'Modeling Biomolecules' (S. Mielke, Mar 2016)
Notes

Slide 2:
One of my main research interests is the functional role of biomolecular structure and dynamics, particularly with respect to the biopolymers, DNA and proteins. Biopolymers—other examples of which include RNA and polysaccharides—are biological macromolecules comprised of long chains of covalently bonded, monomeric subunits, which, in the case of nucleic acids and proteins, contain reactive side-groups. That DNA, RNA, and proteins are, fundamentally, robust, highly flexible polymers characterized by incredible biophysical versatility, is central to their role as the basic molecules of life.

(Double-stranded) DNA consists of two interwound nucleotide chains in which base pairs embodying the genetic code are formed according to their chemical complementarity.

Proteins consist of a long peptide-chain, whose component amino acids, through a variety of chemically diverse side-groups, determine the precise three-dimensional structure of the folded protein, which is intimately related to its biological function.

Example of protein structure - function relationship: Insulin / receptor binding

Slide 3:
Other roles of molecular structure in biology include the influence of protein environments on electrochemistry, and the formation of higher-order geometries in order to efficiently pack (in humans) a meter or more of DNA into the nucleus of a cell.

Protein electrochemistry: My research in this area involves, for example, using known protein sequences and structures to obtain homology-based structural models of protein complexes whose 3-D geometries are unknown. With this information, one can calculate, e.g., redox and binding energies of electron-transport cofactors, and thereby assess the influence of specific amino-acid substitutions on redox chemistry.

'DNA Packaging':
http://www.nature.com/nrneph/journal/v6/n6/images/nrneph.2010.55-f2.jpg

Slide 4:
Another function of higher-order geometries -- and the polymeric qualities -- of DNA (supercoiled or superhelical DNA) is to facilitate localized transformations of the double-helix to alternative forms essential for genetic regulation. An important example of this is torsional stress-driven strand separation (double-strand 'melting') localized at gene promoters or transcription-factor binding sequences (see region indicated by arrow in Figure "d").

My work in this area (e.g., Tanaka, et al., 2008) includes equilibrium statistical mechanics calculations predicting the locations, extents, and destabilization energies (probabilities) of these transitions, given a DNA base sequence and the level of superhelical stress.

'DNA Superhelicity':
http://www.nature.com/nrneph/journal/v6/n6/images/nrneph.2010.55-f2.jpg

Slide 5:
In a living cell, processes involving DNA superhelicty are dynamic, typically driven by interactions between DNA and proteins that unfold in real time. Such a process – the 'twin supercoiling domain' model – was proposed by J. Wang and co-workers ca.1987. In this model (see figure on lower left), transcription by RNA polymerase(s) (RNAP) generates torques that actively supercoil the template DNA.

Many biological examples of this process have since been discovered. A notable example is regulation of the *c-myc* proto-oncogene in humans. The product of this gene is a multifunctional transcription factor involved in cell cycle progression, apoptosis, and cellular transformation [http://www.ncbi.nlm.nih.gov/gene/17869]. Its disregulation has been implicated in various cancers. Expression of *c-myc* by RNAP (see figure on lower right), as in the twin domain model, dynamically supercoils the DNA template, generating superhelical stress sufficient to induce strand separation localized 1,700 base pairs upstream at the 'far-upstream element' (FUSE), which promotes binding of single strand-specific regulatory proteins – FUSE binding protein (FBP) and FBP interacting repressor (FIR) – that help control expression of the gene [e.g., Kouzine, et al., 2008].

Twin Domain: http://www.cell.com/cell/pdf/0092-8674%2888%2990163-8.pdf

Slide 6:
One way to model the dynamics of biomolecular processes is computer simulation, which, in general, employs numerical methods to solve equations of motion describing the time evolution of systems characterized by various forces (potentials).

My research efforts have also applied simulation methods to investigate DNA superhelicity-mediated regulatory processes, such as *c-myc* expression. Because these processes occur on relatively large space- and time-scales (typically milliseconds and longer), and are subjected to thermal fluctuations characterized by the finite temperature of the intracellular environment, a coarse-grained, Brownian dynamics formulation is used. In the equations shown, which are iteratively evaluated to obtain the time course of particle positions and rotation angles, the terms $\boldsymbol{R}_i$ and f_i capture thermal fluctuations as stochastic (random, Gaussian) translational and rotational displacements over each time-step.

Slide 7:
Mielke, et al., 2004: 24-microsecond trajectory of twin supercoiling in torsionally constrained, 150 base-pair DNA domains (single-chain model)

Mielke, et al., 2005: 0.5-millisecond trajectory of supercoiling and localized strand separation in a closed-circular, 150 base-pair domain (double-chain model)

Slide 8:
Continuing and planned studies include simulations incorporating explicitly represented ions, and periodic boundary conditions, to capture electrostatic effects of the high ionic-strength intracellular environment (Mielke, et al., 2008); and simulations investigating the propagation of superhelical stress in chromatin (nucleosome-bound DNA), ultimately in periodic systems with explicit ions.