Sequence-Dependent DNA Curvature and Flexibility from Scanning Force Microscopy Images

Anita Scipioni,* Claudio Anselmi,* Giampaolo Zuccheri,† Bruno Samori,† and Pasquale De Santis*
*Dipartimento di Chimica, Università "La Sapienza," Roma I-00185, Italy; and †Dipartimento di Biochimica G. Moruzzi, Università di Bologna, Istituto Nazionale per la Fisica della Materia, Bologna I-40126, Italy

ABSTRACT This paper reports a study of the sequence-dependent DNA curvature and flexibility based on scanning force microscopy (SFM) images. We used a palindromic dimer of a 1878-bp pBR322 fragment and collected a large pool of SFM images. The curvature of each imaged chain was measured in modulus and direction. It was found that the ensemble curvature modulus does not allow the separation of static and dynamic contributions to the curvature, whereas the curvature, when its direction in the two dimensions is taken into account, permits the direct separation of the intrinsic curvature contributions static and dynamic contributions. The palindromic symmetry also acted as an internal gauge of the validity of the SFM images statistical analysis. DNA static curvature resulted in good agreement with the predicted sequence-dependent intrinsic curvature. Furthermore, DNA sequence-dependent flexibility was found to correlate with the occurrence of A·T-rich dinucleotide steps along the chain and, in general, with the normalized basepair stacking energy distribution.

INTRODUCTION

Superstructural properties of DNA chain are involved in the management of the informational content of DNA. Protein-DNA association, transcription, replication, and recombination, as well as packaging and writhing transitions, are controlled by the static (intrinsic) and dynamic curvature of the DNA sequences involved. This seems to be largely accepted at present. However, a better knowledge of these mechanical properties and of their physical origin in terms of DNA sequence is crucial to improve our understanding of the molecular biology of DNA.

Many biophysical properties of DNA are influenced by its local curvature and flexibility, and different experimental techniques were designed to study them, such as gel electrophoresis, circularization kinetics, electric dichroism and, more recently, scanning force microscopy (SFM) visualization.

Polyacrilamide gel electrophoresis was the first simpler and commonly used technique to investigate the curvature and flexibility of DNA tracts. Curved DNAs were found to migrate anomalously so that their molecular weights appear to be higher or, in a few cases, lower than the real ones by a retardation factor (Marini et al., 1982; Wu and Crothers, 1984; Koo et al., 1986; Hagerman, 1985, 1986; Diekmann, 1987; Chastain II et al., 1995; Chastain and Sinden, 1998). Different models have been proposed to relate electrophoretic retardation to DNA curvature and flexibility (Trifonov, 1980; Lumpkin et al., 1985; De Santis et al., 1986, 1988; Trifonov et al., 1987; Koo and Crothers, 1988; Levene and Zimm, 1989; Bolshoy et al., 1991; Olson and Zhurkin, 1996). The experimental evidence is that intrinsic curvature plays the major role while the flexibility is generally not well defined or controversially considered. Therefore, models based on intrinsic curvature are capable of successfully predicting gel electrophoresis anomalies, also neglecting the sequence-dependent flexibility.

Information about DNA curvature and flexibility can be obtained also from circularization kinetics experiments of selected DNA tracts. A large pool of DNA sequences was investigated by different authors, and their propensity to circularization was experimentally determined and predicted by theoretical approaches (Jacobson and Stockmayer, 1950; Yamakawa and Stockmayer, 1972; Flory et al., 1976; Shimada and Yamakawa, 1984; De Santis et al., 1996) or simulated by Monte Carlo methods (Levene and Crothers, 1986; Sprous et al., 1996; Kahn and Crothers, 1998; Hockings et al., 1998; Roychoudhury et al., 2000).

Although curvature and flexibility cooperate in DNA circularization and looping by reducing the distortion energy cost (De Santis et al., 1996), the entropy reduction in nucleosome formation, where a free DNA tract is frozen by the interaction with the histone octamer, does not appear to favor flexible DNAs (Anselmi et al., 1999, 2000).

However, while the main determinants of DNA curvature are generally recognized, the evaluation of DNA flexibility in terms of the sequence appears to be controversial and still remains an open problem. For instance, the same dinucleotide steps are considered flexible in certain cases, rigid in others. The dispersion of the orientational parameters of x-ray crystal structures of double helix oligonucleotides with different sequences produces a scale of flexibility where the AA·TT step belongs to the rigid class, while GG·CC and GC·GC dinucleotides result in being more flexible (Olson et al., 1993).

However, this approach could be invalidated by the systematic choice, made by the crystallographers, of double-stranded (ds)-oligonucleotides characterized by G·C-rich terminals, whereas the A·T steps were segregated in the

Submitted February 28, 2002; and accepted for publication June 25, 2002.
Address reprint requests to Pasquale De Santis, Università "La Sapienza," P. le A. Moro 5, Rome I-00185, Italy. Tel.: 39-06-49913228; Fax: 39-06-4453827; E-mail: pasquale.desantis@uniroma1.it.

0006-3495/02/11/2408/11 $2.00

central regions. The terminals are more sensitive to the different crystal packing effects, so that the average basepair orientational parameters and the relative dispersion result to be significantly influenced. In fact, a different scale was obtained from the statistical analysis of x-ray crystal structures of protein-oligonucleotide complexes (Olson et al., 1998). Recently, Langowski and co-workers investigated short DNA fragments containing AA·TT, AT·AT, GC·GC, or GG·CC dinucleotide steps by means of molecular dynamics simulations (Lankas et al., 2000). They tried an evaluation of the twisting and bending persistence lengths and the bending anisotropy and cross-term elastic constants. AT·AT stands out as the most flexible dinucleotide, whereas GC·GC results the most rigid. AA·TT and GG·CC show comparable bendability. In all cases they found a pronounced bending anisotropy.

However, the thermodynamic differential stability of ds-oligonucleotides (Sugimoto et al., 1996; SantaLucia, 1998) and of double helical polynucleotides (Gotoh and Tagashira, 1981) suggests a quite different scale of flexibility. This correlates with the quantum-mechanical evaluation of the stacking energy of the 16 dinucleotide steps (Ornstein et al., 1978) and was successfully adopted in a theoretical model of the sequence-dependent nucleosome stability (Anselmi et al., 1999, 2000).

It is therefore very important to define the differential flexibility of the DNA chain in terms of thermodynamic parameters directed by its sequence. SFM (or AFM) has been recently proved to be a very useful tool to this end (Zuccheri et al., 2001).

Mapping the curvature and flexibility along the DNA sequence of a DNA molecule that is imaged just as a thin trace by EM (Muzard et al., 1990; Bednar et al., 1995) or by SFM (Rivetti et al., 1996, 1998; Cognet et al., 1999), requires the identification of which of the two imaged molecular ends corresponds to the starting point of the sequence between the two alternative choices. This problem was commonly solved by labeling (e.g., with ferritin-avidin-streptavidin complex or by gold balls) one terminal of DNA chain or, more recently, by adopting the strategy of palindromic constructs (Zuccheri et al., 2001). This latter approach overcomes the problem because the two alternatives are equivalent and provides the doubling of observations of the curvature versus the sequence and an internal gauge of the statistical validity of the curvature evaluation. This approach is less prone to artifacts due to modifications of the curvature and flexibility of the terminal DNA tract where the label had been introduced.

When a sufficiently large pool of SFM images is collected, the curvature of each imaged chain can be easily measured in modulus and direction. Different authors adopted the curvature modulus to study and test the relation between sequence and curvature (Muzard et al., 1990; Bednar et al., 1995; Rivetti et al., 1996, 1998; Cognet et al., 1999). Incidentally, some authors used approximate evaluation of the curvature as related to the ratio between the chain contour length and the corresponding cord, which, actually, converges to the squared curvature modulus for real DNA chains. Nevertheless, in the cases investigated an apparent qualitative agreement was found with the theoretical predictions of DNA intrinsic curvature (Muzard et al., 1990).

In a previous paper (Zuccheri et al., 2001) we reported a first approach to the statistical analysis of the curvature and flexibility distribution along DNA molecules. However, we did not consider the segmental nature of the SFM data and its implications in the evaluation of curvature and flexibility. This does not substantially modify the flexibility and curvature trends along the sequence, but introduces changes in their amplitudes.

To achieve a deeper insight into the mechanism by which the DNA sequence drives the local curvature and flexibility along the chain we further improved the theoretical model previously proposed (Zuccheri et al., 2001) and implemented the statistical analysis on a large pool of SFM images of a palindromic dimer (1878 bp of a pBR322 DNA tract).

The present paper reports an explicit treatment of the statistical analysis and the interesting relations between the average curvature modulus and the associated standard deviation. The results reported here show that intrinsic curvature and flexibility can be not only experimentally mapped along the chain of a DNA molecule, but also very effectively predicted on the basis of a theoretical framework based on a sequence-dependent model of the DNA curvature and flexibility we first advanced (De Santis et al., 1986; Anselmi et al., 1999, 2000).

A strict correlation of the flexibility with the differential thermodynamic stability of DNAs was also confirmed. Finally, a validity test of the short-range equilibration of DNA molecules on flat surfaces is proposed as related to the predicted proportionality between the ensemble (or time) average of the curvature modulus and the corresponding standard deviation.

MATERIALS AND METHODS

SFM imaging, image processing, and molecule measurements

The dimer of the *Eco*RV/*Pst*I fragment of pBR322 DNA, having a length of 1878 bp, was obtained as reported in our previous paper using standard molecular biology techniques (Sambrook et al., 1989; Zuccheri et al., 2001). This molecule, whose ends are the two *Eco*RV half-sites, consists of two copies of the *Eco*RV/*Pst*I fragment of 937 bp plus the four bases of the annealed *Pst*I staggered ends.

DNA molecules have been deposited on freshly cleaved ruby mica (B & M Mica, New York) from a nanomolar DNA solution containing 4 mM Hepes buffer (pH 7.4), 10 mM NaCl, and 2 mM $MgCl_2$. The Mg(II) ions are added to promote DNA adsorption on mica (Hansma and Laney, 1996; Bustamante et al., 1992). 10–15 μl of DNA solution is deposited on a 1- to 1.5-cm^2 mica disk and left there for ~2 min. The solution then is rinsed with 2–3 ml of Milli Q deionized water (Millipore,) added by drops and dried under a gentle flow of nitrogen gas.

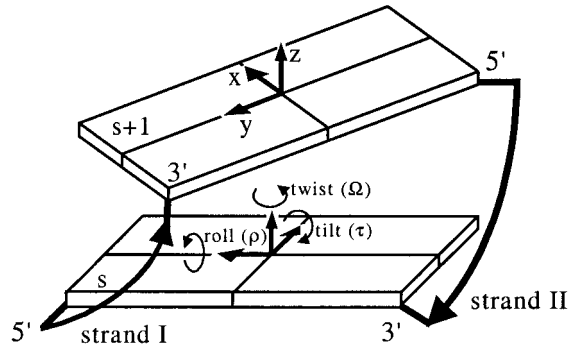

FIGURE 1. Orientational parameters of the basepair average plane in a dinucleotide step.

Imaging was performed in tapping-mode with PointProbe noncontact silicon probes (Nanosensors, Wetzlar-Blankenfeld, Germany) on a NanoScope IIIa SFM system equipped with a Multimode head and a type E piezoelectric scanner (Digital Instruments, Santa Barbara, CA). Images have been recorded with a 10–15 μm/s linear scanning speed at a sampling density of 4–9 nm^2/pixel. Raw SFM images have been processed only for background removal (flattening) using the microscope manufacturer's image-processing software. DNA molecule profiles have been measured from the SFM images using ALEX, a software package written for probe microscopy image processing (Rivetti et al., 1996), by semiautomatically tracking the molecule contours on the SFM images.

The distribution of the molecule contour lengths is clustered around a value very close to that expected for B-DNA and a 6% relative standard deviation on lengths. An approximate evaluation of the error in the determination shows that 0.5% of the standard deviation could be due to the semiautomatic digitalization of the molecule profiles, while 1% could be due to the recording of the image itself (microscope instrumental imprecision). In digitizing the molecule contours we have left out molecules with suspiciously short contour length (probably fragments), and a few that had suspiciously long contour lengths.

The model

According to the classical formulation by Landau and Lifshitz (1970), the curvature of a space line is defined as the derivative $C = \mathrm{d}\mathbf{t}/\mathrm{d}\ell$ of the tangent versor, $\mathbf{t}$, along the line, ℓ. Its modulus is the inverse of the curvature radius and its direction is that of the main normal to the curve. In the case of DNA, the line corresponds to the helical axis and the curvature is a vectorial function of the sequence. It represents the angular deviation between the local helical axes of the nth and $(n+1)$th helix turns centered on the nth and $(n+1)$th dinucleotide step, respectively (Fig. 1).

The DNA curvature is a superstructural sequence-dependent property, which is continuously changed by the thermal energy of the environment. Assuming first-order elasticity corresponding to a linear response of the DNA to such energy exchanges, the ensemble average value of curvature is:

$$\langle \mathbf{C}(n) \rangle = \langle \mathbf{C}_o(n) + \boldsymbol{\chi}(n) \rangle = \mathbf{C}_o(n) + \langle \boldsymbol{\chi}(n) \rangle \qquad (1)$$

where $\mathbf{C}(n)$ is the observed curvature per basepair at position n, $\mathbf{C}_o(n)$ is the corresponding intrinsic curvature, and $\boldsymbol{\chi}(n)$ represents the dynamic fluctuation. Angle brackets mean averaging is over the statistical ensemble. The latter term, which corresponds to the dynamic contribution to the curvature, is obviously zero. Therefore, the intrinsic curvature, $\mathbf{C}_o(n)$, is the statistical average of the curvature at the nth sequence position or alternatively, the time average of the DNA superstructure.

Different models were proposed to evaluate the intrinsic curvature from the sequence. The first, advanced by Crothers and co-workers (Wu and Crothers, 1984; Koo et al., 1986), was based on the finding that the curvature is generally associated with repeated AA·TT tracts in phase with the period of B-DNA structure. The curvature should be a consequence of the helical symmetry breakdown at the boundary between B-DNA and a modified helical form that would characterize the repeated AA·TT sequences. The other models are based on the hypothesis that the differential interactions between the nearest-neighbor basepairs are the main determinant of the DNA curvature, first advanced by Trifonov (1980) as a model of sequence-dependent bendability. Therefore, values of the mutual orientational angles for each of the 10 independent dinucleotide steps were derived starting from conformational energy calculations, fitting electrophoresis retardations of a large pool of ds-multimeric oligonucleotides or from x-ray crystal structures of ds-oligonucleotides. Such models were critically commented on by Crothers (1998).

We first advanced (De Santis et al., 1986) a simple model based on conformational energy minimization of the 10 independent dinucleotide steps that provided an evaluation of the differential deviation angles from the standard B-DNA. Such a model coherently explained the retardation factors of a large number of multimeric oligonucleotides published by Crothers and co-workers (Wu and Crothers, 1984; Koo et al., 1986), Hagerman (1985, 1986), and Diekmann (1987), as well as the origin of differential curvature of kinetoplast DNA from *Leishmania tarentolae* (Marini et al., 1982), and introduced a useful diagrammatic representation of the curvature along the sequence. The model was later refined to best fit the gel electrophoresis retardation of a large pool of DNAs (De Santis et al., 1988). The roll (ρ), tilt (τ), and twist (Ω) parameters (Table 1) were defined for each dinucleotide step (Fig. 1). These parameters represent a self-consistent set, useful to predict the large-scale structure and statistical mechanics of DNAs, the related thermodynamic properties, and their biological implications.

We introduced a simple evaluation of the differential curvature along the DNA sequence as

$$\mathbf{C}_o(n) = \frac{1}{\nu_n} \sum_{\text{nth turn}} \mathbf{d}_s \exp\left(\frac{2\pi i s}{\nu_n}\right) \qquad (2)$$

where $\mathbf{C}_o(n)$ is the intrinsic curvature per basepair at the position n of the sequence; $\mathbf{d}_s = \rho_s - i\tau_s$ is the complex representation of the roll and tilt angles at the sth position, and $\nu_n = 360/\langle\Omega\rangle_n$ is the corresponding helical

TABLE 1 Roll, ρ, tilt, τ, and twist, Ω, angles in degrees for each dinucleotide step

	A			T			G			C		
	ρ	τ	Ω	ρ	τ	Ω	ρ	τ	Ω	ρ	τ	Ω
T	8.0	0.0	34.5	−5.4	0.5	36.0	6.8	−0.4	34.1	2.0	1.7	34.6
A	−5.4	−0.5	36.0	−7.3	0.0	35.3	1.0	−1.6	34.4	−2.5	−2.7	33.7
C	6.8	0.4	34.1	1.0	1.6	34.4	4.6	0.0	33.5	1.3	0.6	33.1
G	2.0	−1.7	34.6	−2.5	2.7	33.7	1.3	−0.6	33.1	−3.7	0.0	33.3

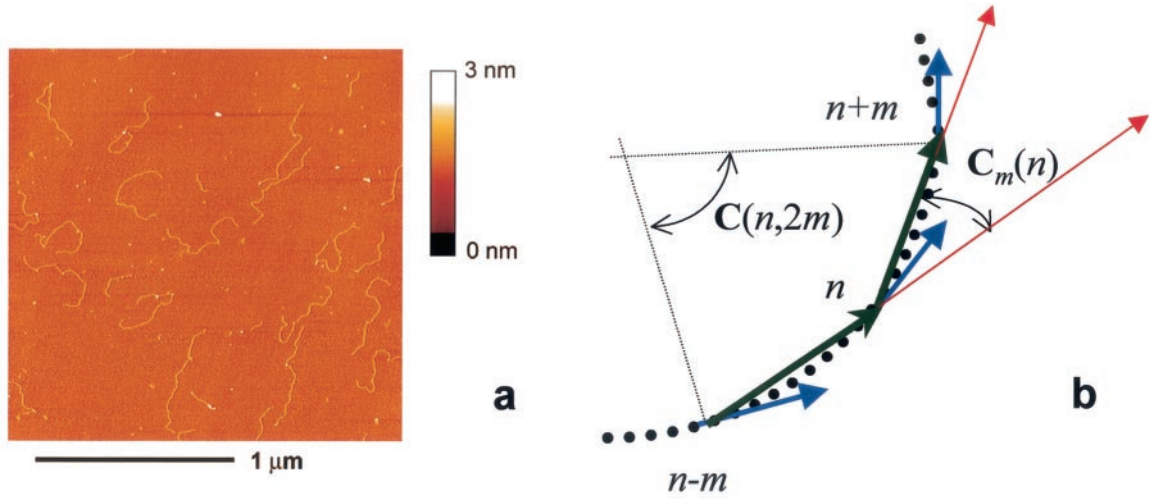

FIGURE 2. (*a*) An example tapping-mode SFM image of DNA palindromic dimers deposited on a surface of fresh ruby mica. (*b*) The sequence of black dots represents the DNA trace; the black arrows represent the directional segmental chain spanning *m* bp; the red arrows define the angle between the virtual segments at a sequence position *n*, $C_m(n)$. This angle is half of the curvature $C(n, 2m)$, which represents the angular deviation of the local helical axis (*blue arrows*, tangent to the DNA trace) pertinent to nucleotide steps separated by $2m$ bp.

periodicity. The summation is iteratively extended to each turn of double helix and assigned to the central basepair. Such a formulation of the intrinsic sequence-dependent curvature function in the complex plane represents in modulus and phase the local deviation of the DNA helical axes from the straight direction.

Adopting first-order elasticity and according to Landau and Lifshitz (1970), we define the bending distortion energy of an *m* bp DNA tract as:

$$\Delta E_b = \frac{1}{2}\sum_1^m b(n)|\mathbf{C}(n) - \mathbf{C}_o(n)|^2 = \frac{1}{2}\sum_1^m b(n)|\boldsymbol{\chi}(n)|^2 \quad (3)$$

where $b(n)$ is equal to the product of Young's modulus and the inertia moment of an isotropic rod-like DNA chain and represents the apparent harmonic force constant at the position *n* of the sequence. It can be conveniently factorized as the product of an average DNA force constant b^* times the sequence-dependent rigidity factor (or divided by the flexibility factor $f(n)$). Such a factor is 1 for a DNA chain with a random distribution of the basepairs and represents the deviation of the dinucleotide step at position *n* from the average DNA flexibility. If $f(n)$ is higher or lower than 1, the flexibility at position *n* is proportionally higher or lower than the average, respectively; $b^* = RTP/\ell$ is related to the average persistence length *P*, where $\ell = 0.34$ nm is the helix rise of the standard B-DNA structure.

Therefore, adopting the first-order elasticity for DNA axis deformations and setting $B = b(n)/2RT$, we define the elastic partition function as

$$Q(n) = \int \exp(-B|\boldsymbol{\chi}(n)|^2)d\boldsymbol{\chi}(n) \quad (4)$$

and the curvature dispersion at position *n* becomes

$$\langle|\boldsymbol{\chi}(n)|^2\rangle = \frac{1}{Q(n)}\int |\boldsymbol{\chi}(n)|^2 \exp(-B|\boldsymbol{\chi}(n)|^2)d\boldsymbol{\chi}(n) = \frac{1}{B} \quad (5)$$

according to Hagerman's result in solution (Hagerman, 1988). It should be noted that in 3D, $\boldsymbol{\chi}(n) = \mathrm{C}(n) - \mathbf{C}_o(n)$ is a complex quantity (defined in modulus and phase) according to our definition of curvature (Eq. 2) and the integration is made in the complex plane. However, when DNA is forced from a 3D space to a 2D surface, as in SFM microscopy, the curvature phase is restricted to 0 and $\boldsymbol{\pi}$ and the integral in Eq. 5 is made on the real axis. As a consequence, the lower dimensionality of the chain halves the curvature dispersion.

The large variety of shapes, assumed by DNA molecules under the thermal stochastic perturbation by their molecular environment, is well illustrated by the example of their SFM images reported in Fig. 2 *a*. Due to SFM resolution limits, DNA images are generally fitted by segmental chains. Therefore, we introduced a new parameter characterizing the curvature, $C_m(n)$, corresponding to the angle between the *m* bp virtual segments at a sequence position, *n*, as well as the corresponding deviation, $\chi_m(n)$. Such a segmental curvature $C_m(n)$ is practically half of the curvature $C(n, 2m)$, which represents the angular deviation of the local helical axis pertinent to nucleotide steps separated by $2m$ bp as illustrated in Fig. 2 *b*. The latter can be theoretically calculated while the first one is what we experimentally measure. Thus, we introduce the curvature as a function of both the position *n* and the length of virtual segments, *m*.

Therefore, the average curvature deviation after $2m$ bp is

$$\langle\chi(n, 2m)\rangle = \langle C(n, 2m) - \langle C(n, 2m)\rangle\rangle = 0 \quad (6)$$

and the relative dispersion is

$$\langle\chi^2(n, 2m)\rangle = \frac{1}{Q(n, 2m)}\int_{-\infty}^{\infty} \chi^2(n, 2m)\exp \times\left(-\frac{B}{2m}\chi^2(n, 2m)\right)d\chi(n, 2m)$$

$$= \frac{m}{B} = \frac{2mRT}{b(n)} \quad (7)$$

where

$$Q(n, 2m) = \int_{-\infty}^{\infty} \exp\left(-\frac{B}{2m}\chi^2(n, 2m)\right)d\chi(n, 2m) \quad (8)$$

Therefore, the experimentally measured $\chi_m^2(n)$ is $\chi^2(n,2m)/4$ because it depends on the curvature fluctuations of $2m$ bp, and its square root is the corresponding standard deviation. Therefore, from Eq. 7, recalling that $B = b(n)/2RT$

$$SD(C_m(n)) = \left(\frac{m}{4B}\right)^{1/2} = \left(\frac{mRT}{2b(n)}\right)^{1/2} \quad (9)$$

Therefore, the standard deviation of the curvature only depends on the differential flexibility along the DNA chain.

It is interesting to note the difference between the virtual segmental dispersion, observable on the SFM images along the DNA chain, and the dispersion of the relative orientational angles among dinucleotide steps. Therefore, the experimental standard deviation of the curvature is a function of the sequence-dependent flexibility and m, the length of the DNA virtual segments in basepairs, and is independent of the intrinsic curvature.

On the contrary, as shown in Appendix A, the average curvature modulus $\langle|C(n, 2m)|\rangle$, generally adopted to characterize the 2D DNA structure in EM and SFM images, contains both the static and the dynamic curvature contributions, i.e., it is related to both the intrinsic curvature, $C_o(n, 2m)$, and the curvature fluctuations, which involve $2m$ bp (as illustrated in Fig. 2 *b*). In fact, we have shown that

$$\langle|C_m(n)|\rangle = \left[\frac{2}{\pi}\left(\langle C_m(n)\rangle^2 + \frac{mRT}{2b(n)}\right)\right]^{1/2} \quad (10)$$

This result clearly points out that the average curvature modulus depends on two terms: one is related to the intrinsic curvature and the other one represents the sequence-dependent flexibility of the chain. The differential flexibility function along the sequence, as a function of the average value of the curvature modulus, can be obtained only for an intrinsically straight DNA, where the static curvature approaches zero (see Appendix A).

Adopting the same approximation, we obtain the standard deviation of curvature modulus:

$$SD(|C_m(n)|) = \left[\left(\frac{\pi - 2}{\pi}\right)\left(\langle C_m(n)\rangle^2 + \frac{mRT}{2b(n)}\right)\right]^{1/2} \quad (11)$$

Unexpectedly, both the average curvature modulus and the relative standard deviation have the same dependence on the intrinsic curvature and flexibility. They are proportional and their ratio is $(2/(\pi - 2))^{1/2}$, contrary to the general feeling that associates the flexibility to the curvature dispersion. This is correct, provided that the curvature phases are also taken into account, as shown by Eq. 9.

RESULTS AND DISCUSSION

The large variety of shapes assumed by the palindromic DNA molecules under the thermal stochastic perturbation, as illustrated by the example of their SFM images reported in Fig. 2 and described by the above theoretical section, allows two possibilities for the statistical analysis of their curvatures. We can analyze them in terms of the statistical ensemble average curvature either with or without taking into account the relative phases of the curved tracts along the chain. This latter approach, in terms of average curvature modulus, has been so far generally adopted to analyze and characterize the 2D structures of DNAs both in EM and SFM images (Muzard et al., 1990; Bednar et al., 1995; Rivetti et al., 1996, 1998; Cognet et al., 1999).

The profiles of the average curvature modulus and the corresponding standard deviation along the chain obtained from the SFM images are reported in Fig. 3. The distribution of the pixel file, which interpolates the DNA chain, was normalized via Fourier transform operations. These convert the nonuniform pixel sequence of the DNA images into a uniform coordinate distribution along the contour length. After this transformation the number of points, which interpolate the DNA traces, remains practically invariant and corresponds to the average value of pixels per molecule. The curvature angles were evaluated from the vectorial product of overlapping directional chain m bp segments. The resulting curvature functions were averaged and the corresponding standard deviation calculated for different segment lengths. In this way the number of considered segments per DNA chain is practically equivalent, except for the terminal effects.

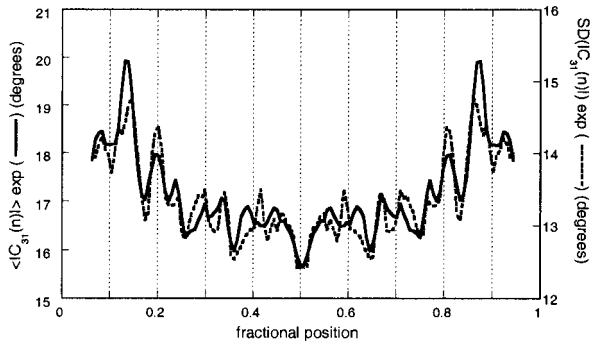

FIGURE 3. Comparison between the profiles of the experimental average curvature modulus and the corresponding standard deviation along the DNA chains, obtained from the SFM images for recurrent three turns long DNA tracts ($m = 31$ bp).

The two profiles show very similar trends, and a nearly constant ratio close to the theoretical value of $(2/(\pi - 2))^{1/2}$ is in good agreement with Eqs. 10 and 11. This result strongly confirms that the average curvature modulus and the related standard deviation contain the same dependence on the static and dynamic curvature contributions, independently of the segment length.

Furthermore, as this result was obtained assuming the canonical ensemble distribution, the existence of a strict relation between the average curvature modulus and the related standard deviation confirms the condition of local thermodynamic equilibrium of the DNA tracts on 2D surfaces.

However, the average curvature modulus and the related standard deviation contain terms of intrinsic curvature and flexibility with a different dependence on the segment length, m. The static curvature increases linearly with m, whereas the dynamic contribution changes with $m^{1/2}$. An evaluation of $\langle b(n)\rangle = b^*$ equal to 85 kcal/bp rad^{-2}, corresponding to a rather standard persistence length of 48 nm in the 3D state of the solution, was obtained plotting the lowest values of the average curvature modulus against $m^{1/2}$.

Fig. 4 shows the satisfactory agreement between the experimental data of the average curvature modulus and their standard deviation along the chain, and their theoreti-

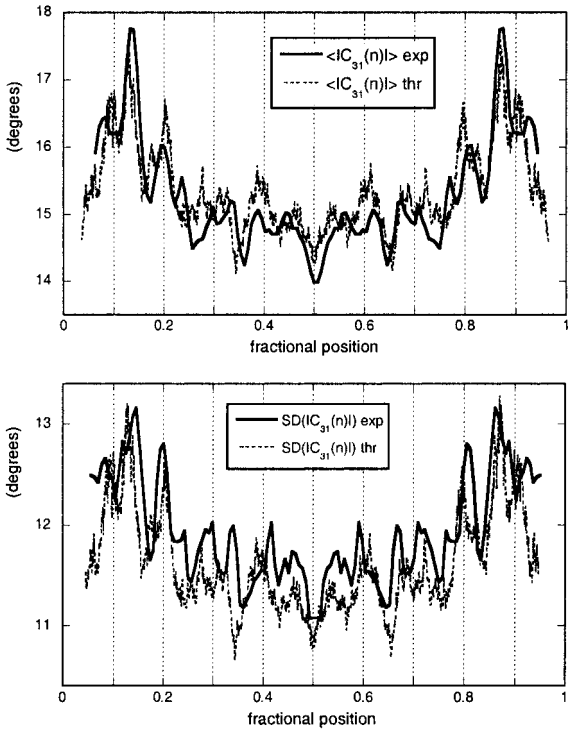

FIGURE 4. Comparison between the profiles of the experimental and theoretical average curvature modulus (*top*) and the related standard deviation (*bottom*) along the chain. The theoretical profiles were obtained according to Eqs. 10 and 11 with the average elastic constant $b^* = 85$ kcal/bp rad^{-2}. The diagrams refer to a value of $m = 31$ bp.

cal value calculated by Eqs. 10 and 11, adopting for $b(n)$ the average value of 85kcal/bp rad^{-2}. The diagrams refer to $m = 31$ bp (about three DNA turns, namely a DNA length adequate to the SFM resolution power).

These results demonstrate that the statistically averaged curvature modulus is a quantity depending on the intrinsic curvature and flexibility of the chain and does not allow the separation of the two contributions. Only in the case of an intrinsically straight DNA, when the intrinsic curvature vanishes, can the differential flexibility function along the sequence be obtained. In general, the possibility of separating the two terms and directly obtaining the differential flexibility, as shown by Eq. 9, should require that the statistical ensemble average curvature analysis is made by taking into account the relative phases of the single curvatures along the chain. However, it has not escaped our attention that this approach should lead the statistical averaging of the curvature along the chain to zero. This expectation can be inferred on the basis of the following considerations.

Unlike the case of generally straight molecules, large-scale curvature provides a palindromic molecule with a neat distinction between two faces with which it can flatten on a surface and interact with it. The statistical conformational equivalence of the two halves of a palindromic sequence

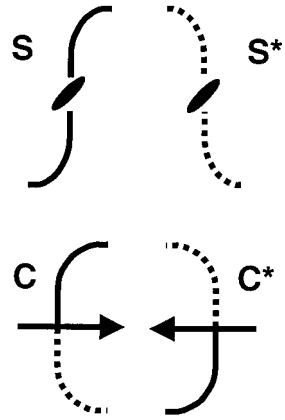

FIGURE 5. Pictorial representations of the four possible symmetry classes a palindromic molecule can assume when it is flattened into a 2D surface. They result in two different prochiral S-shaped forms (S and S*) and two equivalent C-shaped forms (C and C*). Solid and dashed lines indicate the different faces they expose to the surface after deposition; arrows and filled ovals represent the dyad axes parallel and perpendicular to the plane, respectively.

results in the presence of a dyad axis in the 3D superstructure of the DNA double helix centered at the inversion point of the sequence. When a palindromic 3D molecule is flattened into a 2D object upon deposition onto the mica surface, the conformational equivalence of the two halves results in one among two different prochiral forms with the dyad axis perpendicular to the substrate plane (which we called S and S*), or two equivalent forms with the dyad axis lying on the substrate plane (C and C*), as sketched in Fig. 5. In fact, using a palindromic molecule gives way to peculiar statistical symmetry classes. The two distinct S-shaped forms only differ in the face they present to the surface and to the solvent after deposition; they are chemically different because they expose complementary sequences and are expected to interact differently with the mica. The two C-shaped forms expose to the substrate different faces for each of the two halves they are constituted of (which correspond to the two faces of the S forms). Since one of the two alternative terminals is chosen as the starting points in the SFM images processing, they are equivalent and are expected to be equally probable.

Therefore, the profiles of the statistical ensemble curvature of palindromic DNA molecules are characterized by an inversion center between the two oppositely signed halves of the two S-shaped forms or single-signed (either positive or negative) for the two C-shaped forms.

The transition from a 3D to a 2D space plausibly takes place at the expenses of the local twist, leaving the main curved tracts practically unchanged. In fact, the S- or C-like shapes require a maximum value of 0.2 kcal/bp (estimated by adopting a torsional rigidity modulus equal to $2.1 \cdot 10^{-19}$ erg·cm), because of the long DNA tract bridging the two peripheral curved tracts.

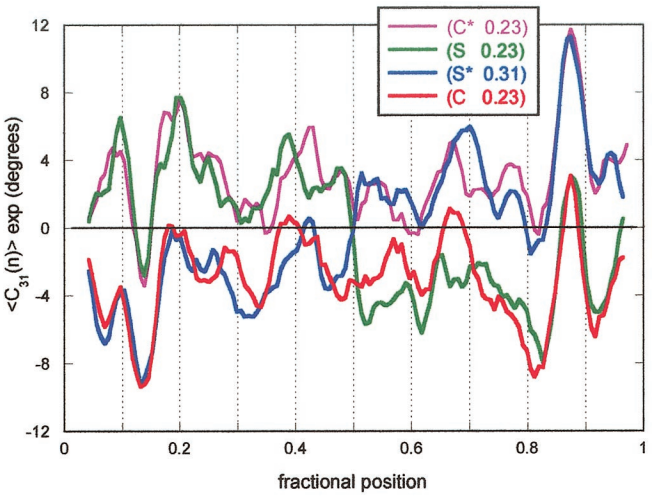

FIGURE 6. Experimental normalized average curvature profiles for the four different symmetry classes of the palindromic sequence versus the fractional sequence position. The relative populations are indicated in the legend. The diagram refers to $m = 31$ bp.

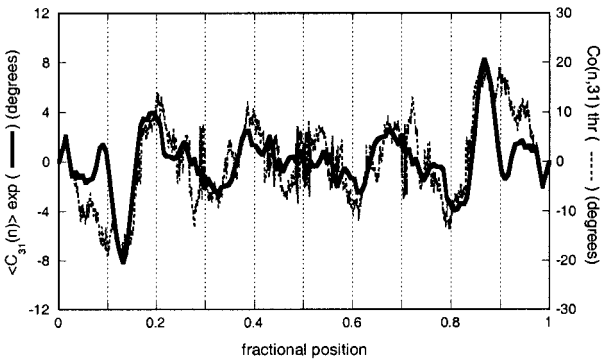

FIGURE 7. Comparison between the experimental average curvature of the molecules of the S* class and the theoretical intrinsic curvature. The profiles refer to $m = 31$ bp.

However, in the hypothesis of the absence of any differential recognition from the mica surface, the stochastic deposition of both the 2D DNA faces results in an ensemble average curvature equal to zero. But, contrary to these expectations, the curvature profile did not average to zero. Its shape was that of an odd function, with a sign inversion at the middle point. This would indicate that the molecules had predominantly been deposited on the surface in an S-like shape.

Therefore, in the case of palindromic DNAs, the intrinsic curvature information from the 2D molecule profiles can be obtained only through a classification of all the recorded images in the four classes, and analyzing their relative curvature profiles.

In Fig. 6 the experimental normalized average curvature profiles for the different classes and their relative populations are reported along the fractional sequence position. It must be emphasized that the different symmetries characterizing the four profiles and the cross-equivalence of C- and S-shaped forms prove the statistical validity of the ensemble examined. Furthermore, it is apparent that the S- and S*-shaped forms significantly deviate from the expected prochiral symmetry. In fact, the pairs of S forms does not show mirror curvature profiles. The same evidence was found for the C forms where the two halves are not mirror images.

This result strongly suggests some capability of the mica surface of recognizing the two different faces of the S-shaped molecules and favoring the deposition of one of the prochiral forms. Experiments expressly tailored for the identification of the chemical interactions that can lead to this kind of recognition process are in progress. Very recent results provide a rather sound basis for such suggestions. In fact, the mica surface shows surprisingly high selectivity (up to 9 to 1) toward palindromic constructs of well-known curved tracts of *Crithidia fasciculata* for one of the prochiral faces of S-like shapes (paper submitted for publication).

Fig. 7 shows the good agreement between the experimental average curvature of the molecules of the S* class and the theoretical intrinsic curvature, whose phases are polarized in the two opposite directions, which characterizes the two equivalent curved regions. Such an agreement proves that DNA on the mica surface retains for this class a large part of the curvature that characterizes its superstructure in solution; therefore, the 3D → 2D transformation should occur at the expense of the local twist, as aforementioned.

As theoretically predicted, the standard deviation of curvature depends linearly on the flexibility of DNA along the sequence, $f(n)$, and on the square root of the segment length, m. Fig. 8 shows the expected linear trend of the average standard deviation of curvature along the sequence versus $m^{1/2}$. This allows an evaluation of the average bending force constant and the corresponding persistence length ($P = 48$ nm, related as $P = b\ell/RT$ in the 3D state) equal to the value already obtained from the analysis of the average curvature modulus.

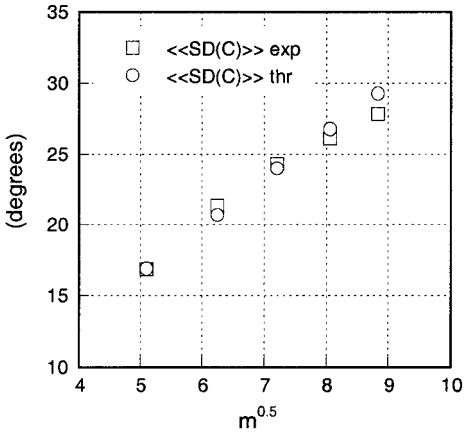

FIGURE 8. Comparison between the profiles of the expected and experimental values of the average standard deviation of curvature along the sequence versus $m^{1/2}$.

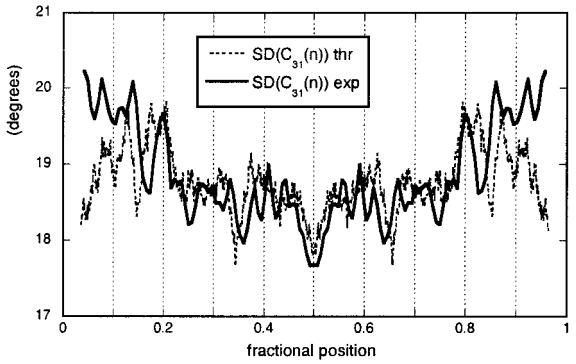

FIGURE 9. Comparison between the experimental and theoretical curvature standard deviation (evaluated according to Eq. 9) for $m = 31$ bp. Flexibility is represented by the normalized theoretical stacking energy of the dinucleotide steps (see Appendix B), while b^* is set equal to 85 kcal/bp rad^{-2}.

The theoretical standard deviation of curvature, calculated as in Eq. 9 for $m = 31$ bp, is compared with the experimental standard deviation in Fig. 9. The flexibility, $f(n)$, was represented by the normalized theoretical stacking energy of the dinucleotide steps (see Appendix B), and b^* assumes the value of 85 kcal/bp rad^{-2}. The satisfactory agreement between the two profiles satisfactorily proves the model we propose for the differential flexibility. The significant larger values of the experimental standard deviation with respect to the theoretical one at the DNA ends could be plausibly due to enhancement of dynamic effects at the terminals where the thermal fluctuations result in a more effective chain diffusion (Perico, 1989).

Finally, Fig. 10 shows the correlation between the experimental standard deviation of the curvature and the frequency of AT·AT+AA·TT+TA·TA and for comparison, with the frequency of GC·GC+GG·CC+CG·CG dinucleotide steps along the fractional sequence position. This further confirms that A·T-rich sequences are more flexible than G·C-rich sequences. This is particularly relevant because at contrast with a current opinion mainly based on the structural dispersion of the different dinucleotide steps observed in the x-ray crystal structures of a large set of ds-oligonucleotides (Olson et al., 1993, 1998). It must be pointed out that the model proposed here for DNA flexibility has predictive power for the differential thermodynamic stability of nucleosomes (Anselmi et al., 1999, 2000) and gel electrophoresis anomalies of a number of intrinsically straight multimeric ds-oligomers (Anselmi et al., 2002).

CONCLUSIONS

We have shown that SFM images can provide an appropriate data basis for mapping intrinsic curvature and flexibility of DNAs along the sequence. This is obtained by averaging the internal coordinates, namely the observed local DNA curvature, over an adequate pool of SFM images.

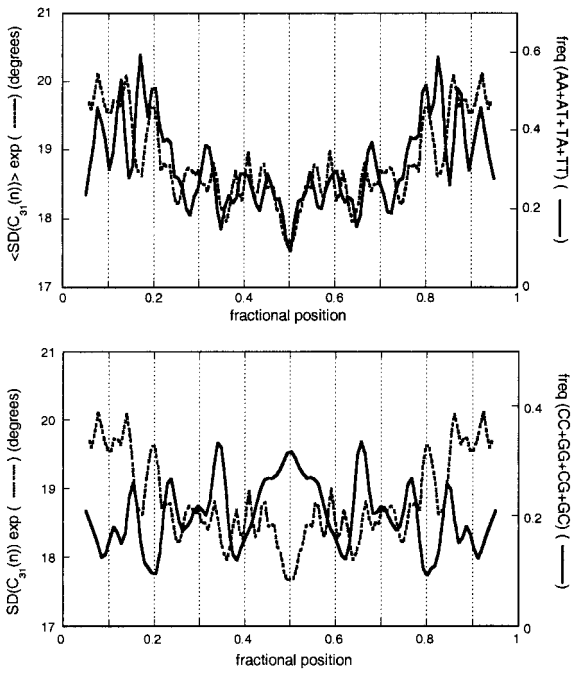

FIGURE 10. Correlation between the experimental standard deviation of curvature ($m = 31$ bp) with the frequency of AT·AT+AA·TT+TA·TA (*top*) and GC·GC+GG·CC+CG·CG (*bottom*) dinucleotide steps, versus the fractional sequence position.

The profiles of both the ensemble average curvature and the corresponding standard deviations along the sequence are in very satisfactory agreement with the theoretical predictions, which convincingly proves the model for predicting sequence-dependent curvature and flexibility we advanced several years ago (De Santis et al., 1986, 1990; Cacchione et al., 1989; Anselmi et al., 1999, 2000).

An interesting result is that, surprisingly, the average curvature modulus is nearly proportional to the related standard deviation, contrary to the immediate feeling, and, as a consequence, does not allow the separation of static and dynamic curvature contributions, which is only possible when the phases of curvature are taken into account. The proportionally between the ensemble average curvature modulus and the related standard deviation represents a stringent condition of the local thermodynamic equilibrium of the DNA molecules on the 2D surface.

Finally, the results provide evidence about the existence of sequence-dependent interactions between DNA and the mica surface.

APPENDIX A

Derivation of the ensemble curvature modulus and the corresponding standard deviation for a segmented DNA

The average curvature modulus $\langle|C_m(n)|\rangle$ (generally adopted to characterize the 2D DNA structure in EM and in SFM images) is the angle between the

virtual bonds spanning m bp at the nth position along the sequence (Fig. 2 *b*). It contains both static and dynamic curvature contributions, i.e., it is related to the intrinsic curvature of two consecutive m bp virtual bonds, $\langle C(n, 2m)\rangle$, and to the corresponding curvature fluctuations.

Setting $C = C(n, 2m)$, $\langle C\rangle = \langle C(n, 2m)\rangle$, $\chi = C(n, 2m) - \langle C(n, 2m)\rangle$, recalling that $B = b(n)/2RT$ and setting the partition function

$$Q = Q(n, 2m) = \int_{-\infty}^{\infty} \exp\left(-\frac{B}{2m}|\chi|^2\right)dC \qquad (12)$$

we obtain the average curvature modulus for a $2m$ bp segment

$$\langle|C|\rangle = \frac{\displaystyle\int_{-\infty}^{\infty} |C| \exp\left(-\frac{B}{2m}|\chi|^2\right)dC}{Q} = \left|\frac{\displaystyle\int_{-\infty}^{\infty} C \exp\left(-\frac{B}{2m}|\chi|^2\right)dC}{Q}\right| - 2\frac{\displaystyle\int_{-\infty}^{0} C \exp\left(-\frac{B}{2m}|\chi|^2\right)dC}{Q} \qquad (13)$$

By a simple substitution,

$$\left|\frac{\displaystyle\int_{-\infty}^{\infty} (\chi + \langle C\rangle)\exp\left(-\frac{B}{2m}\chi^2\right)d\chi - 2\int_{-\infty}^{-\langle C\rangle} (\chi + \langle C\rangle)\exp\left(-\frac{B}{2m}\chi^2\right)d\chi}{Q}\right| = |\langle C\rangle| + \frac{1}{Q}\left[\frac{2m}{B}\exp\left(-\frac{B}{2m}\langle C\rangle^2\right) - 2|\langle C\rangle|\left(\int_0^{\infty}\exp\left(-\frac{B}{2m}\chi^2\right)d\chi - \int_0^{\langle C\rangle}\exp\left(-\frac{B}{2m}\chi^2\right)d\chi\right)\right] \qquad (14)$$

Expanding the latter integral and the term $\exp(-(B/2m)\langle C\rangle^2)$ to the second order, considering that the curvatures of interest are small, we obtain

$$\langle|C|\rangle = \frac{1}{Q}\left(\langle C\rangle^2 + \frac{2m}{B}\right) = \left(\frac{B}{2m\pi}\right)^{1/2}\langle C\rangle^2 + \left(\frac{2m}{\pi B}\right)^{1/2} \qquad (15)$$

To separate the flexibility and curvature terms, we squared $\langle|C|\rangle$ and neglected the term higher than the second order. As a result, we obtain the average curvature modulus for a $2m$ bp segment in a compact formulation valid for small curvature (up to 0.6 rad)

$$\langle|C(n, 2m)|\rangle = \left[\frac{2}{\pi}\left(\langle C(n, 2m)\rangle^2 + \frac{2mRT}{b(n)}\right)\right]^{1/2} \qquad (16)$$

The average curvature modulus, which refers to a $2m$ bp DNA tract, is two times the corresponding experimental value for a segmented chain (see Fig. 2):

$$\langle|C_m(n)|\rangle = \frac{\langle|C(n, 2m)|\rangle}{2} = \left[\frac{2}{\pi}\left(\langle C_m(n)\rangle^2 + \frac{mRT}{2b(n)}\right)\right]^{1/2} \qquad (17)$$

This result clearly points out that the average curvature modulus depends on two terms: one corresponds to the intrinsic curvature and the other represents the sequence-dependent flexibility of the chain. The differential flexibility function along the sequence can be obtained only for an intrinsically straight DNA, where the static curvature is zero. It is to be noted that the formulation obtained for the average curvature modulus appears different from the one we previously published (Zuccheri et al., 2001); here we have taken into account the effects of the segmentation of the chain not considered in the previous paper. Moreover, the approximation adopted to evaluate the average curvature modulus was extended to the second order to obtain a more adequate approximation for curvatures of interest.

The corresponding standard deviation can be easily obtained by considering that $\langle C_m{}^2(n)\rangle = \langle C_m(n)\rangle^2 + m/2B$ (see Eqs. 7–9); therefore

$$SD(|C_m(n)|) = (\langle C_m^2(n)\rangle - \langle|C_m(n)|\rangle^2)^{1/2} = \left[\frac{\pi - 2}{\pi}\left(\langle C_m(n)\rangle^2 + \frac{mRT}{2b(n)}\right)\right]^{1/2} \qquad (18)$$

Therefore, the ensemble curvature modulus and the corresponding standard deviation are expected to be proportional because they contain the same dependence from both the static and dynamic curvature contributions. Alternatively, this proportionality can be adopted as an internal gauge of the existence of the small-scale thermodynamic equilibrium of the DNA molecule ensemble on the mica surface.

APPENDIX B

A model of the sequence-dependent DNA flexibility

The definition of the DNA sequence-dependent flexibility implies the knowledge of the statistical parameters necessary to evaluate the average flexibility of any DNA tract. This would require the determination of the persistence length of a large number of sequences. In the present lack of these data, we connected the sequence-dependent DNA flexibility of a DNA tract to its double helix thermodynamic stability that was evaluated on the basis of the differential thermodynamic dinucleotide stability data proposed by different authors (Gotoh and Tagashira, 1981; Sugimoto et al.,

1996; SantaLucia, 1998). This approach was based on the following considerations.

Assuming a first-order elastic behavior of DNA, the thermal energy can be related to the isotropic bending variance (Hagerman, 1988)

$$\tfrac{1}{2}\, b\langle|\boldsymbol{\chi}|^2\rangle = RT \tag{19}$$

where b is the apparent elastic force constant and $\langle|\boldsymbol{\chi}|^2\rangle$ is the average dinucleotide-step bending fluctuation. Similarly, for a standard DNA, namely a straight chain with random sequence

$$\tfrac{1}{2}\, b^*\langle|\boldsymbol{\chi}^*|^2\rangle = RT^* \tag{20}$$

Thus, for the same fluctuation amplitudes:

$$b = b^*\left(\frac{T}{T^*}\right) \tag{21}$$

Due to the catastrophic character of the helix-coil transition, until the temperature is a few degrees lower than that corresponding to the melting point, the double helix still represents the thermodynamic state of DNAs (in fact, spectroscopic and electrochemical properties are only slightly modified). The small difference between such "pre-melting" temperatures and the melting ones does not significantly affect their ratio.

The assumption that the helix-coil transition starts at the temperature when the basepair librations reach a critical value, which corresponds to unstacking the basepairs, implies the practical equivalence of the curvature fluctuations at the melting point independently of the sequence.

In fact, assuming an elastic rod-like DNA model, we considered the normalized melting temperature of a DNA tract as a factor modulating the average elastic force constant that quantifies the sequence-dependent rigidity, $f^{-1}(n)$. It is indeed the reciprocal of the sequence-dependent flexibility $f(n)$.

Therefore, the differential flexibility of a DNA tract can be represented by the ratio of the dinucleotide melting temperatures (in thermodynamic scale), averaged over the tract considered. We successfully adopted such a representation in the theoretical prediction of nucleosome thermodynamic stability (Anselmi et al., 1999, 2000), as well as in the first statistical-mechanical analysis of the SFM images of DNA (Zuccheri et al., 2001).

The correlation of the flexibility with the thermodynamic stability was obtained by extending the harmonic model of the curvature fluctuations until the pre-melting temperatures. However, the model can be improved by introducing an anharmonic potential to describe the basepair librations of the dinucleotide steps. According to this hypothesis, the bending energy, E_b, assumes a Morse-like formulation:

$$E_b = E_{st}(1 - \exp(-D|\boldsymbol{\chi}|))^2 \tag{22}$$

where E_{st} is the theoretically evaluated stacking energy (Ornstein et al., 1978) and D a constant. Equation 22 describes the dinucleotide step libration with a symmetric potential; for sufficiently small curvature deviations, the exponential term can be expanded and Eq. 22 assumes the harmonic form $E_b = D^2E_{st}|\boldsymbol{\chi}|^2$. This corresponds to assuming the bending energy linearly dependent on the stacking energy for any small deformation.

In this case, when the basepair librations are harmonic, the term $D^2E_{st}|\boldsymbol{\chi}|^2$ can be considered equivalent to $\frac{1}{2}b|\boldsymbol{\chi}|^2$, where b is the apparent elastic force constant. For a standard DNA, namely a straight chain with random sequence, we can write $\frac{1}{2}\, b^*|\boldsymbol{\chi}|^2 = D^2E^*_{st}|\boldsymbol{\chi}|^2$; as a consequence

$$\frac{b}{b^*} = \frac{E_{st}}{E^*_{st}} \tag{23}$$

TABLE 2 Rigidity parameters, $f^{-1}(n)$, for each dinucleotide step, expressed in terms of the normalized melting temperatures (Gotoh and Tagashira, 1981) and stacking energies (Ornstein et al., 1978)

Dinucleotide Step	T_{melt}/T^*_{melt}	E_{stk}/E^*_{stk}
AA·TT	0.950	0.703
AC·GT	1.065	1.323
AG·CT	0.961	0.780
AT	0.947	0.854
CG	0.992	1.124
GA·TC	1.032	1.230
GC	1.185	1.792
GG·CC	1.031	0.984
TA	0.899	0.615
TG·CA	0.951	0.790

In this way, the stacking energy for each dinucleotide step modulates the corresponding force constant introducing a sequence-dependent flexibility along the DNA chain.

The ratios E_{st}/E^*_{st} correlate well with the corresponding T_m/T^*_m ($R = 0.97$) previously adopted to modulate the force constant (Anselmi et al., 1999, 2000; Zuccheri et al., 2001; see Table 2).

This correlation can hold only if the entropy, the conformational energy, and the water and counterions interaction changes at the DNA melting transition are constant for all the dinucleotide steps. Such a hypothesis, not plausible at the melting point, seems to be acceptable at pre-melting temperatures. The good linear correlation proves that the unstacking of dinucleotide steps occurs immediately before melting and strengthens the hypothesis that stacking energies are the main factor of DNA stiffness, as already proposed by Hagerman (1988).

Therefore, the ratio b^*/b so obtained for the 10 independent dinucleotide steps is the basis set to characterize the sequence-dependent flexibility parameters, $f(n)$. Actually, as bending anisotropy is expected (Lankas et al., 2000), during the calculations they are conveniently evaluated over the double helix turns (analogously to the intrinsic curvature), so that the anisotropy results averaged.

This work was supported by "Progetto 60% Ateneo" of University "La Sapienza", Programmi Biotecnologie Legge 95/95 (MURST) 5%, MURST Progetti di Ricerca di Interesse Nazionale 1999-2001, and Istituto Pasteur, Foundazione Cenci Bolognetti.

REFERENCES

Anselmi, C., G. Bocchinfuso, P. De Santis, M. Savino, and A. Scipioni. 1999. Dual role of DNA intrinsic curvature and flexibility in determining nucleosome stability. *J. Mol. Biol.* 286:1293–1301.

Anselmi, C., G. Bocchinfuso, P. De Santis, M. Savino, and A. Scipioni. 2000. A theoretical model for the prediction of sequence-dependent nucleosome thermodynamic stability. *Biophys. J.* 79:601–613.

Anselmi, C., P. De Santis, R. Paparcone, M. Savino, and A. Scipioni. 2002. From the sequence to the superstructural properties of DNAs. *Biophys. Chem.* 95:23–47.

Bednar, J., P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, and A. Stasiak. 1995. Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* 254:579–594.

Bolshoy, A., P. McNamara, R. E. Harrington, and E. N. Trifonov. 1991. Curved DNA without A·A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. USA.* 88:2312–2316.

Bustamante, C., J. Vesenka, C. L. Tang, W. Rees, M. Guthold, and R. Keller. 1992. Circular DNA molecules imaged in air by scanning force microscopy. *Biochemistry.* 31:22–26.

Cacchione, S., P. De Santis, A. Foti, A. Palleschi, and M. Savino. 1989. Periodical polydeoxynucleotides and DNA curvature. *Biochemistry.* 28: 8706–8713.

Chastain II, P. D., E. E. Eichler, S. Kang, D. L. Nelson, S. D. Levene, and R. R. Sinden. 1995. Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry.* 34:16125–16131.

Chastain, P. D., and R. R. Sinden. 1998. CTG repeats associated with human genetic disease are inherently flexible. *J. Mol. Biol.* 275: 405–411.

Cognet, J. A., C. Pakleza, D. Cherny, E. Delain, and E. L. Cam. 1999. DNA bending induced by the archaebacterial histone-like protein MC1. *J. Mol. Biol.* 285:997–1009.

Crothers, D. M. 1998. DNA curvature and deformation in protein-DNA complexes: a step in the right direction. *Proc. Natl. Acad. Sci. USA.* 95:15163–15165.

De Santis, P., M. Fuà, M. Savino, C. Anselmi, and G. Bocchinfuso. 1996. Sequence dependent circularization of DNAs: a physical model to predict the DNA sequence dependent propensity to circularization and its changes in the presence of protein-induced bending. *J. Phys. Chem.* 100:9968–9976.

De Santis, P., A. Palleschi, S. Morosetti, and M. Savino. 1986. Structure and superstructures in periodical polynucleotides. *In* Structures and Dynamics of Nucleic Acids, Proteins and Membranes. E. Clementi and S. Chin, editors. Plenum Press, New York. 31–49.

De Santis, P., A. Palleschi, M. Savino, and A. Scipioni. 1988. A theoretical model of DNA curvature. *Biophys. Chem.* 32:305–317.

De Santis, P., A. Palleschi, M. Savino, and A. Scipioni. 1990. Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry.* 29: 9269–9273.

Diekmann, S. 1987. DNA methylation can enhance or induce DNA curvature. *EMBO J.* 6:4213–4217.

Flory, P. J., U. W. Suter, and M. Mutter. 1976. Macrocyclization equilibriums. 1. Theory. *J. Am. Chem. Soc.* 98:5733–5739.

Gotoh, O., and Y. Tagashira. 1981. Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers.* 20:1033–1042.

Hagerman, P. J. 1985. Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry.* 24:7033–7036.

Hagerman, P. J. 1986. Sequence-directed curvature of DNA. *Nature.* 321:449–450.

Hagerman, P. J. 1988. Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* 17:265–286.

Hansma, H. G., and D. E. Laney. 1996. DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. *Biophys. J.* 70:1933–1939.

Hockings, S. C., J. D. Kahn, and D. M. Crothers. 1998. Characterization of the ATF/CREB site and its complex with GCN_4. *Proc. Natl. Acad. Sci. USA.* 95:1410–1415.

Jacobson, H., and W. H. Stockmayer. 1950. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* 18:1600–1606.

Kahn, J. D., and D. M. Crothers. 1998. Measurement of the DNA bend angle induced by the catabolite activator protein using Monte Carlo simulation of cyclization kinetics. *J. Mol. Biol.* 276:287–309.

Koo, H. S., and D. M. Crothers. 1988. Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc. Natl. Acad. Sci. USA.* 85:1763–1767.

Koo, H. S., H.-M. Wu, and D. M. Crothers. 1986. DNA bending at adenine·thymine tracts. *Nature.* 320:501–506.

Landau, L. D., and E. M. Lifshitz. 1970. Theory of Elasticity. Pergamon Press, Oxford, New York.

Lankas, F., J. Sponer, P. Hobza, and J. Langowski. 2000. Sequence-dependent elastic properties of DNA. *J. Mol. Biol.* 299:695–709.

Levene, S. D., and B. H. Zimm. 1989. Understanding the anomalous electrophoresis of bent DNA molecules: a reptation model. *Science.* 245:396–399.

Levene, S. D., and D. M. Crothers. 1986. Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J. Mol. Biol.* 189:61–72.

Lumpkin, O. J., P. Dejardin, and B. H. Zimm. 1985. Theory of gel electrophoresis of DNA. *Biopolymers.* 24:1573–1593.

Marini, J. C., S. D. Levene, D. M. Crothers, and P. T. Englund. 1982. Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. USA.* 79: 7664–7668.

Muzard, G., B. Theveny, and B. Revet. 1990. Electron microscopy mapping of pBR322 DNA curvature. Comparison with theoretical models. *EMBO J.* 9:1289–1298.

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.

Olson, W. K., N. L. Marky, R. L. Jernigan, and V. B. Zhurkin. 1993. Influence of fluctuations on DNA curvature. A comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* 232: 530–554.

Olson, W. K., and V. B. Zhurkin. 1996. Twenty years of DNA bending. *In* Biological Structure and Dynamics. Proceedings of the Ninth Conversation. R. H. Sarma and M. H. Sarma, editors. Adenine Press, New York. 341–370.

Ornstein, R. L., R. Rein, D. L. Breen, and R. D. Macelroy. 1978. An optimized potential function for the calculation of nucleic acid interaction energies. I. Base stacking. *Biopolymers.* 17:2341–2360.

Perico, A. 1989. Segmental relaxation in macromolecules. *Acc. Chem. Res.* 22:336–342.

Rivetti, C., M. Guthold, and C. Bustamante. 1996. Scanning force microscopy of DNA deposited onto mica: equilibration *versus* kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* 264:919–932.

Rivetti, C., C. Walker, and C. Bustamante. 1998. Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.* 280:41–59.

Roychoudhury, M., A. Sitlani, J. Lapham, and D. M. Crothers. 2000. Global structure and mechanical properties of a 10-bp nucleosome positioning motif. *Proc. Natl. Acad. Sci. USA.* 97:13608–13613.

Sambrook, J., T. Maniatis, and E. F. Fritsch. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

SantaLucia Jr., J. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA.* 95:1460–1465.

Shimada, J., and H. Yamakawa. 1984. Ring closure probability for twisted wormlike chains. Application to DNA. *Macromolecules.* 17:689–698.

Sprous, D., R. K.-Z. Tan, and S. C. Harvey. 1996. Molecular modeling of closed circular DNA thermodynamic ensembles. *Biopolymers.* 39: 243–258.

Sugimoto, N., S. Nakano, M. Yoneyama, and K. Honda. 1996. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acid. Res.* 24:4501–4505.

Trifonov, E. N. 1980. Sequence-dependent deformational anisotropy of chromatin DNA. *Nucleic Acids Res.* 8:4041–4053.

Trifonov, E. N., R. K.-Z. Tan, and S. C. Harvey. 1987. DNA bending and curvature. *In* Structure and Expression, DNA Bending and Curvature, Vol. 3. W. K. Olson, M. H. Sarma, and M. Sundaralingam, editors. Adenine Press, New York. 243.

Wu, H.-M., and D. M. Crothers. 1984. The locus of sequence-directed and protein-induced DNA bending. *Nature.* 308:509–513.

Yamakawa, H., and W. H. Stockmayer. 1972. Statistical mechanics of wormlike chains. II. Excluded volume effects. *J. Chem. Phys.* 57: 2843–2854.

Zuccheri, G., A. Scipioni, V. Cavaliere, G. Gargiulo, P. De Santis, and B. Samori. 2001. Mapping the intrinsic curvature and the flexibility along the DNA chain. *Proc. Natl. Acad. Sci. USA.* 98:3074–3079.