Potential Function Describing the Folding of the 30 nm Fiber

Thomas C. Bishop and John E. Hearst*

Department of Chemistry, University of California, Berkeley, California 94720-1460

Received: January 27, 1998

Most of the DNA in a eucaryotic cell nucleus is accurately described as an extremely long thin fiber which is smoothly folded by histones into a thicker fiber called the 30 nm fiber. For the DNA to remain biologically functional, the 30 nm fiber must be only marginally stable. Therefore, we assume that the 30 nm fiber has fluidlike properties under physiologic conditions. It does not possess a single well-defined structure, but rather at any given instant it will possess a structure that lies within specified limits in an energy landscape. We predict an energy landscape for the 30 nm fiber by modeling the fiber as an elastic rod (DNA) subjected to an external potential (histones). For this purpose, closed-form helix-on-a-linear-helix solutions for the equilibrium configuration of an elastic rod are parametrized to fit the dimensions of the so-called solenoid model of the 30 nm fiber. We introduce a method of including a Lennard-Jones type potential function into the elastic rod model that represents histone−DNA interactions, and we parametrize the potential according to known thermodynamic data. The resulting energy landscape for the 30 nm fiber allows for structural irregularities under physiologic conditions and for unfolding of the fiber when the DNA is sufficiently over- or underwound. Our model predicts that forcing a change in the amount of twist in DNA of −6.5% or +4.5% from an intrinsic twist of 10.5 base pairs per turn is sufficient to unfold the 30 nm fiber. These results describe a self-activating mechanism which enables RNA polymerase to unfold chromatin during transcription. According to our model, formation of the open complex itself is sufficient to unfold over 200 base pairs of DNA.

Introduction

In a typical eucaryotic cell, over 1 m of DNA resides in the nucleus which has a diameter of approximately 10 μm. The DNA within the nucleus is found to be organized in successive levels of folding, each level representing a different path for the double helix.[1a] The zero level of folding is completely unfolded, linear, double-stranded DNA. The first level is the 11 nm fiber, which relies on a specific protein−DNA structure called the nucleosome.[2] The 11 nm fiber folds into a fiber with an outer diameter of 30 nm; this 30 nm fiber represents the second level of folding. During most of the cell cycle, the 30 nm fiber is folded into a third level, which is characterized by random walk/giant-loop structures.[3] Transcription mechanisms have evolved to interact specifically with the second level of folding;[4] however, primary features of the 30 nm fiber and a mechanism of transcription have yet to be determined. We present a model of the second level of folding that reproduces the known thermodynamic data and enables us to propose a mechanism of transcription.

At the first level of folding, the core of each nucleosome is well-defined. It is a complex of eight histones and DNA. The histone octamer is a three-part construction of two H2A−H2B dimers bound to opposing faces of a wedge-shaped $(H3-H4)_2$ tetramer;[5] the octamer can dissociate from DNA piece by piece.[6] First, the two dimers dissociate, leaving the tetramer bound to DNA, and then the tetramer dissociates from DNA. X-ray crystallography of stable nucleosome core particles indicates that the octamer wraps 146 base pairs (bp) of DNA into 1.75 turns of a left-handed helix; this helix has a diameter of 11 nm and a pitch of 2.8 nm.[5,7] The DNA in these nucleosome core particles exhibits nonuniform bending and twisting which define three regions of the nucleosome. The average helix repeat of DNA in a nucleosome is 10.2 bp/turn and is overwound compared to the value of 10.5 bp/turn observed for free DNA.

The segment of DNA located between two nucleosome core particles is called the linker, and it is considered part of a nucleosome. The length of the linker varies; thus, nucleosomes contain an unspecified length of DNA. The nucleosome repeat is defined as the length of the linker plus the length of DNA in one core particle (146 bp). We assume a nucleosome repeat of 167 bp, representing two turns of the DNA around the histone octamer (146 bp/1.75 turns × 2 turns/nucleosome repeat). The absolute minimum nucleosome repeat of 146 bp has been observed for yeast chromatin.[8] Longer nucleosome repeats, ranging from 175 to 230 bp, are typical.[9a]

The 11 nm fiber is an array of nucleosomes with a diameter of approximately 11 nm. Depending on the environment, the fiber is either saturated with nucleosomes and has a minimal linker and maximal histone content or subsaturated with less than optimal linker and histone content.[10] Scanning force microscopy[11a] and electron microscopy[11b,c] reveal that nucleosomes in the 11 nm fiber have an elliptic, as opposed to a circular, cross section. The aspect ratio ranges from 1.1 to 1.6. This suggests that while individual nucleosomes in an array have the same gross structure as isolated or crystallized nucleosome particles, the details of the structure may differ. In particular we propose that the tripartite structure of the octamer allows for macromolecular rearrangements.

The second level of folding, from an 11 nm fiber to a 30 nm fiber, can be induced by sufficiently high salt concentrations or by linker histones. Several models with a regular structure have been suggested for the folding of the 30 nm fiber,[1b] but

* Corresponding author.

S1089-5647(98)00909-2 CCC: $15.00
Published on Web 07/25/1998

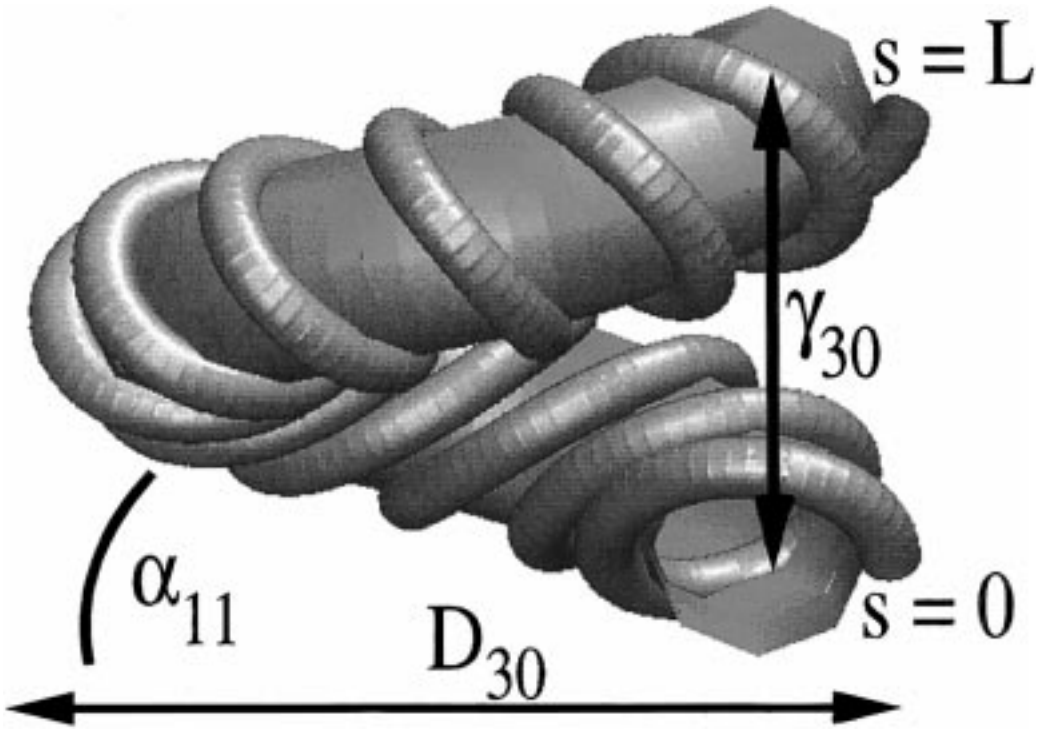

Figure 1. Helix-on-a-linear-helix. The helix-on-a-linear-helix solution is represented by a thin tube. The linear part of the solution is represented by a thick tube. Biologically, the solution describes the path of DNA in the solenoid model of the 30 nm fiber, and the linear helix represents a continuous protein core formed by the histone octamers. The diameter of the 30 nm fiber is indicated by D_{30}, the pitch by γ_{30}, and the pitch angle by α_{11}. One turn of the 30 nm fiber contains a length, *L*, of DNA. Arc position, denoted *s*, is measured from an origin located on the outer edge of the 30 nm fiber.

the fiber appears to be regular only for short segments.[4] Various features of each model are the subject of debate.[4,10,12] We interpret the variety of observed states and the lack of a definite structure as an indication that the 30 nm fiber is a dynamic entity possessing fluidlike properties.

In our model the 11 nm fiber is saturated, and the "ideal 30 nm fiber" is an absolute minimum in a relatively flat energy landscape. The absolute minimum energy must be negative for the fiber to fold spontaneously. Thermal fluctuations and a nonhomogeneous environment, i.e., physiologic conditions, are sufficient to cause the fiber to possess an irregular structure and a higher energy than the absolute minimum. A partial unfolding of the fiber corresponds to an even higher energy, but one that is still negative. Peaks in the landscape that achieve a positive energy are interpreted as complete unfolding.

The actual energy landscape is determined by three types of interactions: histone−histone, histone−DNA, and DNA−DNA interactions. These interactions are in turn influenced by the state of the histones (e.g., phosphorylated or acetylated), the state of the DNA (e.g., over/underwound or methylated), and the state of the local environment (e.g., the pH or ionic strength). Of these interactions, the DNA−DNA interactions have the longest range. The long range is a consequence of a mechanical linkage that is created by the chemical bonds of the DNA backbone. The linkage can transmit forces along the axis of the DNA unless restricted by "clamps" such as DNA binding proteins or inherently stable regions of the histone−DNA complex. The existence of topoisomerases is an indication that torsional forces are transmitted along strands of cellular DNA, and topological studies emphasize the importance of twist as a primary determinant of conformation.[13] We therefore seek to determine the effects of twist on the 30 nm fiber.

Elastic rod models are ideal for describing the mechanical linkage[14] and are applicable to DNA wherever it remains double-stranded and does not contain sharp kinks.[15] Shi and Hearst utilized an integration technique developed by Langer and Singer[16] to obtain general closed-form solutions for an isotropic elastic rod in the absence of external forces.[17] The higher-order solutions describe a helix-on-a-linear-helix and are elliptic in nature. They denote the most smoothly bent conformations of a non self-touching, uniformly twisted elastic rod, subject to specified boundary conditions. These stationary-state solutions represent a mechanical equilibrium, not a thermal equilibrium, and in the absence of external forces, these solutions are unstable. Analysis of the elastic energy associated with various conformations of the rod indicates that a left-handed helix will preferentially fold into a right-handed, as opposed to a left-handed, helix-on-a-linear-helix.[18]

These solutions provide a physical basis for the so-called solenoid model of the 30 nm fiber.[19] In this model nucleosomes are arranged in a helical fashion, and the path of DNA is described as a helix-on-a-linear-helix (Figure 1). Since this arrangement of nucleosomes avoids DNA self-contact, all of the DNA−DNA interactions are described by the elastic rod model. However, an external potential is required in the elastic rod model to represent the histone−DNA and the histone−histone interactions that stabilize the 30 nm fiber.

In eucaryotic cells, approximately 80% of the DNA is wrapped up in nucleosomes,[20] indicating that the details of the histone−DNA interactions are of secondary consideration for the structure of the 30 nm fiber in vivo. In our model the histone−histone interactions and the histone−DNA interactions are combined into a single octamer−DNA potential. This octamer−DNA potential is a first-order approximation of the complex histone−histone and histone−DNA interactions that are required to stabilize the 30 nm fiber. We interpret the sum of the octamer−DNA and the DNA−DNA interactions as the free energy of the 30 nm fiber. The model enables us to

determine the free energy of the 30 nm fiber as a function of the amount of over- or underwinding of the DNA.

Modeling the 30 nm Fiber

The helix-on-a-linear-helix stationary-state solutions are most naturally represented in cylindrical coordinates. The complete mathematical form of the solutions is presented in Shi and Hearst.[17] Here, we express the solutions as

$$\rho(s) = \rho_l + \rho_p(s) \tag{1a}$$

$$\theta(s) = \theta_l s + \theta_p(s) \tag{1b}$$

$$z(s) = z_l s + z_p(s) \tag{1c}$$

where s represents the arc position measured from a designated origin along the length of the rod. The variables ρ_l, θ_l, and z_l are constants so the first terms in eqs 1 describe a linear helix. The functions $\rho_p(s)$, $\theta_p(s)$, and $z_p(s)$ are periodic functions of elliptic functions and integrals, and the solution requires that they have the same period, τ. The conformation described by the solution is a helix-on-a-linear-helix, or simply a superhelix. In biologic terminology, the solution represents the path of DNA in the 30 nm fiber, the linear helix represents the axis of the 11 nm fiber within the context of the 30 nm fiber, and the periodic components represent the path of the DNA around the axis of the 11 nm fiber (Figure 1).

The above expression for $\rho(s)$ is different from the expression $\rho(s) = (a + bf(s))^{1/2}$ in Shi and Hearst,[17] where a and b are constants and $f(s)$ is a periodic function with period 2τ. The terms in eq 1a are related to the expression in ref 17 by the relations

$$\rho_l = (a + bf(\tau/4))^{1/2} \tag{2a}$$

$$\rho_p(s) = (a + bf(s))^{1/2} - \rho_l \tag{2b}$$

Equation 2a locates the linear helix beneath maxima in $z(s)$ and allows us to identify the linear helilx as the axis of the 11 nm fiber. The choice of ρ_l also defines an origin for the cross sections described below.

The pitch and the pitch angle of the linear helix are observable quantities of the 30 nm fiber. The pitch of the 30 nm fiber, γ_{30}, is the same as the pitch of the linear helix (Figure 1), and it is determined by the ratio $2\pi z_l/\theta_l$. If the 30 nm fiber is right-handed, then γ_{30} must be positive; if it is left-handed, then γ_{30} must be positive. The pitch angle of the linear helix is the same as the angle of inclination of the 11 nm fiber with respect to the axis of the 30 nm fiber and is denoted by α_{11} (Figure 1). It satisfies the relation

$$\tan(\alpha_{11}) = \gamma_{30}/2\pi\rho_l \tag{3}$$

Combinations of the linear and periodic components are necessary to describe other properties of the 30 nm fiber. The diameter of the 30 nm fiber, denoted D_{30}, is typically measured from the outer edges of the DNA (Figure 1), so 2 nm (the width of DNA) is added to the diameter of the superhelix.

$$D_{30} = 2(\rho_l + \max[\rho_p(s)]) + 2 \tag{4}$$

Nucleosomes, in our model, consist of two loops of DNA; thus the number of loops (n) in the superhelix for every turn of the linear helix is twice the number of nucleosomes per turn of the 30 nm fiber. The ratio $2\pi/\theta_l\tau$ is equivalent to n. The handedness of the 11 nm fiber in our model is determined by the relative phase of $\rho_p(s)$ and $z_p(s)$. If $\rho_p(s)$ and $z_p(s)$ are in phase, then the 11 nm fiber has the same handedness as the 30 nm fiber; if they are out of phase, then the handedness of the 11 nm fiber is opposite that of the 30 nm fiber.

Equations 1 provide a convenient means of locating octamers within the context of the 30 nm fiber. The centers of mass of neighboring octamers are located at successive points on the linear helix. In a uniform solenoid model of the 30 nm fiber,[19] the octamers are discrete, cylindrical objects evenly spaced along a linear helix. They are inclined by an amount α_{oct} with respect to the axis of 30 nm fiber. In our model of the 30 nm fiber the octamers are a fluid entity, so the discrete particles are replaced by a continuous octamer core. This requires $\alpha_{oct} = \alpha_{11}$.[21] The linear helix is then the axis of this octamer core, and the wrapping of DNA around the octamer core is described by the path of the superhelix around the linear helix (Figure 1).

We assume that a first-order approximation of the octamer−DNA interactions depends only on the distance between the octamer core and the DNA. Thus it is sufficient to determine the wrapping of the DNA around a cross section of the octamer core. A plot of $\rho_p(s)$ versus $z_p(s)$ in Cartesian coordinates describes the wrapping of DNA around the linear helix in a vertical plane. Since the linear helix is inclined by an amount α_{11} with respect to the z-axis, the z-component must be scaled by $\cos(\alpha_{11})$ to obtain the wrapping around a cross section of the linear helix. This cross section corresponds to the wrapping of the DNA around the octamer core, which can be envisioned as a disk in a plot of $\rho_p(s)$ versus $\cos(\alpha_{11})z_p(s)$ (Figure 2a). The radius of the cross section, $r_{nuc}(s)$, represents the radius of the nucleosome as a function of arc position:

$$r_{nuc}(s) = [\rho_p{}^2(s) + \cos^2(\alpha_{11})z_p{}^2(s)]^{1/2} \tag{5}$$

If the octamers do not stretch or twist the DNA, i.e., if the forces only act perpendicular to the axis of the DNA, then any function of the radius of the nucleosome can describe octamer−DNA interactions. We define the aspect ratio for a given cross section, denoted ϵ, as the ratio of the width to the height of the cross section:

$$\epsilon = r_w/r_h = \{r_{nuc}(\tau/2) + r_{nuc}(0)\}/\{2r_{nuc}(\tau/4)\} \tag{6}$$

Generating Solutions

The solutions of eqs 1 are a function of four parameters (w, μ, ν, Q), as described in Shi and Hearst.[17] Q is a free parameter in our model, and the remaining parameters (w, μ, ν) are determined by suitable boundary conditions. Q represents the excess twist in the rod measured relative to an intrinsic twist ($\omega_3°$) and scaled by the ratio of the twisting to the bending stiffnesses (Λ):

$$Q = \Lambda(\omega_3 - \omega_3°) \tag{7}$$

The boundary conditions require that the length of the rod (L), the pitch of the superhelix (γ_{30}), and the number of loops

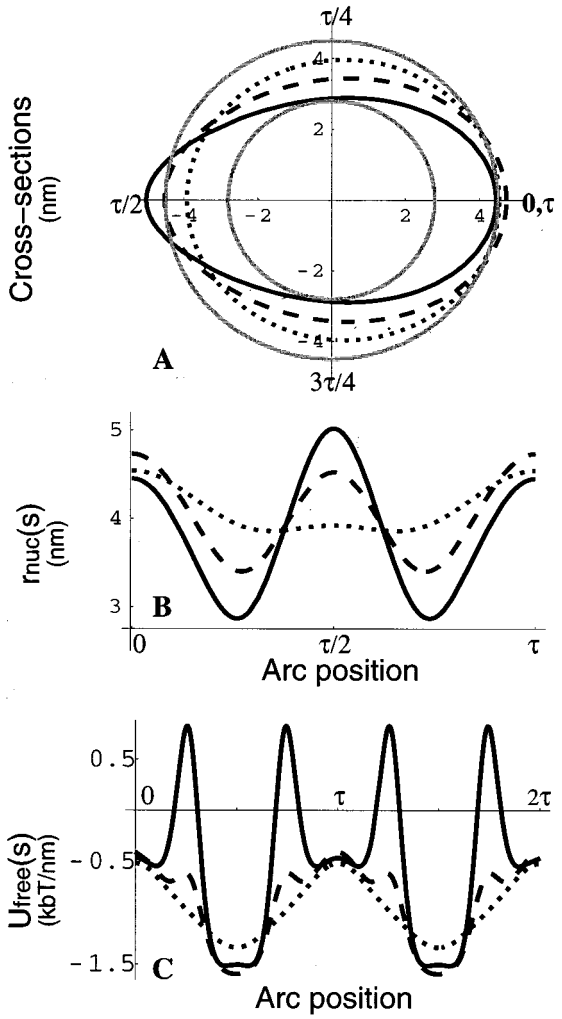

Figure 2. Cross sections. (A) Cross sections that represent the wrapping of the DNA (helix-on-a-linear-helix) around the octamer core (linear helix). The outer surface of the 30 nm fiber corresponds to arc positions $s = n\tau$ ($n = 0, 1, 2, 3, ...$). Each black ellipse represents a cross section with a specified amount of excess twist: $Q = -4\%$ (dotted), $Q = 0\%$ (dashed), and $Q = 1.5\%$ (solid). The gray circles represent equipotential surfaces of U_{AR}. The smaller circle has radius r_{zero}, and the larger circle has radius r_{min}. (B) The radius of each selected cross section, $r_{nuc}(s)$, as a function of arc position for one turn around each cross section. (C) The free energy, U_{free}, of each selected cross section as a function of arc position for two turns around each cross section.

(n) for one turn of the superhelix be specified. To determine the energy associated with any given configuration of the rod requires that a persistence length (l) be specified as well. Persistence length is a statistical measure of how far a fiber in a Brownian bath extends in a given direction if held at one end.

The material parameters that we use are similar to those in Hearst and Shi[18] and agree with various experimental data.[22] Free DNA has a diameter of 2 nm, a length of 0.34 nm/bp, and a persistence length of 55.0 nm. Our ideal nucleosome has a length of 167 bp (56.7 nm) and corresponds to two loops of the superhelix (2τ). There are six nucleosomes (12τ) per turn of the 30 nm fiber, so the length of one turn of the 30 nm fiber is 340 nm. The pitch of the 30 nm fiber is 11 nm. The average twist of DNA in stable nucleosomes is 10.2 bp/turn ($\omega_3° = 1.81$ radians/nm), and for convenience we assume that this is the intrinsic twist of DNA in the 30 nm fiber. This assumption determines a reference point for our results but does not fundamentally affect them. The ratio of the twisting to the bending stiffnesses is set equal to 1.0, so Q reduces to the excess

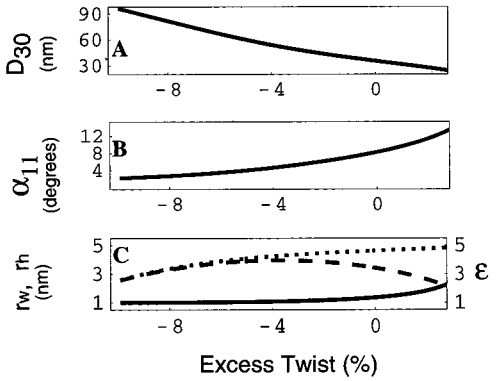

Figure 3. Geometric properties. (A) The diameter of the 30 nm fiber as a function of excess twist. (B) The pitch angle as a function of excess twist. (C) The aspect ratio (solid), width (dotted), and height (dashed) of the nucleosomes in the 30 nm fiber as a function of excess twist.

twist in the rod. It may be either positive or negative. A positive value of Q represents a twist excess, and a negative value of Q represents a twist deficit. Q is reported as a percentage of the intrinsic twist, $\omega_3°$. The actual twist in the rod is ω_3.

We impose boundary conditions on $\rho(s)$, $\theta(s)$, and $z(s)$ that produce a regular right-handed superhelix with a pitch γ_{30} by minimizing the function

$$\phi(w,\mu,\nu) = [\rho(L) - \rho(0)]^2 + [\theta(L) - \theta(0) - 2\pi]^2 + [z(L) - z(0) - \gamma_{30}]^2 \quad (8)$$

In practice the boundary condition for $\rho(s)$ is solved explicitly to obtain a value for w as a function of L, n, and μ. The two remaining parameters are determined numerically.

This boundary value method provides a means of generating a series of solutions for the 30 nm as a function of any or all of the parameters. To preserve the simplicity of the model, we vary the excess twist (Q) while keeping the other parameters and boundary conditions constant. In this manner solutions as a function of Q are obtained for a range of Q values from -10% to $+3\%$. This method produces a near linear relationship between the diameter of the 30 nm fiber and Q; D_{30} ranges from 95 nm at $Q = -10\%$ to 28 nm at $Q = +3\%$ (Figure 3a). The model thus predicts fibers that are wider than typically reported (20−50 nm).[9b] At $Q = 1.5\%$ the canonical measure of 30 nm is achieved. Maintaining a constant pitch as D_{30} decreases requires the pitch angle to increase nonlinearly; α_{11} varies from 2.3° at $Q = -10\%$ to 14.0° at $Q = 3\%$ (Figure 3b).

Since L and n are held constant as D_{30} increases (decreasing Q), the average radius of the nucleosomes must eventually decrease. In other words, since a constant number of loops per turn is maintained as the superhelix becomes wider, the size of the loops must shrink. At $Q = -4\%$ $r_{nuc}(s)$ is nearly constant ($\epsilon = 1.1$) with an average radius for the nucleosome of 4.0 nm. For Q less than -4%, ϵ approaches 1.0 asymptotically, and the average radius becomes shorter. For Q greater than -4%, r_h decreases and r_w increases so the nucleosomes possess a noncircular cross section (Figure 3c).

The elastic energy of each configuration is the sum of a bend energy, U_{bend}, and a twist energy, U_{twist}. It represents the DNA−DNA interaction energy, U_{dna}. These energies are defined in

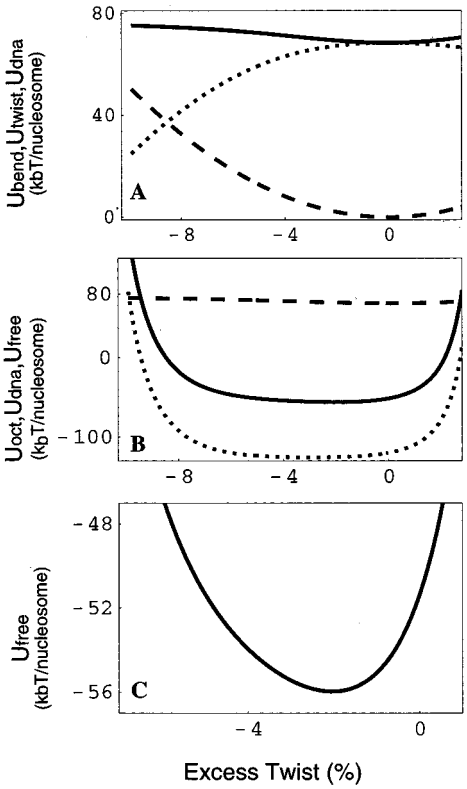

Figure 4. Energy landscape. (A) U_{dna} (solid) is the sum of the elastic energies U_{bend} (dotted) and U_{twist} (dashed) and is plotted here as a function of excess twist. (B) U_{free} (solid) is the sum of U_{dna} (dashed) and U_{oct} (dotted) and is plotted here as a function of excess twist. If $U_{free} < 0$, then the octamers and DNA will spontaneously fold into a 30 nm fiber, if $U_{free} > 0$, the 30 nm fiber unfolds. (C) An enlarged view of U_{free} as a function of excess twist near the absolute minimum.

units of k_bT per nucleosome by the expressions

$$U_{bend} = \frac{l}{2}\int_0^{2\tau} \kappa(s)^2 \, ds = 4lw^2\tau\left(\upsilon - 1 + \frac{E(\mu)}{K(\mu)}\right) \quad (9a)$$

$$U_{twist} = \frac{l}{2\lambda}\int_0^{2\tau} Q^2 \, ds = \frac{Q^2 l\tau}{\lambda} \quad (9b)$$

$$U_{dna} = U_{bend} + U_{twist} \quad (9c)$$

Here $\kappa(s)$ is the curvature of the elastic rod as expressed in Shi and Hearst[17] and $K(\mu)$ and $E(\mu)$ are the complete elliptic functions of the first and second kind, respectively.[23] U_{bend} is a quadratic, increasing function of the curvature of the DNA that is zero for linear DNA. For our solutions of the elastic rod that correspond to the 30 nm fiber, U_{bend} achieves a maximum at zero excess twist and decreases, indicating that the rod "straightens", with excess twist (Figure 4a). U_{twist} is also a quadratic, increasing function of Q that is zero at $Q = 0$ (Figure 4a). Any intrinsic twist value can be associated with the $Q = 0$ structure; we have chosen 10.2 bp/turn as the intrinsic twist. U_{bend} and U_{twist} vary by approximately $50k_bT$/nucleosome over the range of excess twist values studied, and they tend to cancel so that the value of U_{dna} varies by less than $10k_bT$/nucleosome (Figure 4a). The minimum value of U_{dna} is $68k_bT$ and occurs at $Q = 0$.

Stabilizing the 30 nm Fiber

An energy of approximately $70k_bT$/nucleosome is required to bend DNA into a 30 nm fiber. This energy is provided by favorable histone−histone and histone−DNA interactions. These interactions are combined into an octamer−DNA interaction that is a balance between electrostatic attraction, i.e., $N = 1$ in eq 10, and steric repulsion. The attractive forces bend the DNA, but if the DNA is forced into close contact with the octamer, then repulsive forces dominate. The repulsive forces can also be interpreted as disrupting the ideal conformation of the octamer or the histone−histone interactions. We use the expression for the radius of the nucleosome to introduce an external potential that describes these interactions.

These interactions are represented by the attractive−repulsive potential

$$U_{AR}(r_{nuc}) = U_o[(\sigma_1/r_{nuc})^M - (\sigma_2/r_{nuc})^N] \quad (10)$$

Here r_{nuc} is the radius of nucleosome as an implicit function of s, and U_o has units of energy per length of DNA. We define r_{zero} such that $U_{AR}(r_{zero}) = 0$ and r_{min} such that $U_{AR}(r_{min})$ is the minimum of U_{AR}. Integration of $U_{AR}(r_{nuc})$ over two loops of the 30 nm fiber yields the interaction energy between the octamer core and DNA for one nucleosome.

$$U_{oct} = \int_0^{2\tau} U_{AR}(r_{nuc}) \, ds \quad (11)$$

The total energy of the 30 nm fiber per nucleosome is defined as the sum of U_{dna} and U_{oct}. It is interpreted as the free energy of the 30 nm fiber.

$$U_{free} = U_{oct} + U_{dna} \quad (12)$$

The choice of parameters in eq 10 should produce a negative value for U_{free} at $Q = 1.5\%$ since this solution has a diameter of 30 nm. At $Q = 1.5\%$ r_w is 4.7 nm, r_h is 2.9 nm, and U_{dna} is $69k_bT$/nucleosome. If the parameters in eq 10 are such that $r_{zero} < 2.9$ nm, then U_{oct} will be negative, and a value of U_o can be chosen so that U_{oct} is sufficient to overcome U_{dna} and fold the DNA into a 30 nm fiber, i.e., $U_{free} < 0$. If the parameters are chosen to also satisfy 2.9 nm $< r_{min} <$ 4.7 nm, then the behavior of U_{oct} as a function of Q is determined.

For $Q > 1.5\%$, r_h decreases and r_w increases (Figures 2a and 3c); thus U_{oct} is an increasing function of Q. For a sufficiently large twist excess , U_{free} becomes positive and the 30 nm fiber unfolds. For $Q < -4\%$, the nucleosomes are nearly circular with $r_{nuc} < 4.0$ nm. At a sufficiently large twist deficit, $r_{nuc} < r_{zero}$ for all s; thus U_{oct} is positive. The 30 nm fiber unfolds at a smaller twist deficit when U_{free} becomes positive. There may be more than one local minimum in U_{oct} between these extremes, depending on the choice of parameters.

To determine the parameters for U_{AR}, we utilize the expressions

$$r_{zero} = \left(\frac{\sigma_2}{\sigma_1}^M\right)^{1/(1-M)} \quad (13a)$$

$$r_{min} = r_{zero}\left(\frac{1}{M}\right)^{1/(1-M)} \quad (13b)$$

which are valid for $N = 1$. For ideal nucleosome particles the distance from the center of the octamer to the axis of the DNA is 4.5 nm, so r_{min} is chosen to be 4.5 nm. In agreement with our above criteria, r_{zero} is chosen to be 2.8 nm. Using these values and eqs 13, M is determined to be 3.8. Arbitrarily setting σ_1 to 1 nm produces a value for σ_2 of 0.054 nm.

To determine the value of U_o, we require U_{free} to be $-30k_bT$/nucleosome at $Q = 1.5\%$. This choice corresponds to an

average free energy of binding between the DNA and the octamer of $-0.53k_bT$/nm, which is the least stable of the expected range of values from -0.5 to $-0.74k_bT$/nm.[4] The value of U_o obtained in this manner is $255.6k_bT$/nm, resulting in a minimum octamer−DNA interaction, $U_{AR}(r_{min})$, of $-2.3k_bT$/nm or equivalently $-0.78k_bT$/bp. The latter value clearly indicates that the octamer−DNA interactions in our model are fluid under physiologic conditions. The stability of the 30 nm fiber results from the extent of contact between the octamer and the DNA not the stability of individual contacts.

Using these parameters, we evaluate U_{oct} and U_{free} as functions of excess twist (Figure 4b). The results indicate that the value of U_{free} is negative between $Q = -8.5\%$ and $Q = 2.0\%$, and the 30 nm fiber folds spontaneously. A twist excess or deficit beyond this range of values corresponds to positive values of U_{free} and thus unfolding of the 30 nm fiber. The diameter of the 30 nm fiber associated with this range of Q values varies from 85 to 28 nm. The absolute minimum of U_{free} occurs at $Q = -2.0\%$ (Figure 4c) and has a diameter of 44 nm. Thus our model predicts that changes in excess twist, denoted ΔQ, of -6.5% or 4.5% are sufficient to unfold an ideally folded 30 nm fiber. At $Q = -2\%$ U_{free} has a value of $-56k_bT$/nucleosome or $-0.99k_bT$/nm, which predicts tighter binding than expected.[4]

If under typical conditions we assume the free energy of the 30 nm fiber is within $10k_bT$/nucleosome of the minimum (Figure 4c), then the fiber can achieve a range of diameters between 67 and 32 nm (Figure 3a). This energy difference is the equivalent of a few hydrogen bonds or electrostatic interactions per nucleosome, so the energy landscape for the 30 nm fiber in our model allows structural irregularities under physiologic conditions. These irregularities are "fluid" and continually appear and disappear along the length of an actual 30 nm fiber.

When $Q = 1.5\%$, the free energy as a function of arc position indicates three distinct regions for each nucleosome (Figure 2c). For lower values of Q, there remain three regions that can be identified by two local minima ($s = \tau/2, 3\tau/2$) and a maxima ($s = \tau$). These regions arise from the elliptic nature of the cross section and may be interpreted as the three regions observed in isolated nucleosomes or as a preferred orientation of the octamers in the 30 nm fiber. The latter is supported by the observation that linker histones bind on the interior of the 30 nm fiber.[24] In our model, the arc position is defined such that the most stable regions are located on the interior surface of the 30 nm fiber. The choice of p_l affects the difference in free energy between the minima ($s = \tau/2, 3\tau/2$) and maxima ($s = 0, \tau$).

Finally, the external potential will perturb the stationary-state solutions. The results presented here are therefore an approximation to the actual energy and conformation of an ideal elastic rod in such an external potential. Since the external potential is a radially symmetric function about the axis of the linear helix, the actual solutions will tend toward solutions with a circular cross section. Such a solution is the asymptotic limit of twist deficit in our model, so the actual solutions will yield results that are shifted toward a larger twist deficit (lower Q value). Currently these solutions can only be obtained numerically.

Conclusions

We have demonstrated how an attractive−repulsive potential can be incorporated into an elastic rod model of DNA to represent the 30 nm fiber. The resulting energy landscape is a first-order approximation that allows for irregularities in the structure of the 30 nm fiber under physiologic conditions. The actual landscape contains higher-order effects that can be superimposed on our first-order approximation. The nonuniform bending and twisting of DNA observed for isolated nucleosomes,[5,7] which are designed for maximum stability in vitro, represent a path through the actual energy landscape that is routed as nearly as possible to the absolute minimum. These minimal structures have been exploited to achieve nucleosome positioning in vitro[25] and may also be exploited in vivo for specific purposes,[26,27] but do not represent the prevalent conformation of nucleosomes in nature.

The most fundamental result of our model is that a twist excess or deficit can unfold the 30 nm fiber. The model predicts that a difference in excess twist (ΔQ) of -6.5% or $+4.5\%$ from the ideally folded 30 nm fiber ($Q = -2\%$) will result in unfolding of the 30 nm fiber. The unfolding results from a change in the shape of the cross section of the nucleosomes in the 30 nm fiber. The solutions obtained by Shi and Hearst[17] dictate that when the 11 nm fiber is folded into a 30 nm fiber, increasing the excess twist (increasing Q) will produce a more elliptic cross section for the nucleosomes, and decreasing the excess twist (decreasing Q) will produce a more circular cross section for the nucleosomes. This provides a reasonable explanation for the evolution of the histone octamer as a three-part assembly; such an assembly will readily conform to nucleosomes with an aspect ratio that varies. The solutions also dictate that the first level of folding of an elastic rod is helical and that the second level of folding is superhelical, with only the latter possessing elliptic properties. Thus, the experimental observation both of a circular cross section for isolated nucleosomes (first level folding)[5,7] and an elliptic cross-section for arrays of nucleosomes (second level folding)[11] is a tell-tale sign that DNA behaves as an elastic rod.

Our model predicts how a twist excess or deficit may be transmitted over long distances in the 30 nm fiber. For a rather extensive range of twist excess or deficit, there is a small energetic cost and a limited deformation of the 30 nm fiber; however, for a sufficiently large twist excess or deficit the energetic cost rapidly increases, producing a narrow transition region between a folded and an unfolded fiber. In the absence of clamps, a twist excess or deficit will dissipate along the DNA without significantly disrupting the structure of the 30 nm fiber, while appropriately spaced clamps may create a twist excess or deficit that partially or totally unfolds the 30 nm fiber in a local region.

We use these results to propose a mechanism of action of RNA polymerase. Formation of the open complex requires that 10−11 bp are completely unwound by RNA polymerase II.[28] Assuming that the initiation complex serves as a clamp that forces this unwinding of the DNA to be transmitted as a twist excess ahead of the polymerase, then the maximum number of base pairs that can be unfolded downstream from the polymerase is 244 (11 bp/4.5%). Interestingly, there is more than enough twist excess to unfold one nucleosome but not enough twist excess to unfold two nucleosomes. In the absence of a second clamp or other torsional restraints, the twist excess will diffuse and the unfolded nucleosome will spontaneously refold into the 30 nm fiber. However, during transcription, the movement of polymerase along DNA becomes a source of twist excess, continually increasing the amount of excess twist downstream from the polymerase. As long as the dissipation of the twist excess does not "outrun" this source, the unfolding of the 30 nm fiber required for transcription is self-activating.

Acknowledgment. We would like to thank Yaoming Shi for his assistance during the development of this model, Annalisa Calini and Jonathon Widom for reading an early version of the manuscript, and Karen Elliott for editorial assistance. This work was supported by the Office of Basic Energy Science in the Biological Energy Division of the Department of Energy under contract number DE-AC03-76SF00098 and the National Institutes of Health, grant number GM41911.

References and Notes

(1) (a) Tsanev, R.; Russev, G.; Pashev, I.; Zlantanova J. *Replication and Transcription of Chromatin*; CRC Press: Boca Raton, 1993; Chapter 1. (b) *Ibid*., p 22.

(2) Kornberg, R. *Science* **1974**, *184*, 868.

(3) Sachs, R.; Vandenengh, G.; Trask, B.; Yokota, H.; Hearst, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2710.

(4) Felsenfeld, G. *Cell* **1996**, *86*, 13.

(5) Arents, G.; Burlingame, R. W.; Wang, B. C.; Love, W. E.; Moudrianakis, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10148.

(6) Eickbush, T. H.; Moudrianakis, E. N. *Biochemistry* **1978**, *17*, 4955.

(7) Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. *Nature* **1997**, *389*, 251.

(8) Lowary, P. T.; Widom, J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 10004.

(9) (a) van Holde, K. *Chromatin*; Springer-Verlag: New York, 1989; pp 292−301. (b) *Ibid*., pp 327−329.

(10) Fletcher, T.; Hansen, J. *Crit. Rev. Eukaryotic Gene Expression* **1996**, *6*, 149.

(11) (a) Fritzsche, W.; Henderson, E. *Biophys. J.* **1996**, *71*, 2222. (b) Oliva, R. D.; Bazett-Jones, D. P.; Locklear, L.; Dixon, D. H. *Nucleic Acids Res.* **1990**, *18*, 2739. (c) Zabal, M. M. Z.; Czarnota, G. J.; Bazett-Jones, D. P.; Ottensmeyer, F. P. *J. Microsc.* **1993**, *172*, 204.

(12) van Holde, K.; Zlantanova, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10548.

(13) (a) Fuller, F. B. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 815. (b) White, J. H.; Bauer, W. R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 772.

(14) (a) Love, A. *A Treatise on the Mathematical Theory of Elasticity*, 4th ed.; Dover Publications: New York, 1944; Article 254. (b) Landau, L.; Lifshitz, E. *Theory of Elasticity*, 3rd ed.; Permagon: New York, 1986.

(15) (a) Benham, C. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2397. (b) Schlick, T.; Olson, W. *J. Mol. Biol.* **1992**, *223*, 1089.

(16) Langer, J.; Singer, D. A. *J. London Math. Soc.* **1984,** *30*, 512.

(17) Shi, Y.; Hearst, J. E. *J. Chem. Phys.* **1994,** *101*, 5186.

(18) Hearst, J. E.; Shi, Y. *Nonlinear Sci. Today* **1995,** *5*, 1.

(19) Thoma, F.; Koller, Th.; Klug, A. *J. Cell Biol.* **1979,** *83*, 403.

(20) Noll, M. *Nuc. Acids Res.* **1974,** *1*, 1573.

(21) Conversely, inclining the octamers with respect to the linear helix, i.e., α_{oct} not equal to α_{11}, requires that the octamers be discrete entities with a specified location for the center of mass of each octamer and a specified beginning and ending arc position to limit the extent of each octamer. Our approximation avoids these complexities and is consistent with a fluid structure of the 30 nm fiber.

(22) Wolffe, A. *Chromatin Structure and Function*, 2nd ed.; Academic Press: San Diego, 1995.

(23) Bowman, F. *Introduction to Elliptic Functions with Applications*; Wiley & Sons: New York, 1953; Chapter 2.

(24) Graziano, V.; Gerchman, S.; Schneider, D.; Ramakrishnan, V. *Nature* **1994**, *368*, 351.

(25) Travers, A. A. *Trends Biochem. Sci.* **1987**, *12*, 108.

(26) Lowary, P. T.; Widom, J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1183.

(27) Bishop, T. C.; Kosztin, D.; Schulten, K. *Biophys. J.* **1997,** *72*, 2056.

(28) Gamper, H. B.; Hearst, J. E. *Cold Spring Harbor Symp. Quantum Biol.* **1983**, *XLVII*, 455.