Published in final edited form as:

Biopolymers. 2011 April ; 95(4): 254–269. doi:10.1002/bip.21570.

Nucleic-acid Structural Deformability Deduced from Anisotropic Displacement Parameters†

Heather E. Peckham[†] and Wilma K. Olson^{*}

Rutgers the State University of New Jersey, Department of Chemistry & Chemical Biology, BioMaPS Institute for Quantitative Biology, Wright-Riemann Laboratories, 610 Taylor Road, Piscataway, NJ 08854, USA

Abstract

The growing numbers of very well resolved nucleic-acid crystal structures with anisotropic displacement parameters provide an unprecedented opportunity to learn about the natural motions of DNA and RNA. Here we report a new Monte-Carlo approach that takes direct account of this information to extract the distortions of covalent structure, base pairing, and dinucleotide geometry intrinsic to regularly organized double-helical molecules. We present new methods to test the validity of the anisotropic parameters and examine the apparent deformability of a variety of structures, including several A, B, and Z DNA duplexes, an AB helical intermediate, an RNA, a ligand-DNA complex, and an enzyme-bound DNA. The rigid-body parameters characterizing the positions of the bases in the structures mirror the mean parameters found when atomic motion is taken into account. The base-pair fluctuations intrinsic to a single structure, however, differ from those extracted from collections of nucleic-acid structures, although selected base-pair steps undergo conformational excursions along routes suggested by the ensembles. The computations reveal surprising new molecular insights, such as the stiffening of DNA and concomitant separation of motions of contacted nucleotides on opposite strands by the binding of *Escherichia coli* endonuclease VIII, that suggest how the protein may direct enzymatic action.

Keywords

atomic thermal motion; nucleic acid deformability; base-pair parameters

Introduction

The deformability of double-helical DNA is critical to its packaging in cells, recognition by other molecules, and transient opening during biochemical processing. The sequence-specific interactions of DNA with proteins and other ligands call for a sophisticated description of chain flexibility that reflects the ease with which a given sequence bends, twists, stretches, melts, *etc.* in response to forces in its local chemical environment.

[†]This work was taken in part from the dissertation of Heather E. Burden (now Heather E. Peckham) written in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Rutgers University, 2004.

^{*}To whom correspondence should be addressed. Tel: 732-445-3993; Fax: 732-445-5958; wilma.olson@rutgers.edu.

[†]Current address: Life Technologies, Suite 2450, 500 Cummings Center, Beverly, MA 01915

Empirical energy functions suitable for describing the sequence-dependent deformations of DNA have been extracted from the fluctuations and correlations of local structural parameters in ensembles of nucleic-acid structures,^{1–5} initially from known crystallographic examples and later from simulated atomic arrangements. Estimates of the range of local motions are based on the mean values and dispersion of the rigid-body parameters that define the relative orientation and positioning of complementary bases and neighboring base pairs, *i.e.*, dimer steps.

High-resolution X-ray crystallography can be used not only to solve the structures of molecules, but also to define the probability density functions that describe the apparent motions of individual atoms.⁶ Highly refined crystal structures contain information about the observed spatial distributions of each atom, *i.e.*, the principal axes of the so-called thermal ellipsoids that define the range of likely atomic positions (Figure 1). Knowledge of the average atomic coordinates and the associated ellipsoidal distributions makes it possible to describe a crystallized molecule as a flexible rather than a static entity. The incorporation of atomic displacements into the study of nucleic acids accordingly has the potential to enhance our understanding of the relationship between sequence, structure, flexibility, and function in DNA and RNA.⁷

The pioneering research of Holbrook and Kim provided the first atomic-level insights into the natural motions of double-helical nucleic acids.^{8–10} Analysis of the coordinates and thermal ellipsoids of atoms in small (2–3 bp), high-resolution RNA miniduplexes in combination with refinement of DNA dodecamer structures in terms of segmented rigid bodies¹¹ pointed to the distinctly different mobilities of the base, sugar, and phosphate moieties, *i.e.*, the average mobility of the phosphates in a variety of different structures is greater than that of the sugars, which in turn are more mobile than the bases. The work further showed (i) that the atomic vibrations are larger at the ends than in the middle of a double helix, suggesting a partial unstacking of bases at the helical termini, (ii) that base-pair sliding (a general term for the lateral translations of bases) in RNA and B-form DNA arises from the mobility of the base pairs rather than the individual bases given the maintenance of hydrogen bonds between complementary bases, (iii) that the directions of the local double-helical motions are similar regardless of crystalline environment, nucleotide sequence, and sugar type (2'-deoxyribose in DNA or ribose in RNA), (iv) that the intrinsic motions of the bases are sequence-dependent but independent of double-helical conformation, and (v) that changes in DNA mobility due to drug binding are primarily limited to the residues at the binding site and do not propagate to neighboring residues.

The crystallographically based DNA elastic functions do not consider the deformations of base-pair geometry introduced by fluctuations in atomic positions but rather collect sets of so-called 'static' arrangements of bases and base pairs from many known structures.^{1,5} The elastic properties extracted from molecular dynamics simulations, by contrast, depend upon the atomic motions dictated by the assumed force field.^{2–4} Now, a decade after the development of the sequence-dependent energy functions, there are several high-resolution nucleic-acid structures with reported anisotropic displacement parameters. This information, which started to accumulate long after the work of Holbrook and Kim was published, provides the opportunity to investigate the 'real' fluctuations of base-pair arrangements in

6–13-bp double-helical structures and to compare these distortions with the elastic properties of DNA extracted from ensembles of ‘static’ or simulated structures and with the information deduced by Holbrook and Kim from high-resolution RNA miniduplexes and novel refinement of less well-resolved DNA structures.

Here we present a new Monte-Carlo approach that takes direct account of the variation in atomic positions within a nucleic-acid crystal structure in order to simulate the likely distortions of covalent structure and the base and base-pair geometry within the molecular lattice. We present new methods to test the validity of the anisotropic displacement parameters and examine the conformational properties of a broad range of double-helical molecules, including several A and B DNA structures and individual examples of an AB helical intermediate, a Z DNA structure, a broken A RNA chain, a ligand-DNA complex, and an enzyme-bound DNA. We analyze and draw overall conclusions about the flexible nature of nucleic acids and demonstrate new methods to determine the rigid and fluctuating segments within these molecules. The approach is general and the software can be applied to the study of intrinsic deformations in other nucleic-acid molecules as more structures are crystallized and refined at sufficiently high resolution.

Methods

Displacement parameters

The anisotropic displacement parameters of each atom in a given structure are collected in a 3×3 matrix called **U**, the elements of which represent the likely spatial displacements of the atom from its mean Cartesian coordinates. The eigenvectors and eigenvalues of **U** respectively describe the directions and relative magnitudes of the principal axes of these movements,^{12,13} and the square roots of the eigenvalues, σ_i ($i=1-3$), correspond to displacements of one standard deviation.¹³ A 3σ displacement along each of the axes encloses ~99.7% of the normal Gaussian density distribution of the atomic positions.

Rigid-bond factor and rigid-bond matrix

Rigid-bond factor—One needs to check whether the reported thermal parameters represent real atomic motion before attempting to study the associated deformations of covalent structure and base-pair geometry. We make use of the Hirshfeld ‘rigid-bond test’¹⁴ to determine whether two bonded atoms have mean-square displacements that are nearly the same along the direction of the bond and whether the actual bond length is not changed appreciably during vibration. The difference $\langle u^2 \rangle_{A-B}$ in the mean-square displacements $\langle u^2 \rangle_{A \rightarrow B}$ and $\langle u^2 \rangle_{B \rightarrow A}$ of two atoms, A and B, in the direction of the bond A–B should be near zero or else the anisotropic displacement parameters may be indicative of disorder rather than apparent motion.¹⁵ The mean-square displacements of individual atoms are obtained from the eigenvalues and eigenvectors of the **U** matrix.

We define the *rigid-bond factor* f_{AB} as the percentage increase of the larger of $\langle u^2 \rangle_{A \rightarrow B}$ and $\langle u^2 \rangle_{B \rightarrow A}$ to the smaller (see illustration in Figure 2):

$$f_{AB} = \left(\Delta_{A-B} / \langle u^2 \rangle_{A \leftrightarrow B(\min)}^{1/2} \right) 100,$$

where $\Delta_{A-B} = \langle u^2 \rangle_{A \rightarrow B}^{1/2} - \langle u^2 \rangle_{B \rightarrow A}^{1/2}$ and $\langle u^2 \rangle_{A \leftrightarrow B(\min)}^{1/2}$ is the smaller of $\langle u^2 \rangle_{A \rightarrow B}$ and $\langle u^2 \rangle_{B \rightarrow A}$. If the two terms are equal (*i.e.*, $\Delta_{A-B} = 0$), then the rigid-bond factor is zero and atoms A and B move in apparent synchrony. As Δ_{A-B} increases in magnitude, the rigid-bond factor also increases and atoms A and B are less in sync with each other. A significantly lower rigid-bond factor of bonded atoms compared to non-bonded atoms shows that the atomic displacement parameters truly represent coordinated motion rather than random or systematic error.

Rigid-bond-factor matrix—Rigid segments within a molecule are identified by computing the Δ_{A-B} value for every pair of atoms, bonded and non-bonded, within the segment. If all of the atoms in a segment have Δ_{A-B} values near zero, then the motion of the atoms in that segment of the molecule is correlated and the segment behaves as a rigid body.¹⁵ If the rigid-bond factors between two segments are high, then the atoms in the two segments fluctuate independently of each other. We create rigid-bond-factor matrices to represent the rigid-bond factors of all pairs of atoms in both the base rings and all parts of a given molecule and use them to identify rigid and fluctuating segments within the molecule, *i.e.*, to display the segmented flexibility of nucleic-acid structures.

Simulation methods

We introduce a new, computationally efficient Monte-Carlo approach to generate representative arrangements of crystalline duplexes. We modify a standard Gaussian random-number generator¹⁶ and collect a Boltzmann distribution of atomic displacements without the necessity of using traditional Metropolis sampling.¹⁷ This approach is computationally more efficient and also does not suffer from the correlations between sample points or incomplete coverage of phase space associated with the Metropolis method. We repeatedly displace the atoms in one of two ways and assess the current state of the system after each displacement. Specifically, we consider displacements of individual atomic coordinates within 1σ or 3σ of the normal ellipsoidal distributions about the reported atomic positions in what we call the *Atom* method of simulation (Figure 3(a)).

We also introduce rigid-body movements of the base as a whole in which each base atom is transformed to a new position and orientation within 1σ or 3σ of its observed position (Figure 3(b)). In this *Base* method three of the ring atoms of each base are randomly selected and then displaced according to the *Atom* method of simulation. A fitting procedure, taken from the 3DNA suite of programs,¹⁸ is used to determine the transformation matrix that fits the observed coordinates of the ring atoms of the base to the plane formed by the displacement of the three selected atoms. The base as a whole is then displaced by applying the transformation matrix to the observed coordinates of each ring atom and the conventional Metropolis Monte-Carlo test¹⁷ is used to assess the new position and orientation of each base atom against the thermal ellipsoids. If the Monte-Carlo test is

satisfied for the base and at least two-thirds of the ring atoms, the movement of the base is accepted. The procedure is then repeated with the next simulated move defined, as at the start, with respect to the observed atomic coordinates.

The algorithm for thermal-motion analysis is implemented in a C++ program for Windows and is available for academic use.

Structures selected for analysis

We perform simulations based on the coordinates and thermal ellipsoids of 16 well-resolved, double-helical structures from the Nucleic Acid Database (NDB).¹⁹ Each structure meets the following criteria: (i) refinement at a resolution of 1.3 Å or better, (ii) chain length of at least six nucleotides on each strand, (iii) anisotropic displacement parameters reported for both strands, and (iv) no major chemical modifications or structural abnormalities. The data include three right-handed A-DNA decamers (NDB_IDs AD0007, AD0027, AD0028),^{20–22} six right-handed Dickerson-Drew B-DNA dodecamers (NDB_IDs BD0007, BD0012, BD0030, BD0041, BD0054, BD0060),^{20,23–27} one B-DNA decamer with 3'-terminal overhanging bases (NDB_ID BD0018),²⁸ one right-handed AB-DNA intermediate dodecamer (NDB_ID BD0026),²⁹ two left-handed Z-DNA hexamers (NDB_IDs ZD0005, ZDF060),^{30,31} one drug-DNA complex with a cytotoxic platinum(II) compound associated through a novel backbone-binding motif (NDB_ID DD0086),³² one protein-bound DNA structure in which the DNA is bound to *Escherichia coli* endonuclease VIII (NDB_ID PD0254),³³ and one RNA structure containing two coaxially stacked double-stranded fragments (NDB_ID AR0013).³⁴

Conformational parameters

Base-pair-step parameters—The simulated DNA structures are described in terms of the three translational (Shear S_x , Stretch S_y , Stagger S_z) and three rotational (Buckle κ , Propeller π , Opening σ) base-pair parameters that define the relative position and orientation of the complementary bases in a Watson-Crick base pair, and the three translational (Shift D_x , Slide D_y , Rise D_z) and three rotational (Tilt τ , Roll ρ , Twist ω) base-pair-step parameters that define the relative position and orientation of the two adjacent base pairs in a base-pair-step.³⁵ (See Figure S1 in the Supplementary Materials for descriptive images of the parameters.)

Phosphorous displacement—The structures are also characterized in terms of z_p , a parameter introduced by El Hassan and Calladine,³⁶ with the capability to distinguish between A and B DNA.^{37,38} The parameter corresponds to the displacement of the phosphorus atoms from the xy -plane of the 'middle' frame between neighboring base pairs.^{36–39} The displacements of the phosphorus atoms on individual strands, z_{p1} and z_{p2} , are averaged to find z_p (see Figure S2 in the Supplementary Materials for descriptive images). Dimer steps with z_p values greater than 1.5 Å and Slide less than −0.8 Å are classified as A type, and those with values less than 0.5 Å and Slide greater than −0.8 Å as B type, following earlier analyses.^{18,38} Dimer steps with z_p values in the range 0.5 Å < z_p < 1.5 Å are termed AB intermediates along the B→A conformational transition pathway.⁵

Results

Validity of anisotropic displacement parameters

Analysis of the rigid-bond factors of all pairs of atoms shows that the reported anisotropic displacement parameters accurately represent the thermal motions of 14 of the 16 selected structures (Table I). That is, the average values of f_{AB} for the covalently bonded atoms in each base and the averages for all covalently bonded atoms in these structures are much smaller than the averages for the non-bonded pairs of atoms. The averages of the bonded base atoms and those of all bonded atoms, however, are close to those of the non-bonded atoms in one of the B-DNA structures, NDB_ID BD0041,²⁵ suggesting that the anisotropic displacement parameters may not specify thermal motion since the bonded atoms do not move in sync significantly more than the non-bonded atoms. The three rigid-bond-factor averages in one of the Z-DNA's, NDB_ID ZDF060,³¹ are high compared to those in other structures. The higher averages may signify disorder in the crystal structure as well as thermal motion among the atoms. Whereas atomic thermal motion increases the rigid-bond factor among non-bonded pairs of atoms significantly more than among bonded atoms, disorder in the crystal structure increases the rigid-bond factor between all pairs of atoms.¹⁵ The anisotropic displacement parameters within this Z-DNA structure are also small relative to those of other structures in the study, suggesting that the displacement parameters may be less accurate than those in the other structures.

Comparative atomic displacements

The different types of atomic simulations offer useful insights into the intrinsic motions of crystalline duplexes. Whereas all approaches account satisfactorily for the atomic coordinates in the sense that the mean (x , y , z) values closely match the observed values (data not shown), the range of atomic motion, as measured by the computed mean-square displacement matrix $\mathbf{U}$, depends upon the method of simulation (Table II). As expected from the definition of $\mathbf{U}$, the computed atomic variance closely matches the reported vibrations if the atoms are allowed to fluctuate independently and at random within 3σ of the normal distributions (compare the observed $\mathbf{U}$ matrix for a representative atom of the 53-d(CpGpCpGpApApTpTpCpGpCpG)-33 B-DNA dodecamer duplex (NDB_ID BD0054)²⁶ with the values in Table II generated by an *Atom3 σ* simulation of 10^4 random moves). The magnitude of atomic displacement decreases by roughly a third but the preferred directions of motion persist if the atoms fluctuate independently within the 1σ limit, *i.e.*, the eigenvectors of $\mathbf{U}$ are roughly coincident but the eigenvalues are reduced in value by a factor of $\sim 3^2$ (see the *Atom1 σ* entry in the table, where the eigenvectors are stored in the matrix $\mathbf{Q}$ and the eigenvalues in the matrix $\mathbf{\Lambda}$).

The discrepancies in the $\mathbf{U}$ matrices generated in the *Base1 σ* and *Base3 σ* simulations compared to observation are indicative of the extent to which the bases behave as ideal, rigid bodies. For example, the observed motion of the N9 atom of guanine 10 in the dodecamer duplex, the atom described in Table II, is more limited and directed in a different sense with base-level moves compared to simulated atomic motions. Because the observed fluctuations of N9 are not perfectly coupled to those of the other guanine 10 atoms in this structure, only a fraction of the possible moves of N9 within its allowed 3σ limit concomitantly

accommodates the allowed 3σ fluctuations of all other guanine atoms. The accepted rigid-body motions of the base thereby limit N9 to a smaller ellipsoidal volume (spanning ~25% that of the observed thermal ellipsoid in the case of the *Atom3 σ* calculations), with principal axes realigned (by 40–60 deg) with respect to the observed axes. The restriction of rigid-base geometry on acceptable atomic motions is clear from Figure 3. Although the proportions and alignment of the nine thermal ellipsoids in the image are suggestive of comparable atomic motions, with the largest atomic variations occurring in the plane of the base along axes that are slightly inclined with respect to the edges of the page, the subtle differences in ellipsoidal size and orientation restrict the allowed range of rigid-body motions. Whereas the long axes of the guanine 10 ellipsoids are roughly parallel to one another, the secondary axes are not so well aligned. That is, the long axis of the N9 atom forms angles of 10–21 deg with the long axes of the eight other guanine 10 atoms, but the secondary axes form angles of 21–60 deg with the corresponding axes of the other atoms. Thus, the simulated motions of the base as a whole are more limited than those described by random, independent atomic motions.

Simulated bond lengths

Although the *Atom3 σ* simulation accurately reproduces the thermal ellipsoids, the mean bond lengths determined by such motions substantially exceed the observed values ($\langle d \rangle$ vs. d_0 values in Figure 4). The computed mean values, reported here for the B-DNA dodecamer duplex (NDB_ID BD0054)²⁶ (and in Figure S3 in the Supplementary Materials for six other nucleic-acid structures), also exceed the observed values, but to a much lesser extent, if the atoms of the structure are allowed to fluctuate independently and at random within their 1σ limits. These differences follow from the theory of Busing and Levy,⁴⁰ which predicts the average length of a fluctuating bond to be greater than the observed value. The instantaneous distance between two atoms is derived from both the distances between the centers of the ellipsoids used to define their likely displacement and the correlation of the motion between the ellipsoids, with the lower and upper limits on the length of a bond determined respectively by the parallel and anti-parallel correlated moves of the constituent atoms.⁶ These limits on the bond length, represented schematically in Figure 4(a), are the bounds in which the average value of the parameter should be found. The lower bound d_L always exceeds the observed length of the bond,⁴⁰ and the upper limit d_U corresponds to the theoretical limit on the bond length if the atoms remain within their mean-square displacements. It should be noted that the mean-square bond displacement corresponds to one standard deviation of the displacement of each atom. Displacing the atoms beyond one standard deviation, as done in the *Atom3 σ* simulation, may thus be unrealistic. The mean bond length matches the observed value more closely if the atoms move both synchronously, *i.e.*, with only minor occurrences of anti-parallel motion, and to a small extent. The simulated average bond lengths of the relatively ‘rigid’ bases, with smaller thermal ellipsoids, accordingly lie closer than those of the more widely deformed backbone atoms to the observed bond lengths (Figure 4). The simulated bond lengths will equal the observed bond lengths only if the thermal ellipsoids are congruent and the atomic moves are perfectly synchronized.

Simulated base parameters

Regardless of the method of simulation, the average values of the Watson-Crick base-pair parameters, the sequential differences in base-pair parameters in adjacent residues, and the base-pair-step parameters very closely match those in the reported experimental reference structures (see Tables SI and SII in the Supplementary Materials). Thus, the base-pair and base-pair-step parameters of a nucleic-acid crystal structure are an accurate representation of the average values of these parameters when thermal motion is taken into account.

As expected, the range of base-pair and base-pair-step parameters depends on the simulation method, *i.e.*, greater variation of derived parameters with increasing range of allowed atomic motions. The relative variation of structural parameters, nevertheless, persists at different levels of computation and, significantly, in different helical forms. Figure 5 illustrates the common deformational patterns in two representative structures: (i) the B-DNA dodecamer duplex (NDB_ID BD0054)²⁶ and (ii) the complex of a 13 bp oligodeoxynucleotide duplex with the *Escherichia coli* endonuclease VIII protein, also termed Nei (NDB_ID PD0254).³³ The very large range of conformational variation in these molecules — as much as ± 1.5 Å in translational variables (images on left) and ± 23 deg in rotational variables (images on right) — reflects the assumed 3σ limit of atomic movement, *i.e.*, results based on *Atom3 σ* simulations. Although the range of variation is smaller when more restrictive atomic simulations are performed, *i.e.*, deviations up to ± 0.4 Å in translations and ± 7 deg in rotations for *Atom1 σ* simulations, the range of sampled conformations remains large for base-level simulations, with respective translational parameters varying as much as ± 2.0 and ± 2.9 Å and rotational parameters as much as ± 26 and ± 55 deg for *Base1 σ* and *Base3 σ* calculations, respectively. The fluctuations in five of the six complementary base-pair parameters generally exceed those in the base-pair-step parameters (points connected by solid vs. dotted lines in Figure 5). The structures, however, deform less easily via Stretch (*Sy*), the base-pair parameter that most directly reflects the lengths of the Watson-Crick hydrogen bonds. Among base-pair-step parameters, Slide (*Dy*), the translational motion of successive base pairs analogous to Stretch, and Twist (ω) consistently show the least variation. As discussed below, these deformational propensities differ substantially from those deduced from analyses of ensembles of crystal structures.^{1,5} Nevertheless, given that Slide is related to the change of helical phase,^{38,41} the limited variation in Slide suggests that the atomic displacements in the crystal do not generally induce substantial deformations to other double-helical forms.

Whereas the fluctuations in structural parameters are relatively insensitive to base-pair position in the B-DNA dodecamer and in most other crystalline duplexes (see Figure 5 and Figure S4 in the Supplementary Materials), the base pairs on either end of the Nei-bound DNA fluctuate much more than those in the middle of the structure where the protein, a DNA repair enzyme that excises oxidized pyrimidines, binds. In taking a firm hold on DNA and thereby stiffening the base pairs on either side of the repair site, the enzyme appears to let the ends of the molecule flex. The current analysis, however, excludes the abasic pentane-3,4-diol-5-phosphate modification used to capture the enzyme-DNA complex structure and thus sheds no light on the deformability of DNA at the cleavage site (in the center of the crystallized DNA fragment). The distinct asymmetry in motions at the 5'- and

3'-ends of the DNA sugar-phosphate backbone may reflect the loose contact of protein to the backbone near the stiffer end and the absence of close contacts at the other end (see interatomic contact data presented below).

Remarkably, the motions at the ends of the Nei-DNA complex exceed those found in all 'pure' ligand-free double-helical structures considered here. The redistribution of motion may reflect a redistribution of the thermal energy that is apparently more evenly apportioned in an unbound duplex. The Z-DNA structure, by contrast, stands out as being substantially more rigid than all other structures, with smaller deviations in all derived base-pair and base-pair-step parameters. Despite its alternating 'zig-zag' conformational pattern, the variability of base-pair motion is essentially constant along the length of the molecule. Two base-pair-step parameters, Shift (D_x) and Rise (D_z), however, exhibit an alternating, zigzag pattern of variability in the A-DNA, AB-DNA, and RNA structures. Moreover, the degree of (D_x , D_z) variability is coupled in these structures with respective correlation coefficients of 0.91, 0.80, and 0.72. The analysis omits consideration of factors in the crystal lattice, which may limit or enhance the motions of base pairs in these double helices. There is no apparent connection of the alternating pattern of fluctuations to base sequence in the few structures examined here.

Phosphorus displacement

The broad range of atomic fluctuations allowed in the *Atom3σ* simulations of A- and B-DNA structures appears to induce changes of double-helical state at selected base-pair steps. The computed values of Slide and z_p , the two structural parameters capable of distinguishing A-form and B-form DNA,^{18,38} occasionally fall outside the characteristic helical ranges. That is, 17,581 of the 110,000 simulated structures of the B-DNA dodecamer (NDB_ID BD0054)²⁶ adopt intermediate AB states and 165 assume A-like states and 15,452 of the 90,000 simulated structures of the A-DNA decamer (NDB_ID AD0027)²¹ adopt AB states and 431 assume B-like states. The number of such conformational excursions naturally decreases in simulations with more restrictions on atomic displacements. The majority of *Atom3σ*-simulated configurations (99,526 of 110,000) of the right-handed AB-DNA intermediate structure (NDB_ID BD0026)²⁹ assume values of z_p characteristic of A DNA, suggesting that the structure may be an A-type helix rather than the intermediate conformational form proposed in the literature.²⁹ The other types of simulations lead to the same conclusion; 106,570 of the 110,000 *Atom1σ*-simulated configurations, 96,016 of the 110,000 *Base3σ*-simulated configurations, and 106,064 of the 110,000 *Base1σ*-simulated configurations of the given structure are A-like

Examination of the z_{p1} and z_{p2} components of phosphorus positions in the simulated structures (Figure 6) reveals notable differences in the helical states of the two strands of the Nei-bound duplex. The leading strand I of DNA, which bears a modified abasic site, adopts a more A-like form than the complementary strand II carrying an unpaired stacked adenine, *i.e.*, $\langle z_{p1} \rangle > \langle z_{p2} \rangle$. The phosphorus atoms in the reactive strand also undergo excursions to larger A-like values of z_{p1} at the flexible, 3'-end of the visualized duplex, *i.e.*, the G₄pA₋₃ and A₃pA₋₂ dimers at positions -4 and -3 with respect to the excised base X₀ (data not shown). The close atomic contacts between DNA and protein at these steps may contribute

to these deformations. (Atomic contacts are defined as pairs of atoms separated by distances of less than or equal to 3.4 Å in which one atom is part of the nucleic acid and the other atom is part of the protein molecule.)

Comparable differences in z_{P1} and z_{P2} do not occur in the 'pure', ligand-free structures and the drug-DNA complex considered here (Figure 6). The positioning of phosphorus atoms in the simulated A-DNA structures clearly stands out from that in the B-DNA and TriplatinNC-bound structures. The distributions of z_{P1} and z_{P2} in both the right-handed AB-DNA intermediate structure and the double-stranded RNA structure resemble those in the A-DNA structures. The two parameters exhibit bimodal distributions in the Z-DNA structures, with lower values of both z_{P1} and z_{P2} at pyrimidine-purine steps and higher values at purine-pyrimidine steps. The distinctive positioning of phosphates seemingly reflects the known dinucleotide repeating unit of Z DNA.⁴²

Correlations between structural parameters

The simulations produce correlated changes in base-pair and base-pair-step parameters as atoms are moved in their crystal lattices. The effects are naturally more discernable when the atoms fluctuate over broader ranges. The *Atom3σ* correlation pattern of base-pair-step parameters and sequential differences in complementary base-pair parameters in the B-DNA dodecamer (NDB_ID BD0054)²⁶ in Figure 7 is representative of the correlations found in each of the seven B-DNA structures that have been analyzed. (Corresponding images of the correlation patterns in other double-helical forms and in the TriplatinNC-DNA and Nei-DNA complexes are reported in Figure S5 in the Supplementary Materials.) Here a color-coded gradient is used to represent correlations ranging from -1.0 (red) to 0.0 (white) to +1.0 (blue). The small boxes within each of the 11² submatrices of parameter correlations represent the base-pair steps along the sequence of the structure. The number of squares in the rows and columns of the submatrices is therefore equal to the number of dimers in the structure. Adjacent dimers share a common base-pair with the dimer of interest, *i.e.*, the 5'-base M of the MpN step is the 3'-base of the preceding LpM step and the 3'-base N is the 5'-base of the succeeding NpO step in the hypothetical LMNO nucleotide sequence.

Each of the base-pair-step parameters, except for Slide (Dy), and all of the sequential differences in base-pair parameters are negatively correlated with the corresponding parameter in adjacent dimers (red boxes on either side of the blue diagonals of relevant submatrices). Such correlations serve to balance the overall structure of the molecule. The slightly positive correlation of Slide values in adjacent dimers in combination with accompanying positive correlations between Slide and Roll (Dy, ρ) and between Slide and Twist (Dy, ω) in the same steps is suggestive of a global translational mode that displaces and reorients base pairs with respect to the helical axis.

The color-coded pattern in Figure 7 also reveals the four correlations of base-pair-step parameters and the differences in rigid-body parameters of successive base pairs known to be intrinsic to the calculations:^{39,43} (Dx, s), (Dz, k), (τ, Sz), (ω, Sx). The specified differences in base-pair parameters perturb the 'middle' frame used to define the base-pair-step parameters and consequently affect the computed values of the latter quantities. The effects are most pronounced when base-pair and base-pair-step parameters are coupled

within the same residue but are also strong when the variables are in adjacent dimers. By contrast, some of the rigid-body parameters are fairly independent of others, *e.g.*, the relative translational moves of complementary bases appear to be independent (white spaces associated with S_x , S_y , S_z boxes in Figure 7).

Notably, all of the parametric correlations occur within the same residue or between adjacent residues. Thus, reduced base-pair models of DNA based on the observed fluctuations in crystal structures need not extend beyond the trimeric level.

Atomic-level fluctuations

The color-coded schematics in Figure 8 illustrate the simulated fluctuations of the B-DNA dodecamer (NDB_ID BD0054)²⁶ and the Nei-DNA complex (NDB_ID PD0254).³³ Corresponding images of the molecular motions in other double-helical nucleic-acid structures are reported in Figure S6 in the Supplementary Materials. The bases in the reference structures are represented by block images that are color-coded by chemical identity,¹⁸ and the networks of atomic thermal ellipsoids, shown in the same orientation as the block images, are based on the reported **U** matrices. The thermal ellipsoids of DNA atoms are shown at the 50% probability level (*i.e.*, 0.67σ level, where there is a 50% chance that an atom lies within the ellipsoid at any instant). The color-coding and relative sizes of the thermal ellipsoids depict the differences in base-pair deformability shown in Figure 5. Whereas the allowed fluctuations of base atoms and the consequent variation in base-pair parameters are nearly uniform along the B-DNA structure, the base atoms in the center of the Nei-DNA complex are rigid compared to those at the ends of the duplex. The images also show the greater motion of the sugar-phosphate backbone atoms compared to those in the base-pair core of the double-helical structure, clarifying the large differences in average base and backbone bond lengths reported in Figure 4.

The rigid-bond-factor matrices of the two molecules, shown in the lower half of Figure 8, take account of all possible ‘bonds’ between base ring atoms in each molecule, *i.e.*, both direct chemical bonds and indirect virtual bonds between nonbonded atom pairs. The components of the grids correspond to the individual atoms (the smallest color-coded elements) and bases (the outlined subarrays of ~50 elements) that make up the two structures. (Corresponding grids for other double-helical nucleic-acid structures are reported in Figure S6 in the Supplementary Materials.) A gradient from white to blue is used to represent the average rigid-bond factors. The white spaces indicate atoms in synchronized motion (values near zero) and the blue spaces represent higher degrees of asynchronous movement. The asynchronous motions within the structures correspond to the darkened bands and blocks within the color-coded grids.

The white diagonal line of boxes running from the upper left-hand corner to the lower right-hand corner of the B-DNA matrix (Figure 8(a)) shows that each base in the structure behaves as a rigid body. The diagonal line of boxes running from the lower left-hand corner to the upper right-hand corner of the figure, while not as clearly white as the opposing diagonal, shows that the base pairs are more rigid than other parts of the molecule. All of the base pairs and all but one of the B-DNA bases fluctuate approximately as rigid bodies. The cytidine at the 5'-end of Strand II in the structure, however, fluctuates out of sync from the

other bases (note the darkened vertical and horizontal bands in the center of the image). The coordination of the sugar to a thalium cation at this site may contribute to the apparent conformational freedom of the uncontacted base.

The binding of *Escherichia coli* endonuclease VIII to DNA introduces a very different pattern of asynchronous motion in the double helix. The Nei-DNA rigid-bond-factor matrix, shown in Figure 8(b), includes the pairwise motions of bonds involving the unpaired adenine on strand II but does not take consideration of the motions of the modified abasic site on the pyrimidine-depleted strand I. As is clear from the color-coded diagonals, the bases of the protein-bound duplex exhibit the rigid-body behavior expected from their aromatic chemical structures but the base-pair atoms, particularly those in the terminal G-C pair (formed from the 3'-guanine on strand I and the 5'-cytosine on strand II) and the adjacent A-U pair (located near the center of the grid), show notable (up to 61%) asynchronous motion. The different purine-pyrimidine composition of the two strands, combined with the incorporation of the stacked, unpaired adenine on Strand II, introduces a jaggedness in the (lower-left to upper-right) diagonal associated with atoms in complementary bases. Most of the contacts of protein with DNA occur in the immediate vicinity of the repair site, *i.e.*, there are seven or more close contacts of 3.4 Å or less of Nei with DNA atoms in the 5'-A₊₁pX₀pG₋₁pA₋₂pA₋₃-3' fragment of Strand I and in the 5'-A₀pU₋₁pC₋₂-3' fragment of Strand II, where complementary bases in the vicinity of the repair site are italicized and defined in a positive and negative sense, as in the literature,³³ with respect to the excised base at X₀ and the stacked, unpaired adenine at A₀. Interestingly, the tightly contacted and rigid nucleotides on either side of the repair site fluctuate asynchronously of one another (note the darkened boxes associated with A₊₁ and G₋₁ on Strand I and C₋₂ and G₋₅ on Strand II). The motions of the unpaired adenine on Strand II (A₀) are also uncoupled from those on A₊₁ and G₋₁, but to a lesser extent than C₋₂ and G₋₅. The enzyme not only stiffens its DNA target but also separates the motions of contacted nucleotides on opposite strands, suggesting the way the protein is able to direct enzymatic action. As anticipated from the variation of base-pair parameters in Figure 5, atoms in the contacted residues also move out of sync from those in uncontacted or loosely contacted residues at the 3'-end of Strand I and the 5'-end of Strand II.

Discussion

For the majority of crystalline macromolecules, the 'static' information defined by the three-dimensional atomic coordinates only partially describes the structure. The methods developed in this study allow us to look at the intrinsic flexibility of individual nucleic-acid structures and lead us to a deeper understanding of their three-dimensional folding and variation.

Local mobility of nucleic acids

Our findings confirm and add to the pioneering studies of Holbrook and Kim,⁸⁻¹⁰ who provided early insights into the intrinsic mobility of DNA and RNA. We take a simple, direct approach to the treatment of nucleic-acid motion, using the thermal ellipsoids of the constituent atoms as the potential functions that guide molecular deformations (Figure 1).

We examine long double-helical structures that have become available in recent years (Table I) rather than the short, miniduplexes originally considered. We check the rigid-bond factors between atoms to determine that the thermal ellipsoids are a true measure of molecular deformability (Figure 2). We also characterize the motions quantitatively in terms of (i) the chemical bond lengths (Figures 4 and S3), (ii) the rigid-body parameters that relate the positioning and orientations of complementary bases and successive base-pair steps (Figures 5 and S4),^{18,35} (iii) the locations of phosphorus atoms in dimeric units (Figure 6), and (iv) the relative eigenvalues and eigenvectors of atomic thermal factors generated by random atomic fluctuations *vs.* pure rigid-body moves (Figure 3, Table II). The composite information allows us to obtain a more detailed understanding of nucleic-acid motions than could be gleaned from the original work of Holbrook and Kim.

Like Holbrook and Kim, we find that the backbone atoms are typically more mobile than the base atoms and that the ends of the double helix are often more flexible than the center (Figures 8 and S6), although there are structures, such as the Z DNA duplex (NDB_ID ZD0005),³⁰ where all atoms are uniformly rigid and certain A-DNA structures (NDB_ID AD0027),²¹ where many of the notable fluctuations occur in the center of the duplex. Our computations of rigid-bond-factor matrices (Figures 8 and S6) provide new evidence that the base pairs in crystalline B DNA behave approximately as rigid bodies but give no indication that the base pairs fluctuate as rigid bodies in the one double-stranded RNA structure included in our analysis, *i.e.*, the simulated mobility of the RNA oligonucleotide duplex fragments differs from that of the 2–3 bp miniduplexes examined by Holbrook *et al.*⁸ Although we see differences in the degree of local mobility among different types of helices, we see trends in the relative variation of base-pair and base-pair-step parameters and in the correlations between structural parameters that support the idea of Holbrook and Kim, that local nucleic-acid mobility is independent of helical form, at least for right-handed A- and B-DNA structures and AB-DNA helical intermediates. Many of the parametric correlations also persist in the Z-DNA, drug-DNA, protein-bound DNA, and double-stranded RNA structures considered here. Finally, we see bimodality in the motions of the phosphorus atoms in Z DNA that is consistent with the different modes of mobility found by Holbrook and Kim.⁹

The effects of *Escherichia coli* endonuclease VIII binding on DNA, however, differ strikingly from the localized changes in duplex mobility detected by Holbrook and Kim in the crystal complex of DNA with the minor-groove-binding ligand, daunomycin.¹⁰ Although the protein, like the drug, limits the mobility of DNA at its binding site, changes in mobility occur at the ends of the protein-bound duplex away from the site of binding. By contrast, changes in mobility do not propagate to residues beyond the daunomycin binding site. The partial unstacking of bases at helical termini found by Holbrook and Kim⁸ in the Dickerson-Drew dodecamer occurs in the protein-bound duplex but does not occur in the ligand-free structures examined here. The fluctuations of complementary base-pair parameters are roughly constant along the contours of the ligand-free duplexes (Figure 5). Large variations only occur at the ends of the protein-bound structure.

In the drug-bound DNA considered here (TriplatinNC + DNA) (NDB_ID DD0086),³² the amino-acid-mimicking ligand binds to the phosphate oxygen atoms at one end of the duplex

and thus associates with the backbone rather than intercalates between the base pairs or binds in the grooves. The presence of the drug contributes to an increase in axial bending and a perturbation of the helical structure compared to the native structure.³² Our analysis finds additional changes in the rigid-body motions of the bases and base pairs, most notably enhanced base-pair opening σ and dimeric twisting ω , with the greater perturbations of these rotational parameters occurring near the center of the complex (Figure S4), at sites where the ligand-free B-DNA duplex of the same sequence is comparatively stiff (Figure 5). A corresponding enhancement occurs in the displacement of bases and base pairs, namely in the base-pair shear S_x and the dimeric shift D_x in the same region of the drug-DNA complex. The clamping of the drug to the phosphate limits the thermal motions of the backbone atoms (Figure S6) compared to the ligand-free duplex (Figure 8). This stiffening occurs not only near the site of drug binding toward the 3'-end of the leading strand but also at the other end of the duplex near the sites of coordinated metals. This behavior differs from the changes in DNA mobility associated with daunomycin binding, which do not propagate to neighboring residues.

Comparison with structural ensembles

The motions of complementary base pairs detected in individual high-resolution double-helical structures resemble the patterns of deformation deduced from statistical analyses of ensembles of 'static' structures with no information on the likely positions of individual atoms.^{44,45} Specifically, the rigid-body parameters that determine the length and linearity of hydrogen bonds, Stretch S_y and Opening σ , vary to a lesser extent than other base-pair parameters of the same type, *i.e.*, the two remaining rigid-body translations and the two other rigid-body rotations. Open questions remain, however, regarding the ease of out-of-plane base-pair distortions via Buckle κ and Propeller π rotations. Whereas π is slightly more variable than κ in the isolated duplexes treated here (Figures 5 and S4), the standard deviation of κ exceeds that of π in collections of well-resolved A- and B-DNA structures. There is similar uncertainty in the relative ease of complementary base displacement via Shear S_x or Stagger S_z , *i.e.*, in-plane *vs.* out-of-plane 'melting' of Watson-Crick base pairs.

Despite the approximate similarity of base-pair motions, the fluctuations of dimeric fragments induced by atomic thermal motions in individual structures differ substantially from the deformations of DNA deduced from the conformational variability of many protein-bound DNA duplexes.^{1,5} Whereas the shearing of base pairs along their long axes via Slide D_y spans a broad range of states and the 'vertical' (van der Waals') separation of base-pair planes via Rise D_z spans a very narrow range of values among base-pair steps in the diverse set of 'static' structures, D_y is the stiffest and D_z the most deformable translational parameter in simulations guided by atomic thermal ellipsoids (Figures 5 and S4). Likewise, the bending of dimers via Roll ρ , about the long axes of the dimers and into the grooves of DNA, and the twisting of consecutive base pairs via ω span broader ranges of values than the bending of base pairs via Tilt τ , toward the backbone and about the short dimeric axes, in ensembles of 'static' structures,^{1,5} but the twist fluctuates to a lesser extent than both bending components in the isolated crystal structures studied here. The bending

anisotropy detected in structural ensembles and long predicted to occur in DNA,⁴⁶ however, persists in the simulated crystallographic systems.

These different perspectives on base-pair-step deformability reflect that the fact that the standard deviation of the observed, *i.e.*, average, values within an ensemble of ‘static’ structures is a very different measurement from the standard deviation of the movements of the dimers in a single structure. For example, Rise could adopt a wide range of values within the base-pair steps of a DNA molecule (as suggested by the present simulations), but if the average value of Rise is similar within the dimers of all structures in an ensemble, Rise would not be identified as having much variability. In other words, the standard deviation of a parameter within an ensemble of structures is a measure of the variation among the average values of the parameter in all of the dimers in all of the structures, whereas the standard deviation resulting from a simulation guided by the thermal ellipsoids is a measure of the actual variation of a parameter within a single dimer in a specific crystalline context.

As discussed elsewhere,⁴⁷ the statistical approach^{1,5} used to deduce information about DNA deformability reflects a large number of closely related crystal structures, each of which can be thought of as a distinct low-energy conformational substate. Optimization of DNA conformation on the basis of the detailed interactions of all atoms typically yields a large number of closely spaced minimum-energy substates of similar structural character.⁴⁸ Each of the structures treated here is analogous to one of these minimum-energy states, but transitions between closely related, competing minima/structures are not considered. In other words, the present study extracts the apparent motions of nucleic acids within individual substates whereas the statistical approach smoothes the ‘dimpled’ surface of conformational states and identifies larger-scale motions of neighboring base-pair steps. The correspondence between the mechanical constants of DNA and values extracted from experiment suggests that transitions between conformational substates have a significant influence on the global properties of DNA.⁴⁷

The narrowed ranges of dimeric motions found in isolated structures miss some of the ‘universal’ correlations of base-pair-step parameters identified in structural ensembles^{1,5,49} and also captured in atomic-level simulations,^{2–4} specifically the negative correlations of Twist and Roll (ω - ρ) and Roll and Slide (ρ - Dy) and the positive correlations of Tilt and Shift (τ - Dx) and Twist and Slide (ω - Dy). The present simulations fail to capture the ω - ρ and ω - Dy correlations, reveal a positive ρ - Dy correlation in all structures except Z DNA, but find a positive τ - Dx correlation in right-handed double-helical structures (Figures 7 and S5). Rise (Dz) is also more closely coupled to other parameters (Tilt, Roll, Slide) in the present simulations compared to the ensembles, where Rise varies almost independently of all other parameters save for the coupling to Tilt (τ - Dz) at AG·TC steps.^{1,5} The different pattern of correlations reflects the different motions of DNA within a narrow minimum-energy state *vs.* a broad, smoothed energy surface like that discussed above.

The present study also reveals couplings of parameters in successive base pairs and adjacent base-pair steps that balance the overall crystallized structure (Figures 7 and S5), *e.g.*, the tendency of one step to untwist is compensated by the tendencies of adjacent steps to overtwist. Such correlations also occur in ensembles of DNA structures, for which there are

now sufficient examples of trimeric steps to identify meaningful correlations (A.V. Colasanti & W.K. Olson, unpublished data).

Molecular motions and protein binding

Our investigation of the apparent motions of the Nei-DNA complex adds a new perspective on protein-nucleic acid interactions, showing how the association of the enzyme perturbs the intrinsic structure and deformability of the base-pair steps. The segments of DNA that are bound to protein fluctuate independently from and to a lesser extent than the segments that are unbound (Figure 8). Moreover, the motions of atoms on strand I bearing the enzyme-cutting site are uncoupled from those on strand II with the unpaired adenine and are directed by protein along a unique conformational route. Whereas the phosphorus atoms along strand II fluctuate about positions characteristic of B DNA, the phosphorus atoms on strand I occupy more A-like states (Figure 6). Similar changes of DNA helical state accompany the binding of other enzymes that perform cutting or sealing operations at the phosphodiester linkage, allowing access to the O3'-P bond.³⁸ Thus, we find that *Escherichia coli* endonuclease VIII orchestrates the stiffening of the highly contacted nucleotides at its target site while simultaneously enhancing functionally relevant distortions of flanking base-pair and backbone atoms and decoupling the motions on complementary strands.

Information of this sort demonstrates the promise of direct, knowledge-based simulations of atomic motion in deciphering the important role of molecular fluctuations in the recognition, packaging, and biological processing of nucleic acids by drugs and/or proteins. Such calculations can be performed routinely in the future as new high-resolution DNA and RNA structures with anisotropic thermal factors become available. The information will in turn provide useful benchmarks for the development of atomic-level simulations of nucleic acids.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Professor Helen M. Berman for guidance and helpful discussions on nucleic-acid crystallography, Dr. Xiang-Jun Lu for the use of the source code of 3DNA, Dr. Yun Li for sharing his database of protein-DNA contacts, Dr. David Swigon for guidance with the Monte-Carlo simulations, and Dr. Fei Xu for the Python scripts used to generate the correlation matrices and rigid-bond-factor matrices. The U.S. Public Health Service under research grants GM20861 and GM34809 has generously supported this research. HEP gratefully acknowledges support from a U.S. Department of Education GAANN Fellowship.

References

1. Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB. Proc Natl Acad Sci, USA. 1998; 95:11163–11168. [PubMed: 9736707]
2. Lankas F, Spomer J, Hobza P, Langowski J. J Mol Biol. 2000; 299:695–709. [PubMed: 10835278]
3. Fujii S, Kono H, Takenaka S, Go N, Sarai A. Nucleic Acids Res. 2007; 35:6063–6074. [PubMed: 17766249]
4. Pérez A, Lankas F, Luque FJ, Orozco M. Nucleic Acids Res. 2008; 36:2379–2394. [PubMed: 18299282]
5. Balasubramanian S, Xu F, Olson WK. Biophys J. 2009; 96:2245–2260. [PubMed: 19289051]

6. Willis, BTM.; Pryor, AW. Thermal vibrations in crystallography. Cambridge University Press; London; New York: 1975.
7. Neidle, S. Nucleic Acid Structure and Recognition. Oxford University Press; Oxford; New York: 2002.
8. Holbrook SR, Kim SH. J Mol Biol. 1984; 173:361–388. [PubMed: 6199506]
9. Holbrook SR, Wang AH, Rich A, Kim SH. J Mol Biol. 1986; 187:429–440. [PubMed: 3701869]
10. Holbrook SR, Wang AH, Rich A, Kim SH. J Mol Biol. 1988; 199:349–357. [PubMed: 3351928]
11. Holbrook SR, Dickerson RE, Kim SH. Acta Crystallogr B. 1985; 41:255–262.
12. Busing WR, Levy HA. Acta Crystallogr. 1958; 11:450–451.
13. Sands, D. Vectors and Tensors in Crystallography. Addison-Wesley; Reading, Mass: 1982.
14. Hirshfeld F. Acta Crystallogr A. 1976; 32:239–244.
15. Glusker, JP.; Lewis, M.; Rossi, M. Crystal Structure Analysis for Chemists and Biologists. VCH; New York: 1994.
16. Press, WH. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press; Cambridge [Cambridgeshire]; New York: 1988.
17. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller A, Teller E. J Chem Phys. 1953; 21:1087–1092.
18. Lu XJ, Olson WK. Nucleic Acids Res. 2003; 31:5108–5121. [PubMed: 12930962]
19. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh H, Srinivasan AR, Schneider B. Biophys J. 1992; 63:751–759. [PubMed: 1384741]
20. Egli M, Tereshko V, Teplova M, Minasov G, Joachimiak A, Sanishvili R, Weeks CM, Miller R, Maier MA, An H, Dan Cook P, Manoharan M. Biopolymers. 1998; 48:234–252. [PubMed: 10699842]
21. Teplova M, Wilds CJ, Wawrzak Z, Tereshko V, Du Q, Carrasco N, Huang Z, Egli M. Biochimie. 2002; 84:849–858. [PubMed: 12458077]
22. Prakash TP, Manoharan M, Kawasaki AM, Fraser AS, Lesnik EA, Sioufi N, Leeds JM, Teplova M, Egli M. Biochemistry. 2002; 41:11642–11648. [PubMed: 12269806]
23. Tereshko V, Minasov G, Egli M. J Am Chem Soc. 1999; 121:470–471.
24. Tereshko V, Minasov G, Egli M. J Am Chem Soc. 1999; 121:3590–3595.
25. Sines CC, McFail-Isom L, Howerton SB, VanDerveer D, Williams LD. J Am Chem Soc. 2000; 122:11048–11056.
26. Howerton SB, Sines CC, VanDerveer D, Williams LD. Biochemistry. 2001; 40:10023–10031. [PubMed: 11513580]
27. Wilds CJ, Wawrzak Z, Krishnamurthy R, Eschenmoser A, Egli M. J Am Chem Soc. 2002; 124:13716–13721. [PubMed: 12431101]
28. Minasov G, Tereshko V, Egli M. J Mol Biol. 1999; 291:83–99. [PubMed: 10438608]
29. Ng HL, Kopka ML, Dickerson RE. Proc Natl Acad Sci, USA. 2000; 97:2035–2039. [PubMed: 10688897]
30. Dauter Z, Adams DA. Acta Crystallogr D, Biol Crystallogr. 2001; 57:990–995. [PubMed: 11418767]
31. Harper NA, Brannigan JA, Buck M, Lewis RJ, Moore MH, Schneider B. Acta Crystallogr D, Biol Crystallogr. 1998; 54:1273–1284. [PubMed: 10089504]
32. Komeda S, Moulai T, Woods KK, Chikuma M, Farrell NP, Williams LD. J Am Chem Soc. 2006; 128:16092–16103. [PubMed: 17165762]
33. Zharkov DO, Golan G, Gilboa R, Fernandes AS, Gerchman SE, Kycia JH, Rieger RA, Grollman AP, Shoham G. Embo J. 2002; 21:789–800. [PubMed: 11847126]
34. Mueller U, Muller YA, Herbst-Irmer R, Sprinzl M, Heinemann U. Acta Crystallogr D, Biol Crystallogr. 1999; 55:1405–1413. [PubMed: 10417408]
35. Dickerson RE, Bansal M, Calladine CR, Diekmann S, Hunter WN, Kennard O, von Kitzing E, Lavery R, Nelson HCM, Olson WK, Saenger W, Shakked Z, Sklenar H, Soumpasis DM, Tung CS, Wang AHJ, Zhurkin VB. J Mol Biol. 1989; 208:787–791.
36. El Hassan MA, Calladine CR. Phil Trans R Soc Lond. 1997; A 355:43–100.

37. Lu XJ, El Hassan MA, Hunter CA. *J Mol Biol.* 1997; 273:668–680. [PubMed: 9356255]
38. Lu XJ, Shakked Z, Olson WK. *J Mol Biol.* 2000; 300:819–840. [PubMed: 10891271]
39. Lu XJ, Olson WK. *J Mol Biol.* 1999; 285:1563–1575. [PubMed: 9917397]
40. Busing WR, Levy HA. *Acta Crystallogr.* 1964; 17:142–146.
41. Calladine, CR.; Drew, HR. *Understanding DNA: The Molecule and How It Works.* Academic Press; San Diego: 1997.
42. Wang AHJ, Quigley GJ, Kolpak FJ, Crawford JL, van Boom JH, van der Marel G, Rich A. *Nature.* 1979; 282:680–686. [PubMed: 514347]
43. Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM. *J Mol Biol.* 2001; 313:229–237. [PubMed: 11601858]
44. Olson, WK.; Colasanti, AV.; Lu, X-J.; Zhurkin, VB. *Wiley Encyclopedia of Chemical Biology.* John Wiley & Sons; New York: 2008.
45. Srinivasan AR, Sauers RR, Fenley MO, Boschitsch AH, Matsumoto A, Colasanti AV, Olson WK. *Biophys Revs.* 2009; 1:13–20. [PubMed: 21218180]
46. Zhurkin VB, Lysov YP, Ivanov VI. *Nucleic Acids Res.* 1979; 6:1081–1096. [PubMed: 440969]
47. Matsumoto A, Olson WK. *Biophys J.* 2002; 83:22–41. [PubMed: 12080098]
48. Poncin M, Hartmann B, Lavery R. *J Mol Biol.* 1992; 226:775–794. [PubMed: 1507226]
49. Packer MJ, Hunter CA. *J Mol Biol.* 1998; 280:407–420. [PubMed: 9665845]
50. Merritt EA, Murphy ME. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:869–873. [PubMed: 15299354]
51. Lu XJ, Olson WK. *Nature Protocols.* 2008; 3:1213–1227.
52. Merritt EA, Murphy ME. *Acta Crystallogr D, Biol Crystallogr.* 1994; 50:869–873. [PubMed: 15299354]

Figure 1. Representation of a 50% probability ellipsoid showing the observed coordinate **O** and the principal axes v_1 , v_2 , v_3 of an arbitrary atom. Illustration adapted from Figure 4.4 in Willis and Pryor.⁶

Figure 2.

Illustration of the rigid-bond factor used to measure the difference in displacement between two atoms (A, B) in the direction of the line between them. We define the *rigid-bond factor* f_{AB} as the percentage increase of the larger of $\langle u^2 \rangle_{A \rightarrow B}$ and $\langle u^2 \rangle_{B \rightarrow A}$ to the smaller:

$f_{AB} = \left(\Delta_{A-B} / \langle u^2 \rangle_{A \leftrightarrow B(\min)}^{1/2} \right) 100$, where $\Delta_{A-B} = \langle u^2 \rangle_{A \rightarrow B}^{1/2} - \langle u^2 \rangle_{B \rightarrow A}^{1/2}$ and $\langle u^2 \rangle_{A \leftrightarrow B(\min)}^{1/2}$ is the smaller of $\langle u^2 \rangle_{A \rightarrow B}$ and $\langle u^2 \rangle_{B \rightarrow A}$.

Figure 3. Schematic illustration of the two ways in which atomic thermal ellipsoids guide the simulations of DNA base deformations. Structural changes induced by random fluctuations of (a) individual atoms or (b) the base as a whole using the probability ellipsoids of the

constituent atoms as the determinant of deformation. The color-coded superimposed images represent the observed variable (red) in the high-resolution structure and the displaced variable (blue) in the simulated structure. The base fragment and the corresponding green ellipsoids in the figure are taken, with the use of Rastep,⁵⁰ from the atomic coordinates and anisotropic displacement parameters of the ring atoms of guanine 10 in the high-resolution Dickerson-Drew B-DNA dodecamer structure, 5'-d(CpGpCpGpApApTpTpCpGpCpG)-3' (NDB_ID BD0054).²⁶

Figure 4.

Average chemical bond lengths $\langle d \rangle$ in base and backbone fragments of the fluctuating Dickerson-Drew dodecamer structure (NDB_ID BD0054)²⁶ based on random displacement of the constituent atoms within the observed probability ellipsoids. (a) Schematic illustration, adapted from the textbook of Willis and Pryor,⁶ of the correlated motion of two atoms determining the limits on the length of a chemical bond. Atoms A and B are displaced respectively from their equilibrium positions by u_A and u_B . The distances d_L and d_U correspond respectively to the lower and upper limits of the bond length, obtained when the two atoms are displaced perpendicular to the observed chemical bond of length d_0 . (b) Comparison of simulated average bond lengths, found when atoms are displaced independently and at random within one or three standard deviations (1σ , 3σ) of the

observed positions, versus observed bond lengths d_0 . See Figure S3 in the Supplementary Materials for the corresponding bond lengths in six other structures.

Figure 5.

Standard deviations of complementary base-pair and consecutive base-pair-step parameters from the simulated 3σ distributions of atomic positions in representative nucleic-acid structures. (a) Translational base-pair (solid points connected by thick solid lines) and base-pair-step (open points connected by finely dotted lines) parameters, in Å, and (b) rotational base-pair and base-pair-step parameters, in deg, in the B-DNA dodecamer structure (NDB_ID BD0054, upper pair of images)²⁶ and the Nei-DNA crystal complex (NDB_ID PD0254, lower pair of images).³³ The “PED” refers to the abasic pentane-3,4-diol-5-phosphate modification used to capture the enzyme-DNA complex structure. (See Figure S1 in the Supplementary Materials for definitions of rigid-body parameters and Figure S4 for corresponding images of five other duplex structures.)

Figure 6.

Simulated *Atom3σ* distribution of the z_P values characterizing the positions of P atoms in individual base-pair steps in different types of double-helical structures. Histograms of z_{P1} and z_{P2} values collected for different types of nucleic-acid structures: A DNA (AD0027);²¹ AB DNA (NDB_ID BD0026);²⁹ B DNA (NDB_ID BD0054);²⁶ Z DNA (NDB_ID ZD0005);³⁰ drug-bound DNA (TriplatinNC + DNA) (NDB_ID DD0086);³² protein-bound DNA (Nei + DNA) (NDB_ID PD0254);³³ and RNA (NDB_ID AR0013).³⁴

Figure 7.

Correlation patterns of local conformational variables (base-pair-step parameters and differences in successive complementary base-pair parameters) along the sequence of the fluctuating Dickerson-Drew B-DNA dodecamer structure (NDB_ID BD0054).²⁶ Data based on an *Atom3σ* simulation of molecular deformation (10^3 accepted moves). Each box in the grid is divided into an 11×11 array of sub boxes, each of which represents the correlation between the designated parameters at sequential dimers in the structure. Red squares denote a correlation of -1 and blue squares a correlation of $+1$. Lighter shades of red and blue represent lesser degrees of correlation and white spaces indicate that there is no correlation. Rigid-body parameters labeled as in Figure S1 in the Supplementary Materials. The symbol Δ denotes the difference in parameters at neighboring base pairs, *e.g.*, $\kappa = \kappa_{i+1} - \kappa_i$. Plot generated with software developed by Fei Xu. See Figure S5 in the Supplementary Materials for corresponding images of other duplex structures.

(a) B DNA

(b) Nei + DNA

Figure 8.

Atomic-level fluctuations and simulated asynchronous motions in representative nucleic-acid structures. (Top) Color-coded schematics of the anisotropic thermal factors of (a) the Dickerson-Drew B-DNA dodecamer (NDB_ID BD0054)²⁶ and (b) the DNA crystallized in complex with *Escherichia coli* endonuclease VIII (NDB_ID PD0254).³³ (See Figure S6 in the Supplementary Materials for corresponding images of other duplex structures.) Block images of molecules (top left) generated with 3DNA^{18,51} and networks of atomic thermal ellipsoids (top right) with Rastep.⁵² The red, blue, green, yellow, and turquoise blocks in the block images represent adenine, thymine, guanine, cytidine, and uracil bases, respectively.

Strand I is colored red and Strand II yellow. The minor-groove edges of the blocks are shaded black. Small-molecule ligands are represented by ball-and-stick models. In the networks of atomic thermal ellipsoids the thermal ellipsoids of the atoms are shown at the 50% probability level (*i.e.*, there is a 50% chance that an atom is within the ellipsoid at any instant). The colors of the ellipsoids range from blue to red with blue representing small atomic fluctuations and red representing large fluctuations. The same color scale ($B_{\min} = 9$, $B_{\max} = 25$) is used in both structures so that the relative fluctuations may be compared. (Bottom) Rigid-bond-factor matrices illustrating asynchronous movements within these structures. Each box in the grids is divided into an array of sub boxes, each of which represents the asynchronous movements of ring atoms in each base of the designated structure. White spaces indicate synchronized motion and shades of blue represent higher degrees of asynchronous movement. Rigid-bond factors based on the mean-square displacement of individual atoms along each chemical bond in *Atom3σ* simulations. The ring atoms of each base are presented from left to right in the following order: (purines) N9, C8, N7, C5, C6, N1, C2, N3, C4; (pyrimidines) N1, C2, N3, C4, C5, C6. Color-coded plots generated with software developed by Fei Xu.

Table I

Average rigid-bond factors of selected parts of representative, high-resolution nucleic-acid structures*

Structure type	NDB_ID	Reference	Bonded ring atoms	Bonded atoms	Non-bonded atoms
A DNA	AD0007	20	3.9	5.6	13.6
	AD0027	21	1.8	2.3	13.2
	AD0028	22	6.1	6.5	12.0
B DNA	BD0007	23	1.8	2.5	15.8
	BD0012	24	2.2	2.7	17.7
	BD0030	20	8.3	9.1	18.4
	BD0041 [†]	25	13.2	14.3	17.1
	BD0054	26	1.5	1.8	18.7
AB DNA	BD0060	27	1.2	2.0	17.7
	BD0018	28	2.4	2.4	23.2
	BD0026	29	1.4	1.7	13.0
	ZD0005	30	3.3	3.5	11.2
drug-DNA	ZDF060 [‡]	31	11.1	10.9	32.0
	DD0086	32	3.4	3.7	22.3
protein-DNA	PD0254	33	2.8	3.1	25.0
A RNA	AR0013	34	8.3	6.7	26.3

* Results are broken down into three columns containing the averages of the rigid-bond factors $\langle AB \rangle$ amongst (i) all of the covalently bonded ring atoms in each base, (ii) all of the covalently bonded atoms in the entire structure, and (iii) all of the non-bonded pairs of atoms in the entire structure. Data based on anisotropic displacement parameters that represent deformations in 3σ .

[†] Similar bonded and non-bonded values suggests that anisotropic displacement factors may not specify thermal motion.

[‡] High values are suggestive of disorder in the crystal structure.

Table IIComparative atomic displacements based on different simulation methods.[†]

U	Q	Λ
<i>Observation</i>		
$\begin{bmatrix} 0.2895 & 0.0475 & 0.0107 \\ 0.0475 & 0.1790 & 0.0197 \\ 0.0107 & 0.0197 & 0.1580 \end{bmatrix}$	$\begin{bmatrix} 0.9278 & -0.3421 & 0.1491 \\ 0.3559 & 0.6911 & -0.6291 \\ 0.1122 & 0.6367 & 0.7629 \end{bmatrix}$	$\begin{bmatrix} 0.3090 & 0 & 0 \\ 0 & 0.1736 & 0 \\ 0 & 0 & 0.1438 \end{bmatrix}$
<i>Atom3σ</i>		
$\begin{bmatrix} 0.2922 & 0.0500 & 0.0089 \\ 0.0500 & 0.1868 & 0.0189 \\ 0.0089 & 0.0189 & 0.1618 \end{bmatrix}$	$\begin{bmatrix} 0.9204 & -0.3555 & 0.1628 \\ 0.3777 & 0.7010 & -0.6049 \\ 0.1009 & 0.6182 & 0.7795 \end{bmatrix}$	$\begin{bmatrix} 0.3137 & 0 & 0 \\ 0 & 0.1781 & 0 \\ 0 & 0 & 0.1490 \end{bmatrix}$
<i>Atom1σ</i>		
$\begin{bmatrix} 0.0325 & 0.0056 & 0.0010 \\ 0.0056 & 0.0208 & 0.0021 \\ 0.0010 & 0.0021 & 0.0180 \end{bmatrix}$	$\begin{bmatrix} 0.9195 & -0.3572 & 0.1639 \\ 0.3797 & 0.6998 & -0.6051 \\ 0.1015 & 0.6187 & 0.7791 \end{bmatrix}$	$\begin{bmatrix} 0.0349 & 0 & 0 \\ 0 & 0.0208 & 0 \\ 0 & 0 & 0.0180 \end{bmatrix}$
<i>Base3σ</i>		
$\begin{bmatrix} 0.1064 & 0.0337 & 0.0125 \\ 0.0337 & 0.0557 & 0.0221 \\ 0.0125 & 0.0221 & 0.1216 \end{bmatrix}$	$\begin{bmatrix} 0.5852 & -0.7044 & 0.4016 \\ 0.3983 & -0.1817 & -0.8991 \\ 0.7063 & 0.6861 & 0.1742 \end{bmatrix}$	$\begin{bmatrix} 0.1444 & 0 & 0 \\ 0 & 0.1029 & 0 \\ 0 & 0 & 0.0364 \end{bmatrix}$
<i>Base1σ</i>		
$\begin{bmatrix} 0.0122 & 0.0040 & 0.0016 \\ 0.0040 & 0.0064 & 0.0029 \\ 0.0016 & 0.0029 & 0.0155 \end{bmatrix}$	$\begin{bmatrix} 0.4869 & -0.7731 & 0.4065 \\ 0.3706 & -0.2386 & -0.8976 \\ 0.7909 & 0.5877 & 0.1704 \end{bmatrix}$	$\begin{bmatrix} 0.0178 & 0 & 0 \\ 0 & 0.0122 & 0 \\ 0 & 0 & 0.0040 \end{bmatrix}$

[†]Data simulated at different levels for the N9 atom of guanine 10 in the Dickerson-Drew dodecamer (NDB_ID BD0054)²⁶ are compared with (i) the mean-square displacement, or **U**, matrix, (ii) the eigenvectors of **U** stored in **Q**, and (iii) the eigenvalues of **U** stored in **Λ**. The number of accepted moves in the *Base 3σ* and *Base 1σ* simulations was 3,575 and 3,532, respectively, out of 10,000. The displacement volume *V* is given by $(\lambda_1\lambda_2\lambda_3)^{1/2}$, where the λ_i ($i = 1-3$) are the eigenvalues, and $\mathbf{U} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$.