Collective variable modelling of nucleic acids

Ingrid Lafontaine* and Richard Lavery†

Collective variable models continue to contribute to our knowledge of nucleic acids. The past year has seen considerable progress both in modelling sequence-dependent effects on nucleic acid conformation and in understanding how proteins or external stresses influence nucleic acid structure. Algorithmic developments have also allowed collective models to be applied to studies of thermal fluctuations and dynamics. For larger systems, models with varying degrees of resolution are being refined and applied to nucleic acids containing hundreds or thousands of nucleotides.

Addresses
Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France
*e-mail: ingrid@ibpc.fr
†e-mail: richard@ibpc.fr

Current Opinion in Structural Biology 1999, **9**:170–176

http://biomednet.com/elecref/0959440X00900170

 ISSN 0959-440X

Abbreviations

BD	Brownian dynamics
CAP	catabolite activator protein
JUMNA	junction minimisation of nucleic acids
LEF-1	lymphoid enhancer binding factor-1
MC	Monte Carlo
MC-SYM	macromolecular conformation by symbolic generation
NLPB	nonlinear Poisson–Boltzmann
PNA	peptide nucleic acid
PurR	purine repressor
SELEX	simulated evolution of ligands by exponential enrichment
SRY	sex-determining region Y
TBP	TATA-box-binding protein
WLC	worm-like chain

Introduction

Although collective variable models of nucleic acids have been around for about 25 years (see, for example, [1–3]), they are far from having reached the end of their usefulness or the end of their algorithmic development. The principal reason for this is an explosion in the variety of nucleic acid structures. During the past two decades, experimental studies of nucleic acids have brought to light many new structural families that are important from biological and, in some cases, pharmacological points of view [4]. Triple helices are a good example, where antigene strategies are powering the development of noncanonical oligonucleotide chains. During the same period, the complex folded structures of natural RNAs have come to light and have been joined more recently by 'aptamers' generated with SELEX (simulated evolution of ligands by exponential enrichment) *in vitro* evolution techniques. Ribozymes have also demonstrated that RNA is not simply a passive information carrier. At the finest level of detail, we now have access to the atomic contacts and the sequence-dependent heterogeneity that contribute to protein–nucleic acid recognition, while at the other extreme, we are beginning to probe packaging and control mechanisms involving thousands of nucleotides within chromatin or complex molecular machines, such as the ribosome.

As structural information grows, so does the data coming from atomic-scale simulations, which, although still limited to a few tens of nucleotides and to nanosecond timescales, are notably able to deal with explicit solvent and counterion environments [5]. The challenge today is to put this information together to build simplified, but realistic, models that will be able to overcome the barriers opposing the modelling of systems with hundreds or thousands of nucleotides or systems involving slow dynamic processes, subtle intermolecular recognition or complex folding pathways. We review below some of the recent steps taken towards these goals.

Construction algorithms

Moving from Cartesian to internal coordinates and freezing both bond lengths and a subset of valence angles already allows an order of magnitude reduction in the number of variables necessary to represent a nucleic acid fragment. This gain can be pushed further using pseudorotational sugar representations or helical symmetry restraints. Such models have been widely used for building and energy minimising nucleic acid structures.

There are, however, still difficulties in building helical DNA when taking into account sequence dependence. A partial solution to this problem has been proposed by Hunter and Lu [6], who used a technique based on a library of dinucleotide steps constructed from crystallographic data. DNA fragments are built by plugging together dinucleotides with chosen helical step parameters and averaging the positions of the common backbone atoms. This method, which has been applied to building both oligomers and long DNA fragments, has the advantage of generating reasonable backbones for any helical conformation; however, it requires that the user choose the parameters defining the helix and suffers from the limitations of the current structural database, in which the scatter of parameters for each of the 10 unique dinucleotide steps often shows large overlaps.

For RNA, construction problems are even more severe. One now has to deal not only with helical segments, but also with single-stranded elements, such as loops and bulges. In this field, Lemieux *et al.* [7•] made an interesting refinement by checking whether different base sequences known to be compatible with a given structure (and possibly with a given catalytic activity) are also compatible with a given model conformation. This approach, termed ICS (intersection of conformational space), has been tested within the MC-SYM (macromolecular conformation by

symbolic generation) program [8] on a six nucleotide loop belonging to the active site of a leadzyme. Sequence data enabled the selection of a single structural model from starting conformations that covered all possible hydrogen-bonding arrangements for the loop bases. It is interesting to note, however, that although the MC-SYM approach can often produce structures with good hydrogen-bonding arrangements, it has more difficulty with stacking interactions, apparently because of current limitations in the number of backbone geometries available in its nucleotide library. These limitations are also seen to lead to increased base-pair deformations [9].

Building RNA loops was also the subject of recent work by Tung [10], who has developed a loop-closure algorithm based on setting a pseudo torsional angle (O3′-C5′-O3′-C5′) so that it satisfies the distance between two flanking nucleotides. Conformations satisfying this constraint are improved by Monte Carlo (MC) sampling and energy minimisation. An iterative version of the procedure can be applied to the loops of several nucleotides. Tung [11] has used this approach to break the construction of simple RNA fragments into two steps: firstly, positioning helical segments and, secondly, constructing the linking loops. The scanning of relative helix positions can be limited to some extent by the observed range of loop lengths, but the computational expense increases rapidly with loop size and a two-nucleotide loop already requires the generation of 5000 structures in order to scan the available conformational space.

Helical conformations and dynamics

Sequence-induced deformations of duplex DNA have recently been investigated using a rigid base model by Hunter and Lu [12], who found that rise, roll and tilt depend largely on base stacking and can be successfully predicted from van der Waals' energy calculations. On the other hand, slide and shift are both flexible and dependent on electrostatics, whereas twist is directly coupled to backbone geometry. These results have been extended by analysis of existing oligomer conformations and the formulation of a virtual bond model of the nucleic acid backbone that is flexible with respect to slide, shift and propeller, but rigid with respect to all other degrees of freedom [13]. This model helps to clarify the correlations between base and backbone geometries.

In the realm of triple helices, Srinivasan and Olson have carried out a systematic search of backbone geometries for conventional DNA and RNA backbones [14], and peptide nucleic acid (PNA) and 2′-5′ modified backbones [15]. The results, which can be criticised for being limited to fixed, or only partially mobile, base positions, nevertheless provide a comprehensive view of possible backbone conformations and, in the case of PNA, help to explain the movement towards A-like geometries.

Conformational scanning of sugar puckers within the JUMNA (junction minimisation of nucleic acids) internal and helical coordinate model has been used as a tool to find the most stable conformations of a single adenine bulge within a DNA double helix [16]. Both extruded and stacked conformations of the bulged base were considered and it was shown that, with a nonlinear Poisson–Boltzmann (NLPB) electrostatic solvent correction, the most stable state was stacked and led to axis curvature, in agreement with solution data. The conformational scanning approach also led to the discovery of a new bulged conformation that has interactions with the C•G base pair on the 5′ side. This bulged conformation is related to the base triad structure proposed by Kuryavyi and Jovin [17] and is found within folded RNA structures.

More extensive deformations that are related to DNA–protein interactions have been the subject of a number of recent publications. In the case of the TATA-box-binding protein (TBP), Lebrun *et al.* [18•] have shown that local stretching of the distance between the 3′-phosphate groups bracketing the TATA sequence reproduces the highly distorted DNA conformation induced by the protein. This study also showed that stretching led to rolls at the base-pair steps at which partial protein sidechain intercalation occurs. This work was subsequently extended to three other minor-groove-binding proteins — sex-determining region Y (SRY), lymphoid enhancer binding factor-1 (LEF-1) and the purine repressor (PurR) — with similar results, suggesting that a common DNA deformation mechanism is shared by this family of proteins [19]. It was also found that phosphate neutralisation facilitated DNA deformation, although the mechanism is different from that seen with major-groove-binding proteins, for which DNA bending is towards and not away from the neutralised face of the duplex.

The conformation of a large multimeric protein–DNA complex has been studied by Nagaich and co-workers [20•,21] using an internal coordinate approach and symmetry constraints to further reduce the number of degrees of freedom. This work involved building the tetrameric complex of p53 using the available experimental data on a single DNA-binding domain interaction. Using a specifically developed set of helical coordinates to scan the relative positions of the protein monomers, it was found that DNA curvature was necessary to avoid intermonomer steric clashes. The results suggest that sequence-dependent DNA flexibility appears to play an important role in p53 binding.

Collective variable models can also be used to go beyond conformational optimisations. As a first step, normal mode analysis offers an inexpensive way of studying the dynamics around a stable conformational state. Using internal coordinates implies the necessity of removing the overall translations and rotations of the molecule. This development, made initially for proteins [22], has been extended to valence angles and fully flexible sugar rings, and applied to DNA by Ha Duong and Zakrzewska [23]. The resulting methodology has been used to study indirect recognition in protein–nucleic acid complexes and, more specifically, the

role of the central base sequence within the 434-repressor-binding site. These bases influence binding, but do not directly contact the protein [24]. Although it is generally thought that this effect is due to sequence-dependent flexibility, normal mode calculations do not support this idea. Rather, it was found that sequence-dependent electrostatic properties appear to dominate. It was nevertheless noted that the lowest frequency normal mode of the target base sequence induced bending in the same direction as that seen within the protein–DNA complex. This suggests that normal modes may be a useful guide to large-scale DNA deformations, in analogy with their use in the study of protein domain movements.

Similar studies of the catabolite activator protein (CAP) binding to DNA showed that large amplitude vibrations in the low-frequency normal modes were correlated with kink sites in the complexed DNA conformation. In this study, it was found that phosphate neutralisation, via the formation of salt bridges with cationic protein sidechains, was not only responsible for DNA curvature towards the protein, but also led to DNA rigidification. Both the static and dynamic effects of neutralisation were found to be strongly sequence dependent [25•].

MC simulations that allow more extensive scanning of conformational space have recently been performed using an internal coordinate model of DNA [26]. This approach used the scaled collective variable approach developed by Noguti and Go [27]. This method enables more efficient scanning of conformational space via multiple variable moves based on the eigenvectors of the Hessian matrix derived from an initial energy minimised conformation. The results obtained for an oligomer containing a target site for TBP showed deformations that are very similar to those of Cartesian coordinate molecular dynamics simulations with explicit solvent and counterions.

Despite the promise of long duration molecular dynamic simulations with internal coordinates, the development of such approaches has been hindered both by difficulties in formulating efficient algorithms and by unexpected problems with integration techniques, thermal equilibration and the role of atomic collisions. An original proposition for dealing with these problems, based on an O(n) robotics algorithm, has recently been developed by Brünger and colleagues [28] and used to refine nucleic acid and protein structures. Another approach has been implemented by Guntert *et al.* [29], again with macromolecular structure refinement in view. In both cases, fundamental physical difficulties are avoided by the use of highly simplified force-fields and increased atomic masses and, in the latter case, by using a space-fixed base for the initial nucleotide. The algorithms, which yield good final structures for both proteins and nucleic acids, are faster than Cartesian coordinate approaches and, incidentally, avoid problems caused by chirality inversion, which have been shown to be surprisingly common in published nucleic acid structures [30].

The first real solution to the problem of internal coordinate dynamics has been proposed by Mazur [31••], who has developed an algorithm that includes a refined and stable integration technique, an efficient method for solving the equations of motion and an appropriate treatment of flexible rings. This method allows 10 fs time steps at room temperature when all bond lengths are frozen and valence angles are only considered for atoms belonging to the sugar rings. It is nevertheless necessary to increase the momenta associated with light rigid bodies (–OH and $-CH_3$) and with sugar ring bonds (C–C and C–O) in order to achieve this result. A hybrid solvent model, combining a distance-dependent dielectric constant, reduced phosphate charges and explicit water molecules in the minor groove, has enabled stable trajectories to be generated for a B-DNA dodecamer. The resulting conformations are remarkably close to the corresponding crystallographic results.

DNA under stress

Over the past few years, single-molecule experiments on DNA have led to the discovery of unexpected conformational transitions that are induced by external stress. Modelling with JUMNA has enabled the conformational impact of such constraints to be reproduced. Adiabatic mapping was first used to understand polymeric DNA stretching, leading to models of so-called S-DNA [32,33]. The same approach has now been applied to an oligomer, mimicking an atomic force microscope experiment by Noy *et al.* [34]. The results, which, in common with polymer stretching, led to a force plateau, also reveal similar conformations — a narrow fibre with inclined bases when the traction involves the 3′ termini and a flat unwound ribbon when the 5′ termini are pulled [35]. The same oligomer has also been 'unzipped' by separating the strands at one end. This led to forces that vary as a function of the AT/GC content, in line with recent results on polymeric DNA obtained by Heslot's group [36].

Ribbon-like DNA, obtained by 3′ stretching, has also been used as the starting point for modelling strand exchange induced by the RecA protein [37•]. In contrast to earlier suggestions, this model involves attack by the single-stranded DNA–RecA filament on the minor-groove side of the stretched duplex. Stretching increases the width of the minor groove to allow the bases of the single strand to form interactions with the duplex base pairs. This leads to a triple helix that is stable in its extended form. A similar model [38•] has also been proposed by the authors of a recent NMR spectroscopy study of a RecA monomer–DNA complex. Strand exchange can subsequently take place through a simple base rotation, yielding a triple helix in which the leaving strand lies in the major groove.

Another aspect of DNA deformation has been addressed in an experiment in which magnets are used to rotate a tiny paramagnetic bead attached to a single DNA molecule. Over or undertwisting DNA generally leads to the formation of plectonemes through axial writhing. When this is prevented by pulling the DNA, a structural transition is seen

that corresponds to local denaturation for undertwisting and to the creation of a previously unknown structure (P-DNA) for overtwisting [39••]. Modelling with JUMNA suggested that this structure had unpaired bases surrounding tightly interwound phosphodiester backbones (twist ≈ 137° and rise ≈ 5.8 Å) and, to some extent, resembled the early DNA structure proposed by Pauling and Corey [40]. This result was supported by the observation that the overtwisted and stretched DNA becomes reactive towards glyoxal, a reagent that is specific for unpaired bases.

Low-resolution models

We now turn to collective variable models that no longer resolve individual atoms and that are aimed at studying the properties of large nucleic acids and, in particular, linear or circular DNA. Such models exist with various degrees of resolution: nucleotides, nucleotide pairs, helical segments or, even, chromosome fragments. Given the computational difficulties and the lack of sufficient data for parametrisation, such models must often abandon some characteristic aspects of DNA structure (bending anisotropy, sequence inhomogeneity, electrostatic interactions, etc). Their formulation is still, to some extent, *ad hoc*, but it is worth noting that a recent study has investigated the validity of various rigid-body models through a detailed analysis of the correlated movements observed within Cartesian coordinate dynamic trajectories [41•].

Although most models in this area do not take base sequence or DNA bending anisotropy into account, a number of attempts have been made recently to include such effects. One approach, proposed by Huertas *et al.* [42], uses a single bead per nucleotide, with intrastrand and interstrand spring interactions between neighbouring beads. This construction, however, appears to be less easy to manipulate than models that use segmented rods with one segment per nucleotide pair [43,44•]. An example of this is the model of Munteanu *et al.* [44•], which includes inhomogeneous and anisotropic flexibilities based on DNAse I digestion data. This model leads to a rationalisation of the experimentally observed sequence dependence of Cro repressor binding free energy.

Sequence effects (based on dinucleotide wedge angles) have also been included in an averaged sense within a segmented rod model, in which each segment represents roughly one turn of the double helix. The results show that, although sequence inhomogeneity does not affect the radius of gyration of circular DNA, it does have an important impact on the equilibrium linking number distribution, which may become bimodal [45••]. This result may be important for other studies, such as the recent work of Swigon *et al.* [46], which uses a homogeneous rod model to relate nucleosome binding energies to this property.

Ignoring sequence effects, there are many models using segmented rods whose applications overlap with continuous elastic rod treatments of DNA. Such models, which have been the subject of a number of reviews [47,48], are continuously being improved in terms of both their physical basis (anisotropy, self-contact, electrostatic and hydrodynamic effects, etc.) and their computational efficiency. We mention here only a few recent extensions. Within the segmented rod approach, Langowski's group [49] have developed and applied a model that is applicable to both linear and circular DNA, including the treatment of intrinsically curved segments, and using an efficient second-order Brownian dynamics (BD) algorithm. This model has been used to study the cyclisation reaction of DNA fragments with 120–470 base pairs in good agreement with experiment [50]. A related approach has been used to study the dynamics of site juxtaposition within circular DNA and shows that simple slithering, although accelerated by supercoiling, is not an efficient mechanism for bringing sites together [51]. A segmented rod model [52] has also been used to describe the behaviour of supercoiled DNA under tension, in conjunction with recent micromanipulation experiments [39••,53], a problem that has also been approached analytically within the WLC (worm-like chain) model [54]. It is also remarked that the WLC model has also been recently extended to allow segments exhibiting intrinsic curvature or differing flexibilities to be taken into account [55].

In the field of continuous rod models, a number of simplified representations exist that enable circular DNA to be manipulated with a small number of variables. In this area, Liu *et al.* [56•] have improved on an earlier B-spline approach by introducing a finite Fourier series representation. This technique, applied to Langevin dynamic simulations, maintains the advantage of a reduced number of variables, but describes the DNA path more precisely and automatically satisfies ring closure constraints. Work from Martino and Olson [57] also replaces the B-spline formulation for a description of the equilibrium conformations of circular minichromosomes. This approach uses Newton–Raphson minimisation to relax elastic rod segments joining rigid supercoils around phantom histone cores, whose positions are scanned by MC sampling. This model can be contrasted with the simpler model that the Langowski group [58] used for BD simulations of a linear chromatin fibre. Their model links chromatin beads with straight, extensible rod segments and limits bending to the bead junctions between segments. We finally mention a hybrid WLC plus bead model, introduced by Jian *et al.* [59••], which leads to an improved treatment of hydrodynamic effects. In connection with the passage from discrete to continuous models, it is worth noting the analysis of Klapper and Qian [60•], which shows that the energy of a segmented rod does not yield the continuous elastic rod expression in the limit of infinitely short segments.

A last step in decreasing the resolution takes us to models that attempt to describe complete chromosomes. One such model examining dynamic behaviour (MC/BD) has been proposed by Münkel and Langowski [61•], in which the chromatin fibre is represented by a chain of beads

divided into looped domains, with each bead representing roughly 30 kilobase pairs. The size of the system treated nevertheless requires many approximations (only short-range self-contacts, no electrostatics and no hydrodynamics). Such models can form the basis of studies of nuclear structure, as in recent work, which finds evidence for territorial organisation using static models to interpret localised radiation damage [62].

It is clear that the size of many biologically important objects will impose low-resolution models for the foreseeable future. As Harvey and colleagues [63,64] pointed out in their modelling of ribosome particles, however, this does not exclude the possibility of locally improving the resolution, when available experimental data permits.

Conclusions

Despite the progress achieved with atomic-scale simulations of nucleic acids, collective variable approaches are still making important contributions. Their advantages are not only seen for large systems, but are also in cases in which large conformational changes are involved or in which complex conformations need to be generated. With only a small loss of resolution, they will soon be able to extend the range of dynamic simulations, while, with greater simplification, they offer ways of looking at very large nucleoprotein complexes. How can such models be improved? The treatment of solvent and salt effects via continuum electrostatic and/or surface terms is one active area. Fast and precise NLPB solutions are still awaited, but progress is being made with more approximate models [65•]. Automated prediction of RNA folds, even given secondary structure and experimental constraints, remains a major challenge for anything beyond the smallest systems. Similarly, detecting subtly defined protein-binding sites within genomic sequences will become a more and more pressing goal. Lastly, although low-resolution DNA models continue to incorporate more realistic features, ways need to be found to create a homogeneous hierarchy of models that will allow us to move up and down in resolution as required.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Zhurkin VB, Lysov YP, Ivanov VI: **Computer analysis of conformational properties of double-helical DNA.** *FEBS Lett* 1975, **59**:44-47.
2. Tumanyan VG, Esipova NG: **Investigation of fibrous structures. II. General method for the elucidation of the conformations of the complementary nucleic acids.** *Biopolymers* 1975, **14**:2231-2246.
3. Miller KJ: **Interactions of molecules with nucleic acids. I. An algorithm to generate nucleic acid structures with an application to the B-DNA structure and a counterclockwise helix.** *Biopolymers* 1979, **18**:959-980.
4. Lebrun A, Lavery R: **Unusual DNA conformations.** *Curr Opin Struct Biol* 1997, **7**:348-354.
5. Auffinger P, Westhof E: **Simulations of the molecular dynamics of nucleic acids.** *Curr Opin Struct Biol* 1998, **8**:227-236.
6. Hunter CA, Lu XJ: **Construction of double-helical DNA structures based on dinucleotide building blocks.** *J Biomol Struct Dyn* 1997, **14**:747-756.
7. • Lemieux S, Chartrand P, Cedegren R, Major F: **Modeling active RNA structures using the intersection of conformational space: application to the lead-activated ribozyme.** *RNA* 1998, **4**:739-749.

The authors describe a method for selecting amongst possible models of folded RNA by checking their compatibility with known experimental data on sequence variants.

8. Major F, Turcotte M, Gautheret D, Lapalme G, Fillion E, Cedergren R: **The combination of symbolic and numerical computation for three-dimensional modeling of RNA.** *Science* 1991, **253**:1255-1260.
9. Leclerc F, Srinivisan J, Cedergren R: **Predicting RNA structures: the model of the RNA element binding Rev meets the NMR structure.** *Fold Des* 1997, **2**:141-147.
10. Tung CS: **A computational approach to modeling nucleic acid hairpin structure.** *Biophys J* 1997, **72**:876-885.
11. Tung CS: **A reduced-coordinate approach to modeling RNA 3D structures.** In *Structure, Motion, Interaction and Expression of Biological Macromolecule. Proceedings of the Tenth Conversation.* Edited by Sarma RH, Sarma MH. New York: Adenine Press; 1998:117-123.
12. Hunter CA, Lu XJ: **DNA base-stacking interactions: a comparison of theoretical calculations with oligonucleotides X-ray crystal structures.** *J Mol Biol* 1997, **265**:603-619.
13. Packer MJ, Hunter CA: **Sequence dependent DNA structure: the role of the sugar-phosphate backbone.** *J Mol Biol* 1998, **280**:407-420.
14. Srinivasan AR, Olson WK: **Molecular models of nucleic acids triple helixes I. DNA and RNA backbone complexes.** *J Am Chem Soc* 1998, **120**:484-491.
15. Srinivasan AR, Olson WK: **Molecular models of nucleic acids triple helixes II. PNA and 2′-5′ backbone complexes.** *J Am Chem Soc* 1998, **120**:492-499.
16. Zacharias M, Sklenar H: **Analysis of the stability of looped-out and stacked-in conformations of an adenine bulge in DNA using a continuum model for solvent and ions.** *Biophys J* 1997, **73**:2990-3003.
17. Kuryavyi VV, Jovin TM: **Triad-DNA: a model for trinucleotide repeats.** *Nat Genet* 1995, **9**:339-341.
18. • Lebrun A, Shakked Z, Lavery R: **Local DNA stretching mimics the distortion caused by the TATA box-binding protein.** *Proc Natl Acad Sci USA* 1997, **94**:2993-2998.

The authors show that locally increasing the distance between two phosphate groups at the 3′ ends of the TBP-binding site leads to minor-groove opening and bending, creating a conformation that is very like that induced by the protein.

19. Lebrun A, Lavery R: **Modeling DNA deformations induced by minor groove binding proteins.** *Biopolymers* 1999, in press.
20. • Nagaich AK, Zhurkin VB, Sakamoto H, Gorin AA, Clore GM, Gronenborn AM, Appella E, Harrington RE: **Architectural accommodation in the complex of four p53 DNA binding domain peptides with the *p21/waf1/cip1* DNA response element.** *J Biol Chem* 1997, **272**:14830-14841.

In accord with cyclisation data, simplified modelling of the p53 tetramer complex with DNA suggests that the target site must bend by roughly 50° in order to avoid clashes between the protein subunits.

21. Durell SR, Jernigan RL, Appella E, Nagaich AK, Harrington RE, Zhurkin VB: **DNA bending induced by tetrameric binding of the p53 tumor suppresser protein: steric constraints on conformation.** In *Structure, Motion, Interaction and Expression of Biological Macromolecule. Proceedings of the Tenth Conversation.* Edited by Sarma RH, Sarma MH. New York: Adenine Press; 1998:277-295.
22. Sunada S, Go N: **Small amplitude protein conformational dynamics: second-order analytic relation between Cartesian coordinates and dihedral angles.** *J Comput Chem* 1995, **16**:328-336.
23. Ha Duong T, Zakrzewska K: **Calculation and analysis of low-frequency normal modes for DNA.** *J Comput Chem* 1997, **18**:796-811.
24. Ha Duong T, Zakrzewska K: **Sequence specificity of bacteriophage 434 repressor-operator complexation.** *J Mol Biol* 1998, **280**:31-39.

25. Gurlie R, Zakrzewska K: **DNA curvature and phosphate**
• **neutralisation: an important aspect of specific protein binding.** *J Biomol Struct Dyn* 1999, in press.
Neutralising the phosphate groups involved in salt bridges with cationic sidechains of bound proteins often leads to bending. This study reveals that sliding the neutralisation 'pattern' along the DNA, however, shows that this effect is strongly sequence dependent.

26. Gabb HA, Prevost C, Bertuc :t G, Robert CH, Lavery R: **Collective variable Monte Carlo simulation of DNA.** *J Comput Chem* 1997, **18**:2001-2011.

27. Noguti T, Go N: **Efficient Monte Carlo method for simulation of fluctuating conformations of native proteins.** *Biopolymers* 1985, **24**:527-546.

28. Stein EG, Rice LM, Brünger AT: **Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation.** *J Magn Res* 1997, **124**:154-164.

29. Guntert P, Mumenthaler C, Wüthrich K: **Torsion angle dynamics for NMR structure calculation with the new program DYANA.** *J Mol Biol* 1997, **273**:283-298.

30. Schultze P, Feigon J: **Chirality errors in nucleic acid structures.** *Nature* 1997, **387**:668.

31. Mazur AK: **Accurate DNA dynamics without accurate long range**
•• **electrostatics.** *J Am Chem Soc* 1998, **120**:10928-10937.
The approach proposed by the author enables stable multinanosecond trajectories of DNA oligomers to be obtained using internal coordinates and a time step of 10 fs. The results, which model solvent using a hybrid discrete/continuum model, yield conformations close to crystallographic data.

32. Cluzel P, Lebrun A, Heller C, Lavery R, Viovy JL, Chatenay D, Caron F: **DNA: an extensible molecule.** *Science* 1996, **271**:791-794.

33. Smith SB, Cui Y, Bustamante C: **Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules.** *Science* 1996, **271**:795-799.

34. Noy A, Vezenov DV, Kayyem JF, Meade TJ, Lieber CM: **Stretching and breaking duplex DNA by chemical force microscopy.** *Chem Biol* 1997, **4**:519-527.

35. Lebrun A, Lavery R: **Modeling the mechanics of a DNA oligomer.** *J Biomol Struct Dyn* 1999, in press.

36. Essevaz-Roulet B, Bockelmann U, Heslot F: **Mechanical separation of the complementary strands of DNA.** *Proc Natl Acad Sci USA* 1997, **94**:11935-11940.

37. Bertucat G, Lavery R, Prevost C: **A model for parallel triple helix**
• **formation by RecA: single-strand association with a homologous duplex via the major groove.** *J Biomol Struct Dyn* 1998, **16**:1-13.
This paper reports how stretching the 3′ ends of DNA opens the minor groove sufficiently to allow complexation with a single-stranded DNA–RecA filament and suggests that strand exchange may occur by base rotation, leading to a leaving strand situated on the major groove.

38. Nichinaka T, Shinohara A, Ito Y, Yokoyama S, Shibata T: **Base pair**
• **switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination.** *Proc Natl Acad Sci USA* 1998, **95**:11071-11076.
The authors describe a very similar model to that mentioned in [37•], based on NMR data for a RecA–dinucleotide complex. This mechanism stresses the importance of sugar pucker changes.

39. Allemand JF, Bensimon D, Lavery R, Croquette V: **Stretched and**
•• **overwound DNA forms a Pauling-like structure with exposed bases.** *Proc Natl Acad Sci USA* 1998, **95**:14152-14157.
Positive supercoiling of stretched DNA leads to base-pair breaking and transition to a conformation with the phosphodiester backbones inside and the bases outside. A similar structure has already been proposed to exist within a filamentous bacteriophage [66].

40. Pauling L, Corey RB: **A proposed structure for the nucleic acids.** *Proc Natl Acad Sci USA* 1953, **39**:84-96.

41. Gaudin F, Lancelot G, Genest D: **Search for rigid subdomains in**
• **DNA from molecular dynamics simulations.** *J Biomol Struct Dyn* 1997, **15**:357-367.
This paper describes a method for finding the atoms of DNA that can be grouped into rigid bodies for a given error in the root mean square fluctuations of interatomic distances derived from all-atom dynamic trajectories.

42. Huertas ML, Navarro S, Lopez Martinez MC, Garcia de la Torre G: **Simulation of the conformation and dynamics of a double-helical model for DNA.** *Biophys J* 1997, **73**:3142-3143.

43. Westcott TP, Tobias I, Olson WK: **Modeling self contact forces in the elastic theory of DNA supercoiling.** *J Chem Phys* 1997, **107**:3967-3980.

44. Munteanu MG, Vlahovicek K, Parthasarathy S, Simon I, Pongor S: **Rod**
• **models of DNA: sequence-dependent anisotropic elastic modelling of local phenomena.** *Trends Biochem Sci* 1998, **23**:341-346.
This paper describes a segmented rod model of DNA that only bends towards the major groove at each step and includes sequence-dependent flexibility based on nuclease digestion results.

45. Katritch V, Vologodskii AV: **The effect of intrinsic curvature on**
•• **conformational properties of circular DNA.** *Biophys J* 1997, **72**:1070-1079.
A dinucleotide wedge model of DNA bending was used to build a segmented rod model including averaged sequence effects. The results of Monte Carlo simulations show that intrinsic curvature can strongly influence the equilibrium distribution of linking number for circular DNAs of roughly 300–1000 base pairs.

46. Swigon D, Coleman BD, Tobias I: **The elastic rod model for DNA and its application to the tertiary structure of DNA minicircles in mononucleosomes.** *Biophys J* 1998, **74**:2515-2530.

47. Olson WK: **Simulating DNA at low resolution.** *Curr Opin Struct Biol* 1996, **6**:242-256.

48. Schlick T: **Modeling superhelical DNA: recent analytical and dynamic approaches.** *Curr Opin Struct Biol* 1995, **5**:245-262.

49. Klenin KV, Merlitz H, Langowski J: **A Brownian dynamics program for the simulation of linear and circular DNA and other wormlike chain polyelectrolytes.** *Biophys J* 1998, **74**:780-788.

50. Merlitz H, Rippe K, Klenin KV, Langowski J: **Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation.** *Biophys J* 1998, **74**:773-779.

51. Jian H, Schlick T, Vologodskii A: **Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition.** *J Mol Biol* 1998, **284**:287-296.

52. Vologodskii AV, Marko JF: **Extension of torsionally stressed DNA by external force.** *Biophys J* 1997, **73**:123-132.

53. Strick TR, Allemand JF, Bensimon D, Bensimon A, Croquette V: **The elasticity of a single supercoiled DNA molecule.** *Science* 1996, **271**:1835-1837.

54. Bouchiat C, Mézard M: **Elasticity model of a supercoiled DNA molecule.** *Phys Rev Lett* 1998, **80**:1556-1559.

55. Rivetti C, Walker C, Bustamante C: **Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility.** *J Mol Biol* 1998, **280**:41-59.

56. Liu GH, Schlick T, Olson AJ, Olson WK: **Configurational transitions**
• **in Fourier series-represented DNA supercoils.** *Biophys J* 1997, **73**:1742-1762.
A reduced variable representation of circular DNA elastic rods that improves on earlier B-spline models is described.

57. Martino JA, Olson WK: **Modeling chain folding in protein-constrained circular DNA.** *Biophys J* 1998, **74**:2491-2500.

58. Ehrlich L, Münkel C, Chirico G, Langowski J: **A Brownian dynamics model for the chromatin fiber.** *CABIOS* 1997, **13**:271-279.

59. Jian H, Vologodskii AV, Schlick T: **A combined wormlike-chain and**
•• **bead model for dynamic simulations of long linear DNA.** *J Comput Phys* 1997, **136**:168-179.
The authors extended the worm-like chain model for DNA by the addition of beads, as well as stretching and electrostatic interactions. Brownian dynamics simulations lead to results that are in good agreement with experimental hydrodynamic data.

60. Klapper I, Qian H: **Remarks on discrete and continuous large-scale**
• **models of DNA dynamics.** *Biophys J* 1998, **74**:2504-2514.
A mathematical analysis points to an inconsistency between discrete and continuous models of DNA bending. This implies that an extra tension-like term needs to be added to the limiting case of discrete models to obtain compatible energy expressions.

61. Münkel C, Langowski J: **Chromosome structure predicted by a polymer model.** *Phys Rev E* 1998, **57**:5888-5896.
A dynamic simulation of a chromatin fibre using a 'multiloop subcompartment' model based on linked loops composed of rod segments is described.

62. Kreth G, Münkel C, Langowski J, Cremer T, Cremer C: **Chromatin structure and chromosome aberrations: modeling of damage induced by isotropic and localized irradiation.** *Mutat Res* 1998, **404**:77-88.

63. Harvey SC, VanLoock MS, Rasterwood TR, Tan RKZ: **Molecular modeling studies on the ribosome.** In *Molecular Modeling of Nucleic Acids. ACS symposium 682.* Edited by Leontis NB, SantaLucia J Jr. Washington DC: American Chemical Society; 1998:369-378.

64. Vanloock MS, Easterwood TR, Harvey SC: **Major groove binding of the tRNA/mRNA complex to the 16S ribosomal RNA decoding site.** *J Mol Biol* 1999, **285**:2069-2078.

65. Schaefer M, Bartels C, Karplus M: **Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solution model.** *J Mol Biol* 1998, **284**:835-848.
Dynamics of two small peptides using the ACE continuum solvent model and the CHARMm force-field were shown to lead to folded conformations that are compatible with NMR data. Poor results are obtained using distance-dependent dielectric functions.

66. Liu DJ, Day LA: **Pf1 virus structure: helical coat protein and DNA with paraxial phosphates.** *Science* 1994, **265**:671-674.