

## Research, Teaching and Service Craig Benham

### Research

My primary research focus throughout my career has been on understanding the mechanical properties of DNA, and the roles they play in normal and pathological processes. I have championed the view that DNA is an active participant in the mechanisms by which its functions are regulated, and that understanding these mechanisms requires knowledge of the molecular mechanics of DNA and of its interactions with other molecules.

The importance of DNA mechanics can be appreciated by realizing that stresses are imposed on it in highly controlled ways *in vivo*. Specifically, DNA within living systems is held in looped domains so that its linking number, which is determined by the interlinking of the two strands of the duplex, is held fixed. This topological constraint, often called DNA superhelicity, allows the stresses imposed on each domain to be individually regulated. A variety of biological processes act to regulate these stress levels. They act either by introducing transient cuts, by forming, releasing or moving domain boundaries, or by over- or undertwisting the DNA helix. Changing the level of imposed stress affects patterns of gene expression, the initiation of replication, and many other important physiological activities.

My interests in modeling and thereby understanding these phenomena and the roles they play in normal and pathological processes has several components, which will be briefly described below. (Papers are numbered according to their order in my list of publications.)

### DNA Mechanics

This investigator was the first to propose ascribing mechanical properties to DNA, and analyzing its response to constraints in that framework (4). This area has grown in importance over the years until now there are many research groups working in it worldwide.

My initial interest was in analyzing the structural equilibria of topologically (i.e. superhelicity) stressed DNA. This problem presents a variety of intriguing mathematical aspects amalgamating mechanics, geometry and topology. The founding paper (4) proposed to calculate the mechanical equilibrium conformations of a linearly elastic, circular DNA molecule constrained by the constancy of its linking number. This project was completed in paper (21). Paper (29) presented a bifurcation analysis of the onset of non-planar deformations, which was the first proven result regarding the stability of these equilibria. A complete analysis of the stability of equilibria has been developed over the years by several other research groups.

However, DNA torsional stresses *in vivo* are not static; events such as gene expression produce wakes of undertwist and bow waves of overtwist. So fully dynamic models of the deformations experienced by stressed DNA are expected to be more biologically relevant than are equilibrium results. Recently, one part of my research program has focused on developing dynamic models of the structural deformations that occur in DNA in response to dynamically varying stresses. We are using three approaches to address this problem.

First, I have been working with Steve Mielke, originally a graduate student and then a postdoc in my group, to develop the first computational simulations of DNA under dynamically changing levels of stress (81, 85, 93). This program applies Brownian dynamics methods to increasingly detailed models of the phenomenon. Second, in her PhD thesis research Eva Strawbridge is developing a fully rigorous mathematical analysis of the dynamics of stressed DNA, with drag. Cheryl Serksen is pioneering a third approach to modeling this phenomenon that uses non-equilibrium statistical mechanics. While the first

two techniques can handle dynamically imposed stresses, the last one at present only treats the approach to equilibrium from a non-equilibrium initial state.

### **DNA Structural Transitions**

Although DNA *in vivo* occurs primarily in a double helix, its strands must separate to initiate either gene expression or replication, which are the two major jobs of the molecule. This is required because each process needs access to the individual bases along one strand. In the duplex structure these are paired with the other strand, and hence inaccessible. So maintaining the correct patterns of gene expression and replication requires that the occasions and locations where strand separation events occur must be stringently regulated. Any process that affects the local stability of the double helix can thereby exert regulatory effects on both gene expression and DNA replication.

It has long been known that the untwisting stresses consequent on negative superhelicity (i.e. a deficiency of the linking number relative to its relaxed value) can alter the stability of the DNA, causing destabilization of the local helical structure at thermodynamically susceptible sites. Paper (8) by this investigator presented the first quantitative analysis this phenomenon. There it was shown that local strand separation could be driven by DNA stress levels that are attained under normal physiological conditions. Experimental verification of this prediction came in 1988.

Subsequently, my research group has developed methods that predict with quantitative accuracy the precise locations and extents of destabilization experienced by a DNA domain on which an untwisting torsional stress has been applied, as occurs within superhelical domains (31, 43, 57, 78, 79). This analysis is greatly complicated by the fact that imposed stresses globally couple together the states of all the base pairs that experience them. This makes the destabilization properties of stressed DNA highly context dependent, which gives the phenomenon a rich repertoire of behaviors that are exploited in a variety of regulatory mechanisms.

Our methods perform statistical mechanical calculations of the equilibrium probability of separation of each base pair in a superhelically constrained domain of known sequence. They also calculate the extent to which the energy required for full separation of each individual base pair is decreased by the imposed stress. Although fractional destabilization will not entirely open a site, it is biologically important because it can facilitate opening by other processes (*viz.* interactions with proteins). The values of all geometric and thermodynamic parameters required for this analysis are known from experiments, so there are no free (i.e. tunable) parameters. Yet in all cases where experimental information has become available, our theoretical predictions are in quantitatively precise agreement with experimental measurements, both *in vitro* and *in vivo* (34, 50, 52, 54, 61, 74, 75, 82, 83). This quantitative accuracy allows these methods to be used to predict the destabilization behavior of any DNA sequence. A large number of DNA sequences have been analyzed in this way, including numerous complete genomes (38, 45, 80, 86).

These studies have produced major breakthroughs in our understanding of a variety of DNA regulatory events. A susceptibility to stress-induced destabilization has been documented to be associated with several types of regulatory regions. These include bacterial promoters (sites controlling initiation of gene expression; 80, 88), replication origins (83), transcription terminators (45), putative scaffold attachment regions (52, 87), and FIV integration sites (90).

Collaborations with various experimental groups have explicitly documented roles of stress-induced DNA duplex destabilization in numerous biological processes. Examples involving gene expression include the control of an important human oncogene (61), and of several other types of human genes (68, 71, 74), and the initiation of expression of several families of bacterial genes (45, 54, 58, 77, 80). Stress-destabilized sites have been implicated in the functioning of replication origins in a variety of organisms, including Epstein-Barr virus (82), yeast (91), and in a specific human disease process (75). Sites where DNA attaches to the nuclear scaffold (one process that partitions chromosomes into domains) have characteristic destabilization patterns. Attachment at these sites is known to affect the expression of nearby genes (52,

53, 59, 65, 76, 87). Sites where yeast gene expression terminates were predicted to be highly susceptible to SIDD (45). Subsequent experiments showed that strand separation actually does occur at these sites *in vivo* when the gene involved is active (50). Moreover, the normal termination requires *only* the presence of this site. Stress-destabilized regions have been shown to provide important chemotherapeutic drug target sites (70). Inclusion of such regions in episomal gene therapy vectors has been shown to greatly increase their efficiency (61, 69). These and several other associations are providing unprecedented insights into the mechanisms of function of the DNA regulatory sites involved. In several cases they have elucidated previously unknown strategies for regulating such biologically important events as the expression of specific genes, the initiation of DNA replication, control of the insertion and function of transgenes, bacterial and viral infectivity, and the genesis of certain cancers and neurodegenerative diseases.

To illustrate the character and scope of these investigations, I describe one example in greater detail.

### **Analysis of a Coordinately Regulated Gene Family**

The bacterial *ilvGMEDA* promoter is known to be activated by binding of the IHF protein 100 bp upstream, but only when the DNA is sufficiently negatively supercoiled. Predictions made by this investigator suggested that SIDD may be involved in the mechanism of this regulation. In the absence of IHF, untwisting stresses imposed on the DNA were predicted to destabilize the IHF binding site. But when IHF was bound to this site it could not open, so the stress must be relieved by opening elsewhere. Once the IHF binding site has been removed from the competition, the next most favorable site for opening is in the promoter, precisely where opening is required to initiate expression. So its destabilization facilitates this initiation.

My collaborator, Dr. Wes Hatfield of UC Irvine, has performed a series of experiments that have fully verified this model (54, 58, 73). In particular, an experimental procedure originally proposed by this investigator was used to show that it is specifically the destabilization properties of the region, and not its base sequence, that mediate this regulation (45, 58). We have recently shown that there are 109 IHF regulated genes in the *E. coli* genome with the attributes needed to be regulated by this mechanism. We also have documented use of the same strategy that uses a different regulatory molecule, FIS instead of IHF (77). Taken together, these results suggest that this strategy could be widely employed in the regulation of many genes by the binding of several different regulatory molecules.

This extremely productive collaborative research program is providing the first global view of an entirely new class of bacterial regulatory mechanisms. It is a model for how synergistic a carefully designed collaboration between experimentalists and theoreticians can be, bootstrapping rapidly from an investigation of a single promoter to a global regulatory system.

More generally, this work has discovered a new DNA regulatory strategy, whereby the effects of molecular binding at one position are transduced to remote sites through its influence on the destabilization properties of the DNA involved. It also has shown that DNA superhelicity itself is an important global regulator of gene expression, in this case acting to transmit regulatory information along the molecule through its effects on duplex stability.

### **Service**

The position I accepted in coming to UC Davis had three components – Professor in the departments of Mathematics and of Biomedical Engineering, and Associate Director for Bioinformatics of the UC Davis Genome Center. At that time the Genome Center was in its earliest stages. It had no space, no budget, no staff and no structure, and I was its first employee. However, the position I found myself in when I arrived was rather different. The candidate for Center Director had declined that position, so I became the interim

Director, charged with establishing the Center. I negotiated its initial budgets, recruited its administrative staff and leadership, and supervised the development of all its space, both in a temporary building and in its permanent home. I also led the effort to recruit a permanent Director. With the exception of developing computational resources, none of these responsibilities had been anticipated at the time I was recruited.

The present Director came on board in the fall of 2003, after two years of my five year term were completed. So I had three years to fulfill the obligations I undertook in accepting the position as Associate Director for Bioinformatics. In that time I led the recruitment of seven Center computational faculty. I initiated the development of the Bioinformatics Core, lead the searches for its Core Manager, System Administrator and three programmers. I spearheaded the design and acquisition of computer infrastructure suited to the needs of the faculty and this Core. When my term as Associate Director ended in 2006 I had completed all the obligations that I undertook in accepting this position, as well as those of interim Director. The Center being well established and in good hands, I decided to step down at that time from my administrative roles in it.

In the fall of 2007 I was named Vice Chair of the Biomedical Engineering Department. In that capacity I am chair of the departmental Space, Facilities and Safety committee. I also am responsible for all shared computational resources of the department, and oversee the staff that operate and maintain it. And I am responsible for the Machine room, where faculty fabricate a wide variety of components for their apparatus.

In addition to these major assignments, I have performed a wide variety of other service tasks. These include membership in the CoE Awards Committee, in the BME departmental Executive Committee, and in the departmental Appointments, Merits, Awards and Promotions (AMAP) Committee.

A full list of services is available in my submitted materials.

## **Teaching**

Although the terms of my administrative positions in the Genome Center specified that I need not teach, I insisted on teaching at least one course in each department per year. Information covering most aspects of my teaching record are contained in the attached materials. However, there are a few additional points I wish to raise.

With the exceptions of MAT17 and MAT147, every course I have taught was custom designed by myself. None have been standard courses, and most were not taught directly from a book. The BME graduate core course I teach, BIM204 *Physiology for Bioengineers*, was designed from scratch to include both descriptive physiology and mathematical modeling of physiological systems. As there is no textbook that spans this range of material, I have had to develop all aspects of this course on my own. The lecture notes are in the form of Powerpoint slides, which are available for viewing at the course MyUCDavis sites and (starting this year) at its SmartSite. These evolve from year to year as I continually refine the course; last year there were 992 slides shown in the course. This is a core course, required of all entering graduate students in the BMEGG, and is the only five unit course in that graduate program.

I also developed an undergraduate course on DNA Mechanics (BIM151). And I designed the syllabus for the undergraduate mathematical biology course, MAT124.

The only aspect of my record with which I am not satisfied is MAT17, which is taught in a large lecture format. I am determined to master this challenging pedagogical situation, and am taking steps to do so.