. However, there is evidence that inclusion of bond length and bond angle flexibility contributes only a few percent to the total atomic fluctuations (Lin et al., 1997). Note also that the calculations are performed on isolated molecules and not in the context of the crystal.

An 8-bp dsRNA, $(CCCCGGGG)_2$, which was solved at room temperature to a resolution of better than 1.5 Å (Egli et al., 1996), was used as an example. Note that although the two strands have the same sequence, the x-ray structure is asymmetrical (presumably because of crystal contacts). The energy-minimized crystal structure (rmsd $<$ 0.8 Å from x-ray structure) is still asymmetrical, and the B-factor plots (both the experimental and the calculated plots) are also asymmetrical with respect to the two strands (Fig. 1).

Although no quantitative agreement between calculated and experimental atomic B-factors was observed (Fig. 1), there are some interesting qualitative similarities. Some of the fine structure in the experimental atomic B-factor curve can also be seen in the calculated curves. Atoms belonging to the phosphate groups have the largest B-factors (peaks in the B-factor plots, Fig. 1), followed by sugar atoms and nucleobase atoms with the smallest mobilities. For the calculation with completely free RNA the calculated B-factors are larger than the observed B-factors at both ends of the structures, whereas the opposite is true for central regions. This may be due to the fact that the experimental structures are continuously stacked in the crystal, which restricts the bending motion. In fact, restricting the mobility of the base pairs at both ends of the RNA with respect to the helical axis (Fig. 1 *B*) to approximately mimic the effect of a continuous stacking in the crystal improves the agreement between experimental and calculated B-factors. In this case average calculated atomic fluctuations are slightly smaller than the results from x-ray crystallography, indicating that the harmonic mode approach may underestimate the flexibility of the RNA molecule. However, the larger offset in the experimental B-factor curve may also be due to whole-body motions of the RNA in the crystal not considered in the calculations (and to many other factors; see Discussion). It should be emphasized that for other DNA and RNA crystal structures the agreement between calculated and experimental B-factors can vary, depending, for example, on the resolution of the x-ray analysis. However, the general qualitative conclusions described above were also found for other RNA and DNA structures (data not shown).

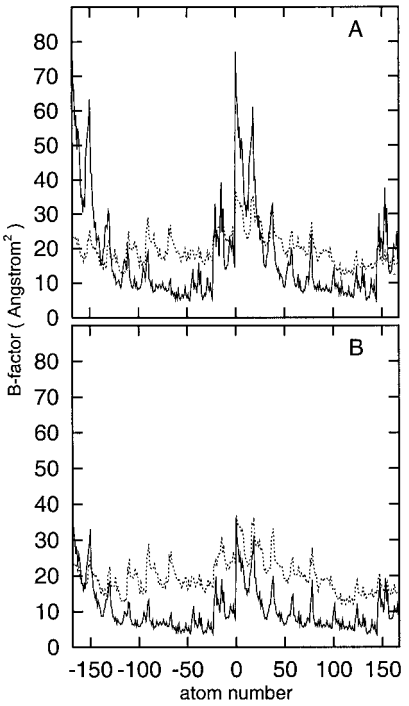

FIGURE 1 Comparison of experimental and calculated atomic Debye-Waller (B) factors. The dashed line corresponds to experimental B-factors derived from the x-ray analysis of a 8-bp RNA $(CCCCGGGG)_2$ (determined at room temperature, resolution 1.46 Å, PDB entry: 259D, NDB entry: ARH074; Egli et al., 1996), and the thin line represents B-factors calculated from the harmonic modes. The numbering (*x* axis) is such that -168 to -1 corresponds to the heavy atom numbers of the first RNA strand (0–167: second strand). The atom sequence is the same as in the x-ray structure, except that the first atom (the O5′ atom of the 5′ nucleotide) of each strand is omitted because it is not included in the calculations. (*A*) B-factors calculated for free RNA. (*B*) B-factors calculated with harmonic restraints on the nucleotides at both ends of the structure, restricting the motion relative to the helical axis of the molecule. The restraining force constants were such that a motion of ~1 Å with respect to the position at the energy minimum is possible for thermal excitation at room temperature.

Helical flexibility of dsRNA

In the present approach the harmonic modes of an RNA molecule are given as a linear combination of helical coordinates for each nucleotide with respect to a linear axis and some internal variables for each nucleotide. The harmonic modes form a complete set of orthogonal coordinates, and it is therefore possible to calculate the variance of a given coordinate by summing over each harmonic mode contribution (Materials and Methods). However, because the placement of the (linear) helical axis with respect to the molecule depends on the user's choice in the Jumna program, the resulting helical description may not be the most useful for the given molecule. A more objective way to characterize helical fluctuations of the RNA molecules is an approach used by Ha Duong and Zakrzewska (1997a). In this method RNA molecules deformed in a given HM are

TABLE 1 Helical coordinate fluctuations in periodic RNA sequences

Step	⟨shift⟩	⟨slide⟩	⟨rise⟩	⟨tilt⟩	⟨roll⟩	⟨twist⟩	⟨buckle⟩	⟨propeller⟩	⟨opening⟩
$(GC)_6$ and $(AU)_6$ sequences									
CG	0.38	0.39	0.37	3.6	4.6	3.3	8.5	6.2	3.0
GC	0.36	0.38	0.30	3.1	3.7	3.1	—	—	—
CG(DNA)	0.35	0.38	0.39	3.5	4.8	3.4	8.6	6.5	3.0
GC(DNA)	0.35	0.39	0.30	2.8	4.0	3.0	—	—	—
UA	0.35	0.29	0.34	4.4	5.0	2.9	8.6	6.9	5.3
AU	0.48	0.35	0.28	4.3	4.5	2.9	—	—	—
$(UUAA)_6$ and $(CCGG)_3$ sequences									
CG	0.40	0.39	0.38	3.2	4.3	3.6	7.7 (6.8)	5.9 (6.3)	2.8 (2.8)
GC	0.38	0.38	0.30	3.3	3.6	3.0	—	—	—
GG (CC)	0.38	0.39	0.36	2.9	3.9	3.2	—	—	—
UA	0.34	0.30	0.30	4.2	4.9	2.9	9.8 (7.9)	7.2 (6.3)	4.7 (4.7)
AU	0.46	0.34	0.29	4.5	4.9	2.9	—	—	—
AA (UU)	0.37	0.35	0.28	4.6	4.6	3.2	—	—	—

Harmonic helical coordinate fluctuations were calculated by summing over contributions from each HM (see Materials and Methods for details). Shift, slide, and rise fluctuations are given in Å; all other fluctuations are given in degrees. Note that buckle, propeller, and opening are specific for a base pair and not a dinucleotide step (G:C and C:G base pairs in $(GC)_6$ are equivalent). In the case of the $(CCGG)_3$ sequence calculated, buckle, propeller, and opening, fluctuations for a C:G base pair in the sequence context GCC differ from those in the CCG context (given in parentheses; same for $(UUAA)_6$ sequence).

analyzed using the CURVES program (Lavery and Sklenar, 1988a,b), and the resulting helical coordinates are compared with the coordinates of the energy minimum structure. The variance of a given helical coordinate can be calculated by analyzing all HM deformed structures (relative to the energy minimum structure), using a thermal excitation of $\frac{1}{2}kT_R$ and summing over all HM contributions to a given coordinate. Note that CURVES offers the possibility of analyzing the nucleotide structure with respect to a local helical description (each dinucleotide step is independent of other steps) or a (possibly curved) global axis description (with respect to an axis for the entire helix; Lavery and Sklenar, 1988a,b). If not indicated otherwise, the fluctuations will be reported for helical coordinates calculated with respect to a global description.

Harmonic mode calculations have been performed on 18-bp RNA molecules with the sequences $(CG)_9$, $(UA)_9$, $C(GGCC)_4G$, and $U(AAUU)_4A$ and on an 18-bp A-DNA ($(CG)_9$). To minimize end effects (fraying of the nucleotides at the ends) only the harmonic coordinate deformabilities of the central 12 base pairs are analyzed (approximately one helical turn). In the following the RNA molecules will be distinguished according to the central 12 bp, $(GC)_6$, $(AU)_6$, $(CCGG)_3$, and $(UUAA)_3$. The calculated Rssd of the helical coordinates shows only a modest sequence dependence (Table 1). In addition, within the harmonic mode approximation little difference between the helical coordinate fluctuations of RNA and A-DNA ($(GC)_6$) has been found. Overall the fluctuation of the translational variable shift is slightly larger than slide and rise, particularly for AU steps. For the angular variables the calculated fluctuation of base pair roll is generally larger than tilt and twist. The calculated tilt deformabilities of U/A-containing sequences are larger than for G/C sequences, whereas the twist flexibility is slightly larger in G/C- compared to U/A-containing sequences. The fluctuation of intra-base pair descriptors follows the (increasing) order opening, propeller, and buckle, with a clearly enhanced opening tendency for A/U sequences. These calculated fluctuations also depend to some degree on the surrounding base pair. The calculated roll flexibility of the pyrimidine-purine step is slightly larger than for other steps (with the exception of the AU step, which had the same calculated base pair roll fluctuation as UA).

Global deformability

The overall bend angle fluctuation of the 12 central base pairs of the RNA molecules summed over all HM contributions (Table 2) is similar for all sequence variants (~12°). Bend angle fluctuation is here defined as the fluctuation of the angular orientation of the helical axis vector of the last base pair with respect to the helical axis vector of the first base pair. This translates to a per base pair bend angle variance ($\langle\theta^2\rangle$) of 13°^{2}. The energy minimum RNA conformation is practically straight, so that twist contributions to

TABLE 2 Global deformability of RNA

Sequence	$\langle \text{bending}_{tot} \rangle$	$\langle \text{twisting}_{bp} \rangle$	$\langle \text{stretching}_{bp} \rangle$
$(GC)_6$	11.8	2.3	0.35
$(GC)_6$(A-DNA)	11.9	2.2	0.37
$(CCGG)_3$	11.6	2.3	0.39
$(AU)_6$	12.4	1.9	0.32
$(UUAA)_3$	12.2	1.9	0.30

$\langle \text{Bending}_{tot} \rangle$ (in degrees) is the angular fluctuation of the helical axis vector of the last base pair with respect to the first base pair of the 12-bp RNA (calculated by summing over all harmonic mode contributions). $\langle \text{twisting}_{bp} \rangle$ (in degrees) and $\langle \text{stretching}_{bp} \rangle$ (in Å) are the per-base pair fluctuations of the total twist and rise of the entire 12-bp RNAs (square root of the total twist/rise variance of the entire helix divided by the number of base pair steps), respectively.

bending can be neglected (Olson et al., 1998). Using an average rise per bp of ~2.6 Å, one can estimate a persistence length (chain stiffness) of ~1300 Å (using $P \approx 2 * \text{rise}/\langle\theta^2\rangle$).

Here the assumption is made that there is no covariation of coordinate deformations of the two base pairs at each end of the helix (see below). The calculated harmonic stretching and twisting flexibility of the whole RNA is slightly larger for the C/G-containing sequences than for the U/A sequences. Such a trend was also observed by Ha Duong and Zakrzewska (1997a) in a NM study of DNA.

Similar to the helical coordinate fluctuations of each base pair, the calculated global flexibilities of A-DNA and RNA of the same sequence are very similar. Interestingly, the twist variance at each dinucleotide step (Table 1) is slightly larger than the average twist variance (variance of the twist between first and last base pairs divided by the number of base pair steps). This indicates an (anti-)correlation between twist deformations of successive dinucleotide steps. Such a correlation between helical coordinate fluctuations of successive steps is also observed for other helical variables (see Fig. 2). For example, the overall per base pair bend angle variance is substantially smaller ($13^{\circ 2}$; see above and Table 2) than the sum of roll and tilt variances at isolated base pair steps ($\sim 3.5^2$(tilt) + $\sim 4.5^2$(roll) = $\sim 32^{\circ 2}$). In the case of independent conformational fluctuations the two variances should be the same. From a simple mechanical point of view it is understandable that, for example, an increase in the roll at one dinucleotide step tends to decrease the roll at adjacent dinucleotide steps. The calculated covariance between a helical coordinate for a given base pair step and the previous step (in 5′ direction) is negative for all helical variables except rise. In this case, the nearest-neighbor covariance is close to zero. Consistent with this observation, the calculated global stretching fluctuation (per bp) is close to the calculated rise fluctuation at each dinucleotide step (compare Tables 2 and 1). The covariance in the case of correlations between helical coordinates describing base pair steps beyond next-nearest neighbors is close to zero.

Only a small number of calculated modes cause significant bending, stretching, and twisting of the RNA helix (Fig. 3). The two largest bending modes also make the largest contribution to Cartesian atomic displacements from the EM structure (Fig. 3). These modes have the smallest force constant with respect to Cartesian displacement. Interestingly, the two largest bending modes make very little contribution to overall twisting and stretching. On the other hand, the mode with the third largest overall atomic displacement (per $\frac{1}{2}kT_R$ excitation) contributes very little to bending but leads to stretching and twisting of the helix (illustrated in Fig. 4). The analysis of the two modes (in the following called mode I and II) that account for most of the helix bending in terms of helical coordinate changes is shown in Figs. 5 and 6. For the $(GC)_6$ sequence the analysis in terms of a local helical description (see above) demonstrates that RNA bending corresponds mainly to a coupled slide and roll motion along the sequence. A stretch of ~4–6 bp with positive change in roll and slide follows a similar stretch with a corresponding negative change of helical variables relative to the energy minimum structure. The distance between these two stretches is such they are approximately on opposite faces of the helix, so that the collective changes add up (at least partially) to an overall bending of the helix. The two bending modes show a phase shift along the sequence such that the two bending modes describe global axis deformations that are approximately perpendicular to each other. Note that Matsumoto and Go (1999) also identified two approximately perpendicular bending modes in a normal-mode study of B-DNA.

In the case of a global helical description, that is, with respect to an almost linear axis (corresponding to the *z* axis

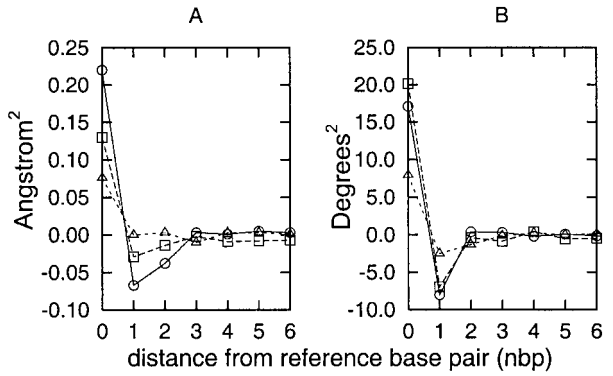

FIGURE 2 Covariation of helical coordinate fluctuations along the sequence. The product of a helical base pair displacement at step 7 (an AU step) in the $(AU)_6$ sequence and the displacement of the same helical variable *n* base pairs (nbp) away from step 7 (in the 5′ direction) was summed over all harmonic mode contributions and plotted versus nbp (nbp = 0 gives the variance of the helical coordinate at step 7). (*A*) Covariation of shift (○), slide (□), and rise (△). (*B*) Covariation of tilt (○), roll (□), and twist (△).

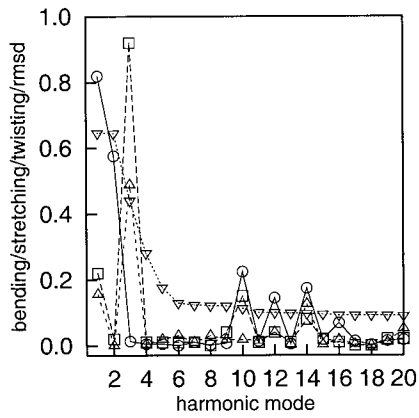

FIGURE 3 Contribution of the first 20 harmonic modes (with smallest eigenvalues) to atomic Cartesian coordinate fluctuations (▽, in Å), stretching (□, in Å) and bending (○, in degrees (× 0.1)), and overall helix twisting (△, in degrees (× 0.1)) for the $(GC)_6$ RNA sequence.

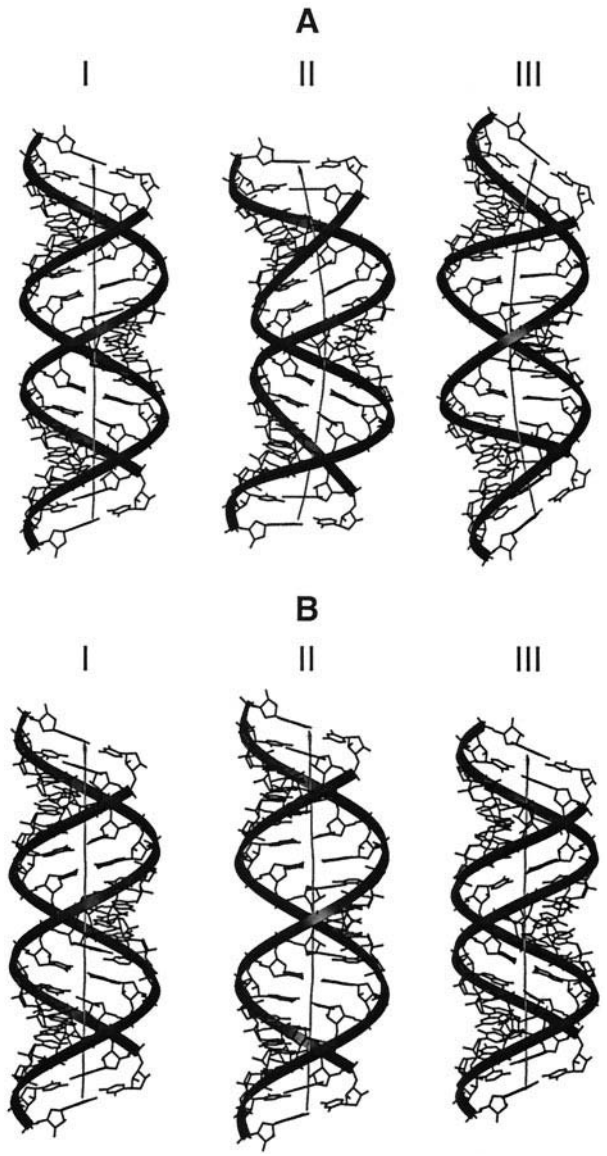

FIGURE 4 (*A*) Deformation of energy-minimized $(GC)_6$ RNA (I) in opposite directions (II, III) of the first bending mode. (*B*) same for the stretching mode of the RNA. A ribbon representation for the RNA backbone has been added (*dark*) to better illustrate the RNA backbone motion. The calculated helical axis for each structure (using CURVES; Lavery and Sklenar, 1988a,b) is shown as a bold line. The excitation energy for each deformation was $5kT_R$.

when the A-helix is aligned along the *z* axis), bending is caused mainly by a collective change in the helical coordinates shift, rise, roll, and twist along the sequence. On top of these collective changes one can also observe a 2-bp periodicity that coincides with the sequence periodicity (CG) of the RNA fragment. Each CG step makes a larger contribution to the helix bending than the GC step in particular in terms of roll. Deformations in both modes significantly alter the width of the major groove along the sequence (periodically with a phase shift along the sequence) by up to 1.4 Å (Fig. 7). The width of the minor groove is also influenced, but to a much lesser degree (changes < 0.3 Å).

Modes I and II in the case of the $(AU)_6$ sequence are similar to the corresponding modes for the $(GC)_6$ sequence. With respect to the local step-to-step helical description (connection of these local axis systems would result in a curved axis), the slide contribution to bending appears to be smaller for the $(AU)_6$ sequence relative to the $(GC)_6$ case. From the global axis viewpoint, $(AU)_6$ RNA helix bending is mostly due to a correlated change in the helical variables rise and roll along the sequence and slightly smaller contributions from slide and twist compared to $(GC)_6$.

Similar bending modes were found for the $(UUAA)_3$ and $(CCGG)_3$ sequences, respectively, with a more pronounced kinking at the pyrimidine-purine steps (data not shown).

Harmonic flexibility of single G:U and G:A and tandem G:A mismatches

For the study of single non-Watson-Crick base pairs, an A-form $(CG)_5$ RNA structure was used as a reference, and mismatch model structures were generated by insertion of G:C, A:U, G:U, or G:A base pairs at the center of the dsRNA. In the case of tandem G:A basepairs two mismatches (G:A, A:G) were inserted in a face-to-face (Wu and Turner, 1996) or sheared arrangement of the G:A mismatches (SantaLucia and Turner, 1993). As outlined in the Materials and Methods section, the resulting energy-minimized structures are in close agreement with structural data determined using NMR-spectroscopy and/or x-ray crystallography.

The calculated bending flexibility of the two single mismatch-containing sequences is not significantly different from that of regular base-paired structures (Table 3). Calculated Cartesian coordinate fluctuations (Table 3) and helical coordinate fluctuations (Fig. 8) of atoms belonging to the mismatch and flanking nucleotides differ only slightly from those calculated for regular Watson-Crick paired RNA. The uridine in the G:U mismatch shows a slightly enhanced Cartesian coordinate fluctuation. In addition, the rise fluctuation of the step consisting of the central G:U mismatch and the 5′ base pair (a C:G) is larger than, for example, the fluctuation in the case of an A:U pair. Because of the placement of the uridine in a G:U mismatch compared to uridine in an A:U pair (larger shear than for A:U; Mueller et al., 1999), it has more freedom for motions in the helical axis direction (practically no stacking interactions with the guanine in the neighboring C:G base pair). Presumably, this space in the z direction allows for larger rise fluctuation (and to a lesser degree also roll) compared to regular RNA. Note that the uridine stacks well on the cytidine of the 3′-flanking base pair (a G:C), and therefore the rise fluctuation of this step is not enhanced.

More significant differences in the calculated harmonic deformability can be observed in case of the G:A tandem

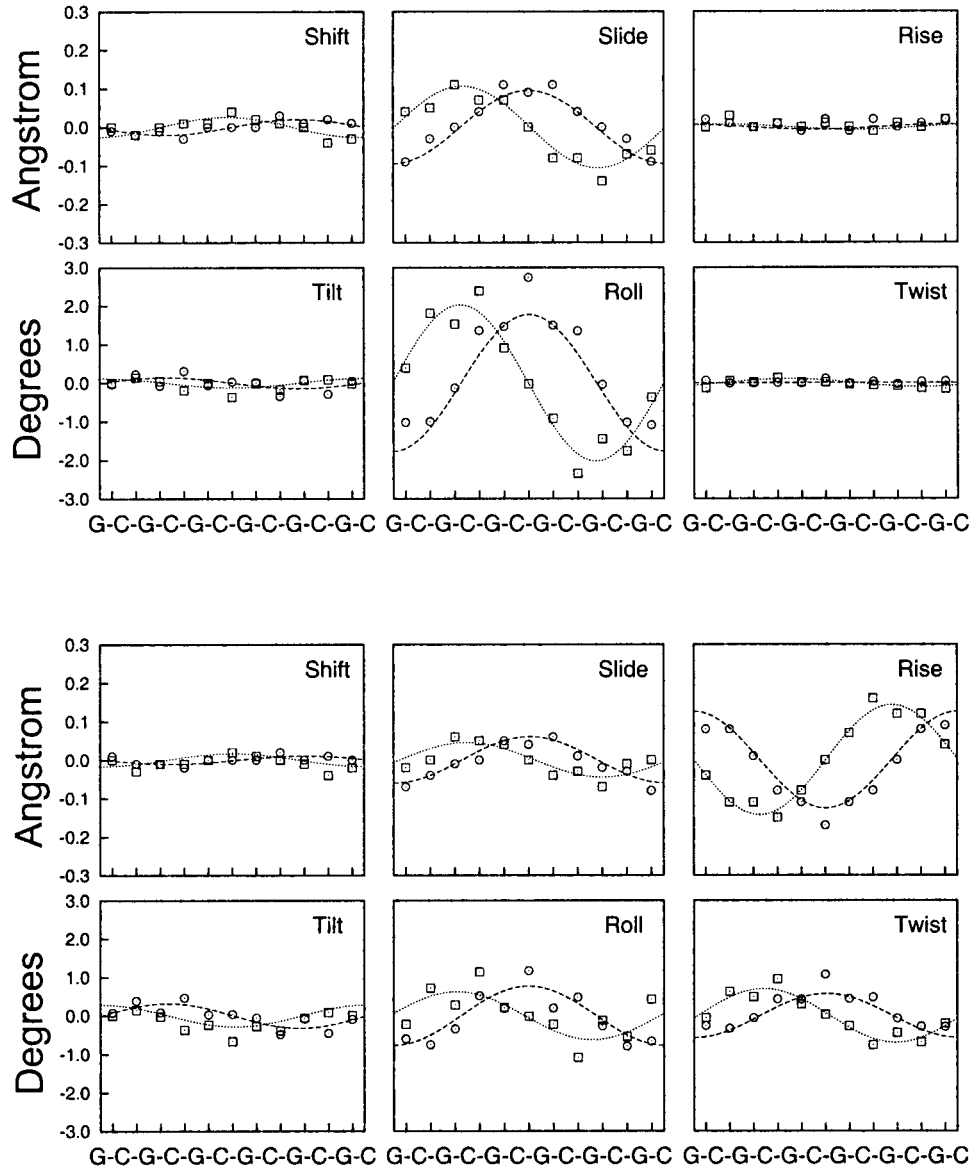

FIGURE 5 Helical coordinate changes (with respect to the minimum energy structure) versus base pair step (x axis) upon deformation of the $(GC)_6$ model structure in the two HMs with the largest bending contribution (○, mode I; □, mode II). A function, $f(n) = A \sin(bn + C)$ (n, base pair step; $b = 2\pi/11$; parameters: A and C), was fitted to the data points of each mode (*broken and dashed lines*). (*A*) Changes in helical coordinates were calculated with respect to a local helical description, that is, each dinucleotide step is independent of other steps. (*B*) Helical coordinate changes were calculated with respect to a global axis (calculated using CURVES, close to the z axis).

mismatches. Adding a second G:A face-to-face (ftf) mismatch increases the bend angle variance from 13.0° (single ftf-GA mismatch; Table 3) to 13.8°. This corresponds to an increase in the bend angle variance of ~22°^{2}, which is relatively close to the increase comparing no central element and the addition of a single face-to-face G:A mismatch (~26°^{2}; see Table 3).

In the case of tandem G:A mismatches in the sheared conformation the calculated bend angle variance is significantly smaller than for the ftf arrangement. The greater stiffness of the sheared tandem mismatches is also reflected in the smaller mean Cartesian atomic fluctuation compared to the tandem ftf mismatches (Table 3). The calculated helical coordinate fluctuations of the ftf tandem mismatch structures around the central site are similar or even slightly larger than for regular RNA (Fig. 8). In the case of the sheared mismatch arrangement the slide, rise, and roll fluctuations of the central base pair steps and flanking steps are significantly reduced relative to regular RNA. Interestingly, the shift and tilt fluctuations for steps involving adjacent base pairs are significantly larger than for regular RNA. Presumably the side-by-side placement of guanine and adenine in a sheared arrangement allows for greater shift and tilt motions that do not disrupt the hydrogen bonding and stacking interactions.

Flexibility of single-base adenine bulges in dsRNA

In the case of single-base bulges, the central base pair in the 11-bp RNA sequence (see previous section) was replaced

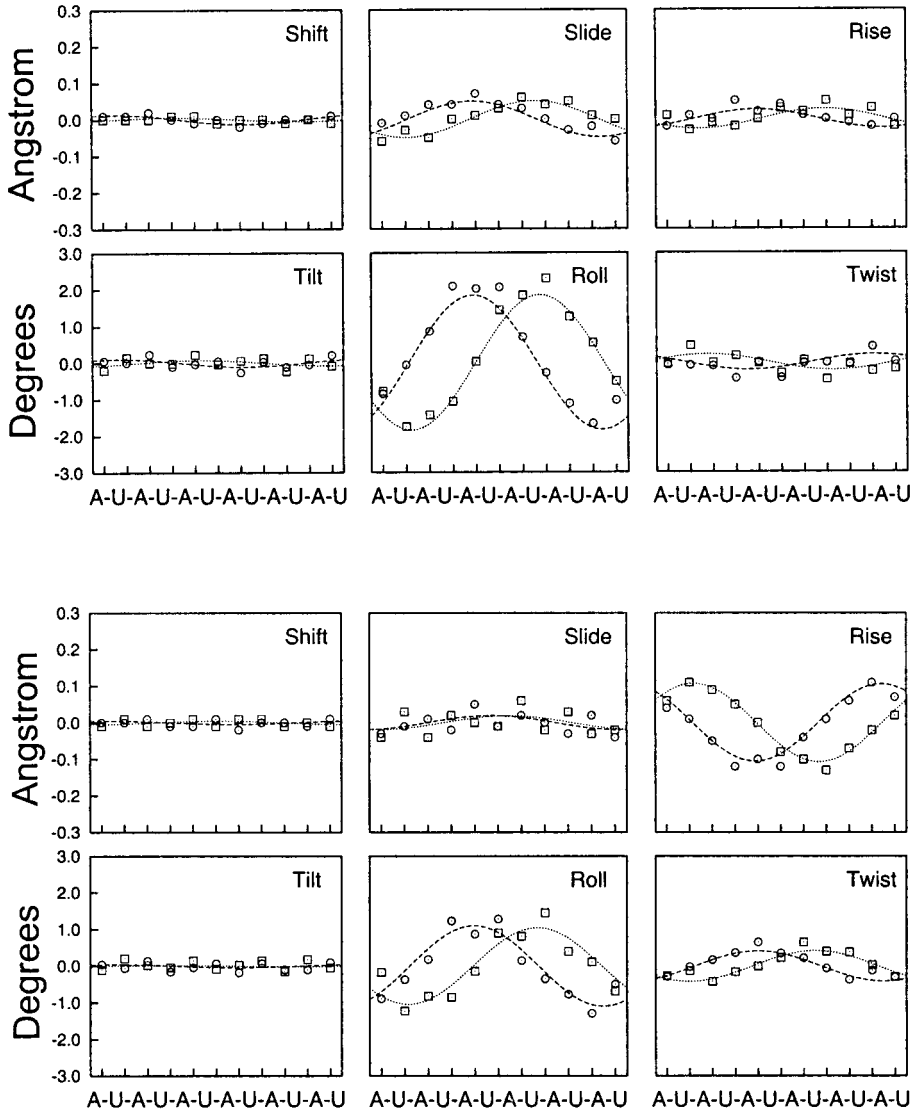

FIGURE 6 Helical coordinate changes upon deformation of a the $(AU)_6$ model structure in the two HMs with the largest bending contribution (see legend of Fig. 5).

by a single unpaired adenine. Two different bulge conformations were considered that represent two main classes of single extra-mismatched nucleotides that have been observed experimentally (Joshua-Tor et al., 1988; Nikonowicz et al., 1990; Portmann et al., 1996). In the stacked bulge conformation the extra base stacks between the two flanking base pairs, and the backbone torsion angle pattern is close to what has been observed for standard A-form RNA. The structure is kinked at the bulge site by ~20° , which is similar to what has been found experimentally (Zacharias and Hagerman, 1995).

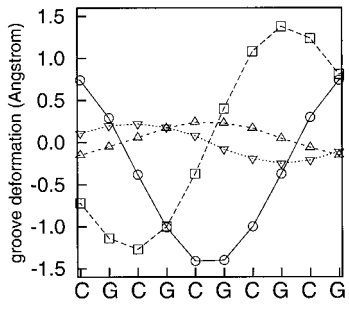

FIGURE 7 Change in major groove (○, mode I; □, mode II) and minor groove width (△, mode I; ▽, mode II) along the sequence relative to the EM structure ($(GC)_6$) for the two largest bending modes, I and II.

The second bulge form considered in the present study contains a single unmatched adenine in a looped-out conformation. This conformation is similar to an x-ray structure of a single adenine bulge in an A-form RNA/DNA chimeric molecule (Portmann et al., 1996). It is characterized by continuous stacking of the sequences flanking the bulge in a nearly A-form geometry.

For the stacked form the HM calculations predict a slightly enhanced bend angle fluctuation compared to a

TABLE 3 Comparison of bending and atomic position fluctuations of mismatch and bulge-containing RNAs

Central bp	$\langle\Delta\theta_{bend}\rangle$	Rmsd_{tot}	Rmsd_{fl}	Rmsd_{cnu1}	Rmsd_{cnu2}
G:C	12.5	1.2	0.54	0.51 (0.45)	0.52 (0.48)
A:U	12.6	1.1	0.52	0.51 (0.44)	0.51 (0.47)
G:U	12.6	1.2	0.58	0.53 (0.46)	0.59 (0.62)
G:A	13.1	1.2	0.57	0.56 (0.48)	0.56 (0.46)
G:A (t-ftf)	13.8	1.4	0.67	0.57 (0.53)	0.68 (0.51)
G:A (t-shear)	12.4	0.92	0.48	0.43 (0.38)	0.45 (0.43)
A-bulge(s)	13.7	1.5	0.73	0.71 (0.8)	—
A-bulge(1)	12.9	1.4	0.57	2.57 (3.7)	—
No central bp	11.9	0.97	0.46	—	—

The first column indicates the base pair or extra nucleotide (X:Y) at the center of a 11- or 10-bp RNA molecule, respectively (sequence CGCGCXGCGCG/CGCGCYGCGCG). No element indicates a 10-bp RNA $(CGCGCGCGCG)_2$, and G:A (t-ftf)/(t-shear) refers to a tandem G:A/A:G mismatch at the center in a face-to-face (t-ftf) or sheared (t-shear) arrangement $(CGCGCGAGCGCG)_2$, respectively. $\langle\Delta\theta_{bend}\rangle$ is the fluctuation of the angle of a linear axis fitted to the last five base pairs with respect to a linear axis fitted to the first five base pairs of each HM deformed structure. Rmsd_Y are the calculated Cartesian coordinate fluctuations of all atoms (Y = tot), belonging to base pairs directly flanking the central element (Y = fl), atoms of the central nucleotide(s) in the first strand (Y = cnu1) or second strand (Y = cnu2). Values in parentheses are the Cartesian coordinate fluctuations of only those atoms belonging to nucleobases. All angular fluctuations are given in degrees; all Cartesian coordinate fluctuations are given in Å.

regular 11-bp RNA or 10-bp RNA (10% and 15% increases in bend angle fluctuation, respectively). Such an increase in bend angle fluctuation from 11.9° (10-bp RNA) to 13.7° (stacked A-bulge) corresponds to what one would expect for a ~3-bp addition to the 10-bp RNA.

In contrast to the stacked bulge base, the bend angle flexibility of the looped-out bulge structure does not differ from the 11-bp reference structure. The harmonic analysis suggests no significant enhancement of the bend angle flexibility in the case of the looped-out bulge structure

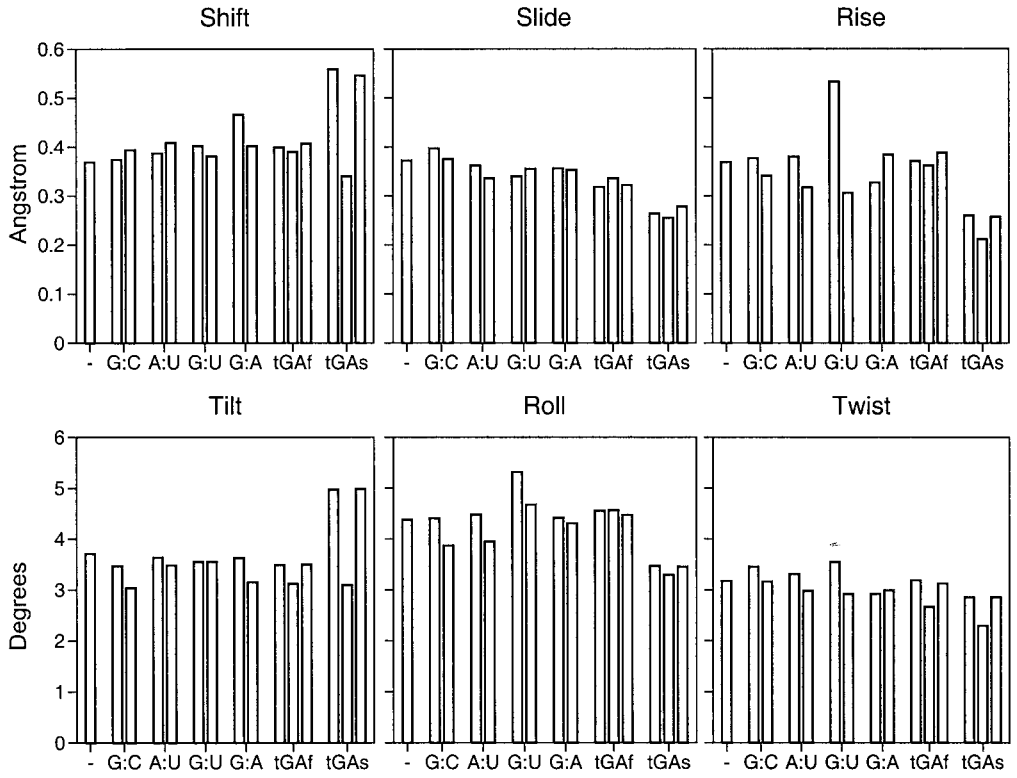

FIGURE 8 Helical coordinate fluctuations at base pair steps involving a central regular base pair or mismatch. At the *x* axis of each panel the central element is indicated (− means no central element; tGAf and tGAs refer to tandem GA mismatches in a face-to-face or sheared arrangement, respectively). The first box in each panel gives the helical coordinate fluctuation of the central step (C:G) of the 10-bp reference RNA $(CG)_5$. Each single mismatch or regular central base pair case (G:C, A:U, G:U, and G:A) involves a base pair step formed by the 5′ base pair and central element (*first bar*) and one between the central element and the 3′ base pair (*second bar*). For the tandem mismatches the middle bar indicates the helical coordinate fluctuations of the step formed by the tandem mismatches.

compared to the addition of a regular base pair. Note that because of the looped-out conformation of the bulge base the flanking sequences stack on each other in a fashion very similar to that of a dsRNA (10 bp) without the bulge base (Portmann et al., 1996). This favorable stacking arrangement presumably restricts the bending mobility of the looped-out bulge conformation.

The calculated local mobility in terms of atomic fluctuations of the stacked adenine base and neighboring nucleotides is larger than, for example, the mobility of adenine in the A:U base pair at the center of the RNA sequence (by ~40%/80%, considering atoms belonging to the bulge nucleotide/nucleobase, respectively; Table 3).

In the case of the looped-out adenine bulge conformation the loop base is predicted to be much more mobile than a base-paired adenine or a stacked extra adenine. The calculated Cartesian fluctuation of atoms of the looped-out base is ~5 times larger than for the stacked bulge base. Interestingly, the calculated atomic mobilities of atoms belonging to flanking nucleotides are not enhanced compared to regular RNA.

The increased mobility of the looped-out bulge base and, to a lesser extent, of the stacked bulge base is also reflected in enhanced fluctuations of the helical coordinates (Table 4). Because bulges contain unmatched (nonpaired) bases an analysis with respect to base pair coordinate fluctuations is not possible. In Table 4 the helical coordinate fluctuations of the central adenine with respect to the 5′ flanking base pair are summarized. In both stacked and looped-out adenine conformations, the helical coordinate fluctuations are significantly larger than for a base-paired adenine. In particular, twist fluctuations are two (stacked bulge) or even three or four times larger (looped-out form) than in the case of a regular base pair.

TABLE 4 Helical coordinate fluctuations of the bulge nucleotide with respect the 5′ flanking base pair

Coordinate	Central A:U	Central A (stacked)	Central A (looped-out)
⟨xdisp⟩	0.34	0.78	2.5
⟨ydisp⟩	0.35	1.2	1.9
⟨rise⟩	0.29	0.61	1.5
⟨inc⟩	5.6	9.8	11.4
⟨tip⟩	4.7	6.5	12.7
⟨twist⟩	3.6	7.9	13.2

The helical coordinate fluctuations were calculated for a description of the central adenine base with respect to the 5′ flanking base pair (C:G). The helical variables describing the adenine axis system with respect to the axis system associated with the 5′ base pair were calculated according to the definitions given by Lavery and Sklenar (1988a,b). Let $\vec{P}$ be the origin of the base pair system, let $\vec{u}$, $\vec{v}$, $\vec{w}$ be the associated axis system (calculated with Curves), and let $\vec{O}$ be the origin of the bulge base (fixed) system, where $\vec{x}$, $\vec{y}$, $\vec{z}$ are the axis vectors (*n* indicates normalized vectors):
xdisp = $(\vec{O} - \vec{P})\ (\vec{y} \times \vec{u})_n$; ydisp = $(\vec{O} - \vec{P})\ (\vec{u} \times (\vec{y} \times \vec{u})_n)$; rise = $(\vec{O} - \vec{P})\ \vec{u}$; inc = arcsin $(\vec{y}\ \vec{u})$; tip = ± arccos $((\vec{y} \times \vec{u})_n\ \vec{x})$, sign is the same as $-(\vec{u}\ \vec{x})$; twist = ± arccos $((\vec{y} \times \vec{u})_n\ \vec{v})$, sign is the same as $-(\vec{y} \times \vec{u})_n \vec{w}$. Helical fluctuations were calculated by summing over all contributions from structures deformed in each HM (excited by $\frac{1}{2}kT_R$).

DISCUSSION

The reduced number of variables needed to describe a nucleic acid in terms of the helicoidal coordinates in Jumna (Lavery et al., 1995) allows a rapid energy minimization down to very small residual gradients and fast solution of the eigenvalue problem to calculate harmonic modes. For example, the energy minimization and harmonic mode calculation of a 11-bp dsRNA molecule require only minutes on a SGI R10000 workstation, compared to up to several hours or days for Cartesian coordinates. A systematic application to several RNA structures is therefore feasible. In contrast to other rapid harmonic mode methods that, for example, model the nucleic acid by pseudo-atoms, each representing one nucleotide (Bahar and Jernigan, 1998), the present method still uses an all-atom model and allows us to interpret global deformations in terms of atomic interactions and helical coordinate changes at the nucleotide level.

A disadvantage of the current variables (similar to torsion angles) is the fact that there is no clear correlation between eigenvalue and Cartesian conformational fluctuation of the structure. However, it has been possible to simply excite each mode with an energy of $\frac{1}{2}kT_R$ and order the harmonic modes according to the desired property (i.e., atomic position fluctuations). This is similar to an ordering with respect to a force constant (eigenvalue) for Cartesian displacements.

As a test case to check the ability of the current approach to reproduce available experimental results on the atomic mobilities in RNA, it has been applied to a dsRNA x-ray structure determined at room temperature to high resolution (Egli et al., 1996). Calculated B-factors showed good qualitative agreement with experimental B-factors. Most notably, a characteristic pattern of the atomic mobilities along the sequence with highest B-factors found for the phosphate groups and lowest mobilities for the base atoms could be reproduced. A similar result has been obtained in a normal-mode analysis of B-DNA using torsion angle coordinates (Lin et al., 1997). It is important to note that a more than qualitative agreement between experimental and calculated B-factors is not expected because of the many factors that affect experimental B-factors and are not included in our calculations (e.g., resolution of the structure, crystal disorder, intermolecular contacts in the crystal, lattice vibrations, and structural substates in the crystal).

The main purpose of the present study was to characterize helical and global motions of regular RNA within the harmonic mode approximation and to study the influence of noncanonical elements. The calculated harmonic modes may give hints on the mechanism and magnitude of these motions and in turn could be helpful for understanding deformations of RNA occurring upon ligand binding, folding, and association.

Only a modest sequence dependence of the helical coordinate fluctuations was observed for the four regular RNA model structures. Notably, the roll and tilt fluctuations for

U/A-containing sequences were larger than for C/G-containing sequences. For DNA there is experimental evidence that the largest roll deformations involve pyrimidine-purine steps. To some degree this can also be seen in our calculations on RNA. The CG step showed a larger roll fluctuation than the GG(CC) or GC step. The calculated harmonic flexibility of A-form DNA was similar to that of RNA of the same sequence.

The calculated helical parameter flexibilities can also be compared to standard deviations of helical descriptors from a statistical analysis of A-DNA crystal structures (Jones et al., 1999). Note that an agreement is not necessarily expected, because alternative structures in crystals can represent different distinct substates of DNA molecules and can be influenced by the crystal environment. The HM calculations do not include these influences but account only for the RNA deformability close to standard A-form geometry. The standard deviation (SD) for the translational descriptors from crystal structures follows the decreasing order shift (0.5 Å), slide (0.4 Å), rise (0.3 Å), which is similar to the calculated order (flexibility of shift: ~0.4 Å, slide: ~0.35 Å, rise: ~0.3 Å; see Table 1). The SD of the orientational base-pair descriptors in the crystal structures follows the order Roll(5.0°), Twist (4.5°), and Tilt (2.7°), which differs from the calculated order of helical flexibilities (Roll: ~4.5–5.0°, Tilt: ~3.0–4.5°, Twist: ~2.9–3.6°).

The helical coordinate fluctuations calculated with the present approach compare well with results on A-DNA obtained using torsion angle variables by Ha Duong and Zakrzewska (1997a) and the Flex force field (Lavery et al., 1995). Ha Duong and Zakrzewska (1997a) found slightly larger rise fluctuations (~0.5–0.8 Å, compared to 0.3–0.45 Å in the present study) and slightly smaller tilt fluctuations for T/A-containing sequences (2.5–3.0° compared to 3–4° in the present study). The calculated helical coordinate fluctuations are generally smaller than what has been found as standard deviations in molecular dynamics simulations on RNA and A-DNA with explicit solvent and ions (Cheatham and Kollman, 1996, 1997). However, most helical coordinate fluctuations of base pairs from MD simulations are only 20–50% larger than the present estimates. For example, standard deviations of ~0.5–0.6 Å have been found for the translational helical coordinate fluctuations in MD (Cheatham and Kollman, 1997), compared to ~0.3–0.45 Å in the HM analysis and 4–8° for angular variables (MD) and 2.9–5° in the current HM study. In addition, the trend that angular helical fluctuations follow the increasing order twist, tilt, and roll has been observed in both the MD (Cheatham and Kollman, 1997) and HM analyses. This indicates that a significant part of the coordinate fluctuations of RNA can be modeled using a harmonic model.

For regular A-form RNA the estimated harmonic chain stiffness or persistence length is almost two times larger than what has been found experimentally (~720 Å; Kebbekus et al., 1995). Note, however, that an experimental measure of the persistence length usually does not distinguish between a "flexibility" due to a mixture of possible distinct substates of a RNA molecule (each may have a slightly different overall bend angle) and thermal fluctuations of one substate around its equilibrium state. It should be emphasized that in the present study only fluctuations around one substate are considered, and not transitions between distinct conformational states separated by barriers. Such transitions may add significantly to the observed flexibility (or heterogeneity) of RNA in solution.

In contrast to normal-mode or harmonic-mode calculations, sufficiently long molecular dynamics simulations in principle include transitions between conformational substates of RNA. In this regard it is interesting to note that molecular dynamics simulations of DNA in the presence of explicit water and ions indicate a persistence length smaller than the experimental value (220 Å, Cheatham and Kollman, 1996, compared to an experimental estimate of ~ 500 Å, Hagerman, 1981; note, however, that these authors may have used average angular fluctuations at each individual base pair step to calculate P, not considering that there is some covariation with adjacent steps (see current results), which lowers the "effective" fluctuation per step. Accounting for this covariation may considerably increase the calculated P).

It has been possible to analyze the sterically allowed motions that lead to global bending of an RNA molecule in terms of the helical motion at the nucleotide level. Similar to normal-mode studies by Ha Duong and Zakrzewska (1997a) and Matsumoto and Go (1999) on DNA, only very few harmonic modes were found to make a significant contribution to helix bending. The HM analysis suggests that the origin of RNA bending motions in terms of the local motion of each base pair (relative to the neighboring base pairs) is mainly a collective change in roll and slide along the sequence in such a way that it adds up to a global curvature.

Bending motions were found to be associated with significant changes of the RNA groove geometry. The calculations indicate that the major groove of RNA can fluctuate around its equilibrium width of ~4.5 Å by up to 1.4 Å at room temperature. The maximum possible width (due to thermal fluctuations) of the major groove of regular RNA is still much smaller than in the case of B-DNA, so that the accessibility by proteins is limited or requires an excess of energy to further open the groove.

To our knowledge no experimental data on the torsional rigidity of dsRNA are available. However, for B-DNA, experimental data on the twist flexibility suggest a twist fluctuation of ~4–6°/bp at room temperature (Shore and Baldwin, 1983; Taylor and Hagerman, 1990; Fujimoto and Schurr, 1990). It is interesting to compare this with our calculated "effective" twist flexibility of ~2°/ bp, which suggests that RNA is more rigid in terms of twist than DNA (harmonic mode calculations on B-DNA indicate that the

calculated harmonic twist flexibility is indeed higher for B-DNA, ~3–4°/bp; data not shown).

It should be noted that because of the present harmonic approximation the (twist) torsional rigidity of the RNA is assumed to be independent of temperature. Experimentally observed changes of the twist flexibility with temperature (for example, a decrease in the torsional rigidity observed by Delrow et al. (1998) for B-DNA) cannot be modeled by the present HM approach.

A negative covariance of helical coordinate deformations at one base pair step and neighboring base pair steps was observed for most helical coordinates. This indicates that, for example, the overall twist variance of a given RNA helix is not simply the sum of the variances at each step (as would be the case for independent twist fluctuations). The analysis of covariances of helical fluctuations in double-stranded nucleic acids could be extended to many more covariances than considered in the present study. Such an analysis could be interesting, in particular, for the study of DNA flexibility with regard to DNA recognition (Gorin et al., 1995; Olson et al., 1998).

By simply adding the contributions of each deformation mode to the covariance of a given pair of helical coordinates, little coupling of deformations was found for base pairs beyond nearest or next-nearest neighbors. We like to point out that such a treatment neglects possible differences in the relaxation times of each mode. Each HM shows a strong covariance of coordinate displacements along the sequence. Under the assumption that distortions in the modes relax with different characteristic relaxation times (due to different characteristic diffusion constants), correlated motions may extend much farther along the sequence (Schurr and Fujimoto, 1999).

Substitution of canonical base pairs at the center of a dsRNA by two common and relatively stable non-Watson-Crick base pairs (G:U and G:A) did not alter the harmonic flexibility significantly (except for an increase in the rise fluctuation in the case of G:U). This result is compatible with the interpretation of crystallographic structures that contain single G:U and G:A mismatches. These elements were found to adopt conformations in the crystal very similar to those of regular A-form RNA with no apparent B-factor increase or structural heterogeneity (Leonhard et al., 1994; Mueller et al., 1999). It should be emphasized that the present study does not exclude the possibility that the mismatches more easily adopt distinct subconformations than regular Watson-Crick base pairs that are separated by small energy barriers in solution, which can add to the conformational heterogeneity of the RNA.

Tandem G:A mismatches have been shown to adopt two very different conformations in RNA, depending on flanking sequences (sheared or face-to-face form; SantaLucia and Turner, 1993; Wu and Turner, 1996; see Materials and Methods). The harmonic-mode analysis provides evidence that these two conformations also show important differences in conformational deformability. The presence of a tandem G:A mismatch in the sheared conformation is predicted to stiffen the RNA compared to regular RNA of the same length or tandem G:A mismatches in the ftf form. A possible reason for this finding is that a sheared G:A base pair forms a more compact structure than a face-to-face G:A or a regular Watson-Crick base pair. Based on this result, one could speculate that one possible function of the sheared tandem G:A mismatch in many biologically important RNA molecules (i.e., hammerhead ribozyme) is to reduce the conformational flexibility and to stiffen the surrounding structure.

The HM analysis indicates that the presence of a single-bulge nucleotide can increase the conformational flexibility of dsRNA and that the effect depends on the bulge conformation. For the bulge base in the looped-out conformation the mobility of the bulge base is strongly enhanced compared to a base-paired nucleobase or a stacked bulge base. This result is consistent with NM calculations on $tRNA_{Phe}$, which indicate significantly enhanced fluctuations of nucleotides not involved in stacking interactions (Matsumoto et al., 1999). In addition to the increased harmonic motion close to one energy minimum, it is likely that the structural heterogeneity of the looped-out base is also increased by the possibility of adopting other distinct substates in solution (see, for example, Zacharias and Sklenar, 1997, 1999b). The flanking nucleotides in the looped-out bulge structure are stacked in a fashion very similar to that of regular A-form, and the calculated global bending flexibility was found to be close to the bending flexibility of a helix with a regular base pair instead of the bulge. Interestingly, in contrast to the HM result, for the crystal structure of a single adenine bulge no strongly enhanced B-factors were reported for the looped-out bulge base (Portmann et al., 1996). However, this looped-out base contacts a neighboring molecule in the crystal that presumably restricts the mobility of the bulge base.

For the stacked bulge form the calculated bending flexibility was slightly larger than that of regular dsRNA, and the bulge base appeared to be more mobile than a paired adenine. However, as expected, the bulge base mobility is much smaller than in case of the looped-out structure.

It is straightforward to apply the present HM approach to other noncanonical structural motifs to get an approximate but very rapid impression of the deformability of these structures in the vicinity of an energy minimum. An application to structures much larger than those considered in the current study, such as tRNAs or even the 160 nucleotide P_4P_6 domain of the group I ribozyme (Cate et al., 1996), is not out of reach of the method. Study of the global motions of such RNAs in terms of nucleotide and atomic motions may add to the understanding of the function of these multidomain structures.

We acknowledge helpful discussions with Wilma K. Olson. MZ was supported in part by a Deutsche Forschungsgemeinschaft Habilitation grant.

REFERENCES

Auffinger, P., and E. Westhof. 1998. Simulations of the molecular dynamics of nucleic acids. *Curr. Opin. Struct. Biol.* 2:227–236.

Bahar, I., and R. L. Jernigan. 1998. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms. *J. Mol. Biol.* 281:871–884.

Beveridge, D. L., K. J. McConell, R. Nirmala, M. A. Young, S. Vijayakumar, and G. Ravishankar. 1994. Recent progress in molecular dynamics simulations of DNA and protein-DNA complexes including solvent. *ACS Symp. Ser.* 568:381–394.

Brooks, B. R., D. Janezic, and M. Karplus. 1995. Harmonic analysis of large systems. I. Methodology. *J. Comput. Chem.* 16:1522–1542.

Brooks, B. R., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.

Case, D. A. 1994. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* 4:285–290.

Cate, J. R., A. R. Gooding, E. Podell, K. Thou, B. L. Golden, C. E. Kundrot, T. R. Cech, and J. A. Doudna. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science.* 273: 1678–1686.

Cheatham, T. E., and P. A. Kollman. 1996. Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* 259:434–444.

Cheatham, T. E., and P. A. Kollman. 1997. Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA and DNA:RNA hybrid duplexes. *J. Am. Chem. Soc.* 119:4805–4825.

Cheatham, T. E., J. L. Miller, T. I. Spector, P. Cieplak, and P. A. Kollman. 1997. Molecular dynamics simulations on nucleic acid systems using the Cornell et al. force field and particle mesh Ewald electrostatics. *In* Molecular Modeling of Nucleic acids. N. B. Leontis, editor. American Chemical Society, Washington, DC. 285–303.

Cornell, W. D., P. Cieplak, C. I. Bayley, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Cadwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Delrow, J. J., P. J. Health, B. S. Fujimoto, and J. M. Schurr. 1998. Effect of temperature on DNA secondary structure in the absence and presence of 0.5 M tetramethylammonium chloride. *Biopolymers.* 45:503–515.

Dickerson, R. E., M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C. S. Tung, A. H.-J. Wang, and V. B. Zhurkin. 1989. Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.* 205:787–791.

Egli, M., S. Portmann, and N. Usman. 1996. RNA hydration: a detailed look. *Biochemistry.* 35:8489–8496.

Feig, M., and B. M. Pettitt. 1998. Structural equilibrium of DNA represented with different force fields. *Biophys. J.* 75:134–149.

Flatters, D., K. Zakrzewska, and R. Lavery. 1997. Internal coordinate modeling of DNA: force field comparisons. *J. Comput. Chem.* 18: 1043–1055.

Fujimoto, B. S., and J. M. Schurr. 1990. Dependence of the torsional rigidity of DNA on base composition. *Nature.* 344:175–178.

Garcia, A. E., and D. M. Soumpasis. 1989. Harmonic vibrations and thermodynamic stability of a DNA oligomer in monovalent salt solution. *Proc. Natl. Acad. Sci. USA.* 86:3160–3164.

Gorin, A. A., V. B. Zhurkin, and W. K. Olson. 1995. B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* 247:34–38.

Ha Duong, T., and K. Zakrzewska. 1997a. Calculation and analysis of low frequency normal modes for DNA. *J. Comput. Chem.* 18:796–811.

Ha Duong, T., and K. Zakrzewska. 1997b. Influence of drug binding on DNA flexibility: a normal mode analysis. *J. Biomol. Struct. Dyn.* 14: 691–701.

Ha Duong, T., and K. Zakrzewska. 1999. Sequence specificity of bacteriophage 434 repressor-operator complexation. *J. Mol. Biol.* 280:31–39.

Hagerman, P. J. 1981. Investigation of the flexibility of DNA using transient electric birefringence. *Biopolymers.* 20:1503–1535.

Irikura, K. K., B. Tidor, B. R. Brooks, and M. Karplus. 1985. Transition from B to Z DNA: contribution of internal fluctuations to configurational entropy difference. *Science.* 229:571–572.

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.

Joshua-Tor, L. F., D. Rabinovich. H. Hope, F. Frolow, E. Appela, and J. L. Sussman. 1988. The three-dimensional structures of a DNA duplex containing looped-out bases. *Nature.* 334:82–84.

Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A.* 34:827–828.

Kebbekus, P., D. E. Draper, and P. J. Hagerman. 1995. Persistence length of RNA. *Biochemistry.* 34:4354–4357.

Kottalam, I., and D. A. Case. 1990. Langevin modes of macromolecules: applications to crambin and DNA hexamers. *Biopolymers.* 29: 1409–1421.

Lavery, R., and H. Sklenar. 1988a. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lavery, R., and H. Sklenar. 1988b. defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6:655–667.

Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimization of nucleic acids). *Comput. Phys. Comm.* 91:135–158.

Leonhard, G. A., K. E. McAuley-Hecht, S. Ebel, D. M. Lough, T. Brown., and W. N. Hunter. 1994. Crystal and molecular structure of r(CGCGAAUUAGCG)$_2$ an RNA duplex containing two G(anti):A(anti) base pairs. *Structure.* 15:483–494.

Levitt, M. A., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* 181:423–447.

Lin, D., A. Matsumoto, and N. Go. 1997. Normal mode analysis of a double-stranded DNA dodecamer d(CGCGAATTCGCG)$_2$. *J. Chem. Phys.* 107:3684–3690.

Matsumoto, A., and N. Go. 1999. Dynamic properties of double-stranded DNA by normal mode analysis. *J. Chem. Phys.* 111:11070–11075.

Matsumoto, A., M. Tomimoto, and N. Go. 1999. Dynamical structure of transfer RNA studied by normal mode analysis. *Eur. Biophys. J.* 28: 369–379.

Mueller, U., H. Schübel, M. Sprinzl, and U. Heinemann. 1999. Crystal structure of acceptor stem tRNA^{Ala} from *Escherichia coli* shows unique G:U wobble base pair at 1.16 Å resolution. *RNA.* 5:670–677.

Nakamura, S., and J. Doi. 1994. Dynamics of transfer RNAs analyzed by normal mode calculations. *Nucleic Acids Res.* 22:514–552.

Nikonowicz, E., R. P. Meadows, and D. G. Gorenstein. 1990. NMR structural refinement of an extrahelical adenosine tridecamer d(CGCAGAATTCGCG)$_2$ via a hybrid relaxation matrix procedure. *Biochem.* 29:4193–4204.

Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 95:11163–11168.

Portmann, S., S. Grimm, C. Workman, N. Usman, and M. Egli. 1996. Crystal structures of an A-form duplex with single adenosine bulges and conformational basis for site-specific RNA self-cleavage. *Chem. Biol.* 3:173–184.

SantaLucia, J., and D. H. Turner. 1993. Structure of (rGGCGAGCC)$_2$ in solution from NMR and restraint molecular dynamics. *Biochemistry.* 32:12612–12623.

Schurr, J. M., and B. S. Fujimoto. 1999. Dynamic twisting correlations in a model DNA with uniform torsion elastic constant. *Biopolymers.* 49: 355–359.

Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* 170: 957–981.

Sklenar, H., R. Lavery, and B. Pullman. 1986. The flexibility of nucleic acids: SIR, a novel approach to the variation of polymer geometry in constraint systems. *J. Biomol. Struct. Dyn.* 3:967–987.

Taylor, W. H., and P. J. Hagerman. 1990. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* 212:363–376.

Tidor, B., K. K. Irikura, B. R. Brooks, and M. Karplus. 1983. Dynamics of DNA oligomers. *J. Biomol. Struct. Dyn.* 1:231–252.

Wu, M., and D. H. Turner. 1996. Solution structure of $(rGCGGACGC)_2$ by two-dimensional NMR and iterative relaxation matrix approach. *Biochemistry.* 35:9677–9689.

Zacharias, M., and P. J. Hagerman. 1995. Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol.* 247: 486–500.

Zacharias, M., and H. Sklenar. 1997. Analysis of the stability of looped-out and stacked-in conformations of an adenine bulge in DNA using a continuum model for solvent and ions. *Biophys. J.* 73:2990–3003.

Zacharias, M., and H. Sklenar. 1999a. Harmonic modes as variables to approximately account for receptor flexibility in ligand-receptor docking simulations: application to a minor-groove DNA binding ligand. *J. Comput. Chem.* 20:287–300.

Zacharias, M., and H. Sklenar. 1999b. Conformational analysis of single-base bulge structures in A-form DNA and RNA using a hierarchical approach and energetic evaluation with a continuum solvent model. *J. Mol. Biol.* 289:261–275.