

Stacked–Unstacked Equilibrium at the Nick Site of DNA

Ekaterina Protozanova, Peter Yakovchuk and
Maxim D. Frank-Kamenetskii*

Center for Advanced
Biotechnology and Department
of Biomedical Engineering
Boston University, 36
Cummington Street, Boston
MA 02215, USA

Stability of duplex DNA with respect to separation of complementary strands is crucial for DNA executing its major functions in the cell and it also plays a central role in major biotechnology applications of DNA: DNA sequencing, polymerase chain reaction, and DNA microarrays. Two types of interaction are well known to contribute to DNA stability: stacking between adjacent base-pairs and pairing between complementary bases. However, their contribution into the duplex stability is yet to be determined. Now we fill this fundamental gap in our knowledge of the DNA double helix. We have prepared a series of 32, 300 bp-long DNA fragments with solitary nicks in the same position differing only in base-pairs flanking the nick. Electrophoretic mobility of these fragments in the gel has been studied. Assuming the equilibrium between stacked and unstacked conformations at the nick site, all 32 stacking free energy parameters have been obtained. Only ten of them are essential and they govern the stacking interactions between adjacent base-pairs in intact DNA double helix. A full set of DNA stacking parameters has been determined for the first time. From these data and from a well-known dependence of DNA melting temperature on G·C content, the contribution of base-pairing into duplex stability has been estimated. The obtained energy parameters of the DNA double helix are of paramount importance for understanding sequence-dependent DNA flexibility and for numerous biotechnology applications.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: DNA melting; nearest-neighbor stability parameters; base-pairing; base stacking; gel electrophoresis

*Corresponding author

Introduction

It has long been realized that two qualitatively distinct modes of interactions contribute to the stability of the DNA double helix: stacking between adjacent base-pairs and pairing between complementary bases.^{1–4} Quantitative estimation of their relative contributions to the overall duplex stability has been attempted both experimentally^{5–8} and theoretically.^{9–12} Two approaches have been employed to evaluate the contribution of base stacking to stabilization of the double helix; both rely on thermal stability measurements of short nucleic acid duplexes. One of them involves extraction of the base stacking contribution by examining the stabilizing effect brought about by stacking of a single dangling (unpaired) terminal

base flanking the double helix.^{13–15} The other approach is based on the fact that stacking entails pronounced mutual stabilization of DNA duplexes formed due to tandem binding of short oligonucleotides to DNA template.^{5,8} Direct stacking interactions of the terminal bases of the oligonucleotides have been implicated in differential stabilizing effects observed for dinucleotide stacks forming the nick interface. Free energy values of coaxial stacking obtained from the data of this kind range from -0.85 kcal/mol for TA dinucleotide stack to -2.76 kcal/mol for GC stack.⁷ Although stacking free energies determined in different studies agree in the range, there are major discrepancies in the values for individual stacks. For example, the study on coaxial stacking contribution to stabilization of gel-immobilized duplexes by in-tandem hybridization of pentamers revealed the adenine stacking with other bases to be significantly stronger than stacking of other bases.⁶ Such pronounced stabilizing stacking efficacy of adenine does not correlate with the parameters of duplex

Abbreviations used: bp, base-pair; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis.

E-mail address of the corresponding author: mfk@bu.edu

stability obtained in the nearest neighbor approximation.^{16–19} The reason for these discrepancies remains unclear.

Here we apply a totally new approach to determine stacking parameters for the DNA double helix which relies on characterization of the energetics of the DNA nick. Previous structural studies concur that the introduction of a break in one of the strands of the DNA double helix hardly affects the B-type DNA conformation.^{20–24} In fact, compared to DNA without lesions, distortions accompanying the single-stranded break in the DNA backbone are limited to the local geometry of the bases immediately adjacent to the nick site. Base stacking and hydrogen bonding between the bases at the site of the nick are conserved keeping the double helix in register. The introduction of a nick however affects the DNA helix rigidity.^{25–28} During gel electrophoresis conducted under native conditions, nicked DNA fragments display small though detectable retardation with respect to the molecules without lesions.^{28–30}

A major hypothesis underlying the present study is that nicked DNA exists in solution (and in gel) in the form of equilibrium between the closed state (the stacking between the two base-pairs flanking the nick is preserved) and the open state (stacking in the nick site is completely lost). Figure 1 shows our basic model. The stacked conformation of the nicked DNA is very close to the conformation of an intact DNA duplex. The loss of stacking between the base-pairs flanking the nick induces a kink in DNA duplex, thus affecting its electrophoretic migration. Fast equilibration between the two states leads to a differential retardation of the nicked DNA molecules during polyacrylamide gel electrophoresis (PAGE) depending on the identity of the nicked dinucleotide.

We use recently introduced nicking endonucleases to site-specifically produce a single break into only one of the two strands of a DNA fragment. To facilitate quantification of otherwise small retardation effects observed for singly nicked DNA fragments, we conduct our PAGE in the presence of urea. By using urea-enhanced PAGE and by extrapolating the stacking free energy values to zero urea concentration, we determine the free energy difference for the stacked–unstacked equilibrium at the site of the nick for all nicked

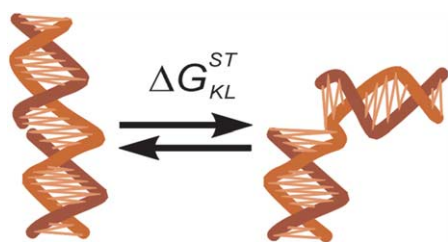


Figure 1. Schematic representation of the stacked-to-unstacked conformational transition of the DNA fragment at the nick site described by the energy difference, ΔG_{KL}^{ST} .

dinucleotide stacks. As a result, we obtain all ten stacking parameters for the DNA duplex. The two remaining parameters, the free energy for A·T and G·C base-pairing, are determined by adjusting the calculated free energy of duplex DNA with random sequence to quantitatively reproduce the classical Marmur–Doty dependence of the DNA melting temperature on G·C content.^{31–33} Base stacking and base-pairing free energy values make it possible to calculate the overall contribution of individual base-pairs into the stability of the DNA double helix within the nearest neighbor approximation. We compare the parameters obtained by us with the nearest neighbor free energy parameters obtained from DNA melting experiments.

Results

Electrophoretic mobility of DNA molecules with single nicks and gaps

All 16 DNA molecules used in this study were identical 300 bp-long fragments differing only in two base-pairs flanking the nick site (the nicked dinucleotide stack KL/L'K') as depicted in Figure 2a. It has been shown before that nicked DNA moves during polyacrylamide gel electrophoresis only slightly slower than the same molecule without a nick.^{28–30} Small retardation of nicked DNA strongly suggests that the molecules predominantly remain straight during PAGE, which indicates that the base-pairs located at the nick interface predominantly retain stacking with one another.^{20–24,28,29} To enhance the effects of the distortion associated with the nick on the electrophoretic migration of DNA fragments we conduct PAGE in the presence of urea.³⁰ By doing so we are able to clearly distinguish the nicked DNA molecules from the intact fragments (Figure 2b). The mobility of nicked molecules decreases noticeably with increasing urea concentration while the mobility of intact fragments remains virtually unaffected. A distinct dependence of the electrophoretic mobility on the identity of base-pairs flanking the nick is indicative of significant interactions between the bases forming the nicked stack (Figure 3a).

We hypothesize that these differential retardation effects originate from the equilibrium between the stacked (straight) and unstacked (kinked) conformations of the DNA nick schematically shown in Figure 1. In the stacked state at the nick site, the DNA double helix remains in register and is very close to the conformation of the DNA duplex without the lesion. The loss of stacking between the two base-pairs flanking the nick prisms the nicked stack open, resulting in a kink in the DNA molecule bringing about additional flexibility to DNA double helix.^{26–28}

For each individual dinucleotide contact we assign the free energy parameter ΔG^{ST} to describe the equilibrium between stacked and unstacked conformations at the nick site. The ratio of

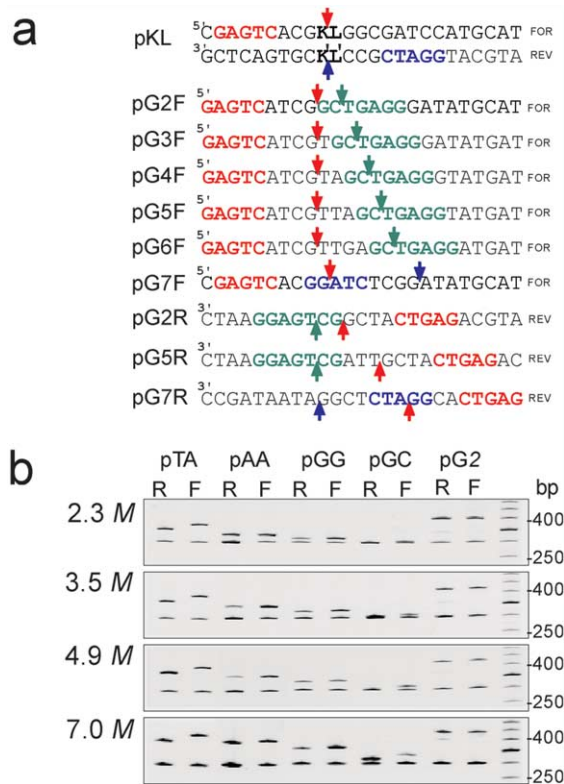


Figure 2. PAGE of DNA fragments with single nicks and gaps. (a) DNA inserts (shown here without the EcoRI/HindIII adaptors) carry the recognition sites for nicking endonucleases N.BstNBI (red), N.AlwI (blue), and N.BbvCIA (green); the locations of the nicking sites are shown by arrows with the same color scheme. For notation see Materials and Methods. (b) The effect of the sequence of the nicked stack on the PAGE migration of DNA fragments. In the presence of urea, 300 bp-long DNA fragments carrying a single nick (slower migrating bands) are readily resolved from the intact fragments (the faster migrating band). Concentration of urea is shown to the left of each panel. The retardation effect depends on the nicked nucleotide stack; the corresponding plasmids are indicated at the top of the gel. Each pKL sample is presented in pairs: the nick is positioned on the forward (F) or reverse (R) strand of the dinucleotide KL. Also shown are DNA fragments with 2 nt gaps positioned on the forward (pG2F) and the reverse (pG2R) strand. A 50 bp DNA molecular mass ladder is shown in the rightmost lane.

occurrences of stacked (N_{closed}) over unstacked (N_{open}) molecules must be governed by the Boltzmann distribution:

$$\frac{N_{\text{closed}}}{N_{\text{open}}} = \exp\left(-\frac{\Delta G^{\text{ST}}}{RT}\right) \quad (1)$$

where T is the absolute temperature, and R is the universal gas constant.

DNA molecules with points of increased flexibility and/or locations of static bending are known to move slower in PAGE when compared to a straight DNA molecule of the same length.^{34–37}

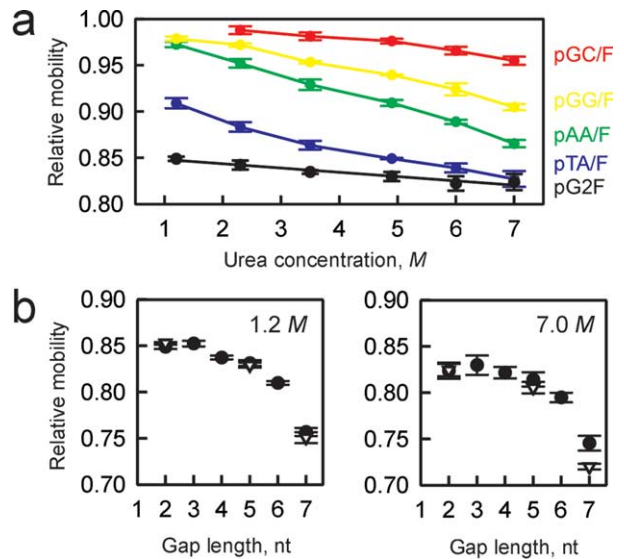


Figure 3. Differential retardation of DNA molecules carrying single nicks and gaps. (a) Relative electrophoretic mobility values of nicked DNA from pKL/F series and DNA fragment with a 2 nt gap, pG2F, are plotted as a function of urea concentration. (b) Relative mobility values of DNA fragments with gaps measured in the presence of 1.2 M and 7.0 M urea decrease with increasing gap. The gaps are produced in the forward (circles) or reverse (triangles) strand of the 300 bp-long fragment. Error bars correspond to variations of electrophoretic mobility measured in three experiments.

Conformational exchange between the straight (stacked) and kinked (unstacked) states gives rise to a single band on the gel. The fact that the band corresponding to the nicked DNA fragment is no broader than the band corresponding to the intact fragment indicates that the exchange is fast with respect to the equilibration in the gel pores, like it takes place in other cases of fast exchange between different DNA conformations during PAGE.^{38–40} The relative mobility of DNA molecules in this band, μ , is a weighted average over the two conformations adopted by the nicked fragment:

$$\mu = \frac{N_{\text{closed}}}{N_{\text{closed}} + N_{\text{open}}} \mu_{\text{closed}} + \frac{N_{\text{open}}}{N_{\text{closed}} + N_{\text{open}}} \mu_{\text{open}} \quad (2)$$

where μ_{closed} is the mobility of the DNA molecule with the nick in the stacked conformation, which is equal to the mobility of the intact DNA fragment; μ_{open} is the mobility of the DNA molecule in fully open unstacked conformation which can be estimated by studying the mobility of gapped DNA molecules (see below). Regrouping equation (2), we obtain:

$$\frac{N_{\text{closed}}}{N_{\text{open}}} = \frac{\mu - \mu_{\text{open}}}{\mu_{\text{closed}} - \mu} \quad (3)$$

Equation (3) allows us to extract the ratio of occupancies of stacked and unstacked conformations of DNA at the site of the nick. According to equation (1), this ratio fully determines the stacking free energy difference between the two states.

Also shown in Figure 2b are 300 bp DNA fragments carrying a single gap in place of the nick. DNA molecules with gaps move considerably slower than any of the nicked fragments of the same length.²⁹ A single-stranded "linker" prevents the two base-pairs flanking the gap from direct interaction forcing the molecule to be fully in the open state. For a series of molecules with gaps ranging in length from 7 nt to 2 nt, we found that their relative mobility values significantly depend on the length of the single-stranded linker (Figure 3b). Fragments with shorter gaps move faster, reaching saturation at about 3–4 nt. Note that this result is consistent with previous observations, reporting similar electrophoretic mobilities for DNA molecules with multiple gaps of 2 nt, 3 nt, and 4 nt in length.²⁹ The gel retardation effect for larger gaps has not been reported before and requires further studies, which are beyond the scope of the present paper. The same dependence of the DNA fragment mobility on the gap length is observed throughout the whole range of urea concentrations we employed (Figure 3b).

For short gaps, the mobility of fragments with a gap positioned in the reverse strand (pGnR series) is undistinguishable from the mobility of the fragments with the same-size gap in the forward strand (pGnF series). As expected, in contrast to the mobility of nicked DNA fragments, mobility of a fragment carrying a 2 nt gap exhibits virtually no dependence on the urea concentration, apparently due to the absence of the direct contact between the two base-pairs flanking the gap (Figure 3a). The independence of the gapped DNA mobility on urea concentration as well as on the location of the gap at one or the other strand strongly indicates against any influence of the single-stranded stacking in the single-stranded linker on the mobility of the fragment.

We assume that the conformation of the molecules with a short single-stranded gap is very close to the open state of the DNA nick. This assumption entails the equivalence of the distributions of the kink angles of the DNA molecules at the site of the lesion in both cases. We will use the relative mobilities of fragments with 2 nt gaps (pG2F and pG2R constructs) to estimate the μ_{open} value in equation (3) for any given urea concentration.

Determination of stacking parameters

We used equations (1) and (3) to calculate the free energy difference between stacked and unstacked conformations for every dinucleotide stack from relative electrophoretic mobility of nicked DNA fragments from the pKL series measured in the presence of urea. Thus obtained ΔG^{ST} values

depend linearly on urea concentration (Figure 4a); the intercept of the linear plot with the ordinate gives the value of the free energy difference, $\Delta G_{\text{KL}}^{\text{ST}}$, of the stacked–unstacked transition at the nick site in the absence of urea. This analysis was performed for the two sets of nicked pKL/PvuII fragments: the nick produced in the forward strand of the KL dinucleotide stack, the pKL/F series, and the nick produced in the reverse strand, the pKL/R series (see schematics in Figure 4b). Our analysis yielded a set of 32 parameters, which are plotted in Figure 4b.

To validate our PAGE-based approach for determining DNA stacking free energies and to assess the accuracy of this determination we need to address the question of how the conditions of PAGE affect the stacked–unstacked equilibrium of the nicked molecule with respect to the same molecule in the free solution. Specifically, such factors include the presence of the gel matrix and the force applied to the molecule during reptation through the gel. The most direct test of such effects of the gel matrix involves the determination of stacking parameters using nicked DNA fragments of different length. For this purpose, we measured apparent $\Delta G_{\text{KL}}^{\text{ST}}$ values for TA and GA contacts using 62, 95, 144 and 197 bp-long molecules with centrally located nicks or gaps. Figure 5a shows that there is a very close agreement between the stacking parameters obtained for all fragments we studied; the scatter of the $\Delta G_{\text{KL}}^{\text{ST}}$ values is within 0.1 kcal/mol.

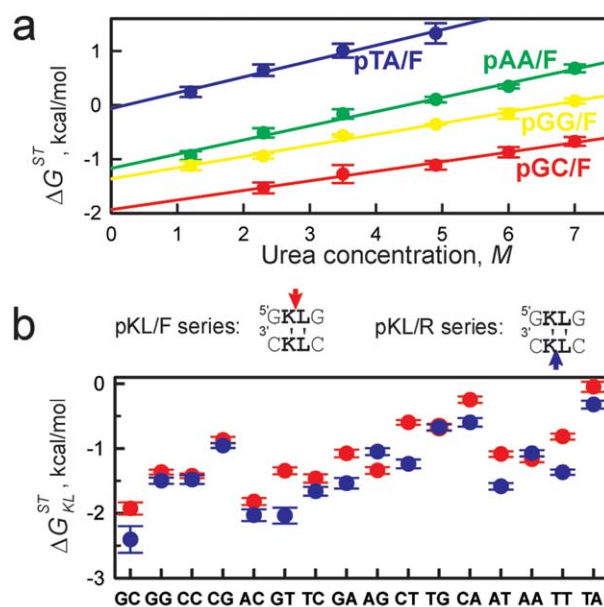


Figure 4. Free energy parameters of stacked–unstacked equilibrium at the DNA nick. (a) Free energy parameters, $\Delta G_{\text{KL}}^{\text{ST}}$, calculated from the electrophoretic mobility data as a function of the concentration of urea (see the text). The intercept with the ordinate gives $\Delta G_{\text{KL}}^{\text{ST}}$, the free energy difference for the stacked–unstacked conformational transition in the absence of urea. (b) All 32 base-stacking parameters for nicked DNA obtained here. For each dinucleotide stack the $\Delta G_{\text{KL}}^{\text{ST}}$ values for the pKL/F series (red) are compared to the $\Delta G_{\text{KL}}^{\text{ST}}$ values for the pKL/R series (blue).

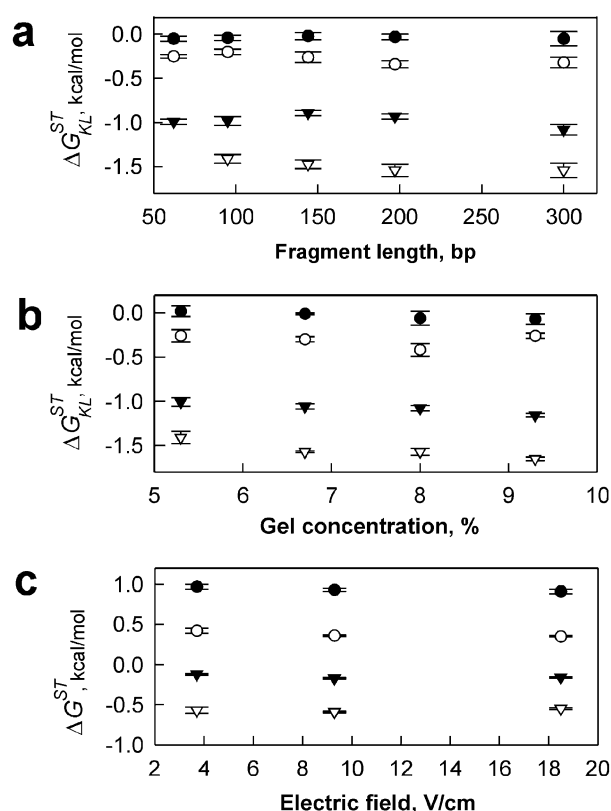


Figure 5. Effect of the PAGE conditions on the apparent stacking parameters. Apparent ΔG_{KL}^{ST} values were obtained for TA (circles) and GA (triangles) dinucleotide stacks with nicks located in the forward (filled symbols) or reverse (open symbols) strands. The length of DNA fragments carrying nicks or gaps (a) and the total concentration of the polyacrylamide gel (b) were varied. Apparent ΔG^{ST} values measured in the presence of 3.5 M urea at different values of electric field applied to the gel are plotted in c.

Another way to assess the effect of the gel matrix involves measuring the apparent stacking free energies using polyacrylamide gels with different pore sizes, which can be achieved by varying the total gel concentration. Apparent ΔG^{ST} values obtained for TA and GA dinucleotide stacks using 300 bp-long molecules decrease only slightly when the gel concentration increases from 5.3% to 9.3% (Figure 5b). Further increase in the gel concentration leads to a significant decrease of apparent ΔG^{ST} values (data not shown). Thus, under standard conditions we chose (see Materials and Methods), which correspond to 8% polyacrylamide gels, the gels are dilute enough to avoid effects of the gel matrix on the measured ΔG^{ST} values. Figure 5c shows comparison of apparent stacking parameters for TA and GA contacts obtained in PAGE conducted at three different values of electric field. No change in ΔG^{ST} values upon a five-time reduction in the applied field is observed. Overall, the data in Figure 5 strongly indicate that our standard conditions make it possible to determine

the true values of stacking parameters in the nick site unperturbed by our specific experiment design. The data also allow estimating the experimental error of individual stacking parameters determined by PAGE, of DNA molecules with nicked dinucleotide stacks not to exceed 0.1 kcal/mol.

The parameter values are scattered from -0.05 kcal/mol for the TA stack to -2.4 kcal/mol for the GC stack. This range is generally consistent with parameter values for coaxial stacking reported before.^{5–8} In fact, similarly to results obtained in melting studies,⁸ we find that GC is the most stable contact, TA is the least stable contact and a tendency towards more efficient stacking for mixed dinucleotides in order $TG < AG < GA < GT$ is observed.

Unlike the coaxial stacking parameters obtained from melting experiments,⁸ our ΔG_{KL}^{ST} values are very close for nicks produced in the forward and in the reverse strand of the given dinucleotide stack (Figure 4b). This is a crucial test of the validity of our two-state stacked–unstacked model of the nicked DNA duplex. Indeed, we assign a single free energy parameter to describe the equilibrium between open and closed states of a given dinucleotide stack $KL/L'K'$. Therefore, the stacking parameters measured for the pKL/F series with the nick in the forward strand (between K and L), should be identical with the energy measured for the pKL/R series with the nick in the reverse strand (between L' and K'). Figure 4b shows that stacking free energy values obtained for the two series agree very well. Indeed, there is no systematic difference between two sets of data. For a half of the stacks the values coincide within error bars. The differences in the other half of stacks mostly do not exceed 0.5 kcal/mol. In this respect our data are self-consistent in contrast to the most comprehensive data set on coaxial stacking obtained from melting experiments by Pyshnyi *et al.*⁸ (see Discussion).

Due to the existence of a dyad axis of symmetry in B-DNA perpendicular to the helix axis there are six pairs of equivalent stacks in intact DNA: AA/TT; GG/CC; GT/AC; GA/TC; TG/CA and AG/CT. Obviously, a nick in one of the two strands violates the dyad symmetry. However, if our model is correct, we expect the stacking parameters obtained for symmetrical stacks to be close. In fact, in our experiments we determine four equivalent ΔG_{KL}^{ST} parameters for each symmetrical dinucleotide stack. For example, results for pGG/F and pCC/R fragments (with nicks between two Gs) and pGG/R and pCC/F fragments (with nicks between two Cs) yield identical ΔG_{KL}^{ST} values (Figure 4b). For the rest five pairs of symmetrical dinucleotide stacks (all pairs are underlined in Figure 4b) the free energy parameters are in fact very close. By averaging corresponding ΔG_{KL}^{ST} values obtained for a given stack (there are four such parameters for symmetrical stacks and two for asymmetrical stacks) we arrive at a set of ten parameters, which fully describe stacking interactions in the nicked DNA duplex (Table 1).

Discussion

We assume that there is an equilibrium between straight and kinked conformations in a nicked DNA molecule and this equilibrium is governed by the stacking–unstacking free energy difference $\Delta G_{\text{KL}}^{\text{ST}}$ of neighboring base-pairs flanking the nick. We have subjected this basic assumption to very strict tests. First, if our assumption were true, the stacking parameters would be the same independently of which strand of the dinucleotide stack is nicked. The data in Figure 4b show that this is indeed the case for a half of the stacks and differences for the other half are not systematic and not large. At least this is the case for sequences studied here when a nicked stack is flanked by G·C base-pairs. Secondly, due to a dyad symmetry axis in duplex DNA perpendicular to the helix axis, 12 out of 16 stacking parameters must form six pairs of equal values. Figure 4b shows that this condition is fulfilled with the same accuracy of 0.5 kcal/mol or better, as the first condition. As a result, we have arrived at a set of ten stacking parameters of DNA presented in Table 1. It is important to note that close $\Delta G_{\text{KL}}^{\text{ST}}$ values obtained for pKL/F and pKL/R series (i.e. the nicked dinucleotide is flanked by two Gs and two Cs, respectively) also reflect the fact that the next-to-nearest neighbors do not significantly affect the energetics of the stacking at the nick site. Further assessment of the contribution of the local sequence on the structure and stability of the nick site deserves separate studies which will be reported elsewhere. Our PAGE experiments have been carried out at 37 °C, resulting in $\Delta G_{\text{KL}}^{\text{ST}}$ values corresponding to this temperature. Our PAGE-based approach makes it possible to conduct isothermal measurements of stacking parameters at temperatures typically ranging from 10 °C to 50 °C. The temperature dependence of stacking parameters will be addressed in our further studies.

Note that unlike our $\Delta G_{\text{KL}}^{\text{ST}}$ values, which are close for pKL/F and pKL/R series supporting the validity of our basic model, parameters of coaxial stacking of Pyshniy *et al.* obtained in melting experiments^{7,8} fail this test. These authors made complexes between DNA hairpins carrying a single-stranded overhang and a short oligonucleotide fully complementary to the overhang. The nicked dinucleotide stack obtained this way was flanked by two G·C base-pairs similarly to our pKL/F and pKL/R series. Pyshniy *et al.* determined

stacking parameters by studying differential stabilization of this short duplex by interaction with the duplex interface of the hairpin in melting experiments. Comparison of parameters obtained for pKL/F and pKL/R series reveals the systematic shift of $\Delta G_{\text{KL}}^{\text{ST}}$ values by 0.7 kcal/mol.⁸ Thus, in contrast to our data, the data of Pyshniy *et al.* do not satisfy the first of the two above criteria. The origin of the systematic shift between the two sets of stacking parameters obtained in melting experiments is unclear and is attributed to the cumulative contribution of what the authors dub structural factors.⁸ Note again that in our case we obtain a very good agreement between the data for pKL/F and pKL/R series.

Stacking parameters in Table 1 are of a great importance by themselves as a measure of stability of the stacked conformation at the nick site. We, however, move further assuming that these stacking parameters are very close to the stacking free energy contribution into the intact DNA duplex stability with respect to strand separation. The total free energy of melting of a long DNA molecule (neglecting the end effects) per one base-pair can be presented as:¹⁶

$$\Delta G_{\text{KL}} = \Delta G_{\text{KL}}^{\text{ST}} + \frac{1}{2} \Delta G_{\text{K}}^{\text{BP}} + \frac{1}{2} \Delta G_{\text{L}}^{\text{BP}} \quad (4)$$

where the ΔG^{BP} terms are the free energy contribution due to hydrogen bonding between complementary bases. There are two ΔG^{BP} parameters, $\Delta G_{\text{A}}^{\text{BP}}$ and $\Delta G_{\text{G}}^{\text{BP}}$ for A·T and G·C pairs, respectively.

It must be emphasized that the melting free energy in equation (4) is an ionic-strength-dependent quantity. Our data for stacking parameters correspond to salt conditions typically used in electrophoretic experiments. In our further analysis we will assume that out of three types of terms in the right-hand side of equation (4), the $\Delta G_{\text{KL}}^{\text{ST}}$ values are salt-independent whereas $\Delta G_{\text{A}}^{\text{BP}}$ and $\Delta G_{\text{G}}^{\text{BP}}$ values are fully responsible for the salt effects. Therefore, adjusting only two parameters, $\Delta G_{\text{A}}^{\text{BP}}$ and $\Delta G_{\text{G}}^{\text{BP}}$, for the DNA melting data collected under specific salt conditions, we obtain the ΔG_{KL} values corresponding to these conditions.

Realistic DNA melting parameters must satisfy a classical Marmur–Doty dependence of DNA melting temperature on the G·C-content (x_{GC}).^{31–33} Our adjustment of $\Delta G_{\text{A}}^{\text{BP}}$ and $\Delta G_{\text{G}}^{\text{BP}}$ parameters is based on the condition that heteropolymeric DNA molecules with random sequence must yield the same

Table 1. Stacking free energy parameters, $\Delta G_{\text{KL}}^{\text{ST}}$

KL	A	T	G	C
A	–1.11	–1.34	–1.06	–1.81
T	–0.19	–1.11	–0.55	–1.43
G	–1.43	–1.81	–1.44	–2.17
C	–0.55	–1.06	–0.91	–1.44

The values of $\Delta G_{\text{KL}}^{\text{ST}}$ in kcal/mol are listed for 16 ⁵KL dinucleotide stacks. Stacking parameters are calculated as an arithmetical average of two values obtained in (i) pKL/F and pKL/R series and, if applicable, (ii) for the pair of dinucleotides with dyad symmetry (all pairs are underlined in Figure 4b). Estimated errors correspond to error bars in Figure 6b.

$T_m(x_{GC})$ dependence as one yielded by stability parameters obtained from melting experiments. The corresponding dependence is described by equations (6) and (8) (see Materials and Methods). For our calculations of the $T_m(x_{GC})$ dependence we have chosen a set of unified nearest neighbor stability parameters corresponding to standard conditions of 1 M Na⁺.¹⁷ Figure 6a shows two calculated $T_m(x_{GC})$ plots, which essentially coincide if the values of $\Delta G_A^{BP} = 0.07$ kcal/mol and $\delta\Delta G^{BP} = \Delta G_G^{BP} - \Delta G_A^{BP} = -0.60$ kcal/mol are assumed. We obtain similar values of adjustable parameters upon fitting the $T_m(x_{GC})$ value calculated from stability parameters reported in dumbbell melting experiments

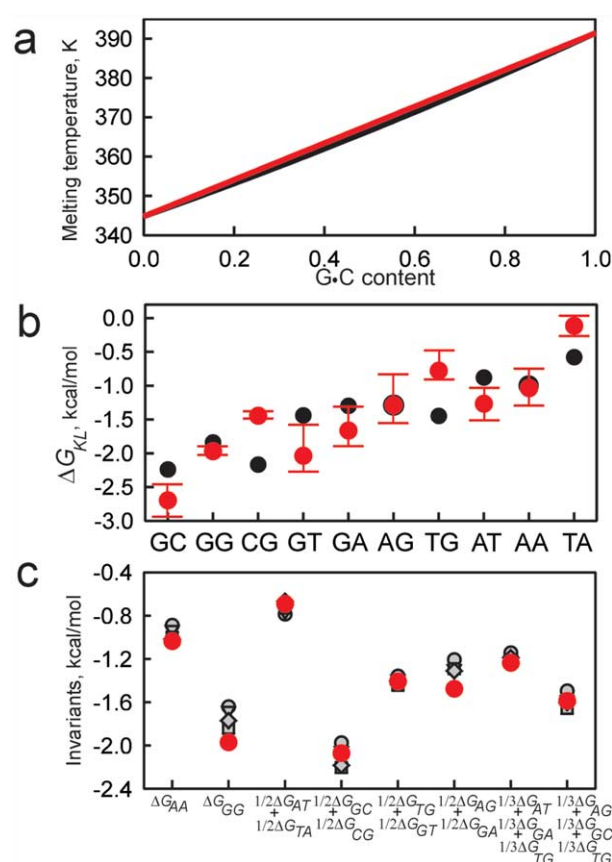


Figure 6. Stability of a DNA polymer. (a) The dependence of melting temperature of a random DNA polymer on the G·C content calculated using equations (6) and (8) for the DNA melting parameters¹⁷ (black) and for the ΔG_{KL}^{ST} values from Table 1 using ΔG_A^{BP} and ΔG_G^{BP} as adjustable parameters (red). The red line corresponds to $\Delta G_A^{BP} = 0.07$ kcal/mol and $\delta\Delta G^{BP} = \Delta G_G^{BP} - \Delta G_A^{BP} = -0.60$ kcal/mol. (b) Comparison of the stability parameters of DNA duplex per bp for all ten unique dinucleotides obtained from melting experiments¹⁷ (black) and from stacking parameters determined in this work (red). Error bars reflect the scatter of individual ΔG_{KL}^{ST} values (Figure 4b) around the mean value calculated for each unique dinucleotide stack (Table 1). (c) Comparison of eight invariants (adopted from Ref. 16) calculated from the stability parameters of DNA melting, ΔG_{KL} values, reported by Vologodskii *et al.*¹⁶ (gray circles), Doktyoz *et al.*¹⁸ (triangles), Santalucia *et al.*¹⁹ (squares), Santalucia¹⁷ (diamonds), and here (red circles).

carried out in 115 mM Na⁺¹⁸ ($\Delta G_A^{BP} = 0.09$ kcal/mol and $\delta\Delta G^{BP} = -0.46$ kcal/mol) and polymer melting studies conducted in 195 mM Na⁺¹⁶ ($\Delta G_A^{BP} = 0.10$ kcal/mol and $\delta\Delta G^{BP} = -0.42$ kcal/mol).

Using equation (4) for the free energy of melting of one base-pair in DNA and assuming that the stacking parameters in Table 1 correspond to the stacking contribution into the melting parameters, we determine, for the first time, base-stacking and base-pairing contributions into the DNA melting parameters. Our set of melting parameters is given in Table 2. Figure 6b compares our ΔG_{KL} values based on equation (4) and Table 1 with ΔG_{KL} values determined from DNA melting experiments.¹⁶ This comparison reveals appreciable difference between black points and red points for all contacts with the exception of GG, GA, AG, and AA.

Nearest-neighbor approximation of DNA stability allows for the sequence effects, i.e. differential contribution of stacking of individual DNA doublets. For an infinitely long DNA polymer, any physical property associated with the nearest-neighbor interactions is fully described by a set of eight invariants, which are linear combinations of the ten parameters for doublets.^{1,16,41} Thus, to verify the validity of our nearest-neighbor stability parameters, it is meaningful to evaluate a set of independent linear combinations (invariants) calculated using our ΔG_{KL} values with those calculated from free energies obtained in melting experiments.^{16–19} The comparison of eight invariants (adopted from Vologodskii *et al.*¹⁶) presented in Figure 6c reveals a remarkable agreement between these sets of data, upholding the validity of our nearest-neighbor stability parameters. Note, that the differences in stability parameters for individual doublets obviously seen in Figure 6b cancel each other when eight invariants are calculated. This agreement also strongly supports our claim that our stacking parameters listed in Table 1 obtained for nicked DNA are actually very close to the stacking parameters for the intact double helix and that our estimate of the $\delta\Delta G^{BP}$ value is quite reliable.

It is imperative to note that, unlike melting experiments, our quantitative analysis of heterogeneous stacking interactions in DNA involves a direct characterization of the stacked–unstacked equilibrium for each DNA doublet measured separately and independently. In this sense our approach is model-free, with no *a priori* constraints imposed on the stability parameters. This is different from the most accurate stability data extracted from the polymer melting where only eight independent parameters can be determined.^{16,42,43} For example, Vologodskii *et al.*¹⁶ complete the set of ten nearest-neighbor free energy parameters by arbitrarily assuming $\Delta G_{AT} = \Delta G_{TA}$ and $\Delta G_{GC} = \Delta G_{CG}$ (an assumption, which proved to be wrong, see Table 2). The entire set of ten free energy parameters have been obtained by thermal denaturation of short DNA duplexes.¹⁹ This approach, however, lacks the accuracy of polymer

Table 2. Standard melting free energy parameters, ΔG_{KL}

KL	A	T	G	C
A	−1.04	−1.27	−1.29	−2.04
T	−0.12	−1.04	−0.78	−1.66
G	−1.66	−2.04	−1.97	−2.70
C	−0.78	−1.29	−1.44	−1.97

The values of ΔG_{KL} in kcal/mol for 16 ^{5'}KL dinucleotide stacks are calculated from $\Delta G_{\text{KL}}^{\text{ST}}$ (Table 1) using equation (8) with $\Delta G_{\text{A}}^{\text{BP}} = 0.07$ kcal/mol and $\delta\Delta G^{\text{BP}} = -0.60$ kcal/mol (see error bars in Figure 6b for estimated errors).

melting due to a large width of the melting transitions typical for short duplexes. Additionally, in melting experiments, the measurable quantity, i.e. stability of duplex with respect to strand separation, includes the contributions from all contacts in the given DNA duplex. This cumulative quantity collected for different sequences is then processed to extract the nearest neighbor energy parameters. In our approach, each energy parameter is measured separately and independently in a strictly defined sequence environment.

We believe that our set of stability parameters in Figure 6b (red points) and in Table 2 presents a substantial revision of sets of stability parameters compiled on the basis of melting data.^{16–19} Indeed, comparing our data with the existing sets reveals that the stability parameter for the TA contact is significantly overestimated in the existing set. In fact, we find that both base stacking and base-pairing contributions of TA virtually are close to zero. This finding may be of a great importance because it supports an assumption often expressed in literature that the TATA box, which is a conserved sequence element in promoters, has anomalously low stability with respect to opening and, as a result, facilitating initiation of the transcription bubble formation.^{44–46}

It is hard to underestimate the importance of a reliable unbiased set of the DNA base stacking and base-pairing parameters for a number of biotechnology applications. Coaxial stacking at the interface of two adjacent DNA helices contributes to stabilization of duplexes formed by in tandem binding of oligonucleotides to complementary single-stranded DNA probe.^{6–8} Thus achieved stabilization of short duplexes has been exercised to improve the efficiency of short primer hybridization for standard sequencing protocols⁴⁷ and within the format of sequencing by hybridization approach.⁴⁸ Another biotechnology approach relying on the efficient stacking interaction at the helical interface at the nick site is associated with the formation of padlock and earring probes *via* ligation of a nick formed by two termini of artificial oligonucleotide hybridized to single-stranded or locally opened duplex DNA.^{49–51}

In conclusion, we have determined a complete set of free energy parameters governing the stacked–unstacked equilibrium at the site of the DNA nick. We have also derived the base stacking (Table 1) and base-pairing ($\Delta G_{\text{A}}^{\text{HP}} \approx 0.1$ kcal/mol and $\delta\Delta G^{\text{HP}} \approx -(0.4 \text{ to } 0.6)$ kcal/mol) contributions to the stability

of B-DNA. We have presented a revised set of DNA melting parameters (Table 2) calculated from totally different data than a traditional set.^{16–19} As it has been long anticipated, the double helix is stabilized by stacking interactions rather than base-pairing.^{4,15,52,53} This result is not surprising since the loss of hydrogen bonding between base-pairs may be easily compensated by hydrogen bonding of separated bases with water molecules. As it has been long anticipated, the double helix is stabilized by stacking interactions.^{4,15,52,53} Our conclusions about the relative contribution of base-pairing and base stacking are qualitatively in accord with theoretical predictions.^{10,12}

Materials and Methods

DNA

DNA constructs used in our study are pUC19 derivatives with inserts cloned in the EcoRI/HindIII site of the polylinker (Figure 2a). Recombinant plasmids were introduced into Dam[−], Dcm[−] strain of *Escherichia coli* (Invitrogen). The plasmids were digested with PvuII yielding two fragments; the smaller, 300 bp-long fragment carried the insert. Single-stranded breaks (nicks) were introduced enzymatically by N.Bst.NBI, N.AlwI or N.BbvCIA nicking endonucleases (all restriction and nicking endonucleases have been purchased from New England Biolabs, Beverly, MA).

The recombinant plasmid with an insert carrying a single nicking site within the KL/L'K' (where K and K' are complementary bases) dinucleotide stack is referred to as pKL. Digestion of pKL with N.Bst.NBI or N.AlwI resulted in a single nick positioned on the forward (pKL/F) or reverse (pKL/R) strand of the KL/L'K' stack, respectively. The nick was located 104 bp apart from a terminus of the 300 bp PvuII/PvuII fragment. DNA fragments with single-stranded gaps were obtained by introducing two nicks *n* nt apart, to the forward (pGnF series) or reverse (pGnR series) strand by sequential digestion with corresponding nicking enzymes as indicated in Figure 2a.

Alternatively, these recombinant plasmids were used as templates to obtain DNA fragments of different lengths by PCR with DNA Taq Polymerase (Promega, Madison WI) and primers designed so that the nick site (and the gap site) is located in the center of the molecule.

Detection of single nicks or gaps

Nicked/gapped DNA fragments were subjected to polyacrylamide gel electrophoresis (PAGE), concurrently with the corresponding intact fragments. Unless stated

otherwise, PAGE conditions were as follow: 8% (w/v) polyacrylamide gel (acrylamide to bis-acrylamide, 29:1, w/w) in 1×TBE (90 mM Tris (pH 8.0), 90 mM boric acid, 1 mM EDTA) with addition of urea (1.2–7 M) run at 37 °C at 18.5 V/cm. DNA band patterns were visualized by ethidium bromide staining and detected by the CCD camera with IS-1000 digital imaging software (Alpha Innotech, San Leandro, CA). Distances migrated by the DNA fragments were measured from the origin to the maximum of the peak profile using the imaging software. The relative mobility of the nicked/gapped fragment, μ , was calculated as a ratio of the distance migrated by the nicked/gapped DNA fragment to the distance migrated by the intact fragment.

Calculation of Marmur–Doty plot for DNA with a random sequence

Let us consider an infinitely long DNA polymer with random sequence, with the G·C content of x_{GC} . The occurrence of G or C is given by $(x_{GC}/2)$, the occurrence of A or T is $(1-x_{GC})/2$. Occurrences of specific dinucleotides are calculated as follows: $(x_{GC}/2)(x_{GC}/2)$ for dinucleotides containing G and C, $((1-x_{GC})/2)((1-x_{GC})/2)$ for dinucleotides containing A and T, and $(x_{GC}/2)((1-x_{GC})/2)$ for mixed dinucleotides containing A or T and G or C. The DNA polymer stability, ΔG , is expressed in terms of free energies of dinucleotide stacks, ΔG_{KL} :

$$\Delta G = \left(\frac{x_{GC}}{2}\right)^2 \sum_{\substack{GG,CC \\ CC,CC}} \Delta G_{KL} + \left(\frac{1-x_{GC}}{2}\right)^2 \sum_{\substack{AA,TT \\ AT,TA}} \Delta G_{KL} + \left(\frac{x_{GC}}{2}\right)\left(\frac{1-x_{GC}}{2}\right) \sum_{\substack{GA,AG \\ CT,TC \\ GT,TC \\ CA,AC}} \Delta G_{KL} \quad (5)$$

Regrouping equation (1), we obtain:

$$\Delta G = \left(\frac{x_{GC}}{2}\right)^2 \left[\sum_{\substack{GG,CC \\ CC,CC}} \Delta G_{KL} + \sum_{\substack{AA,TT \\ AT,TA}} \Delta G_{KL} - \sum_{\substack{GA,AG \\ CT,TC \\ GT,TC \\ CA,AC}} \Delta G_{KL} \right] + \left(\frac{x_{GC}}{2}\right) \left[\frac{1}{2} \sum_{\substack{GA,AG \\ CT,TC \\ GT,TC \\ CA,AC}} \Delta G_{KL} - \sum_{\substack{AA,TT \\ AT,TA}} \Delta G_{KL} \right] + \frac{1}{4} \sum_{\substack{AA,TT \\ AT,TA}} \Delta G_{KL} \quad (6)$$

Using thermodynamic parameters of the nearest neighbor approximation, ΔG_{NN}° ¹⁷ we calculate the standard free energy for the DNA molecule as a function of the G·C content. The ΔG° value may be expressed as:

$$\Delta G^{\circ} = \Delta S^{\circ}(T_m - T) \quad (7)$$

where T is a standard temperature which is 310 K and T_m is a melting temperature of the polymer. Assuming that both ΔS° and T_m are independent of temperature and $\Delta S^{\circ} = -25$ cal/mol K, thermal stability of a DNA molecule in terms of the melting temperature is given by:

$$T_m = T + \frac{\Delta G^{\circ}}{\Delta S^{\circ}} \quad (8)$$

Acknowledgements

This work was supported by the HFSP fellowship to E.P. and NIH grant GM59173 to M.D.F.-K. We thank Heiko Kuhn and Alexander Vologodskii for helpful discussions.

References

1. Cantor, C. R. & Schimmel, P. R. (1980). *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*. W. H. Freeman, San Francisco.
2. Saenger, W. (1984). *Principles of Nucleic Acid Structure*. Springer Advanced Texts in Chemistry, Springer, New York.
3. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. (1999). *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito, CA.
4. Kool, E. T. (2001). Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1–22.
5. Lane, M. J., Paner, T., Kashin, I., Faldasz, B. D., Li, B., Gallo, F. J. & Benight, A. S. (1997). The thermodynamic advantage of DNA oligonucleotide "stacking hybridization" reactions: energetics of a DNA nick. *Nucl. Acids Res.* **25**, 611–617.
6. Vasiliskov, V. A., Prokopenko, D. V. & Mirzabekov, A. D. (2001). Parallel multiplex thermodynamic analysis of coaxial base stacking in DNA duplexes by oligodeoxyribonucleotide microchips. *Nucl. Acids Res.* **29**, 2303–2313.
7. Pyshnyi, D. V. & Ivanova, E. M. (2002). Thermodynamic parameters of coaxial stacking on stacking hybridization of oligodeoxyribonucleotides. *Russ. Chem. B.* **51**, 1145–1155.
8. Pyshnyi, D. V., Goldberg, E. L. & Ivanova, E. M. (2003). Efficiency of coaxial stacking depends on the duplex structure. *J. Biomol. Struct. Dynam.* **21**, 459–468.
9. Danilov, V. I. & Tolokh, I. S. (1984). Nature of the stacking of nucleic acid bases in water: a Monte Carlo simulation. *J. Biomol. Struct. Dynam.* **2**, 119–130.
10. Cieplak, P. & Kollman, P. A. (1988). Calculation of the free energy of association of nucleic acid bases *in vacuo* and water solutions. *J. Am. Chem. Soc.* **110**, 3734–3739.
11. Spomer, J., Leszczynski, J. & Hobza, P. (2001). Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. *Biopolymers*, **61**, 3–31.
12. Dang, L. X. & Kollman, P. A. (1990). Molecular dynamics simulations study of the free energy of association of 9-methyladenine and 1-methylthymine bases in water. *J. Am. Chem. Soc.* **112**, 503–507.
13. Petersheim, M. & Turner, D. H. (1983). Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, **22**, 256–263.
14. Bommarito, S., Peyret, N. & SantaLucia, J., Jr (2000). Thermodynamic parameters for DNA sequences with dangling ends. *Nucl. Acids Res.* **28**, 1929–1934.
15. Guckian, K. M., Schweitzer, B. A., Ren, R. X.-F., Sheils, C. J., Tahmassebi, D. C. & Kool, E. T. (2000). Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.* **122**, 2213–2222.
16. Vologodskii, A. V., Amirikyan, B. R., Lyubchenko, Y. L.

- & Frank-Kamenetskii, M. D. (1984). Allowance for heterogeneous stacking in the DNA helix-coil transition theory. *J. Biomol. Struct. Dynam.* **2**, 131–148.
17. SantaLucia, J., Jr (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
 18. Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J. & Benight, A. S. (1992). Studies of DNA dumbbells. I. Melting curves of 17 DNA dumbbells with different duplex stem sequences linked by T4 endloops: evaluation of the nearest-neighbor stacking interactions in DNA. *Biopolymers*, **32**, 849–864.
 19. SantaLucia, J., Jr, Allawi, H. T. & Seneviratne, P. A. (1996). Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
 20. Pieters, J. M., Mans, R. M., van den Elst, H., van der Marel, G. A., van Boom, J. H. & Altona, C. (1989). Conformational and thermodynamic consequences of the introduction of a nick in duplexed DNA fragments: an NMR study augmented by biochemical experiments. *Nucl. Acids Res.* **17**, 4551–4565.
 21. Snowden-Ifft, E. A. & Wemmer, D. E. (1990). Characterization of the structure and melting of DNAs containing backbone nicks and gaps. *Biochemistry*, **29**, 6017–6025.
 22. Aymami, J., Coll, M., van der Marel, G. A., van Boom, J. H., Wang, A. H. & Rich, A. (1990). Molecular structure of nicked DNA: a substrate for DNA repair enzymes. *Proc. Natl Acad. Sci. USA*, **87**, 2526–2530.
 23. Roll, C., Ketterle, C., Faibis, V., Fazakerley, G. V. & Boulard, Y. (1998). Conformations of nicked and gapped DNA structures by NMR and molecular dynamic simulations in water. *Biochemistry*, **37**, 4059–4070.
 24. Kozerski, L., Mazurek, A. P., Kawecki, R., Bocian, W., Krajewski, P., Bednarek, E. *et al.* (2001). A nicked duplex decamer DNA with a PEG(6) tether. *Nucl. Acids Res.* **29**, 1132–1143.
 25. Hagerman, K. R. & Hagerman, P. J. (1996). Helix rigidity of DNA: the meroduplex as an experimental paradigm. *J. Mol. Biol.* **260**, 207–223.
 26. Zhang, Y. & Crothers, D. M. (2003). High-throughput approach for detection of DNA bending and flexibility based on cyclization. *Proc. Natl Acad. Sci. USA*, **100**, 3161–3166.
 27. Furrer, P., Bednar, J., Stasiak, A. Z., Katritch, V., Michoud, D., Stasiak, A. & Dubochet, J. (1997). Opposite effect of counterions on the persistence length of nicked and non-nicked DNA. *J. Mol. Biol.* **266**, 711–721.
 28. Le Cam, E., Fack, F., Menissier-de Murcia, J., Cognet, J. A., Barbin, A., Sarantoglou, V. *et al.* (1994). Conformational analysis of a 139 base-pair DNA fragment containing a single-stranded break and its interaction with human poly(ADP-ribose) polymerase. *J. Mol. Biol.* **235**, 1062–1071.
 29. Mills, J. B., Cooper, J. P. & Hagerman, P. J. (1994). Electrophoretic evidence that single-stranded regions of one or more nucleotides dramatically increase the flexibility of DNA. *Biochemistry*, **33**, 1797–1803.
 30. Kuhn, H., Protozanova, E. & Demidov, V. V. (2002). Monitoring of single nicks in duplex DNA by gel electrophoretic mobility-shift assay. *Electrophoresis*, **23**, 2384–2387.
 31. Doty, P. (2003). DNA and RNA forty years ago. *J. Biomol. Struct. Dynam.* **21**, 311–316.
 32. Marmur, J. & Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal melting temperature. *J. Mol. Biol.* **5**, 109–118.
 33. Frank-Kamenetskii, M. D. (2002). DNA Structure: Sequence Effects. In *Encyclopedia of Life Sciences*. Macmillan Publishers Ltd, Nature Publishing Group www.els.net.
 34. Levene, S. D. & Zimm, B. H. (1989). Understanding the anomalous electrophoresis of bent DNA molecules: a reptation model. *Science*, **245**, 396–399.
 35. Kerppola, T. K. & Curran, T. (1991). DNA bending by Fos and Jun: the flexible hinge model. *Science*, **254**, 1210–1214.
 36. Kahn, J. D., Yun, E. & Crothers, D. M. (1994). Detection of localized DNA flexibility. *Nature*, **368**, 163–166.
 37. Kuhn, H., Cherny, D., Demidov, V. & Frank-Kamenetskii, M. D. (2004). Inducing and modulating anisotropic DNA bends by pseudocomplementary peptide nucleic acids. *Proc. Natl Acad. Sci. USA*, **101**, 7548–7553.
 38. Wang, H., Yang, Y., Schofield, M. J., Du, C., Fridman, Y., Lee, S. D. *et al.* (2003). DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proc. Natl Acad. Sci. USA*, **100**, 14822–14827.
 39. Grainger, R. J., Murchie, A. I. & Lilley, D. M. (1998). Exchange between stacking conformers in a four-way DNA junction. *Biochemistry*, **37**, 23–32.
 40. Cann, J. R. (1996). Theory and practice of gel electrophoresis of interacting macromolecules. *Anal. Biochem.* **237**, 1–16.
 41. Gray, D. M., Tinoco, I., Jr & Chamberlin, M. J. (1972). The circular dichroism of synthetic ribonucleic acids and the influence of uracil on conformation. *Biopolymers*, **11**, 1235–1258.
 42. Gotoh, O. & Tagashira, Y. (1981). Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers*, **20**, 1033–1042.
 43. Delcourt, S. G. & Blake, R. D. (1991). Stacking energies in DNA. *J. Biol. Chem.* **266**, 15160–15169.
 44. Drew, H. R., Weeks, J. R. & Travers, A. A. (1985). Negative supercoiling induces spontaneous unwinding of a bacterial promoter. *EMBO J.* **4**, 1025–1032.
 45. Davis, N. A., Majee, S. S. & Kahn, J. D. (1999). TATA box DNA deformation with and without the TATA box-binding protein. *J. Mol. Biol.* **291**, 249–265.
 46. Travers, A.A. (2004). The structural basis of DNA flexibility. *Phil. Trans. R. Soc. Lond.* **362**, 1423–1438.
 47. Kotler, L. E., Zevin-Sonkin, D., Sobolev, I. A., Beskin, A. D. & Ulanovsky, L. E. (1993). DNA sequencing: modular primers assembled from a library of hexamers or pentamers. *Proc. Natl Acad. Sci. USA*, **90**, 4241–4245.
 48. Azhikina, T., Veselovskaya, S., Myasnikov, V., Potapov, V., Ermolayeva, O. & Sverdlov, E. (1993). Strings of contiguous modified pentanucleotides with increased DNA-binding affinity can be used for DNA sequencing by primer walking. *Proc. Natl Acad. Sci. USA*, **90**, 11460–11462.
 49. Hardenbol, P., Baner, J., Jain, M., Nilsson, M., Namsaraev, E. A., Karlin-Neumann, G. A. *et al.* (2003). Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nature Biotechnol.* **21**, 673–678.
 50. Kuhn, H., Demidov, V. V. & Frank-Kamenetskii, M. D. (2002). Rolling-circle amplification under topological constraints. *Nucl. Acids Res.* **30**, 574–580.

51. Nilsson, M., Malmgren, H., Samiotaki, M., Kwiatkowski, M., Chowdhary, B. P. & Landegren, U. (1994). Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science*, **265**, 2085–2088.
52. Crothers, D. M. & Zimm, B. H. (1964). Theory of the melting transition of synthetic polynucleotides: evaluation of the stacking free energy. *J. Mol. Biol.* **9**, 1–9.
53. Liu, H., Gao, J., Lynch, S. R., Saito, Y. D., Maynard, L. & Kool, E. T. (2003). A four-base paired genetic helix with expanded size. *Science*, **302**, 868–871.

Edited by Sir A. Klug

(Received 25 February 2004; received in revised form 24 June 2004; accepted 18 July 2004)