Biophysical Journal Volume 95 August 2008 1007–1017

Geometry of the Nucleosomal DNA Superhelix

Thomas C. Bishop
Tulane University, New Orleans, Louisiana

ABSTRACT Nucleosome stability is largely an indirect measure of DNA sequence based on the material properties of DNA and the ability of a sequence to assume the required left-handed superhelical conformation. Here we focus attention only on the geometry of the superhelix and present two distinct mathematical expressions that rely on the DNA helical parameters (Shift, Slide, Rise, Tilt, Roll, Twist). One representation requires torsion for superhelix formation; the other requires shear. To compare these mathematical expressions to experimental data we develop a strategy for Fourier-filtering the helical parameters that identifies necessary and sufficient conditions to achieve a high-resolution model of the nucleosome superhelix. We apply this filtering strategy to 24 high-resolution structures of the nucleosome and demonstrate that all structures have a highly conserved distribution of Roll, Slide and Twist that involves two length scales. One length scale spans the entire length of nucleosomal DNA. The other is associated with the helix repeat. Our strategy also enables us to identify ground state or simple nucleosomes and altered nucleosome structures. These results form a basis for characterizing structural variations in the emerging family of nucleosome structures and a method for further developing structure-based models of nucleosome stability.

INTRODUCTION

In 1974, Kornberg proposed that chromatin is a repeating unit of protein-DNA complexes called nucleosomes (1). The general idea that nucleosomes represent the first level of a hierarchy of folding (2,3) that allows lengths of DNA on the order of one meter to fit inside a cell nucleus with a diameter of $\sim 3\ \mu$m quickly followed. However, it was some time before x-ray crystallographic studies provided detailed insights into the structure of the nucleosome's protein core (4–6), complete nucleosomes (7,8), and even tetranucleosome packing (9). The nucleosome's protein core is an histone octamer assembled as two $H2A-H2B$ dimers bound on opposite faces of a central $(H3-H4)_2$ tetramer. Wrapped around this octameric core is 146 basepairs (bp(s)) of DNA forming ~ 1.7 turns of a left-handed superhelix. Fig. 1 provides a representation of one of the x-ray structures.

The ability of different sequences to assume the requisite superhelical conformation is a primary determinant of the relative stability of isolated nucleosomes containing different DNA oligomers (10). The histone-DNA interactions are not sequence-specific, so nucleosome stability is an indirect measure of sequence based on the physical rather than chemical properties of DNA. Our present interest is only the superhelical conformation of the DNA since this is a necessary component of structure-based nucleosome stability models. The energetics of superhelix formation will be considered elsewhere.

There are ~ 25 high-resolution (defined here as ≤ 3.0 Å) nucleosome structures available in the protein databank (www.rcsb.org). The histones are derived from chicken, mouse, human, yeast, or frog but most contain the same 146-bp sequence of α-satellite DNA with the following exceptions: entry 2nzd is 145-bp long; entries 2cv5 (*Human*), 1kx3 (*Xenopus*), and 1aoi (*Xenopus*) have an identical sequence that differs from the consensus 146-bp sequence at two places; entry 1kx5 (11) has a single basepair insertion so that it is a 147-bp sequence rather than 146-bp; 2fj7 (12) (not high resolution by our criteria), contains a poly(dA.dT) sequence and is also 147-bp. There are also a number with structural variations including: nucleosomes complexed with a minor groove binding ligand ((13), pdb entry 1m18); a series of mutants ((14), the 1p3 series in the pdb); histone variants ((15), pdb entry 1u35); and a structure with a pyrrole-imidazole hairpin polyamide that spans the supergroove ((16), pdb entry 1s32).

The path is similar in all available structures and is presumed to be nearly the same for all DNA sequences. However, there is a growing acceptance that the nucleosome is not a monolithic entity, but rather a family of structures. We must keep in mind that histone variants, modifications, and even different states of association/dissociation may have a substantial impact on the conformation and dynamics of the DNA.

For homogenous arrays of nucleosomes, sequence-dependent nucleosome stability can be associated with nucleosome positioning, but self-intersection and nucleosome-nucleosome interactions must also be accounted for. In vivo the relation between nucleosome positioning and sequence-dependent nucleosome stability is less clear, as the nucleosomes are heterogenous and there are epigenetic effects and external molecular influences, e.g., linker histones and regulatory proteins that bind DNA, that must also be considered.

On a gross structural level, crystallography confirmed what was known about the nucleosome for some time. From simple mathematical reasoning and biochemical studies, we know the DNA must be bent ($\geq 4.2°$/bp) and has a non-

Submitted September 28, 2007, and accepted for publication March 5, 2008.

Address reprint requests to Thomas C. Bishop, Tel.: 504-862-3370; E-mail: bishop@tulane.edu.

Editor: Ruth Nussinov.

0006-3495/08/08/1007/11 $2.00

doi: 10.1529/biophysj.107.122853

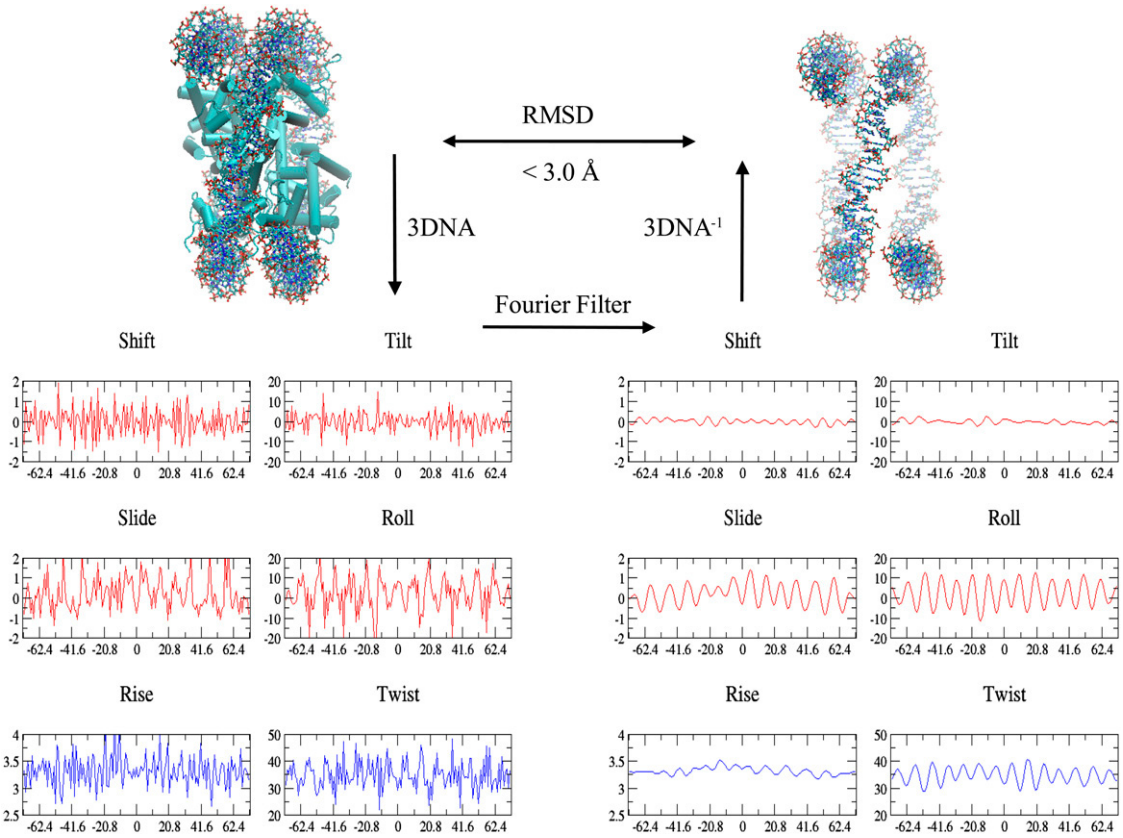

FIGURE 1 Schematic representation of filtering strategy. Beginning with a high-resolution structure of the nucleosome (*top left*) the DNA helical parameters (Shift, Slide, Rise, Tilt, Roll, Twist) are extracted using 3*DNA* (27) (*bottom left*). The helical parameters are subjected to Fourier filtering (33) (*bottom right*) and used to recreate an atomic model of the DNA superhelix, $3DNA^{-1}$ (*top right*). The RMSD between the reconstructed superhelix and the input structure is calculated (34) and utilized to assess the effect of the filter. The filtering has two stages: the first creates knock-outs and the second creates knock-ins, as described in Methods.

uniform overwound conformation with an average helix repeat of 10.3 bp/turn compared to 10.5 bp/turn for B-from DNA. (See, for example, Section 2.2.3 in (17); or see (18).)

An analysis of the conformation of the DNA based on pdb entry 1kx5 (19) compared the distribution of the DNA inter-basepair parameters Roll and Tilt to the distribution expected for a regular superhelix, and identified kinks in the DNA between basepairs 35-36, 47-48, and 57-58. (Distances are measured from the central basepair, identified as basepair 0, in this 147-bp-long oligonucleotide.) As expected from geometric considerations, Roll and Tilt vary sinusoidally as a function of total Twist accumulated (or equivalently position) along the 146 bps of DNA, but the simple superhelical model did not compare well to the atomic structures and the amplitude of Roll was reported to be approximately twice the expected value.

A Fourier analysis technique was subsequently used to quantify the distribution of helical parameters on three different length scales as observed during a molecular dynamics simulation of the nucleosome (entry 1kx3) (20): Long (longer than the helix repeat); Intermediate (less than the helix repeat but >3 bp); and Short (<3 bp). It was concluded that both long and intermediate length-scale distributions of helical parameters are required for proper folding of the nucleosome superhelix. Subsequent analysis by Tolstorukov et al. (21) proposed a novel Roll-Slide mechanism. Here we formalize the method presented in Bishop (20) and systematically apply it to analyze 24 high-resolution structures (23 from x-ray crystallography (8,11–14,16,22–25) and 1 from simulation (20)). We demonstrate that 11 Fourier components are both necessary and sufficient to describe the distribution of DNA inter-basepair helical parameters in a simple or ground-state nucleosomal superhelix and that the Roll-Slide mechanism (21) is incomplete. The method also enables us to characterize complex or activated nucleosome superhelices that vary from this structure. Such a thorough understanding of the distribution of DNA helical parameters is a prerequisite for the development of models of nucleosome stability based on structure as recently proposed in Tolstorukov et al. (21).

The following sections include Theory, the mathematical formulas describing the distribution of helical parameters expected for an ideal superhelix; Methods, our Fourier filtering and reconstruction techniques; Results, and Discussion.

THEORY

DNA inter-basepair helical parameters

X-ray crystallography has provided a number of all-atom Cartesian coordinate description of the nucleosome. Thus, we know the x,y,z coordinates of every atom in the DNA in what we will call a fixed or laboratory reference frame. DNA, especially duplex DNA, can also be described very accurately using an internal or local coordinate system of helical parameters as defined in Dickerson (26). The helical parameters include inter- and intra-basepair descriptors and there are a number of software packages freely available for calculating them (27,28). The key point in the development of our method is that the conversion from Cartesian coordinates to DNA helical parameters is invertible. A complete cycle from Cartesian coordinates to helical parameters and then back to Cartesian coordinates provides a structure that is nearly identical to the original (27).

There are six inter-basepair helical parameters, denoted here as an array $HP =$ (Shift, Slide, Rise, Tilt, Roll, Twist). As the names suggest, the parameters provide information about the relative orientation, (Tilt, Roll, Twist), and the relative position, (Shift, Slide, Rise), between adjacent basepairs. The inter-basepair parameters are thus a description of DNA as a stack of rigid bodies with each rigid body being a single basepair. Rise and Twist are translations and rotations along the longitudinal axis of the DNA itself and are the only two parameters needed to describe an idealized representation of DNA that is straight. Nonzero values of Tilt or Roll produce a bend in the DNA (axis of rotation in a plane that is orthogonal to the longitudinal axis) and nonzero values of Shift or Slide produce a shearing of adjacent basepairs (translations orthogonal to the longitudinal axis).

While the inter-basepair parameters account for the stacking of the basepairs, the intra-basepair parameters account for deformations of the basepairs themselves. Each base is a fairly rigid planar ring system so the intra- and inter-basepair helical parameters describe the bases in duplex DNA very accurately. The basepairing and stacking, along with geometric constraints based on chemical bonding, are then sufficient to determine the structure of the phosphate-deoxyribose backbone. In practice, we have found that using only the inter-basepair helical parameters rather than the complete set of inter- and intra-basepair helical parameters yields an error of <1.0 Å RMSD when the transformation from Cartesian coordinates to helical parameters and inverse transform is applied to the nucleosomal superhelix. For this reason, we will not consider the intra-basepair parameters further. The intra-basepair parameters are all set to zero in our analysis of the structures and so we expect the highest resolution reconstruction we can achieve is $\approx$1.0 Å RMSD of the input structure.

Two superhelix expressions

Using the intra-basepair helical parameters there are at least two sets of simple mathematical expressions that will produce a regular (constant bend and constant pitch) superhelical structure that also has constant Rise and Twist. General expressions that represent both possibilities are given below as continuous functions of $s \in (0, nbp)$, where nbp is the number of basepairs in the superhelix. For the purposes of modeling DNA, the variable s in Eq. 1 can have only integer values, i.e., the helical parameter representation of DNA is naturally basepair-discrete, not continuous as explored in Manning et al. (29). We emphasize in this section that the helical parameters are a mathematical construct and that our expressions for superhelices are not specific to DNA:

$$\begin{array}{cc} Sh = \gamma \sin(Tw_o\, s) & Ti = \kappa \sin(Tw_o\, s) \\ Sl = \gamma \cos(Tw_o\, s) & Ro = \kappa \cos(Tw_o\, s). \\ Ri = Ri_o & Tw = Tw_o + \tau \end{array} \quad (1)$$

The usual superhelix description does not have shear, i.e., $Shift = Slide = 0$. The pitch arises only from torsion. We label this superhelix a torsion helix, TH. Alternatively, if shear is allowed, then a regular helix can be created without torsion. We label this superhelix a shear helix, SH. Both expressions can be parameterized to represent a superhelix with a given pitch and curvature.

We point out that in case $\tau = 0$ (no torsion) and $\gamma = 0$ (no shear), the above expressions describe a structure that is circular. Addition of a phase term to the trigonometric functions only changes which face of the fiber is on the outer edge of the circle.

For the TH $\gamma = 0$, and $\tau \neq 0$ is the torsion. The superhelix arises only from a sinusoidal distribution of Tilt and Roll and has constant Twist and Rise. The value τ controls the superhelical pitch and register of the helix. This description is equivalent to Eq. 4 in Chouaieb et al. (30) with $\phi(s) = (Tw_o)\, s$. Such a helix has register $\phi(s)$, curvature κ, radius $\kappa/(\kappa^2 + \tau^2)$, and pitch $2\pi\tau/(\kappa^2 + \tau^2)$.

To model the nucleosome superhelix, a right-handed helix, and a left-handed superhelix as a TH requires $Tw > 0$ and $\tau < 0$, so that $Tw < Tw_o$. The ideal superhelix model in Richmond and Davey (19) has $\kappa \approx 4.5°$/bp, $Ri \approx 3.4$ Å/bp, radius $\approx$42 Å, and pitch $\approx$26 Å, corresponding to a torsion of $\tau \approx -0.45°$/bp in the TH. This amount of torsion is well within the range of Twist fluctuations (3.9° to 6.9°) associated with any basepair step as determined by analysis of x-ray structures of free DNA (31). The pitch associated with this parameterization of the TH is very sensitive to torsion ($\approx$60 Å pitch per degree of torsion), thus a uniform change in Twist by 0.1°/bp changes the superhelical pitch by 6 Å. By comparison, the pitch is insensitive to changes in κ and the radius is not sensitive to changes in κ or τ.

The TH is employed to describe the mechanics of springs. A compression or extension of the spring (change in superhelical pitch) arises from a rotation of the material cross section. Such a model of the nucleosome was investigated in Bishop and Zhmudsky (32). For an arbitrary material, shear can be set to zero by a suitable choice of internal coordinates. In so doing, one can effectively replace a SH description with

a *TH* description. It is likely for this reason that the *SH* has historically received little or no attention in the literature. Shear is not discussed in a recent review of helices (30); however, it is recognized in the literature (19,21) that shear in the form of Shift and/or Slide is a necessary component of the nucleosome superhelix.

In the case of DNA, we consider the helical parameters as a predefined natural description of DNA geometry that has an intuitive mapping to the physical material. We therefore consider the effects of shear rather than develop a new set of shear free coordinates.

For the *SH*, $\tau = 0$ and $\gamma \neq 0$. The curvature is simply $\kappa = \sqrt{Roll^2 + Tilt^2}$, and shear is $\gamma = \sqrt{Shift^2 + Slide^2}$. The contribution of each basepair step to pitch is γ, and the radius is $1/\kappa$. For a right-handed helix and left-handed superhelix, $Tw > 0$ and $\gamma < 0$. For the nucleosome, once again we set $\kappa \approx$ 4.5°/bp and $Ri \approx 3.4$ Å/bp, yielding a radius of ≈43 Å, and use $\gamma \approx -0.33$ Å/bp to create a superhelical pitch of −26 Å. This amount of shear is within the range of Shift (0.46–0.87 Å) and Slide (0.31–0.89 Å) fluctuations associated with any basepair step in free DNA (31). In the *SH*, superhelical pitch is inversely proportional to κ because $1/\kappa$ determines the number of basepairs in a superhelical turn. For the above parameterization, the pitch is sensitive to shear; even though the relation is linear, the proportionality constant is ≈80 Å pitch per Å shear. The variation in radius as a function of curvature is nearly identical to the *TH*. Thus, in both descriptions of the nucleosome, superhelix pitch is sensitive to the parameterization.

It is particularly relevant to analysis of the nucleosome superhelix that, for the *SH* if *Ro* and *Sl* are as indicated, but *Ti* and *Sh* are set to zero, then the resulting structure is still superhelical. It is not a regular superhelical structure. Graphical analysis indicates that this Roll-Slide helix, *RoSlH*, has almost exactly twice the radius and half the pitch of the *SH*. Thus, if the values of Roll and Slide are double the values expected from the *SH* and the values of Tilt and Shift are zero, then the path of the *RoSlH* is nearly the same as that of the *SH*.

We also point out that changing the relative phase between *Ti*, *Ro*, *Sh*, or *Sl* in either the *TH* or *SH* as described in the expressions in Eq. 1, yields an irregular superhelix with nonconstant pitch and curvature.

Thus, using the inter-basepair DNA helical parameters, we have two methods of creating a regular superhelix. Given a superhelix with pitch, p, and radius, r, each solution can be parameterized to provide the same minimal path for each turn of the superhelix, $l = \sqrt{4(\pi r)^2 + p^2}$. The total path is then just Nl where N is the number of superhelical turns. Variations from this minimal path arise when the Bend, Shear, Twist, and/or Rise are not distributed as indicated. Such a superhelical path is not a minimal length path around the nucleosome. As demonstrated previously, thermal fluctuations require that the DNA superhelical path cannot be minimal (20). The nucleosome superhelix must have some slack built into it to accommodate thermal motion.

METHODS

Given that both *SH* and *TH* involve trigonometric functions, we developed a Fourier-filtering strategy for analyzing the DNA helical parameters as obtained from 24 different high-resolution nucleosome structures. The helical parameters are an internal coordinate description of the DNA for which there is no simple criteria for determining the structural significance of each Fourier component. Thus, we developed a two-stage filtering strategy as represented schematically in Fig. 1. This two-stage strategy allows us to systematically determine a minimal set of Fourier components that are both necessary and sufficient to achieve a high-resolution reconstruction (i.e., an RMSD of <3 Å between the reconstructed and initial structure). We chose 3 Å RMSD as the target for our reconstructions because this is the lowest resolution of any of the initial structures. We have also demonstrated that models of the nucleosome built using these helical parameters and a variety of different sequences are of sufficient accuracy to initiate molecular dynamics simulations (T. Bishop, unpublished result).

In the first stage, a single Fourier component, j, is filtered from each of the crystallographic helical parameters. These knock-outs are calculated as

$$\begin{aligned} KO(p,s,k) &= \Sigma_{j_p \neq k}^{nbps/2} Ae(j_p) e^{-2\pi i j_p s/nbps} \\ &= HP(p,s) - A_p(k) e^{-2\pi i k s/nbps}. \end{aligned} \quad (2)$$

Here *HP* designates an array of helical parameters as obtained from 3DNA, *KO* designates an array of knock-outs, A_p an array of complex amplitudes associated with the helix parameter p = (Shift, Slide, Rise, Tilt, Roll, Twist), and s one of the $nbps = nbp - 1$ basepair steps. The wavelength associated with knock-out k is *nbps*/k. There are *nbps*/2 possible wavelengths ranging from nbps to 2 bps. The average value of any helical parameter is associated with $k = 0$.

For each of the 24 structures, 7**nbps*/2 knock-outs are created: six sets of *nbps*/2 knock-outs in which only one of the six helical parameters is filtered, and a seventh set of *nbps*/2 in which the k^{th} Fourier component is filtered from all six helical parameters simultaneously. For convenience, the six sets are referred to as k_p knock-outs, and the seventh set as a k knock-out. For each structure, the RMSD between the initial structure and the knock-out structures is calculated. Any knock-out for which the RMSD exceeded our 3 Å cutoff criteria is obviously a necessary Fourier component for that structure. The data in Fig. 2 also enables us to rank the effects of each knock-out, k_p, according to the RMSD that it introduces into the reconstruction. In this manner, a sorted list of wavenumbers, denoted k_{pl} with individual elements denoted j_{pl}, is obtained.

For the second stage, our strategy is to gradually add complexity to our representation of the DNA helical parameters. The knock-ins are determined using the sorted list of wavenumbers, k_{pl}, obtained from the first stage, as follows:

$$KI(p,s,k) = \Sigma_{l=0}^{k} A_p(j_{pl}) e^{-2\pi i j_{pl} s/nbps}. \quad (3)$$

Similar to the knock-outs, seven sets of knock-ins are calculated for each structure. Six corresponded to individual helix parameter knock-ins and one in which all helix parameters are knocked-in simultaneously. For the seventh set, we find that, if separate k_{pl} lists are used for each helix parameter, the structure does not converge as rapidly as if a single list, denoted k_l, is used for all six parameters simultaneously. This appears to be due to geometric couplings between the various helical parameters. Thus, for the seventh set, we determine a single, potentially unique, list k_l for each structure obtained from RMSD sorting the knock-outs, as described above.

For the individual parameter knock-ins, e.g., *KI*(*Roll*, *s*, *k*), all other helix parameters are as observed in the initial structure while Roll is made incrementally more complex with each knock-in. The resulting RMSD values enable us to determine how many Fourier components Roll requires to achieve our high-resolution reconstructions and which helical parameters control the RMSD. The results are plotted in Fig. 3. Comparing the k_l values enables us to group the structures into subfamilies with matching k_l. Members of each subfamily have similar length-scale dependencies as indicated by the ordering of Fourier components in k_l.

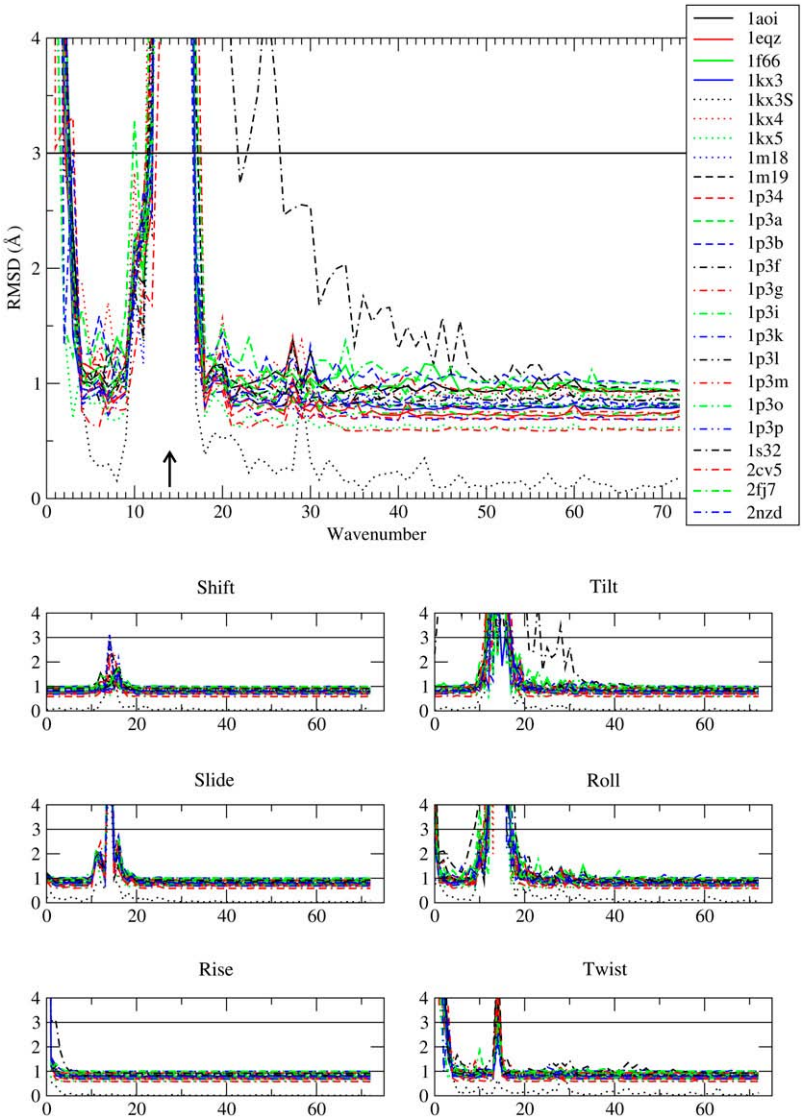

FIGURE 2 Knock-out filter RMSD plots. (*Top*) The RMSD (Å) introduced by the knock-out filter is plotted versus the wavenumber, k, removed from all of the helical parameters simultaneously. The arrow indicates the location of $k = 14$, which corresponds to a length scale of 145/$k = 10.36$ bps. Structures are identified by their Protein Data Bank id code as indicated in the legend. 1kx3S represents an average structure obtained from a 10-ns molecular dynamics simulation of pdb entry 1kx3 (20). The horizontal line at 3.0 Å indicates our criteria for high resolution. The RMSD does not converge to 0.0 Å because we have set the intra-basepair helical parameter values to zero. (*Bottom*) The RMSD (Å) introduced by individual helical parameter knock-outs. Axes have the same scale as in the top graph. In these knock-outs, only the indicated helical parameter is subjected to filtering; the other parameters are obtained from the initial structure.

We conclude Methods with some necessary technical comments. In all cases, the determination of DNA helical parameters from an atomic model and the reconstruction of an atomic model from helical parameters is achieved using 3DNA (27). The misc_3dna.par descriptors (upper H-bond length, maximum distance between paired bases, etc.) are relaxed from the default values until a full complement of Watson-Crick basepairs are identified in all of the initial structures used in this study. Without relaxing the misc_3dna.par criteria all of the structures do not achieve a complete complement of basepairs. However, once values for misc_3dna.par are determined, the same set of values is used for all data analysis. Fourier filtering is achieved with a FORTRAN program that utilizes FFTW Ver. 2.1.5 (33). (Our program is freely available upon request.) The determination of RMSD values utilized the RMSD fit and measure commands available in VMD (34). All heavy atoms are used for determination of RMSD values.

RESULTS

Knock-outs versus RMSD

Consistent with our two-stage strategy, we consider the knock-outs first. As indicated in the main plot of Fig. 2, Fourier components 0, 1, 13, 14, 15, and 16 are necessary for all of the structures. Removal of any one of these components from all six helical parameters introduces an RMSD >3.0 Å in any of the 24 structures studied.

Thus structural variations on two different length scales are a necessary component of the nucleosome superhelix. One length scale spans the entire 146-basepair length of DNA in the nucleosome, $k = 1$, and the other length scale includes wavenumbers near $k = 14$. The latter corresponds to a wavelength of 145/k basepair steps and is equivalent to a helix repeat of 10.4 bp/turn or 34.8°/bp. This wavenumber is the one that most closely corresponds to the average Twist of DNA in the nucleosome, which ranged from 34.6°/bp to 35.3°/bp for the 24 structures. This wavenumber is expected according to either the *TH* or *SH* expressions. (Notice we have chosen to use 145 for simplicity, since it corresponds to the majority of structures. For 2nzd, the wavelength is 144/$k = 10.3$ or 35°/bp. For 2fj7 and 1kx5, the wavelength is 146/$k = 10.4$ or 34.5°/bp.)

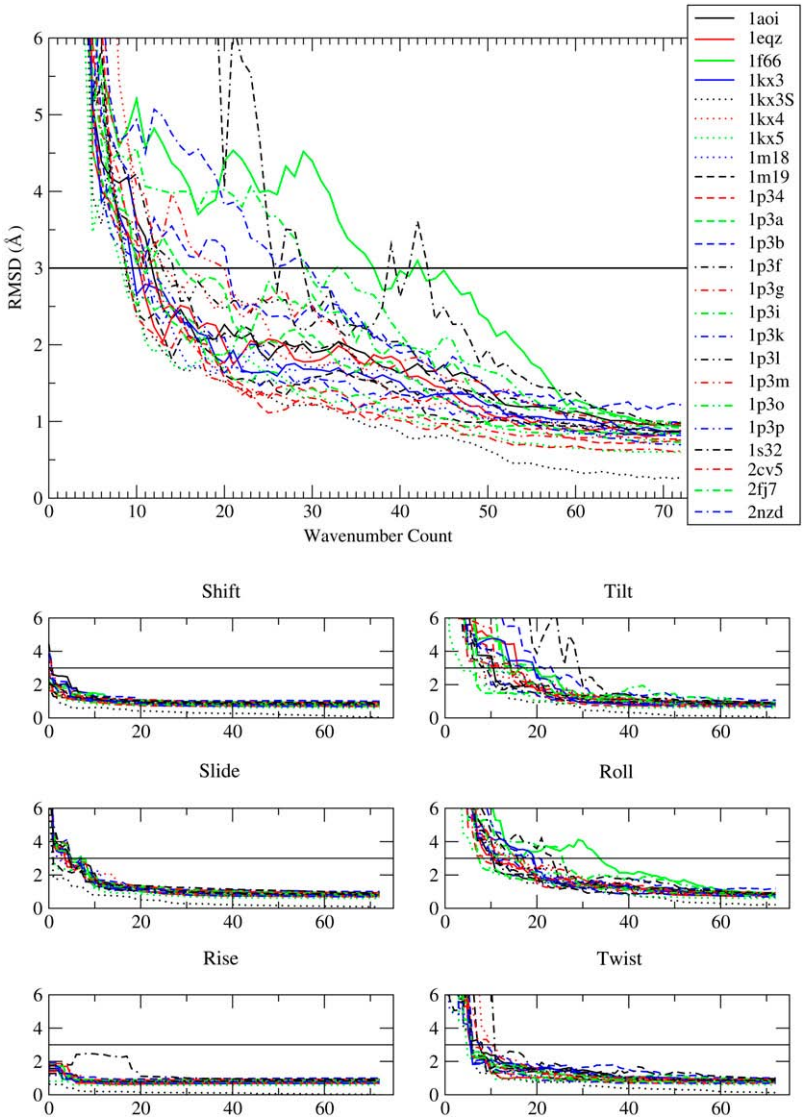

FIGURE 3 Knock-in filter RMSD plots. (*Top*) The RMSD (Å) arising when the knock-in filter is applied to all helical parameters simultaneously. The total number of Fourier components is as indicated but the wavenumbers, k, are not known. The structures are identified in the legend. The horizontal line at 3.0 Å indicates our criteria for high resolution. (*Bottom*) The RMSD (Å) introduced by individual helical parameter knock-ins. Axes have the same scale as in top graph. In these knock-ins, only the indicated helical parameter is subjected to filtering; the other parameters are obtained from the initial structure.

The smaller plots in Fig. 2 indicate the effect of applying the knock-out filter to the individual helix parameters and demonstrate that Roll, as well as Twist and Rise, require nonzero average values. The $k = 0$ knock-out has an RMSD $>$ 3.0 Å for each of these. Twist also requires long length-scale variations, namely $k = 1$ and 2. Shift and Rise knock-outs do not require any single Fourier component that varies on a length scale comparable to the helix repeat.

The structure obtained from an average of helical parameters observed during a 10-ns molecular dynamics simulation, 1kx3S, has the most well-defined spectra. Roll has five components that introduce an RMSD $>$ 3.0 Å, none of which are long length-scale; Tilt has four components, and the RMSD of both is most strongly influenced by $k = 14$. Twist only requires $k = 0$, 1. Rise only requires the average, $k = 0$; and Slide only requires $k = 14$. There is no single component of Shift that when removed introduces an RMSD $>$3.0 Å.

Knock-ins versus RMSD

While only six individual components introduced a deviation $>$3.0 Å, none of the structures converged to within 3.0 Å of the initial structure with $<$10 Fourier components. The six components identified in the previous section are sufficient only to obtain the gross structure ~6.0 Å RMSD, but they are not sufficient to achieve the target resolution of 3.0 Å.

The knock-ins tended to monotonically converge to the initial structure as indicated by data in Fig. 3. More than half the structures achieved the target resolution using only 12 Fourier components: the average helical parameter values plus 11 additional components that varied sinusoidally. Only four structures (1f66, 1p3i, 1s32, 2nzd) required $>$20 components. We classify the former as having a simple or ground-state superhelix conformation and the latter as having a complex nucleosome superhelix conformation because more Fourier components are required to obtain a structure with

the target resolution. It is relevant that the four structures identified as complex include an H2A.Z variant, a Sin mutant, a supergroove spanning ligand, and extreme kinking, respectively.

From the individual knock-ins in Fig. 3, we can identify the complexity required of each helical parameter by the number of Knock-ins required. Care must be taken when interpreting the data in Fig. 3, as only the number of knock-ins is indicated. The wavenumbers are not indicated. Nonetheless we observe that Shift, Rise, and Twist require fewer Fourier components to achieve a high-resolution reconstruction than Slide. Roll and Tilt require the greatest number of components.

To further assess the knock-ins, we directly evaluate the distribution of helical parameters required for high-resolution reconstructions of the simple nucleosomes, as shown in Fig. 4. It is clear that Slide, Roll, and Twist vary little from structure to structure, Rise varies a bit more, while Shift and Tilt exhibit comparatively large variations from structure to structure. Roll and Slide have the overall distribution expected for a *SH* structure. But there is also a long length-scale variation that contributes to the distribution of Rise, Slide and to lesser extents Twist and Roll. The long length-scale variation is confirmed by comparing the amplitude at $k = 1$ to that at $k = 14$ for each of the helical parameters; see plots on right side of Fig. 4.

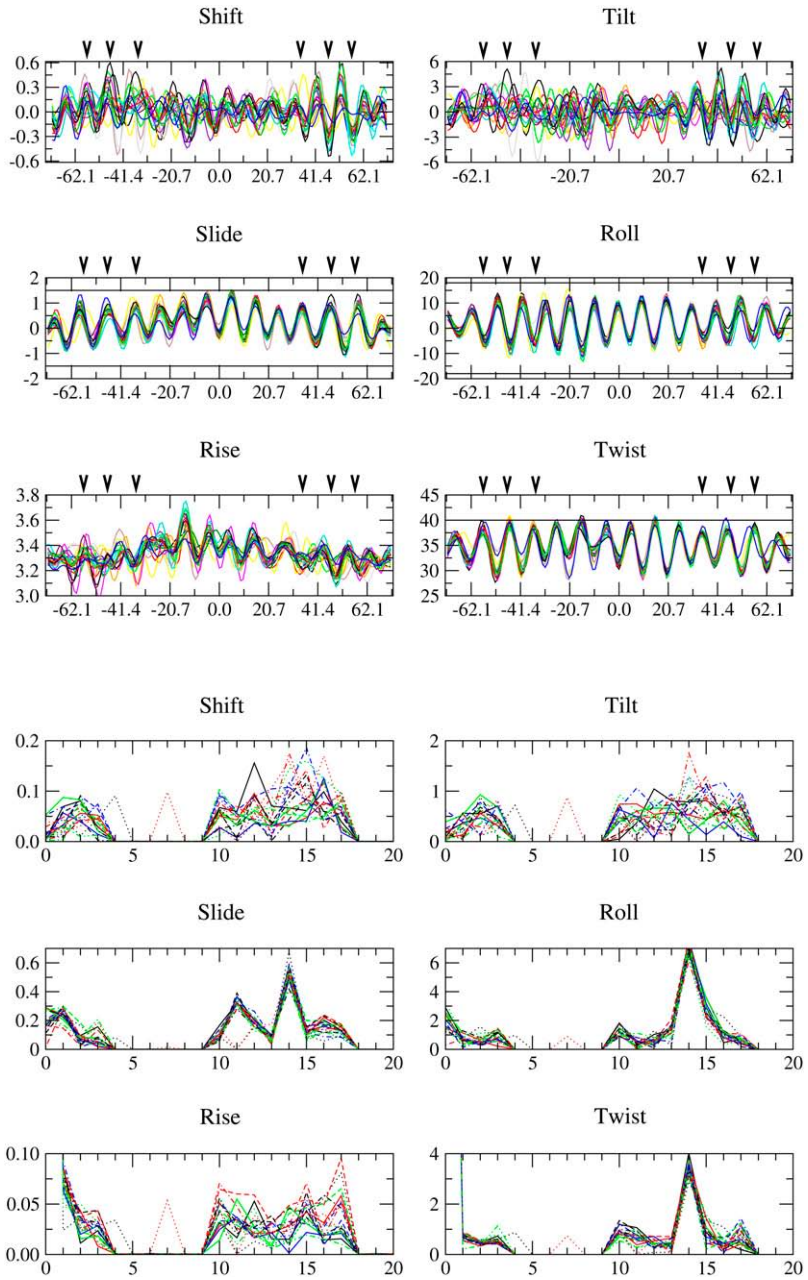

FIGURE 4 Required distribution of helix parameters. (*Top*) The distribution of helical parameters obtained after the two-stage filtering procedure. The indicated distribution is required to achieve 3.0 Å RMSD between the reconstructed superhelix and the initial structure. For Shift, Slide, and Rise the units for the vertical axis are Ångstroms and for Tilt, Roll, Twist the units are degrees. Note that the vertical ranges differ. In all cases the horizontal axis represents the distance from the dyad measured in basepair steps. Inverted triangles indicate the position of kinks identified in Richmond and Davey (19). Horizontal lines on the Slide, Roll, and Twist plots indicate the criteria used to identify the kinks. (*Bottom*) The amplitude (*vertical axis*) associated with each Fourier component is plotted versus wave number (*horizontal axis*). Shift, Slide, Rise are in Ångstroms. Tilt, Roll, Twist are in degrees.

The amplitude plots also indicate some significant results regarding the coupling of Roll, Slide, and Twist. If these three helical parameters were correlated on all length scales, then each plot would exhibit the same profile. They do not. Roll, Slide, and Twist appear to be coupled at $k = 14$, i.e., a large Roll corresponds to a large Slide and a large Twist on this length scale. However, this is not true for $k = 10, 11, 15, 16$. Slide has a large amplitude at $k = 11$, but tends to zero at $k = 10$, while Roll has a large amplitude at $k = 10$, but tends to zero at $k = 11$. On these two length scales, Roll and Slide act independently. Similarly, Slide appears independent of Roll and Twist at $k = 15$ and Twist is independent of Roll and possibly Slide at $k = 17$. We cannot determine from these results whether interactions with the histones provide a coupling only at $k = 14$, break an intrinsic coupling of these helical parameters on length scales other than $k = 14$, or the coupling of the helical parameters in DNA is a length-scale-specific material property of DNA.

The plots in Fig. 4 indicate that the set of 12 wavenumbers common to all of the simple structures is $k = 0, 1, 2, 3, 10, 11, 12, 13, 14, 15, 16$, and 17. Wavenumbers k = 4, 5, 6, 7, 8, and 9 do not appear in any of the structures, except 1m18, which has a contribution from $k = 7$, and 1kx4, which has a contribution from $k = 4$. Wave numbers 4–9 represent variations in structure that have a length scale that ranges from 29 to 16 basepairs. We conclude that variations in structure on this length scale are not characteristic of the nucleosome superhelix. Forcing such variations may be a strategy for destabilizing nucleosomes.

Curvature, pitch, and symmetry

The ordering of the wave numbers, namely k_1, cannot be assessed from data in Fig. 3. However, by inspection of the k_1 and molecular visualization, we can assign a role to the first four Fourier components and identify subfamilies of the nucleosome structures. All structures begin with $k_1 = 0, 14, 15$, and the fourth element in k_1 had three possible values, 1, 13, or 16.

The first knock-in, $k = 0$, provides the average values of each helical parameter. The average Twist ranges from 34.6°/bp to 35.3°/bp, and average Rise is 3.3–3.4 Å/bp in the structures evaluated. The underlying structure of nucleosomal DNA in the nucleosome is nominally B-form DNA.

The second knock-in, $k_1 = 0, 14$, corresponds to variations with a length scale of $145/14 = 10.36 \sim 34.8°$/bp. As indicated in Fig. 5, these knock-ins have enough curvature to achieve more than one turn of the superhelix but none have sufficient pitch to avoid self-intersection upon completion of one turn. Since this knock-in only has $k = 0$ and $k = 14$, we can compare them to our *TH* and *SH* expressions.

For comparison to the *TH* we first consider the average Twist. In seven structures it is <34.8°/bp and therefore $\tau < 0$, in nine structures $\tau > 0$, and in eight structures $\tau = 0$. Thus some structures are expected to have left-handed, some right-handed, and some no superhelical character (respectively) as a *TH*. Only in the case of 2cv5 ($\tau = 0.3°$/bp), 2nzd ($\tau = 0.3°$/bp), and 1kx4 ($\tau = 0.4°$/bp) does the magnitude of τ approach the expected value of $-0.3°$/bp. However, in each of these structures the sign is for a right-handed, rather than left-handed helix. Moreover, we must keep in mind that there is also a variation in Twist with amplitude of ~4°/bp for $k = 14$ (see Fig. 3) that may potentially dominate the effects of τ.

For comparison to the *SH* we consider the distribution of helical parameters as indicated in Fig. 4. None of the structures seems well described by the *SH*, and Fig. 3 confirms that the amplitudes of Roll, Tilt, Shift, and Slide are not such as to yield a constant bend and constant shear helix. However, Roll and Slide are correlated and even of the correct phase to produce a superhelix arising from the Roll-Slide mechanism described qualitatively in Tolstorukov et al. (21) or arising from the *RoSlH* described in Theory. As indicated in Fig. 3, the amplitude of Roll is ~7°/bp and Slide 0.6 Å/bp. For such a parameterization of the *RoSlH*, we expect ~1.4 turns and a pitch of ~31 Å/turn, a close approximation to the superhelix given only two Fourier components.

However, direct comparison of the second knock-in to the *SH*, *TH*, or *RoSlH* expressions ignores the 4°/bp variation of Twist on this length scale, $k = 14$. This variation in Twist apparently works against the development of superhelical pitch because, for the knock-in with wavenumbers $k = 0$ and 14, all structures are nearly flat, see Fig. 5. To confirm the role of Twist at $k = 14$ we constructed knock-ins in which the $k = 14$ variations in Twist were not included. We found that such

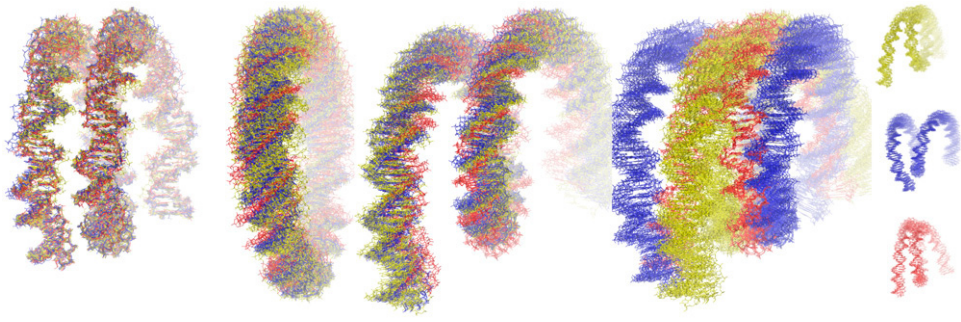

FIGURE 5 Superhelix folding. A series of snapshots represents the effects of filtering on the 20 structures with 146 bps. (*Far left*) Superimposition of the initial high-resolution superhelical structures obtained from the x-ray data. The next three images are the knock-ins $k_1 = 0, 14$; $k_1 = 0, 14, 15$; and $k_1 = 0, 14, 15$; and N where N has the possible values 1 (*yellow*), 13 (*blue*), or 16 (*red*). (*Far right*) The three subfamilies, as identified by the value of N are displayed individually.

knock-ins yielded the curvature and pitch expected from the *RoSlH*. This pitch cancellation effect caused by the variations in Twist has apparently not been reported.

The third knock-in, $k_1 = 0, 14, 15$, produces a structure that has approximately the correct curvature and pitch but only for the central part of the 146 bps. This is the segment of DNA in contact with the tetramer $(H3 - H4)_2$. These structures are all rather symmetric in terms of the overall geometry of the superhelix.

The fourth knock-in includes the initial $k_l = 0, 14, 15$ series but the fourth wavenumber varies, $k = 1, 13$, or 16. This allows us to group them into three families. The curvature of the central segment remains approximately the same regardless of the wavenumber, but the pitch and symmetry differ significantly depending on which wavenumber appears fourth. Nine structures have $k = 1$. In these, the pitch introduced by $k = 15$ is largely removed and the DNA self-intersects. The structures are asymmetric. Seven have $k = 13$. For these structures, the pitch introduced by $k = 15$ is exaggerated in the central region, and the structures are comparatively symmetric. Seven have $k = 16$. This family includes both 147-bp structures, 1kx5 and 2fj7. These structures display the most variability, but in all of them it is primarily the ends of the superhelix that are significantly displaced from proper positioning. These structures are the most symmetric.

We have thus built up an idea of the contributions from each helical parameter as associated with different length scales. For $k = 14$, Roll has an amplitude of $\sim 6°$/bp, Slide 0.6 Å/bp, and Twist 4.0°/bp. Variation on this length scale creates a superhelix with the proper curvature in the central segment but virtually no pitch. This is true for all of the structures. An additional component $k = 15$ is needed to provide pitch. The next component is variable, and tends to produce structures that fundamentally differ in structure, creating three subfamilies. As more Fourier components are introduced, the structure gradually converges to the proper superhelical structure. The total number of wavenumbers required is an indication of the complexity of the distribution of helical parameters, regardless of the subfamily to which the structures belong.

DISCUSSION

None of our superhelix expressions, *SH*, *TH*, or *RoSlH*, is sufficient to model the atomic reality. The nucleosome superhelix appears to be a Roll-Slide-Twist structure in which the curvature arises from Roll as envisioned in Richmond and Davey (19) and the pitch arises primarily from shear, in the form of Slide, as previously reported (21). However, we demonstrate that a variation in Twist ($k = 14$) tends to cancel the superhelical pitch that arises from Slide. Moreover, Roll and Slide are not coupled on all length scales. Pitch and symmetry are strongly influenced by the effects of Twist and its coupling to Roll and Slide on different length scales.

The nucleosome requires a specific distribution of Roll, Slide, and Twist and to a lesser degree, Rise. By this we mean that the distribution identified by our Fourier filtering is highly conserved in the 24 nucleosome structures that we studied. Different realizations of the nucleosome exhibit considerably different distributions of Tilt and Shift. We interpret this result as an indication that the path of DNA on the nucleosome is governed largely by Roll, Slide, and Twist while Rise, Tilt, and Shift are allowed a relative freedom. Since Rise, Tilt, and Shift are known to be stiff helical parameters, this freedom allows the most energetically costly variables to be optimized. If the distribution of all six helical parameters were highly conserved, the superhelical geometry would be completely determined, and the DNA could not accommodate minor variations in the superhelical path that arise, for example, from thermal fluctuations, histone variations, or modification or sequence-specific properties of the DNA. Requiring a specific distribution of Roll and Slide is the least costly, energetically. Twist is a master variable in terms of the superhelix geometry, so it simply must be conserved. Further support of this interpretation is provided by comparison of the structure resulting from simulation to the x-ray structures. In the simulation, the variations associated with the free helical parameters tend to average out over time. The Fourier spectra obtained from the simulation data is comparatively flat for these free helical parameters.

There are no required variations in the helix parameters associated with length scales ranging from 29 to 16 basepairs or shorter than 8.5 basepairs. Twist and Slide each require variations on two length scales. Roll requires a constant value and variations with the helix repeat. Thus, there are two distinct length scales associated with the nucleosome superhelix. One length scale spans the entire nucleosome and the other nominally corresponds to the helix repeat of the DNA.

The long length variation leads to a straightening of the DNA at each end of the superhelix. This region is also demarcated by maxima in Slide that occur at ± 58 basepair steps from the center or ~ 14 basepair step from each end (1.5 turns). This corresponds to the site of kinks as identified by Richmond and Davey (19).

The picture that emerges is that curvature arises largely from the coupling of Roll and Slide for $k = 14$, corresponding to the helix repeat, but that variations in Twist on this length scale prohibit the development of pitch predicted by a constant Twist, constant Rise, Roll-Slide model. Superhelical pitch arises mostly from variations in the helical parameters associated with $k = 15$. On this length scale, and others, Roll, Slide, and Twist are not coupled. The determination of symmetry and finer details of the structure require a more complex representation of the helical parameters than can be achieved with only a few Fourier terms.

Kinks or no kinks

Two sets of three kinks, between basepairs 35-36, 47-48, and 57-58 as measured in each direction from the dyad, have been identified (19). The kinks were defined as having a Roll value

between $-18°$ and $-27°$, Slide > 1.5 Å/bp, and Twist $> 40°$/bp. Kinks, as defined by this criteria, have been removed by the Fourier-filtering method developed here and are not necessary to achieve a high-resolution model of the nucleosome superhelix. We emphasize that if proper formation of the superhelix required a distribution of helical parameters that more closely resembled the data obtained from the x-ray structures, then our Fourier-filtering strategy would simply yield a greater number of knock-ins. The limiting case, no knock-outs allowed, would reproduce the x-ray data exactly. This is not the case.

In an MD study of 1kx3 (20), these kinks healed in time. In a separate molecular dynamics study of 1kx5, the authors reported DNA kinks, but a close inspection of Fig. 3 in Roccatano et al. (35) indicates that the Roll, Slide, Twist criteria is not met at any of the kink sites. It even appears that the fluctuations (average + standard deviation) are not sufficient at any one proposed kink site to simultaneously satisfy the Roll-Slide-Twist criteria for a kink. So either the kinks are transient states that rarely appeared during the dynamics, the kinks have healed as in Bishop (20), or the kinks have dissipated into a longer length-scale deformation that affects nearby basepairs as in Lankas et al. (36). In all three cases, a set of static localized kinks did not persist during dynamics.

The identification of kinks is complicated by the fact that there does not seem to be a consensus on what constitutes a kink. A kink suggests a discontinuity. This discontinuity may be measured along the length of a segment of DNA or may be associated with a single basepair step. Intuitively, a kink measured along a length of DNA occurs whenever there is an abrupt bend at one or two discrete loci of an otherwise straight DNA duplex (37). Identification of such kinks is complicated by the fact that DNA is naturally basepair-discrete, and, therefore, every basepair step represents a kink, if a strict application of mathematical limits is imposed. For this reason, an ad hoc method is utilized in Lankas et al. (36) to identify such kinks.

Alternatively, the conformation of a given basepair step can vary continuously, e.g., as a smooth function of time or exhibit a discontinuity, e.g., as a function of energy. This is the type of kink proposed in Crick and Klug (38) and recently investigated in Wiggins et al. (39). A kink, defined as a discontinuity in the energy landscape associated with a single basepair step, reduces the total energy required to achieve the distorted conformation. Sequences that allow such kinks will certainly have an impact on nucleosome stability and positioning, but the question posed by Crick and Klug in 1975 (38) still remains: At what value of curvature, κ, does the energy landscape change from bending (κ^2 dependency) to kinking (κ dependency)?

Our method of analysis only addresses the geometry of the nucleosome superhelix, not the energetics. We have demonstrated that the specific Roll-Slide-Twist criteria for a kink is not necessary to create a model of the nucleosome superhelix at atomic resolution.

Extension to other states of the nucleosome

The method of filtering does not provide any direct evidence of folding pathways or even guarantee that the filtered structures are physically realizable. See, for instance, the self-interaction that arises in Fig. 5. The strength of this approach is that it provides a means of systematically assigning length scales to deformations that cannot be obtained from a simple RMSD fitting in Cartesian coordinate space. Its weakness is that the length scales are only those accessible through Fourier analysis. In this regard it is important to realize that our method does not optimize the *TH*, *SH*, or Roll-Slide models to fit helical parameter data obtained from x-ray structures. However, now that we have demonstrated that two length scales are necessary and sufficient, a model that is not subject to the length requirements imposed by our Fourier filtering technique could be developed and optimized, based on a generalization of the expressions in Eq. 1.

Nonetheless, our length-scale information has enabled us to quantitatively group different realizations of the nucleosome into subfamilies based on geometric considerations. By extension, we can categorize other states of the nucleosome, for example the tetrasome, which only includes the tetrameric histone core and $\approx$56 bps of contact with the DNA. Assuming the influence of the tetramer does not extend beyond its range of physical contact with the DNA, then the long length variations of 146 and 73 basepairs that we have identified in the octasome simply do not exist in the tetrasome. Thus, according to our Fourier-filtering strategy, the distribution of DNA helical parameters in various nucleosome substates (hexasome, hemisome, or tetrasome) must be fundamentally different than the distribution found in the canonical octasome. Likewise, nucleosome arrays and condensed chromatin allow for longer length-scale variations.

The ability to systematically evaluate length-scale dependencies in the nucleosome and its various states of association/dissociation also enables us to systematically evaluate their effects on nucleosome stability. Sensitivity to these differences may enable drugs and proteins to recognize different states of the nucleosome.

The author gratefully acknowledges Professor Klaus Schulten, the Theoretical and Computational Biophysics Group, the Beckman Institute at the University of Illinois at Urbana-Champaign, and Professor Les Butler in Lousiana State University's Department of Chemistry for support during the evacuation of New Orleans.

This research was supported by National Institutes of Health grant No. R01GM076356 and the Louisiana Board of Regents Research Competitiveness Program contract No. LEQSF 2005-08-RD-A-34.

REFERENCES

1. Kornberg, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science*. 184:868–871.
2. Finch, J. T., and A. Klug. 1976. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA*. 73:1897–1901.

3. Thoma, F., T. Koller, and A. Klug. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* 83:403–427.
4. Burlingame, R. W., W. E. Love, B. C. Wang, R. Hamlin, H. X. Nguyen, and E. N. Moudrianakis. 1985. Crystallographic structure of the octameric histone core of the nucleosome at a resolution of 3.3 Å. *Science.* 228:546–553.
5. Uberbacher, E. C., and G. J. Bunick. 1989. Structure of the nucleosome core particle at 8 Å resolution. *J. Biomol. Struct. Dyn.* 7:1–18.
6. Arents, G., R. W. Burlingame, B. C. Wang, W. E. Love, and E. N. Moudrianakis. 1991. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. USA.* 88:10148–10152.
7. Richmond, T. J., T. Rechsteiner, and K. Luger. 1993. Studies of nucleosome structure. *Cold Spring Harb. Symp. Quant. Biol.* 58:265–272.
8. Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 389:251–260.
9. Schalch, T., S. Duda, D. F. Sargent, and T. J. Richmond. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fiber. *Nature.* 436:138–141.
10. Widom, J. 2001. Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* 34:269–324.
11. Davey, C. A., D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319:1097–1113.
12. Bao, Y., C. L. White, and K. Luger. 2006. Nucleosome core particles containing a poly(dA.dT) sequence element exhibit a locally distorted DNA structure. *J. Mol. Biol.* 361:617–624.
13. Suto, R. K., R. S. Edayathumangalam, C. L. White, C. Melander, J. M. Gottesfeld, P. B. Dervan, and K. Luger. 2003. Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* 326:371–380.
14. Muthurajan, U. M., Y. Bao, L. J. Forsberg, R. S. Edayathumangalam, P. N. Dyer, C. L. White, and K. Luger. 2004. Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions. *EMBO J.* 23:260–271.
15. Chakravarthy, S., S. K. Y. Gundimella, C. Caron, P.-Y. Perche, J. R. Pehrson, S. Khochbin, and K. Luger. 2005. Structural characterization of the histone variant macroH2A. *Mol. Cell. Biol.* 25:7616–7624.
16. Edayathumangalam, R. S., P. Weyermann, J. M. Gottesfeld, P. B. Dervan, and K. Luger. 2004. Molecular recognition of the nucleosomal ''supergroove''. *Proc. Natl. Acad. Sci. USA.* 101:6864–6869.
17. Wolffe, A. 1995. Chromatin Structure and Function, 2nd Ed. Academic Press, New York.
18. van Holde, K. E. 1988. Chromatin. Springer-Verlag, New York.
19. Richmond, T. J., and C. A. Davey. 2003. The structure of DNA in the nucleosome core. *Nature.* 423:145–150.
20. Bishop, T. C. 2005. Molecular dynamics simulations of a nucleosome and free DNA. *J. Biomol. Struct. Dyn.* 22:673–686.
21. Tolstorukov, M. Y., A. V. Colasanti, D. M. McCandlish, W. K. Olson, and V. B. Zhurkin. 2007. A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning. *J. Mol. Biol.* 371:725–738.
22. Harp, J. M., B. L. Hanson, D. E. Timm, and G. J. Bunick. 2000. Asymmetries in the nucleosome core particle at 2.5 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* 56:1513–1534.
23. Suto, R. K., M. J. Clarkson, D. J. Tremethick, and K. Luger. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* 7:1121–1124.
24. Tsunaka, Y., N. Kajimura, S. Tate, and K. Morikawa. 2005. Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle. *Nucleic Acids Res.* 33:3424–3434.
25. Ong, M. S., T. J. Richmond, and C. A. Davey. 2007. DNA stretching and extreme kinking in the nucleosome core. *J. Mol. Biol.* 368:1067–1074.
26. Dickerson, R. E. 1989. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.* 17:1797–1803.
27. Lu, X., and W. K. Olson. 2003. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31:5108–5121.
28. Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6: 655–667.
29. Manning, R. S., J. H. Maddocks, and J. D. Kahn. 1996. A continuum rod model of sequence-dependent DNA structure. *J. Chem. Phys.* 105: 5626–5646.
30. Chouaieb, N., A. Goriely, and J. H. Maddocks. 2006. Helices. *Proc. Natl. Acad. Sci. USA.* 103:9398–9403.
31. Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 95:11163–11168.
32. Bishop, T. C., and O. O. Zhmudsky. 2002. Mechanical model of the nucleosome and chromatin. *J. Biomol. Struct. Dyn.* 19:877–887.
33. Frigo, M., and S. G. Johnson. 2005. The design and implementation of FFTW3. *Proc. IEEE.* 93:216–231.
34. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
35. Roccatano, D., A. Barthel, and M. Zacharias. 2007. Structural flexibility of the nucleosome core particle at atomic resolution studied by molecular dynamics simulation. *Biopolymers.* 85:407–421.
36. Lankas, F., R. Lavery, and J. H. Maddocks. 2006. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure.* 14:1527–1534.
37. Dickerson, R. E. 1998. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* 26:1906–1926.
38. Crick, F. H., and A. Klug. 1975. Kinky helix. *Nature.* 255:530–533.
39. Wiggins, P. A., R. Phillips, and P. C. Nelson. 2005. Exact theory of kinkable elastic polymers. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71:021909.