The structure of DNA in the nucleosome core

Timothy J. Richmond & Curt A. Davey

ETH Zürich, Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich
The authors contributed equally to this work

The 1.9-Å-resolution crystal structure of the nucleosome core particle containing 147 DNA base pairs reveals the conformation of nucleosomal DNA with unprecedented accuracy. The DNA structure is remarkably different from that in oligonucleotides and non-histone protein–DNA complexes. The DNA base-pair-step geometry has, overall, twice the curvature necessary to accommodate the DNA superhelical path in the nucleosome. DNA segments bent into the minor groove are either kinked or alternately shifted. The unusual DNA conformational parameters induced by the binding of histone protein have implications for sequence-dependent protein recognition and nucleosome positioning and mobility. Comparison of the 147-base-pair structure with two 146-base-pair structures reveals alterations in DNA twist that are evidently common in bulk chromatin, and which are of probable importance for chromatin fibre formation and chromatin remodelling.

DNA in eukaryotic cells is packaged repetitively into nucleosomes by means of extensive association with histone proteins[1–3]. The hierarchical chromatin structure formed is the genomic substrate relevant to the vital processes of DNA replication, recombination, transcription, repair and chromosome segregation, and to the pathological progression of cancer and viral disease. Although nucleosomal organization of DNA is essentially ubiquitous throughout genomes and generally repressive to gene expression, it also contributes to gene transcription in a gene-specific manner, suggesting that nucleosome positioning in gene promoter regions is important for genuine gene regulation *in vivo*[4–6]. The question therefore arises of how chromatin structure, in which DNA is normally highly compacted, permits site-specific access to regulatory factors and more extensive exposure to the transcription apparatus.

The answer is likely to require a knowledge of DNA conformation in the nucleosome core. The core comprises 147 base pairs (bp) of DNA and the histone octamer; compared with the nucleosome, it lacks only 10–90 bp of linker DNA envisaged to be naked or bound to histone H1. The histone-fold domains of the octamer organize the central 129 of 147 bp in 1.59 left-handed superhelical turns with a diameter only fourfold that of the double helix. The relatively straight 9-bp terminal segments contribute little to the curvature of the complete 1.67-turn superhelix[7]. So far, the site-specific regulatory factors that have been discovered bind the linker or terminal regions of the intact nucleosome. The lack of binding to the central region of the superhelix might simply be a consequence of bending the double helix or, additionally, of unusual DNA conformations induced by histone binding. At least one protein, HIV-1 integrase, does prefer DNA bent around the nucleosome in contrast to naked DNA[8].

The initiation of DNA-dependent nuclear processes in the context of chromatin implies that nucleosome position is biased by the DNA sequence to facilitate access by initiation factors. Numerous examples of positioned nucleosomes in gene promoter regions have been described both *in vivo* and *in vitro*[9–11]. Preferential positioning could place factor-binding sequences in nucleosome linker or terminal region DNA. Furthermore, nucleosomes are intrinsically mobile and yield access to their DNA *in vitro*, allowing even RNA polymerase to transcribe nucleosomal DNA without causing dissociation of the histone octamer[12–14]. *In vivo*, energy-dependent chromatin remodelling factors, targeted by gene regulatory proteins and acting directly on the nucleosome core, augment nucleosome mobility[15]. Their mechanism of action most probably derives from the innate ability of nucleosomes to 'slide' along DNA without releasing it[3].

An accurate, atomic-level description of DNA conformation in the nucleosome core and comparison with naked DNA is essential to an understanding of chromatin properties such as those resulting from nucleosome position and mobility. Previously, the probable sequence-dependent features of nucleosomal DNA were proposed by extrapolation from high-resolution DNA oligonucleotide structures[16]. Here, a direct analysis of nucleosome DNA is made, permitted by the high quality of the DNA structure in the 1.9-Å-resolution X-ray structure of the nucleosome core particle (NCP147)[17,18], comparable to high-resolution oligonucleotide structures. The persistently tight curvature of nucleosome core DNA causes it to take on conformations not obviously present in oligonucleotides but about which there has been speculation for many years[19]. Furthermore, comparison of the DNA twist values measured for the 147-bp and 146-bp structures with the DNA sequence periodicity found in bulk chromatin indicates that the DNA stretching observed in the nucleosome core is commonplace *in vivo*.

Excess DNA curvature

The ideal DNA superhelix that best fits the NCP147 DNA was constructed from uniformly distributed base pairs by using B-form DNA geometry. It has a radius of 41.9 Å and a pitch of 25.9 Å (Fig. 1a). Each base-pair step comprising two adjacent base pairs contributes 4.53° to the curvature of the ideal superhelix. However, the NCP147 superhelix is not bent uniformly owing to (1) the anisotropic flexibility of DNA, (2) local structural features intrinsic to the DNA sequence, and (3) irregularities dictated by the underlying histone octamer. The NCP147 DNA is generally B-form as judged by the criterion of the phosphate '*Z*-coordinate' with a value of -0.18 ± 0.69 Å (mean ± s.d.; range -2.32 to 1.45 Å). On the basis of oligonucleotide structures, values falling in the range from -1 to 0 are indicative of the B-form, whereas values above 2 signify the A-form[20]. Other standard conformational parameters also indicate that the DNA is predominantly B-form (see Supplementary Table 1).

Remarkably, the NCP147 DNA has double the base-pair-step curvature necessary to produce the DNA superhelix path. The ideal superhelix fit to the NCP147 DNA yields 1.67 superhelical turns for 133.6 bp (1.84 superhelical turns for 147 bp), bending the DNA

through 600.0°, whereas the actual curvature of NCP147 DNA is 1,333.3° (both scalar curvature values were calculated with the program Curves[21]). The use of only the central 129 bp to remove the effect of straightening of the superhelical path at the termini yields values of 579.3° and 1248.6° (9.75 ± 5.13° per base-pair step) for the ideal and actual curvatures, respectively, and thus the terminal regions do not account for the discrepancy. The excess is not adequately explained by the zigzag path of the NCP147 superhelix, as the overall root-mean-square deviation from the ideal path is only 1.42 Å (Fig. 1a). We find that the greater part of the excess curvature is a consequence of alternation of DNA form not revealed by criteria based on oligonucleotide structures. Determining the component of curvature directed towards superhelix formation exposes the origin of the form changes.

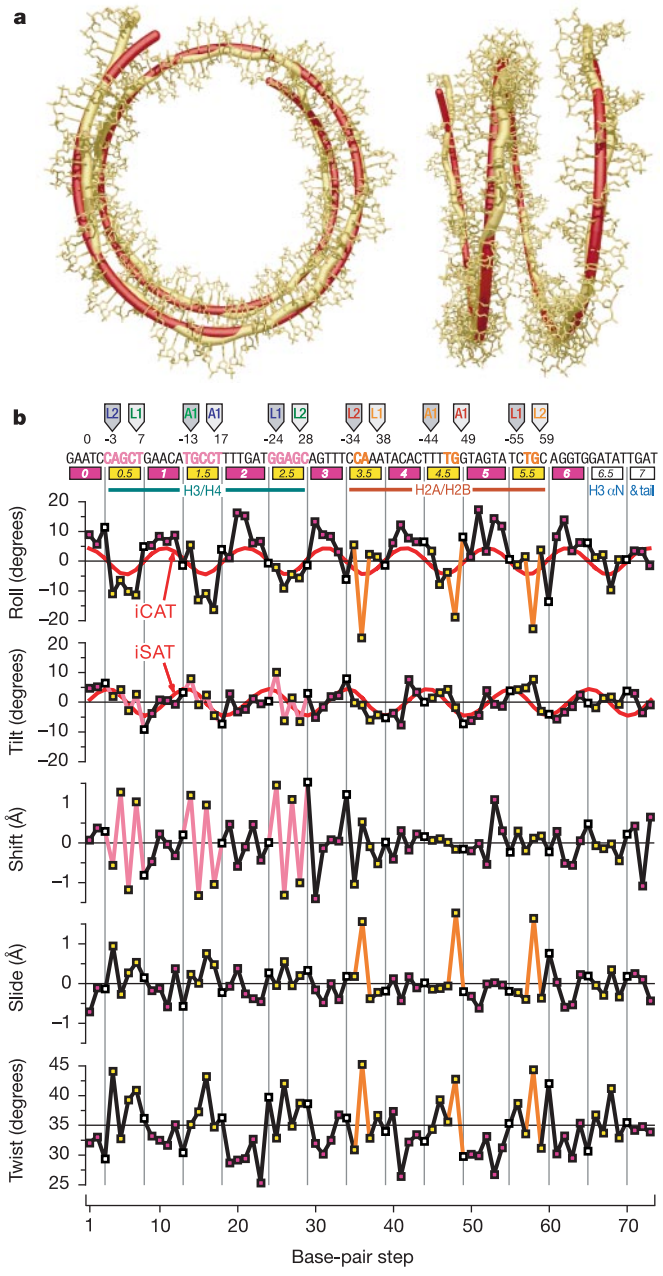

Figure 1 Superhelical path and base-pair-step parameters for NCP147 DNA. **a**, Structural alignment of the NCP147 (gold) and best-fit, ideal superhelix (red) paths. The NCP147 DNA structure is superimposed (gold). The left view is down the superhelix axis and the right view is rotated 90° around the pseudo-two-fold axis (vertical). **b**, Roll, tilt, shift, slide and twist base-pair-step parameters. The base-pair-step values plotted are the means for the two halves of the symmetrical sequence (shown above, with the dyad position labelled '0'). The iCAT and iSAT curves (red) show the roll and tilt contributions to the ideal superhelix curvature. The minor-groove blocks show base-pair-step shift alternation (pink in shift and tilt) or kinks (orange in roll, slide and twist). The primary bound-phosphate groups are indicated above the base sequence by pointers (numbered as the 5′ phosphate of the adjacent base) showing the strand direction (dark, 3′ → 5′; light, 5′ → 3′) and the interacting histone motif (L1, loop 1; L2, loop 2; A1, α1).

The potential contributions to curvature from the roll and tilt (see Supplementary Fig. 1) of each base-pair step vary as sinusoidal functions of the accumulated twist angle (Θ) along the length of the double helix. To evaluate the actual contributions, the roll and tilt values for each base-pair step are multiplied by the corresponding cos Θ (CAT) and sin Θ (SAT), respectively. A convenient origin for the summation of Θ is taken as the central base pair located on the pseudo-two-fold symmetry axis of the particle (base pair 0 or dyad position). For the ideal superhelix, the CAT and SAT values multiplied by 4.53° yield the roll and tilt contributions to superhelix curvature at each base-pair step (iCAT and iSAT; Fig. 1b). The actual roll and tilt values for NCP147 DNA are moderately well correlated with CAT ($R = 0.72$, $P < 0.0001$) and SAT ($R = 0.54$, $P < 0.0001$), respectively. The sum of roll and tilt multiplied by CAT and SAT, respectively, over all base-pair steps yields a curvature free of the effect of path deviations between the ideal and actual superhelices. The resulting value is 897.0°, or 1.5-fold that required to accommodate the full NCP147 superhelix.

Unlike the computer-generated ideal superhelix, for which curvature derives equally from roll and tilt (iCAT and iSAT; Fig. 1b), curvature for oligonucleotide and NCP147 DNA stems predominately from the roll parameter. For DNA oligonucleotide structures, roll is favoured over tilt by 2:1 (ref. 22), presumably because it requires less base-pair unstacking for any particular bend angle. In the nucleosome core, this ratio is virtually identical at 1.9:1, but when the components that contribute to the superhelix only are considered, the ratio is 3.0:1. Curvature into the superhelix is therefore accumulated every approximately five base pairs, where either the major or minor groove faces the histone octamer. Nevertheless, the NCP147 tilt component contributes 77% of that for the ideal superhelix (Fig. 1b; compare tilt and iSAT plots). In contrast, the roll component is 222% of the ideal contribution (Fig. 1b; compare roll and iCAT plots).

Analyses of crystal structures of DNA oligonucleotides and protein–DNA complexes normally equate DNA bending with DNA curvature[23] and rely on base-pair-step parameters to explain bending behaviour; for example, by using roll and tilt[22,24], roll and twist[16,25,26] or roll, tilt, twist, slide, shift and rise[24,27]. This type of approach is inadequate for nucleosome core DNA, given the substantial excess curvature that contributes to superhelix formation. Further analysis indicates that the conformation of nucleosome core DNA is not typical of oligonucleotide DNA and has not previously been described in protein–DNA complexes. Principal-component analysis of DNA structural parameters of both nucleosome core and oligonucleotide DNA show that the two most significant base-pair-step parameter correlations are roll–slide–twist and tilt–shift (see Supplementary Table 2). However, the coefficients for the base-pair-step components are significantly larger for NCP147 DNA (mean magnitudes are 0.88 versus 0.57 for roll–slide–twist and 0.86 versus 0.69 for tilt–shift), indicating that the NCP147 DNA is limited to a smaller regime of conformational space than is occupied by oligonucleotide DNA. When the backbone angles are also included in the analysis, the ε, ζ and β angles are found tightly linked to the roll–slide–twist component. Stronger parameter coupling occurs for blocks of base pairs with the minor groove as opposed to the major groove facing the histone octamer (mean magnitudes of 0.69 versus 0.27 for roll–slide–twist with backbone angles included), indicating differences in modes of

DNA bending or in conformational variability for the two orientations.

Major-groove versus minor-groove curvature

The DNA curvature of the central 129 bp of NCP147 favours the major groove over the minor groove by 1.3:1 per base-pair step in contrast to the approximate 2:1 preference observed for oligonucleotide DNA (evaluated from data in ref. 20). Minor-groove bending is facilitated in NCP147 by the insertion of an arginine side chain into the minor groove at each of the 14 sites at which it faces the histone octamer. The minor groove narrows to a width of 3.0 ± 0.55 Å at these sites[18]. An insightful analysis is made by dividing the DNA along its length into blocks of continuous positive and negative CAT values, defining respectively regions where the major-groove and minor-groove curvatures contribute to superhelix formation (Fig. 1b). A single, junction base-pair step joins the major-groove and minor-groove blocks and cannot contribute significantly to superhelix curvature by base-pair-step roll because these steps have the minimum magnitudes for CAT (−0.10 to 0.26). The major-groove blocks display 'smooth' bending with all roll angles positive (8.47 ± 5.18°) and with systematic underwinding of the DNA indicated by the reduced twist angles and negative slide displacements (Figs 1b and 2a). The minor-groove blocks exhibit 'smooth' bending only over the H3–H4 tetramer, with essentially all base-pair steps having negative roll angles (−7.86 ± 6.02°). In contrast, minor-groove blocks are kinked over the H2A–H2B dimers (−5.13 ± 11.01°; Fig. 1b). This difference in minor-groove-block bending modes associated with H3–H4 and H2A–H2B is most probably a consequence of specific DNA sequences and not underlying differences in the histone binding or degree of curvature (mean values over all base-pair steps: H3–H4, 9.8°; H2A–H2B, 9.9°). Smooth bending in minor-groove blocks is associated with large alternation of shift values roughly between −1 and +1 Å for four base-pair steps, which relieves steric interference between the base edges (Figs 1b and 2b). This mode of bending, which is the prime contributor to the observed shift–tilt coupling, depends most consistently on the GC base-pair step, and on GG = CC or AG = CT base-pair steps, with one and two in each block, respectively. GC steps have been noted previously for their high shift propensity[20]. The GG = CC and AG = CT steps are, of all steps, the only two that can form cross-chain hydrogen bonds in the minor groove as observed for oligonucleotides[28] and for NCP147. The occurrence of two GG = CC/AG = CT base-pair steps in the minor-groove blocks bound to the H3–H4 tetramer, but not the H2A–H2B dimer, is a probable determinant of smooth versus kinked bending.

Kinking in minor-groove blocks always occurs at a single CA = TG base-pair step that has a roll angle in the range −18° to −27°, roughly the total curvature into the superhelix for the entire block (Figs 1b and 2c). These kinked steps have a concomitant large slide value of over 1.5 Å and are also overtwisted to values larger than 40°. They have a conformation that is extreme in roll–slide–twist coupling. Ten of 12 minor-groove blocks have one CA = TG step and two have none. The six CA = TG steps that are kinked have CAT values between −0.6 and −1, whereas the four non-kinked CA = TG steps have values more positive than −0.5. As steps with more negative CAT values have more potential to contribute to superhelix curvature, they are conversely under more stress to kink. The occurrence of CA = TG steps in kinks is consistent with their properties as gleaned from oligonucleotide structures[20]: (1) TA and CA = TG steps are the most flexible as judged from the standard deviation of roll angles, (2) the mean slide for CA = TG is 1.18 Å—the only base-pair step with a value greater than 1 Å (the TA mean value is −0.80 Å), and (3) with a mean twist angle of 37.4°, CA = TG is one of the most overwound base-pair steps. TA and CA = TG are by far the most flexible of the 10 base-pair steps with regard to roll, slide and twist, but the propensity of TA for negative slide makes it much less likely to kink in the manner observed in NCP147 DNA. Clash of the purines across the minor groove in the kinks observed is avoided because of large slide values. Furthermore, TA steps have the largest positive mean roll angle (11.8° versus the next largest of 6.3° for GG = CC), which is consistent with their appearance only in major-groove blocks of NCP147. The importance of base-pair-step flexibility in nucleosome positioning and stability has been suggested previously[29,30]. A survey of protein–DNA complexes shows that only seven base-pair steps are rolled into the minor groove to a degree comparable to that of NCP147 DNA, but none of these is a CA = TG step[31]. Comparison of the mean roll angle for each type of base-pair step within NCP147 DNA and from oligonucleotide structures[20] shows that they are only weakly correlated with each other (see Supplementary Information).

DNA form and superhelix formation

Strikingly for NCP147 DNA, the conformational parameter tip oscillates regularly along the length of the DNA (Fig. 3a). The transitions between negative and positive tip angles represent an alternation of the DNA form that manifests itself in the major and minor blocks as excess roll. By definition, the excess roll of a series of

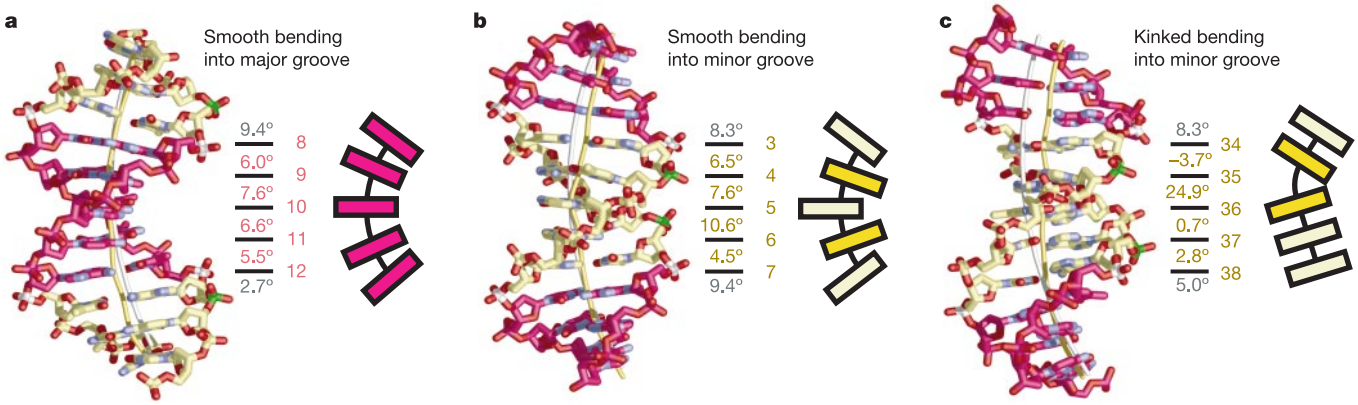

Figure 2 DNA bending in the NCP147 DNA. Structures (left) and schematic representations (right) stress uniformity of curvature in the major-groove blocks (red) (**a**), and alternating shift values (**b**) and localization of curvature in kinks in minor-groove blocks (**c**) (yellow for one representative double-helical turn). Also indicated are the primary bound-phosphate groups (green), the block-junction phosphate groups (white) and the DNA axes for the NCP147 (gold) and ideal (white) superhelices. The contributions of base-pair-step curvature to superhelix bending are listed with base-pair numbers (centre).

base-pair steps is the sum of the change in tip angles over the corresponding base pairs (Fig. 3b). The source of excess roll in the NCP147 superhelix is localized primarily to block junctions and the base pairs immediately adjacent on either side, where the tip is significantly non-zero. The reason for the increased tip-angle magnitudes at these sites is that they accommodate the smaller radius of curvature on the inside surface of the superhelix compared with the outside, about 32 Å compared with 52 Å. In principle, this radius-of-curvature difference should require shorter phosphate–phosphate distances (PP) where the DNA backbone faces the histones than where it faces away. In fact, it is not the absolute distance that is crucial but the component of the PP distance (PP_a) lying parallel to the path of the superhelix. As two adjacent base-pair steps were found compensatory for this distance, PP_a is obtained from three base pairs ($i-1$ to $i+1$), and the direction of the superhelix path is defined by four points on the double-helix axis ($i-1$ to $i+2$). Importantly, it is PP_a that is shortened on the inside and lengthened on the outside of the DNA superhelix by the oscillating tip angles. The difference calculated at each base pair between the values of PP_a for the two phosphodiester chains (ΔPP_a) is strongly correlated with tip angle ($R = 0.91$, $P < 0.0001$; Fig. 3a). The values of ΔPP_a compared with SAT, a direct measure of inside and outside, are somewhat less well correlated ($R = 0.83$, $P < 0.0001$) owing to DNA-sequence-specific effects such as kinking in the minor-groove blocks.

Phosphate–phosphate distances, as opposed to their superhelical path component, are not correlated with SAT, curvature or roll angles, and are only weakly correlated with tip. This is the expected behaviour because the phosphodiester backbone is essentially incompressible, as described previously for oligonucleotide structures[32], having in NCP147 a mean phosphate–phosphate distance of 6.68 ± 0.23 Å between nearest neighbours along the same chain. However, a notable variation in the B-form backbone is the anticorrelation of the ε and ζ backbone angles giving rise to the B_I and B_{II} forms[33]. In NCP147, the population of base-pair steps intermediate between B_I and B_{II} (that is, $B_{I/II}$) is larger than for oligonucleotides, but the distribution is not significantly correlated with other form or backbone parameters. The differences in mean phosphate–phosphate distances for B_I, $B_{I/II}$ and B_{II} forms separately, even taking into consideration the location in major-groove and minor-groove blocks or junctions, are less than 1 s.d. of the mean distance overall. Nevertheless, in minor-groove blocks, a B_{II} to B_I transition occurs at kinks, and B_{II} and B_I alternation occurs at sites of alternating shift. In the latter case, the B_{II} base pair is displaced into the minor groove and has up to −40° of propeller twist. This unusual conformation for GG and AG base-pair steps is stabilized by cross-strand bifurcated hydrogen bonds in the minor groove between guanosine N2 and pyrimidine O2 atoms. Moreover, the variability observed between repeating binding sites seems to stem from a dependence on DNA sequence, not only for the choice of

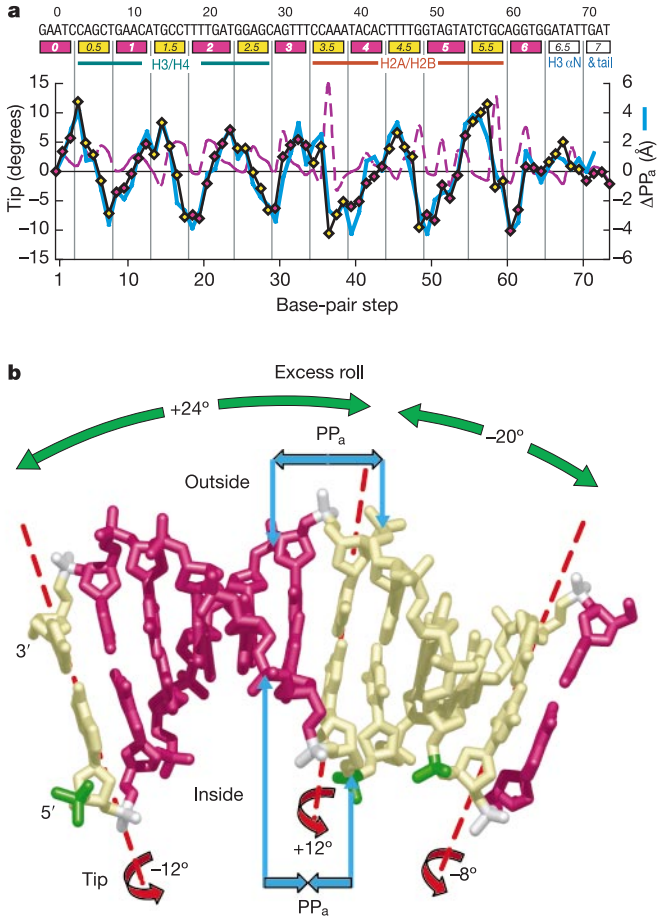

Figure 3 Oscillation of the base-pair-tip parameter for NCP147 DNA. **a**, Correlation of base-pair tip and ΔPP_a, plotted as in Fig. 1b. Excess base-pair-step roll (purple) into the superhelix is a consequence of oscillating base-pair-tip angles (black). ΔPP_a (blue) is the difference between PP_a values for the two phosphodiester chains at a single base pair. **b**, Coupling between PP_a, tip and excess roll. PP_a is the component of the phosphate–phosphate distance that lies in the plane containing the local superhelical path. Rotation of a base pair along its tip axis (red) in the direction indicated decreases PP_a on the inside and increases it on the outside of the superhelix. Excess roll for a base-pair step is the difference between the tip angles for the two contributing base pairs. The DNA blocks shown are from SHL0.0 and SHL0.5.

smooth or kinked bending in minor-groove blocks but also for the precise details of the conformation along the entire DNA.

Histone constraints on DNA conformation

Conformational differences between nucleosome core and oligonucleotide DNA are probably important for the recognition, or lack of it, of nucleosome DNA by nuclear factors, and for the folding of nucleosome arrays into the chromatin fibre. Furthermore, the dependence of nucleosome position and stability on base sequence, as demonstrated convincingly by experiments *in vitro* (for example ref. 11), derives from the energetics of sequence-dependent histone–DNA interactions. Parameterization of an energy function, analogous to that for protein–DNA complexes[34], capable of accurate prediction of sequence-dependent behaviour will require many high-resolution structures of the nucleosome core particle containing different DNA sequences. Currently, comparisons between histones H2A, H2B, H3 and H4 for the three primary types of DNA-binding motif in the nucleosome core give a preliminary indication of the significance of histone restraints on DNA conformational flexibility.

The principal DNA-binding sites in the nucleosome core particle are the eight L1L2 loop and four α1α1 helix structures within the histone-fold domains[7]. Although the histone–DNA interactions at each of these sites are complex (a complete tabulation is given in ref. 17), one phosphate group from each DNA strand shows at least one conserved interaction with protein over the four types of core histone (Figs 1b and 4). These primary bound-phosphate groups are 5′ to the last base pair for both DNA strands in each minor-groove block, with one exception at SHL4.5 (SHL: superhelix location[7]), resulting from a dissimilar packing arrangement of the α1 and α2 helices for H2B in comparison with the other histones[7]. Separate structural alignments, calculated with the histone main-chain segments only, for the L1, L2 and A1 binding motifs (the components of L1L2 loop and α1α1 helix sites) and their associated phosphate groups show the precision with which the selected phosphate groups are localized by binding to histone (see Supplementary Table 3).

For the L1 and L2 binding motifs, the root-mean-square deviations for the primary bound-phosphate group (1.27 and 1.39 Å) and its 3′-adjacent neighbour (1.42 and 1.48 Å) are strikingly low relative to the mean phosphate–phosphate distance of 6.68 Å (see Supplementary Table 3). The values for the phosphate groups are comparable to those for the segments of protein main-chain (0.75 and 0.46 Å) used for the alignments if the phosphates associated with the structurally deviant H2B are not included. Two phosphate groups separated by 21 bases along each DNA strand are therefore well localized on both the H3–H4 and H2A–H2B histone-fold pairs (for example H4:L1 to H4:L2) through histone interaction. Correspondingly, the DNA segments bridging between the halves of the H3–H4 tetramer and H3–H4 tetramer to H2A–H2B dimer are 10 bases in length. Evidently, the histone proteins impose substantial conformational restraints on the DNA, which responds in a sequence-dependent manner, such as the smooth and kinked bending modes observed for the minor-groove blocks. The phosphate deviations for the A1 motif are larger than for L1 and L2 motifs, which might reflect their location between phosphate groups at 10.5 bases. This non-optimal arrangement might decrease sequence-dependent effects on stability by accommodating either 10 or 11 bases between A1 motifs and the adjacent L1 and L2 motifs. A possible example of this multiplicity occurs for H2B A1 at the SHL4.5/5.0 block junction (Fig. 1b). Indeed, the location of A1 relative to L1 and L2 motifs could promote DNA stretching.

DNA stretching

The double-helical twist values for NCP147 DNA range from 23.7° to 47.9° and have a mean value of 34.52 ± 4.75°, or equivalently 10.43 bp per turn, which corresponds to the laboratory reference frame[35]. Values from experiments that probe DNA by enzymatic digestion or hydroxyl radical cleavage correspond to the local reference frame and are calculated from the laboratory frame by removing the geometrical contribution from the superhelix pitch. Assuming a uniform pitch angle α, then Twist(local) = Twist(laboratory) $-(2\pi \sin\alpha)/N$, where N is the number of base-pair steps in one superhelical turn. Using the values from the ideal superhelix of $-5.62°$ and 78.90 steps for α and N, respectively, NCP147 overall has a local frame twist of 10.30 bp per turn. Remarkably, two 146-bp nucleosome core particle structures containing differing DNA sequences have values of 10.23 and 10.15 bp per turn[17]. These reduced values are a consequence of DNA stretching by 1–2 bp to satisfy the crystal packing constraints of

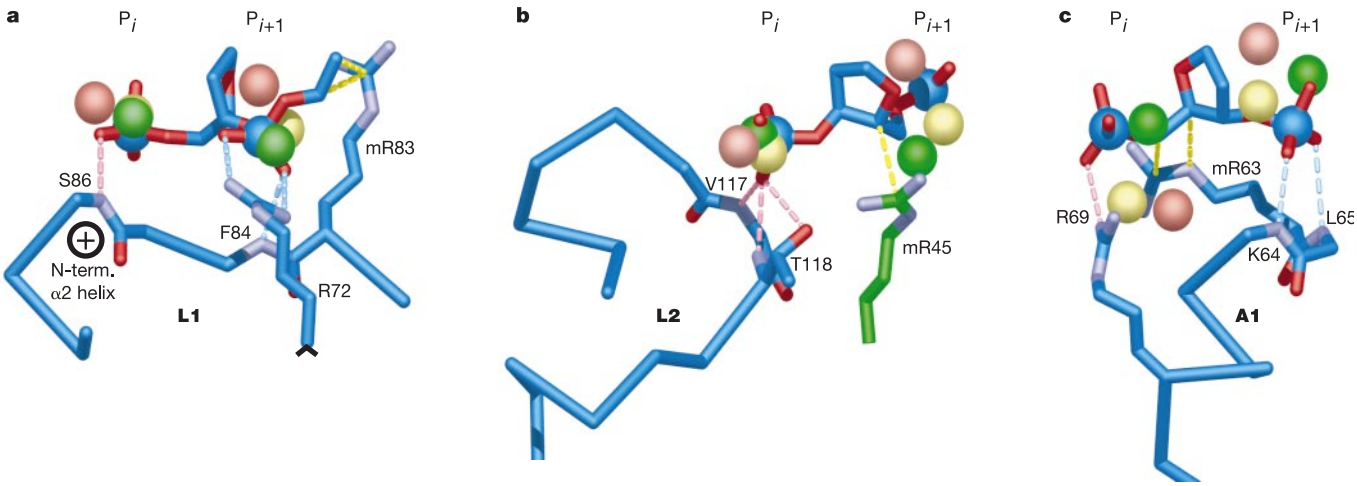

Figure 4 Structural alignments of the histone-fold DNA-binding motifs and bound-phosphate groups. The motifs are L1 (**a**), L2 (**b**) and A1 (**c**). For each, the main-chain C_α atoms of H4, H2A and H2B were aligned with those for H3. The resulting primary bound-phosphate group (P_i) and 3′-adjacent phosphate group (P_{i+1}) positions are shown (H3, blue; H4, green; H2A, yellow; H2B, red) with respect to the H3 main-chain and interacting groups. Completely conserved (pink) and partly conserved (cyan) hydrogen bonds and interactions (yellow) with the minor-groove arginine side chain (mR) are shown.

the DNA termini[7]. The actual 12-bp region stretched is in a different location in the two 146-bp structures, and both regions are distant from the DNA termini. Evidently, stretching at the DNA termini introduces a twist defect in the double helix that then diffuses to predominantly one region in each particle, depending on the DNA sequence. The distorted region is centred on either an H3–H4 or an H2A–H2B α1α1 DNA-binding site and is bound at each end by the adjacent L1L2 sites. These two 146-bp structures therefore contain DNA regions that represent trapped intermediates relevant to a twist-defect-diffusion mechanism of nucleosome mobility[36,37] and twist propagation by chromatin-remodelling enzymes.

DNA isolated from nucleosome core particles prepared from bulk-source endogenous chromatin have a sequence periodicity of 10.17 bp per turn[35], which is significantly closer to the local frame values for the 146-bp versus 147-bp nucleosome core particles. This similarity suggests that nucleosomes in the cell nucleus have evolved to contain DNA stretched on average by 1–2 bp. The difference between 10.17 bp per turn for nucleosomes and 10.5 bp per turn, accepted for the average twist of naked DNA[35], accounts for the discrepancy in the observed and effective values (−1.67 versus −1.2 (ref. 38)) of nucleosome supercoiling. The overtwisting of the DNA double helix on the nucleosome, due in part to DNA stretching, resolves the 'linking number paradox'[39] observed for nucleosomes assembled on closed circular DNA. Unusual twisting or supercoiling of the linker DNA between nucleosome cores in these arrays is not required. Formation of a compact nucleosome higher-order structure could be facilitated by DNA stretching because it would provide a buffer against incompatible DNA linker lengths. Each linker DNA segment running between two nucleosome cores could on average adjust in length by up to four base pairs by stretching or compressing the DNA bound to the neighbouring H2A–H2B and H3–H4 histone-fold domains. Not only would this provide for linker length adjustment, it would also alleviate twist angle restrictions of at least 140°. □

Methods

Crystal preparation, data collection, model refinement and figure preparation were as described previously[18]. The ideal DNA superhelix was calculated by using B-form base-pair geometry[40] and fitted by least squares to the NCP147 DNA (T.J.R., unpublished observations). The twist offset for the central base pair of NCP147 is ±4.99° (negative for strand with A central). It is presumably non-zero because of the asymmetry of the AT base pair lying on the molecular dyad of the particle. DNA curvature, base-pair and base-pair-step parameters, and backbone geometry were calculated with the program Curves[21] (see Supplementary Methods). Excel (Microsoft) and SigmaPlot (SPSS) were used for the analysis of DNA geometry. Correlation coefficients were calculated by using variance weighting. Principal components were calculated with IMSL routines (T.J.R., unpublished observations).

Received 16 January; accepted 12 March 2003; doi:10.1038/nature01595.

1. van Holde, K. E. in *Chromatin* (ed. Rich, A.) (Springer, New York, 1988).
2. Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98,** 285–294 (1999).
3. Elgin, S. C. R. & Workman, J. L. (eds) *Chromatin Structure and Gene Expression* (Oxford Univ. Press, Oxford, 2000).
4. Simpson, R. T. Nucleosome positioning: Occurrence, mechanisms, and functional consequences. *Prog. Nucleic Acid Res. Mol. Biol.* **40,** 143–184 (1991).
5. Wolffe, A. P. & Kurumizaka, H. The nucleosome: A powerful regulator of transcription. *Prog. Nucleic Acid Res. Mol. Biol.* **61,** 379–422 (1998).
6. Wyrick, J. J. *et al.* Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402,** 418–421 (1999).
7. Luger, K., Maeder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).
8. Pryciak, P. M. & Varmus, H. E. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69,** 769–780 (1992).
9. Kornberg, R. D. & Lorch, Y. Chromatin structure and transcription. *Annu. Rev. Cell Biol.* **8,** 563–587 (1992).
10. Shimizu, M., Roth, S. Y., Szent-Gyorgyi, C. & Simpson, R. T. Nucleosomes are positioned with base pair precision adjacent to the alpha 2 operator in *Saccharomyces cerevisiae*. *EMBO J.* **10,** 3033–3041 (1991).
11. Flaus, A. & Richmond, T. J. Positioning and stability of nucleosomes on MMTV 3′ LTR sequences. *J. Mol. Biol.* **275,** 427–441 (1998).
12. Meersseman, G., Pennings, S. & Bradbury, E. M. Mobile nucleosomes—a general behaviour. *EMBO J.* **11,** 2951–2959 (1992).
13. Polach, K. J. & Widom, J. Mechanism of protein access to specific DNA sequences in chromatin: A dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254,** 130–149 (1995).
14. Studitsky, V. M., Kassavetis, G. A., Geiduschek, E. P. & Felsenfeld, G. Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. *Science* **278,** 1960–1963 (1997).
15. Tsukiyama, T. The *in vivo* functions of ATP-dependent chromatin-remodelling factors. *Nature Rev. Mol. Cell Biol.* **3,** 422–429 (2002).
16. Calladine, C. R. & Drew, H. R. Principles of sequence-dependent flexure of DNA. *J. Mol. Biol.* **192,** 907–918 (1986).
17. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319,** 1097–1113 (2002).
18. Davey, C. A. & Richmond, T. J. DNA-dependent divalent cation binding in the nucleosome core particle. *Proc. Natl Acad. Sci. USA* **99,** 11169–11175 (2002).
19. Crick, F. H. C. & Klug, A. Kinky helix. *Nature* **255,** 530–533 (1975).
20. El Hassan, M. A. & Calladine, C. R. Conformational characteristics of DNA: Empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Phil. Trans. R. Soc. Lond. A* **355,** 43–100 (1997).
21. Lavery, R. & Sklenar, H. The definition of generalised helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dynam.* **6,** 63–91 (1988).
22. Young, M. A., Ravishanker, G., Beveridge, D. L. & Berman, H. M. Analysis of local helix bending in crystal structures of DNA oligonucleotides and DNA–protein complexes. *Biophys. J.* **68,** 2454–2468 (1995).
23. Olson, W. K. Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* **6,** 242–256 (1996).
24. Dickerson, R. E. DNA bending: The prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* **26,** 1906–1926 (1998).
25. Goodsell, D. S. & Dickerson, R. E. Bending and curvature calculations in B-DNA. *Nucleic Acids Res.* **22,** 5497–5503 (1994).
26. El Hassan, M. A. & Calladine, C. R. Two distinct modes of protein-induced bending in DNA. *J. Mol. Biol.* **282,** 331–343 (1998).
27. Packer, M. J. & Hunter, C. A. Sequence–structure relationships in DNA oligomers: A computational approach. *J. Am. Chem. Soc.* **123,** 7399–7406 (2001).
28. Yanagi, K., Prive, G. G. & Dickerson, R. E. Analysis of local helix geometry in three B-DNA decamers and eight dodecamers. *J. Mol. Biol.* **217,** 201–214 (1991).
29. Shrader, T. E. & Crothers, D. M. Effects of DNA sequence and histone–histone interactions on nucleosome placement. *J. Mol. Biol.* **216,** 69–84 (1990).
30. Anselmi, C., Bocchinfuso, G., De Santis, P., Savino, M. & Scipioni, A. Dual role of DNA intrinsic curvature and flexibility in determining nucleosome stability. *J. Mol. Biol.* **286,** 1293–1301 (1999).
31. Dickerson, R. E. & Chiu, T. K. Helix bending as a factor in protein/DNA recognition. *Biopolymers* **44,** 361–403 (1997).
32. Packer, M. J. & Hunter, C. A. Sequence-dependent DNA structure: The role of the sugar-phosphate backbone. *J. Mol. Biol.* **280,** 407–420 (1998).
33. Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG. *J. Biol. Chem.* **257,** 14686–14707 (1982).
34. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M. & Zhurkin, V. B. DNA sequence-dependent deformability deduced from protein–DNA crystal complexes. *Proc. Natl Acad. Sci. USA* **95,** 11163–11168 (1998).
35. Travers, A. A. & Klug, A. The bending of DNA in nucleosomes and its wider implications. *Phil. Trans. R. Soc. Lond. B* **317,** 537–561 (1987).
36. Richmond, T. J. & Widom, J. in *Chromatin Structure and Gene Expression* (eds Elgin, S. C. R. & Workman, J. L.) 1–23 (Oxford Univ. Press, Oxford, 2000).
37. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* **34,** 269–324 (2001).
38. Zivanovic, Y., Goulet, I., Revet, B., Le Bret, M. & Prunell, A. Chromatin reconstitution on small DNA rings II. DNA supercoiling on the nucleosome. *J. Mol. Biol.* **200,** 267–290 (1988).
39. Klug, A. & Lutter, L. C. The helical periodicity of DNA on the nucleosome. *Nucleic Acids Res.* **9,** 4267–4283 (1981).
40. Arnott, S., Dover, S. D. & Wonacott, A. J. Least-squares refinement of the crystal and molecular structures of DNA and RNA from X-ray data and standard bond lengths and angles. *Acta Crystallogr. B* **25,** 2192–2206 (1969).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank I. Berger and D. Sargent for comments on the manuscript. This study was supported by the Swiss National Science Fund.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.J.R. (richmond@mol.biol.ethz.ch). The coordinates of the X-ray structures are available from the Protein Data Bank as files 1kx3, 1kx4 and 1kx5.

