The Free Energy Landscape of Inter-Nucleosome Interactions and its Relation to Chromatin Fiber Structure

Joshua Moller,† Joshua Lequieu,‡* and Juan J. de Pablo†,¶

†Institute for Molecular Engineering
University of Chicago, Chicago, IL 60637 USA
‡Chemical Engineering
University of California Santa Barbara, Santa Barbara, CA 93106 USA
¶Materials Science Division
Argonne National Laboratory, Argonne, IL 60439 USA

E-mail: depablo@uchicago.edu

3SPN.2C Model

The 3SPN.2C model is a coarse-grained model of DNA that represents each nucleotide as three beads: one for the phosphate, one for the base, and one for the sugar. We restate the model parameters to make this work self-contained. We use a combination of bond, angle,
and dihedral forces to preserve the B-DNA structure. These bonded forces are

\[ U_b = U_{bond} + U_{ang} + U_{dihe} \]

\[ = \sum_{i} k_b (r_i - r_{eq,i})^2 + 100k_b (r_i - r_{eq,i})^4 \]

\[ + \sum_{i} k_\theta (\theta_i - \theta_{eq,i})^2 \]

\[ + \sum_{i} -k_\phi \exp \left( \frac{-(\phi_i - \phi_{eq,i})^2}{2\sigma_{\phi,i}^2} \right), \]

where \( k_b \) and \( r_{eq,i} \) are the force constant and the equilibrium bond length for the \( i^{th} \) bond, \( k_\theta \) and \( \theta_{eq,i} \) are the force constant and equilibrium angle for the \( i^{th} \) angle, and \( k_\phi, \phi_{eq,i}, \) and \( \sigma_{\phi,i} \) are the well depth, equilibrium angle, and well-width for the \( i^{th} \) dihedral. In this model, dihedrals are only applied to the backbone of the system, which represents the sugars and the phosphates.

There is also an added dihedral function to the DNA backbone in the form of

\[ U_{\phi,periodic} = k_{\phi,periodic} [1 + \cos(\theta - \theta_0)] \]

The non-bonded potentials can be broken down into electrostatic interactions, excluded volume interactions, and base-pairing interactions. The full non-bonded set of contributions is:

\[ U_{nonbond} = U_{exclude} + U_{elec} + U_{bp} + U_{cs} + U_{bs} \]

The excluded volume interaction, \( U_{exclude} \), follows a purely repulsive Lennard-Jones interaction of the form

\[ U_{exclude} = \sum_{i\neq j} \begin{cases} 
\epsilon_r \left[ \frac{(\sigma_{r_{ij}})}{r_{ij}} \right]^{12} - 2 \left[ \frac{(\sigma_{r_{ij}})}{r_{ij}} \right]^{6} & r < r_{Cut} \\
0 & r \geq r_{Cut} \end{cases} \]
In this case, the cutoff, \( r_{cut} = \sigma_{ij} \) is the arithmetic average of the size of the interacting particles. The potential assumes interactions between all sites that are not bonded together or interacting by a base-pair non-bonded potential. The electrostatic interactions occur between the charged phosphates and residues on the proteins. We use the implicit, Debye-Hückel screened electrostatic potential of the form:

\[
U_{\text{elec}} = \sum_{i<j} \frac{q_i q_j e^{-r_{ij}/\lambda_D}}{4\pi\epsilon_0 \epsilon(T,C) r_{ij}},
\]

where \( q_i \) and \( q_j \) are the charges on the \( i^{th} \) and \( j^{th} \) sites, \( \lambda_D \) is the Debye length, and \( \epsilon(T,C) \) is the dielectric permittivity of the solution.

The base-pairing interactions can be broken into three interactions: base-pair \( (U_{bp}) \), base-stacking \( (U_{\text{base-stack}}) \), and cross-stacking interactions \( (U_{\text{cross-stack}}) \). All three rely on a Morse potential of the form

\[
U_{\text{Morse}}(\epsilon_{ij}, \alpha_{ij}, r_{ij}) = \epsilon_{ij}(1 - e^{-\alpha_{ij}(r_{ij} - r_{eq,ij})})^2 - \epsilon_{ij}
\]

which is broken down into a repulsive component,

\[
U_{\text{Morse}}^r(\epsilon_{ij}, \alpha_{ij}, r_{ij}) = \begin{cases} 
\epsilon_{ij}(1 - e^{-\alpha_{ij}(r_{ij} - r_{eq,ij})})^2 & r < r_{eq,ij} \\
0 & r \geq r_{eq,ij}
\end{cases}
\]

and an attractive component,

\[
U_{\text{Morse}}^a(\epsilon_{ij}, \alpha_{ij}, r_{ij}) = \begin{cases} 
-\epsilon_{ij} & r < r_{eq,ij} \\
\epsilon_{ij}(1 - e^{-\alpha_{ij}(r_{ij} - r_{eq,ij})})^2 - \epsilon_{ij} & r \geq r_{eq,ij}
\end{cases}
\]

For these functions, \( \epsilon_{ij} \) is the well depth of the interaction, \( r_{eq,ij} \) is the equilibrium distance of the interaction, and \( \alpha_{ij} \) is a parameter that controls the range of the attraction. We
also incorporate a modulating function $f$ to the angles of interaction in the cross-stacking, base-stacking, and base-pairing interaction to smoothly scale the interactions between non-hydrogen bonded and hydrogen bonded base pairs. The modulating function is of the form

\[
f(K, \Delta \theta) = \begin{cases} 
1 & -\frac{\pi}{2K} < \Delta \theta < \frac{\pi}{2K} \\
1 - \cos^2(K \Delta \theta) & -\frac{\pi}{K} < \Delta \theta < -\frac{\pi}{2K} \quad \text{or} \quad \frac{\pi}{2K} < \Delta \theta < \frac{\pi}{K} \\
0 & \Delta \theta < -\frac{\pi}{K} \quad \text{or} \quad \Delta \theta > \frac{\pi}{K}
\end{cases}
\] (8)

where $K$ is a modulating constant depending on the type of interaction. With these definitions in place, we can fully describe the force field of interaction of a base pair, which is

\[
U_{bp} = \sum_{n_{BP}} \left\{ \begin{array}{l}
U_{\text{rep}}^{\text{Morse}}(\epsilon_{ij}, \alpha_{BP}, r_{ij}) \\
\frac{1}{2} (1 + \cos(\Delta \phi_1)) f(K_{BP, \Delta \theta_{1,ij}}) f(K_{BP, \Delta \theta_{2,ij}}) U_{\text{Morse}}^a(\epsilon_{ij}, \alpha_{BP}, r_{ij}) & r_{ij} < r_{eq,ij} \\
\frac{1}{2} (1 + \cos(\Delta \phi_1)) f(K_{BP, \Delta \theta_{1,ij}}) f(K_{BP, \Delta \theta_{2,ij}}) U_{\text{Morse}}^a(\epsilon_{ij}, \alpha_{BP}, r_{ij}) & r_{ij} \geq r_{eq,ij}
\end{array} \right.
\] (9)

In this case, $\theta_1$ is defined as the angle between the sense-strand sugar, base and the anti-sense base, while $\theta_2$ is the same angle but on the opposite strand of DNA, $\phi_1$ is the dihedral between the sugar and base on the sense and anti-sense strands of DNA. When two base pairs are hydrogen bonded, we also consider the pi-stacking interactions in the form of base-stacking and cross-stacking interactions.

The base-stacking interaction is

\[
U_{bs} = \sum_{n_{BS}} \left\{ \begin{array}{l}
U_{\text{rep}}^{\text{Morse}}(\epsilon_{ij}, \alpha_{BS}, r_{ij}) + f(K_{BS, \Delta \theta_{BS,ij}}) U_{\text{Morse}}^a(\epsilon_{ij}, \alpha_{BS}, r_{ij}) & r_{ij} < r_{eq,ij} \\
f(K_{BS, \Delta \theta_{BS,ij}}) U_{\text{Morse}}^a(\epsilon_{ij}, \alpha_{BS}, r_{ij}) & r_{ij} \geq r_{eq,ij}
\end{array} \right.
\] (10)
where the angle, \( \theta_{BS} \), is defined by the angle between the vector connecting the 5’ direction sugar and base and the vector connecting the two bases in the 3’ direction.

The cross-stacking interaction is

\[
U_{cs} = \sum_{i,j} f(K_{BP}, \Delta\theta_{3,ij}) f(K_{CS}, \Delta\theta_{CS,ij}) U_{Morse}^{a}(\epsilon_{ij}, \alpha_{CS}, r_{ij})
\]  

(11)

where \( \theta_{3} \) is the angle between the vectors connecting the sugars to the bases in a W-C base pair and \( \theta_{CS} \) is the vector connecting the sugar to the base, and the vector connecting the base on the anti-sense strand in the 5’ to the base in the 3’ direction on the sense strand. For completeness, we use the following parameters for the DNA model.

Table 1: 3SPN.2C Parameters

| Parameter | Value |
|-----------|-------|
| \( k_{b} \) | 6.0 kJ/mol/A^2 |
| \( k_{\theta} \) | 200 kJ/mol/rad^2 |
| \( k_{\phi} \) | 6.0 kJ/mol |
| \( \epsilon_{r} \) | 1.0 kJ/mol |
| \( K_{BS} \) | 6.0 |
| \( \alpha_{BS} \) | 3.0 |
| \( K_{CS} \) | 8.0 |
| \( \alpha_{CS} \) | 4.0 |
| \( K_{BP} \) | 12.0 |
| \( \alpha_{BP} \) | 2.0 |
| \( \sigma_{AT} \) | 5.82 Å |
| \( \sigma_{GC} \) | 5.52 Å |
| \( \epsilon_{AT} \) | 16.37 kJ/mol |
| \( \epsilon_{GC} \) | 20.73 kJ/mol |
Table 2: Base-stacking and cross-stacking energies for 3SPN.2C from Freeman et al. Section (a) describes base-stacking energy scales. Sections (b) and (c) describe cross-stacking energy scales. Upward-pointing arrows denote the sense strand while downward-pointing arrows denote the anti-sense strand (for cross-stacking interactions).

|                | $\epsilon_{ij}$ (kJ/mol) | A   | T   | G   | C   |
|----------------|--------------------------|-----|-----|-----|-----|
| **Base $5' \uparrow$** |                          |     |     |     |     |
| **Base $3' \uparrow$** |                          |     |     |     |     |
| A               | 13.82                    | 15.05 | 13.32 | 15.82 |
| T               | 9.15                     | 12.44 | 9.58  | 13.11 |
| G               | 13.76                    | 14.59 | 14.77 | 15.17 |
| C               | 9.25                     | 12.42 | 8.83  | 14.01 |

|                | $\epsilon_{ij}$ (kJ/mol) | A   | T   | G   | C   |
|----------------|--------------------------|-----|-----|-----|-----|
| **Base $5' \uparrow$** |                          |     |     |     |     |
| **Base $3' \uparrow$** |                          |     |     |     |     |
| A               | 1.882                    | 2.388 | 2.439 | 1.680 |
| T               | 2.388                    | 1.882 | 2.187 | 2.566 |
| G               | 2.439                    | 2.187 | 3.250 | 0.972 |
| C               | 1.680                    | 2.566 | 0.972 | 4.135 |

|                | $\epsilon_{ij}$ (kJ/mol) | A   | T   | G   | C   |
|----------------|--------------------------|-----|-----|-----|-----|
| **Base $5' \uparrow$** |                          |     |     |     |     |
| **Base $3' \uparrow$** |                          |     |     |     |     |
| A               | 1.882                    | 2.388 | 2.566 | 2.187 |
| T               | 2.388                    | 1.882 | 1.680 | 2.439 |
| G               | 2.566                    | 1.680 | 4.135 | 0.972 |
| C               | 2.187                    | 2.439 | 0.972 | 3.250 |
AICG Model

The AICG model of the histones employ a Go-like interaction force field. We define the potentials below.

\[
U_{AICG} = U_{\text{bond}} + U_{\text{ang}} + U_{\text{dih}} + U_{\text{natv}}
\]

\[
= \sum_i k_b (r_i - r_{eq,i})^2
\]

\[
+ \sum_i k_{\phi} (\theta_i - \theta_{eq,i})^2
\]

\[
+ \sum_i k_{\phi,1} [1 - \cos (\phi_i - \phi_{eq,i})] + k_{\phi,3} [1 - \cos 3(\phi_i - \phi_{eq,i})]
\]

\[
+ \sum_{i<j-3} \epsilon_{\text{go}} \left[ 5 \left( \frac{r_{ij,0}}{r_{ij}} \right)^{12} - 6 \left( \frac{r_{ij,0}}{r_{ij}} \right)^{10} \right]
\]

\[
+ \sum_{i<j-3} \epsilon_{\text{excl}} \left( \frac{r_{ij,ex}}{r_{ij}} \right)^{12}
\]

The parameters are amino-acid dependent and are calculated using the CafeMol simulation package.

Description of Histone Groups

For a consistent definition of the restraints for the pair potentials, the groups are comprised of residues from the histone core. There are three unique groups of residues and one set of groups for each nucleosome. All sites displayed correspond to both copies of the histone, unless otherwise noted. (Eg. residue 63 references the 63rd residue on both H3 and H3')

Here, the ('') denotes the second instance of that specific histone in the octamer. These residues hold for all simulations, including tail removed/modified runs. The groups are displayed in the following table:

From these groups, we define the vectors for the restraints. The groups were chosen to
create as much of an orthogonal definition as possible. We show in the following table which

groups are involved in each restraint and the values of the angle between those restraints to
hold the orientations for sampling.

Table 4: Definition of Nucleosome Restraint Vectors From Residue Groups

| Vector | Group 1     | Group 2     |
|--------|-------------|-------------|
| \( \mathbf{u}_i \) | Nucl \(_i\) COM | Nucl \(_i\) Dyad |
| \( \mathbf{v}_i \) | Nucl \(_i\) COM | Nucl \(_i\) Ortho |
| \( \mathbf{u}_j \) | Nucl \(_j\) COM | Nucl \(_j\) Dyad |
| \( \mathbf{v}_j \) | Nucl \(_j\) COM | Nucl \(_j\) Ortho |
| \( \mathbf{r}_{ij} \) | Nucl \(_i\) COM | Nucl \(_j\) COM |

The restraints of the two nucleosomes are then characterized by harmonic springs placed
between angles of the above vectors. For the orientations in this study, the restraint angles
were calculated from the initial time step of the nucleosomes at these orientations. The
restraints were then put in place on the following angles:

Table 5: Definition of Harmonic Restraint Angles

| Angle Definition | Vector 1 | Vector 2 |
|------------------|----------|----------|
| \( \theta_1 \)   | \( \mathbf{u}_i \) | \( \mathbf{r}_{ij} \) |
| \( \theta_2 \)   | \( \mathbf{u}_j \) | \( \mathbf{r}_{ij} \) |
| \( \theta_3 \)   | \( \mathbf{v}_i \) | \( \mathbf{r}_{ij} \) |
| \( \theta_4 \)   | \( \mathbf{v}_j \) | \( \mathbf{r}_{ij} \) |

The values of the angles for the four orientations are defined as:

**Acetylation of the H4 Tail**

We considered multiple effects for modifications of the H4 Tail. The effects of the H4 tail
shown in S1 highlight the energetic contributions.
Table 6: Definition of Nucleosome-Nucleosome Orientations for Pair-Potential Calculations

| Orientation | $\theta_{1,eq}$ | $\theta_{2,eq}$ | $\theta_{3,eq}$ | $\theta_{4,eq}$ |
|-------------|-----------------|-----------------|-----------------|-----------------|
| A           | 99.0°           | 99.0°           | 104.0°          | 104.0°          |
| B           | 99.0°           | 99.0°           | 14.0°           | 14.0°           |
| C           | 98.65°          | 8.87°           | 104.48°         | 93.01°          |
| D           | 87.55°          | 89.18°          | 13.74°          | 14.35°          |

Figure S1: The free energy surface for different simulation methods of acetylation.
Acetylation of the lysine residues of the H4 tail and removal of the H4 tail show no quantitative or qualitative change on the free energy surface. For this reason, we choose to use the acetylated tail to represent acetylations and cut-off tails in the manuscript. This result is consistent with the experimental observations of Funke et al.⁵

**Counterion Condensation**

We use the Manning condensation electrostatic definition to quantitatively compare our results to experiment. The interactions drop proportionally for all orientations. The “rotated-stack” orientation exhibits a shift to 87.7Å with an attraction of 1.80$k_B T$, the “side-side” shifts to 124.7Å with a weak interaction of 0.82$k_B T$, and the “rotated-side” shifts to 123.6Å with a weak attraction of 0.28$k_B T$.

![Figure S2: Free energy surface of nucleosome interactions with A) Manning condensation electrostatics and B) Manning condensation with an acetylated H4 tail.](image)

Applying the same analysis for the Debye-Hückel system here, we look at the histone coordination probability of the tails. As can be seen in Fig. S2, the coordination probability decreases for every tail. Comparing to the Debye-Hückel result, the highest probability still
corresponds to the H4 tail at 33 percent. We see no relative change of the tails, except now we note that the relative contribution of the H2A tail is approximately equivalent to that of the H4 tail. In Fig. S3 D, the fraction of H4 contacts with the opposite nucleosome histone is lower than that of the H2A tail. We attribute this to the DNA versus histone contributions of the H4 tail. It can be seen that the relative coordination of the H4 is lowered when considering the fraction of nucleosome contacts, even in the Debye-Hückel case, suggesting a larger contribution of DNA contacts. As the interactions between DNA and histones are reduced in the implementation of Manning condensation, the free energy magnitude is reduced in this range.

Figure S3: A) Tail interactions for all positive residues B) Histone fraction of tail interactions for all positive residues C) Tail interactions with counterion condensation D) Histone fraction of tail interactions with counterion condensation

References

(1) Hinckley, D. M.; Freeman, G. S.; Whitmer, J. K.; De Pablo, J. J. An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynam-
ics, and dynamics of hybridization. *Journal of Chemical Physics* **2013**, *139*, 144903–1–144903–15.

(2) Freeman, G. S.; Hinckley, D. M.; Lequieu, J. P.; Whitmer, J. K.; De Pablo, J. J. Coarse-grained modeling of DNA curvature. *Journal of Chemical Physics* **2014**, *141*, 165103–1–165103–8.

(3) Li, W. F.; Wolynes, P. G.; Takada, S. Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 3504–3509.

(4) Kenzaki, H.; Koga, N.; Hori, N.; Kanada, R.; Li, W.; Okazaki, K. I.; Yao, X. Q.; Takada, S. CafeMol: A coarse-grained biomolecular simulator for simulating proteins at work. *Journal of Chemical Theory and Computation* **2011**, *7*, 1979–1989.

(5) Funke, J. J.; Ketterer, P.; Lieleg, C.; Schunter, S.; Korber, P.; Dietz, H. Uncovering the forces between nucleosomes using DNA origami. *Science Advances* **2016**, *2*, e1600974–e1600974.