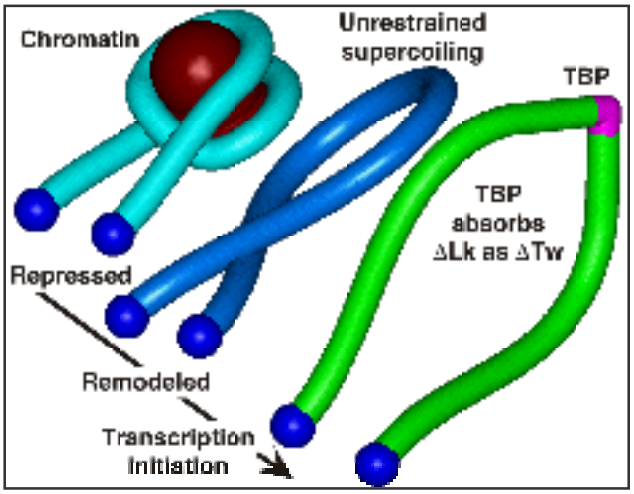
Structure, function, and engineering of large protein-DNA complexes.

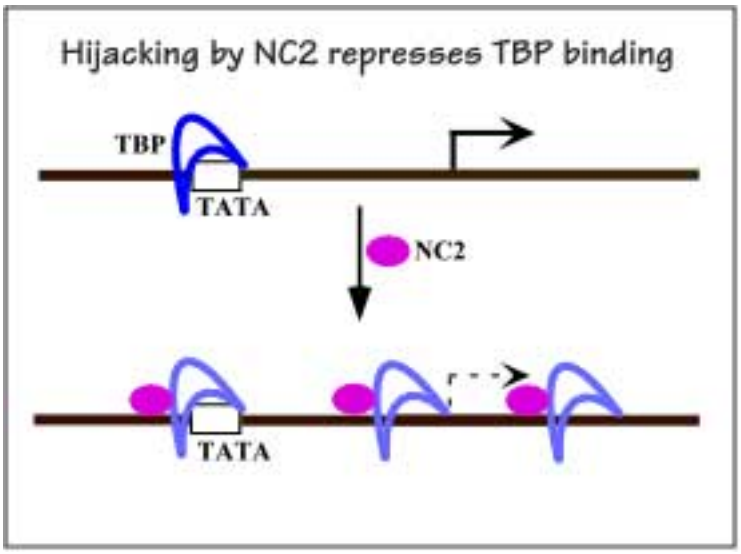
The projects use designed or random DNA molecules with intrinsic curvature to investigate protein-DNA conformation or nucleic acid structure and dynamics. The methods have potential applications in other areas, e.g. recombination. In the future, biophysical characterization of our current systems will continue, and I also anticipate moving into biochemistry on larger systems.

Selected Current Projects and Research Accomplishments

1. DNA deformation and topology in complexes including the TATA box binding protein (TBP). TBP is involved in all eukaryotic transcription. We have characterized the TBP•DNA complex using cyclization kinetics on DNA constructs with TATA boxes helically phased against a sequence-directed bend (Davis *et al.*, 1999). The initial goal was to correlate TBP bending, binding, and function and to establish a baseline for assembly of larger complexes. Two unexpected results emerged. First, the TATA box itself is highly anisotropically flexible. This provides evidence for the role of deformability in protein-DNA recognition and TBP specificity, and it also makes the point that the ground state structure of the DNA need not resemble the final state. Cyclization is uniquely sensitive to occasional dramatic excursions from the equilibrium structure. Second, TBP induces supercoiling upon cyclization. This can be explained by a proposed writhe cancellation mechanism (see the figure; Kahn, 2000), which has implications for the connection between chromatin remodeling and TBP/TFIID recruitment. We have made initial steps toward time-resolved footprinting on minicircles, which could address a number of mechanistic predictions from this work.

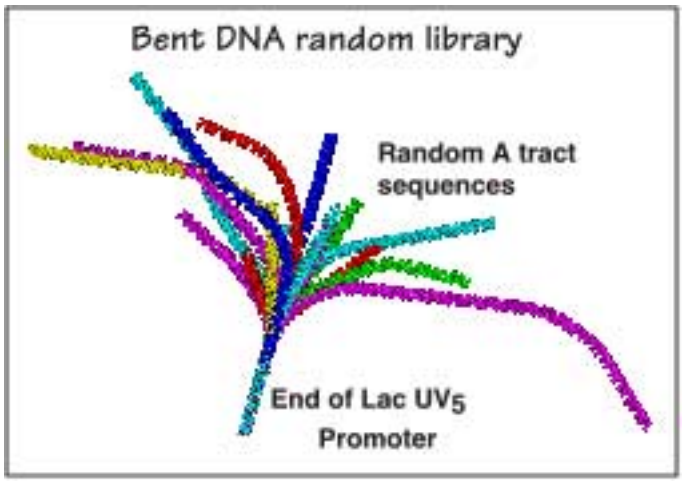


Currently we are extending this project to multi-protein DNA complexes. For example, we have found that the Negative Cofactor 2 (NC2)•TBP•DNA complex does not induce DNA bending substantially different from that of TBP•DNA, but NC2•TBP binds with lower specificity. Repression may be due to "hijacking," whereby NC2 titrates TBP away from the TATA box (see figure). Initial results on holo-TFIID show that induced supercoiling does not differ substantially from that of TBP alone. Structural effects from the histone-like subunits in NC2 and TFIID have yet to appear, suggesting that they may not interact with DNA in the manner of the nucleosome. This work will be extended to TFIID/TFIIA and TFIID/TFIIA/TFIIB complexes. The eventual goal is to map out changes in long-range DNA structure during transcription initiation; these changes are important for regulation and for transcription in chromatin

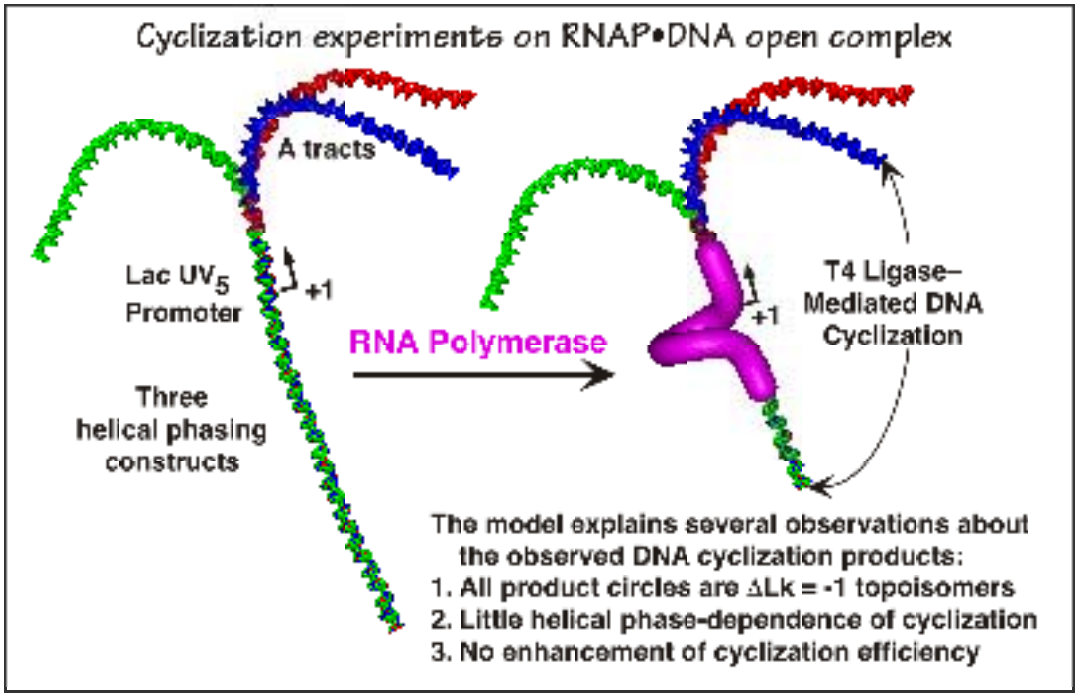


This project illustrates the power of the cyclization methodology. Identical experiments on protein complexes of widely varying sizes can be used to compare their effects on DNA. Careful quantitative analysis was necessary to reach the above conclusions, and this is not a rapid process, but future work on larger complexes will be more qualitative as the questions are less detailed. Funded by an NIH R01.

2. Wrapping, twisting, and bending of DNA in complexes with *E. coli* RNA polymerase (RNAP). It has been claimed for many years that RNAP bends/wraps DNA about itself, but there is little direct evidence as to the direction or magnitude of bending. One of the core projects in the lab has been to tackle this question with a selection-amplification approach based on cyclization. The idea is to make a random ligation library of A-tract containing oligonucleotides, phase the library against a promoter, and select molecules in which the bent portion loops around to meet the DNA exiting the RNAP•promoter complex. The sequence of the A-tract portion can be used to predict its structure and therefore the overall shape of the protein•DNA complex, and minicircles containing optimal A-tracts should provide hyperstable complexes for biophysical studies or crystallography. The figure at the right shows 25 representative library members. This ambitious idea extends the reach of cyclization methods to multi-protein complexes whose structures are in principle entirely unknown. Technical challenges in library construction, cloning, and use have delayed its completion. In the interim we turned to designed constructs as a proof of principle, with the results described below, and now we are returning to the random pool.

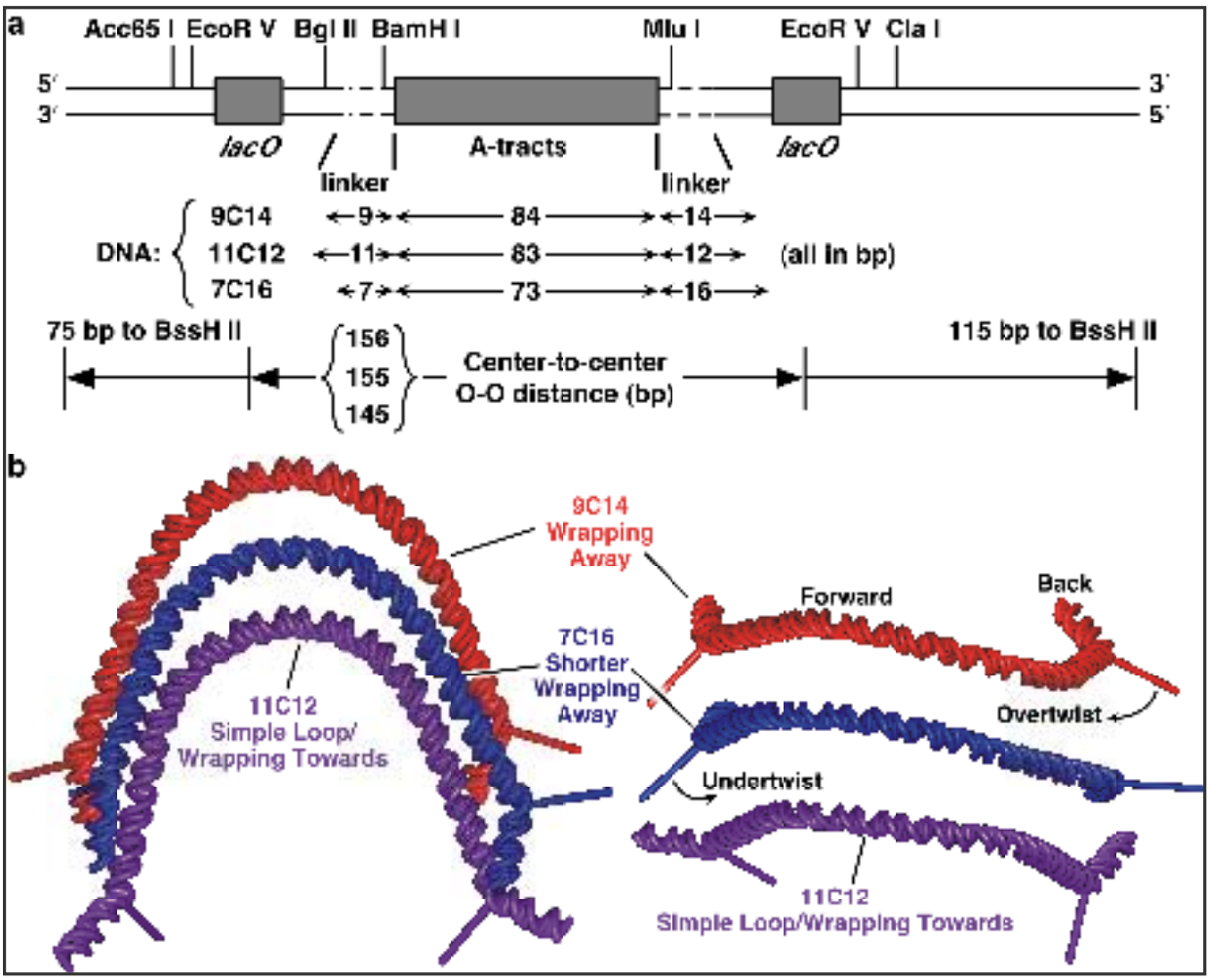


The designed molecules shown below have provided insight into the shape of polymerase complexes. To our surprise, bound RNAP does not affect cyclization efficiency substantially (though it does give a $\Delta Lk = -1$ topoisomer, verifying efficient binding as well as cyclization of the bound DNA). Furthermore, cyclization is relatively insensitive to helical phasing of the promoter and the bend. The model at the right rationalizes the results. It suggests that DNA is wrapped in a nearly 360° bend rather than a 180° or 0° bend. Similar experiments on stalled ternary complexes give more detectable bending. Transcription of supercoiled minicircles identifies promoter escape as the step most affected by bend phasing, consistent with the idea that a bend develops during the transition from initiation to elongation.



An emerging theme from these projects is that larger complexes often have surprisingly subtle net effects on long-range DNA structure. This may allow transcription initiation and elongation without substantial rearrangement of chromosomal DNA. Fleshing out the ideas will require work on a variety of transcription intermediates. The results here have remained outside the range of crystallography, although one of my goals is to change this. Funded by an NSF Career Award.

3. Design and characterization of hyperstable protein-DNA loops including Lac repressor. DNA looping, and more generally the formation of all multi-protein DNA complexes, depends on deforming the intervening DNA. We have successfully designed A tract sequences which dramatically stabilize LacI•DNA loops (Mehta and Kahn, 1999). The results suggest that the most stable loop (formed by the bottom, purple molecule below) is not the model suggested from crystallography, rather it is a more open form. We are in the process of testing this with fluorescence resonance energy transfer and quantitative DNA cyclization. It will also be interesting to test the hyperstable loops *in vivo*.



4. Theoretical work on rod mechanics (in collaboration), Monte Carlo simulation of DNA bending and looping, and writhe in minicircles. The combination of theory and experiment sets my work apart. My focus in theory is always to focus on extracting qualitative insight and experimental predictions from the quantitative analysis, to test whether a proposed model makes correct detailed predictions. For example, Monte Carlo simulation of CAP-induced DNA bending (Kahn and Crothers, 1998) established that out-of-phase bending constructs cyclize via twist deformation, not bend inversion. I took the time to carry forward this work from my post-doc both for its intrinsic interest and because I needed to establish the method for future development: currently we are extending it to the determination of looped shapes, which can be modeled essentially as two separate circles. The rod mechanics work (see Manning *et al.*, 1996) offers the potential for much faster analysis and more straightforward structural predictions, and the two methods complement each other well. Currently we are focusing on understanding TATA box anisotropic flexibility and on completing the extension to protein-induced bending. Both the rod mechanics and Monte Carlo approaches will be critical in quantitatively assessing the writhe cancellation model mentioned above.