2260–2267 *Nucleic Acids Research, 1996, Vol. 24, No. 12*

Modelling extreme stretching of DNA

Anne Lebrun and Richard Lavery*

Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, 13 Rue Pierre et Marie Curie, Paris 75005, France

Received March 15, 1996; Revised and Accepted April 30, 1996

ABSTRACT

Molecular modelling with Jumna is used to study extreme stretching of the DNA double helix. The results, which correlate well with recent nanomanipulation experiments, show how the double helix can be extended to twice its normal length before its base pairs break. Depending on the way the duplex is stretched two types of conformation can occur, either an unwound flat ribbon or a narrow fibre with negatively inclined base pairs. The energetics of both types of deformation are similar and existing structures show that at least the flat ribbon form can exist locally under biological conditions.

INTRODUCTION

DNA is known to be a very flexible molecule and there is already considerable experimental data showing that important deformations of the double helix occur in biological environments. Amongst these one can site helix bending and kinking, base pair opening and strand separation, allomorphic transitions, supercoiling and so on. As structural data on DNA–protein complexes accumulate, it has also become clear that proteins, either in recognising the double helix or as part of their functional role, often induce important local distortions and multiple protein binding, as in the case of RecA (1), can extend such deformations over long tracts of DNA. It is thus important to understand the mechanics of conformational changes which take us into the rarely studied area far from the canonical forms of the double helix.

Before making such studies on specific protein–DNA complexes it seems important to understand the behaviour of the DNA itself. Unfortunately, until recently this was not an easy task. Two developments have however changed this situation. First, the invention of techniques for manipulating individual DNA molecules have made it possible to obtain direct physical data on at least certain types of extreme deformation (2–6). Secondly, development of the Jumna program (7) for modelling nucleic acids has made it possible to simulate large conformational changes in reasonable amounts of computer time. We have taken advantage of both these developments to study one important aspect of DNA deformation, namely helix stretching.

DNA stretching has been studied in two types of nanomanipulation experiment. In the so-called molecular combing experiments of Bensimon *et al.* (2,3), DNA molecules are attached by their ends to a glass surface, while suspended in a droplet of solution. This droplet is then allowed to evaporate, causing the retreating meniscus to exert a traction on the molecule and resulting in its extension as it becomes bound to the glass surface. Maximum extensions of roughly two times the original length of the DNA molecule have been measured. In the case where both ends of a given molecule are attached, the force exerted by the retreating meniscus is sufficient to lead to double-strand breakage (estimated to occur at ~500 picoNewtons). Measurement of DNA stretching in solution rather than on a surface has also been possible through the use of a novel nanomanipulation apparatus which enables a single DNA molecule (~15 μm in length) to be fixed to a glass fibre at one end and to a microbead at the other. Pulling the bead using a micropipette leads to a deflection of the glass fibre. Passing a laser beam through the fibre and onto a photosensitive detector enables the force applied to the DNA molecule to be measured as a function of its extension (4, see also 5,6). This experiment shows that a plateau occurs in the force curve for extensions of ~1.3–1.6 times the original contour length of the DNA molecule, implying the presence of a structural transition.

Such experiments provide very valuable data on the physical properties of DNA molecules, but give little insight into the conformational changes taking place at the molecular level. In order to probe these mechanisms we have used the Jumna program (7). By employing a combination of helical and internal variables, and by optionally including helical symmetry constraints, this program greatly facilitates energy minimisation and allows large scale conformational changes to be mapped out. Preliminary studies in collaboration with one of the groups carrying out the nanomanipulation experiments on DNA have shown that Jumna modelling is capable of reproducing both the plateau observed in the force versus extension curve as well as the limiting extension of the molecule (4). On this basis, we now take a further step and investigate the underlying mechanism of the extreme stretching of DNA. We also provide evidence that the resulting conformations are within the realms of biological possibility and have indeed already been observed for short tracts of DNA.

MATERIALS AND METHODS

DNA modelling has been carried out using Jumna (JUnction Minimisation of Nucleic Acids) (7). This program differs from

* To whom correspondence should be addressed

most molecular modelling approaches in that it represents DNA using internal and helicoidal coordinates, rather than Cartesian atomic coordinates. All bond lengths are taken to be fixed and valence angle changes are limited to the phosphodiester backbone and the sugar rings. These choices enable the total number of variables representing a nucleic acid fragment to be reduced by roughly a factor of 10. If, in addition, helical symmetry is imposed, then a further factor can be gained. Whereas one turn of double helical DNA (~650 atoms) is normally represented by nearly 2000 Cartesian variables, Jumna requires ~300 variables without symmetry and a minimum of only 30 variables with symmetry. This significant reduction, coupled with the use of chemically meaningful variables (single bond rotations and valence angle deformations), greatly improves energy minimisation.

Helicoidal coordinates are introduced by breaking the nucleic acid into a series of 3′-monophosphate nucleotides. Each nucleotide is positioned with respect to the helical axis system, using three translational (Xdisp, Ydisp, Rise) and three rotational (Inclination, Tip, Twist) variables. The internal movements of the nucleotide are represented by sugar puckering (four independent variables), the glycosidic bond and bond rotations and valence angles within the phosphodiester backbone. Inter-nucleotide links (O5′–C5′) are maintained using quadratic distance constraints, causing the bond rotations involving this linkage to become dependent variables.

Helical symmetry can be imposed by simply equating all symmetrically equivalent variables (rather than by introducing extra constraints, as would be the case with Cartesian coordinates). Using symmetry, it becomes possible to model effectively infinitely long polymeric nucleic acids by optimising the energy per monomer unit in the presence of a sufficient number of neighbouring monomers. This procedure saves computational effort and avoids the end-effects associated with studies of oligomeric fragments.

Jumna uses the Flex force field, developed specifically for nucleic acids (7–8), which includes Lennard-Jones and electrostatic terms between non-bonded atoms (including an angle dependent hydrogen bonding term) in addition to valence angle and bond torsion contributions. Solvent damping of electrostatic interactions is treated using a sigmoidal distance dependent dielectric function (8,9) and counterion damping is mimicked by a reduction of the net phosphate charge to –0.5e. Although this is a rather simple model, which ignores detailed solvent and salt effects, it has been found to yield double helical structures, and to predict conformational transitions, in good agreement with experimental data (10,11). Lastly, this representation of DNA enables any structural feature of the model to be controlled very simply. This feature has already been extensively used to map the energy changes associated with changes in sugar pucker and with backbone or helical conformations (8,12). Here we use adiabatic mapping as a function of chosen distance constraints to stretch the DNA double helix, successive energy minimisations being carried out with regular steps of 0.5 Å.

RESULTS

We have used the possibilities offered by Jumna to map the conformational and energetic effects of stretching duplex DNA. Following recent nanomanipulation experiments (2–6) described in the Introduction, we have stretched DNA to its limit when, in our modelling study, strand separation occurs by rupture of the Watson–Crick hydrogen bonds. We also investigate the effects of base sequence and of allomorphic form and, in particular, show that the conformational changes induced are sensitive to the way the duplex is stretched.

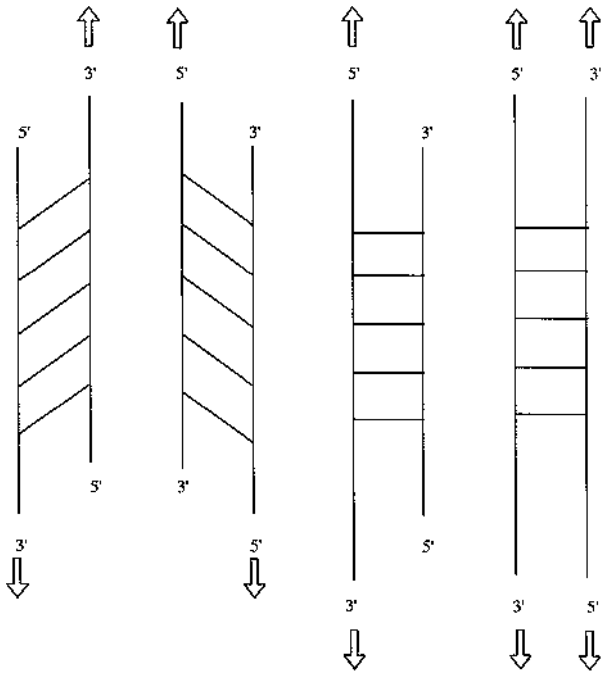

Figure 1. Ways of stretching double-stranded DNA (arrows indicate the extremities which are pulled. The orientation of the duplex corresponds to looking into the major groove and, viewed in this way, positive inclination appears as an anticlockwise rotation of the base pairs).

Ways of stretching duplex DNA

Since the double helix contains two anti-parallel strands (whose direction is conventionally defined as 5′→3′ on the basis of the sugar linkages of each nucleotide), stretching can be carried out in four different ways: (i) using the two 3′-termini; (ii) using the two 5′-termini; (iii) using the 5′ and 3′ termini of one strand or (iv) using the 5′ and 3′ termini of both strands simultaneously (Fig. 1). Each of these techniques is topologically distinct and could lead to a different type of DNA deformation. Notably, it can be seen that the first two pathways should tend to induce positive and negative base inclination respectively (positive inclination is defined as right-handed rotation of the base pairs around a pseudodyad axis pointing into the major groove of the duplex) (13). The result of 5′3′ stretching of one or both strands is less clear since base pair inclination is not necessarily generated. It is also probable that the deformation resulting from each of these pathways will be modulated by the twist of the double helix, which, under physiological conditions, is right-handed.

Stretching polymeric DNA with symmetry constraints

We began our studies of these different stretching processes using polymeric DNA with three regular base sequences: poly(dC–dG)·poly(dC–dG), poly(dT–dA)·poly(dT–dA) and poly(dC–dA)·

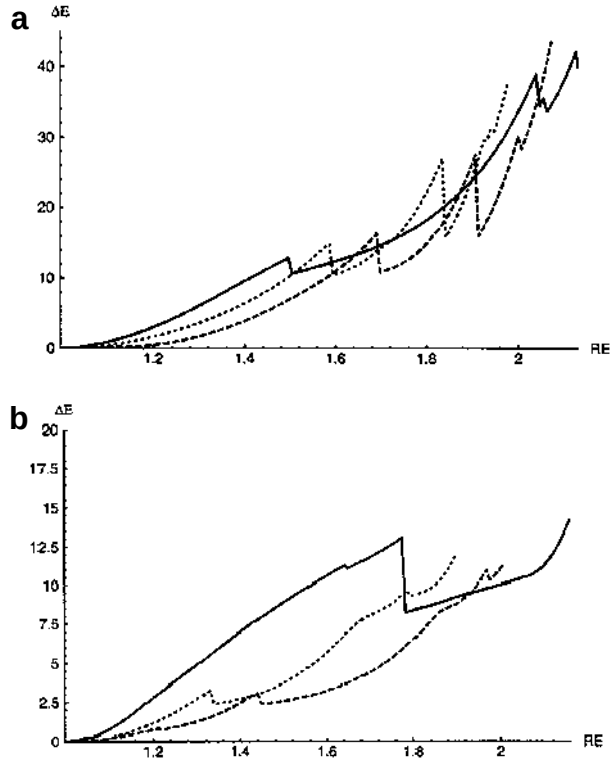

Figure 2. Deformation energy curves (Kcal/mol) as a function of relative extension for B-DNA under dinucleotide symmetry constraints (3′3′, solid line; 5′5′, dashed line; 5′3′, dotted line): (**a**) alternating CG sequence, (**b**) alternating TA sequence.

poly(dT–dG). Due to the choice of repeating sequences we can impose helical symmetry on the DNA duplex and, by considering 16 base pair interactions on either side of a double-stranded dinucleotide unit cell, we can effectively model an infinite length polymer. This approach simplifies the conformational space to be studied by strongly reducing the number of independent variables in the system. A comparison with unconstrained simulations is presented in a following section.

Using this approach, the deformation energy as a function of stretching an alternating CG sequence is shown in Figure 2a. Stretching can be conveniently measured as a relative extension, RE, that is equal to the actual length of DNA divided by the length of the corresponding relaxed conformation. The energy cost of fully stretching the DNA duplex turns out to be roughly similar whichever way the pulling is performed. All energy curves show that ~30 Kcal/mol per nucleotide pair is necessary to stretch the duplex to a little more than twice its normal length, although the 3′3′ pathway is less favourable for extensions $\lesssim 1.5$. Beyond an extension of ~2.1 times the initial length, the base pairs are completely broken and the two strands slip with respect to one another. All the energy curves in Figure 2 show sharp jumps in the energy during the stretching process (most pronounced for the 5′5′ and 5′3′ curves). These jumps correspond to sudden conformational rearrangements induced during stretching. This effect is necessarily amplified by using regular repeating base sequences and imposing symmetry constraints on DNA (*vide infra*).

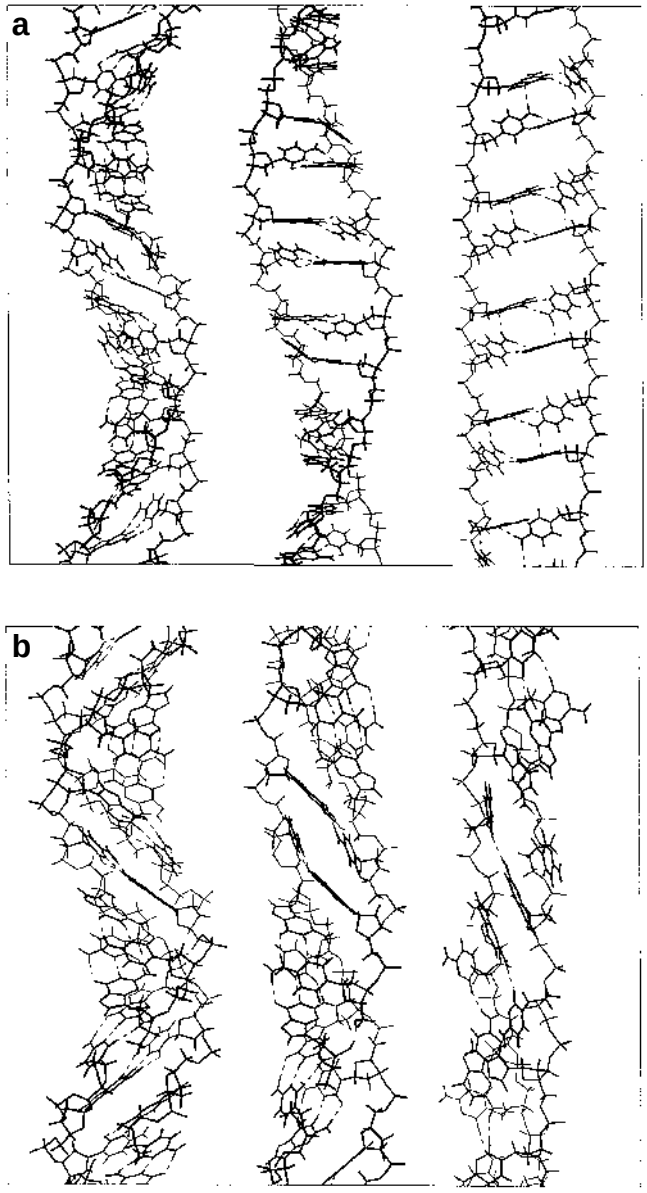

Figure 3. Conformations of B-DNA with an alternating CG sequence (DNA is oriented with its major groove side facing the viewer in the centre of the figure and dotted lines indicate hydrogen bonds) at relative extensions of 1.3, 1.6 and 1.9 (left to right) along the various stretching pathways: (**a**) 3′3′, (**b**) 5′5′.

Stretching the 3′ ends of duplex DNA

Although the energy variations do not depend strongly on the way of pulling, this is not true for the conformational changes that are induced. Figure 3A shows three steps along the 3′3′ stretching pathway, at relative extensions of 1.3, 1.6 and 1.9. Since the helical rise between successive base pairs necessarily increases as the duplex is stretched, the conformation must change to maintain vertical stacking interactions. At the initial stage of stretching this leads principally to base pair buckling, however, as stretching progresses, the minor groove opens and buckling is replaced by the formation of intercalation sites at each CpG step and restored stacking at the GpC steps. Note that stacking is sacrificed

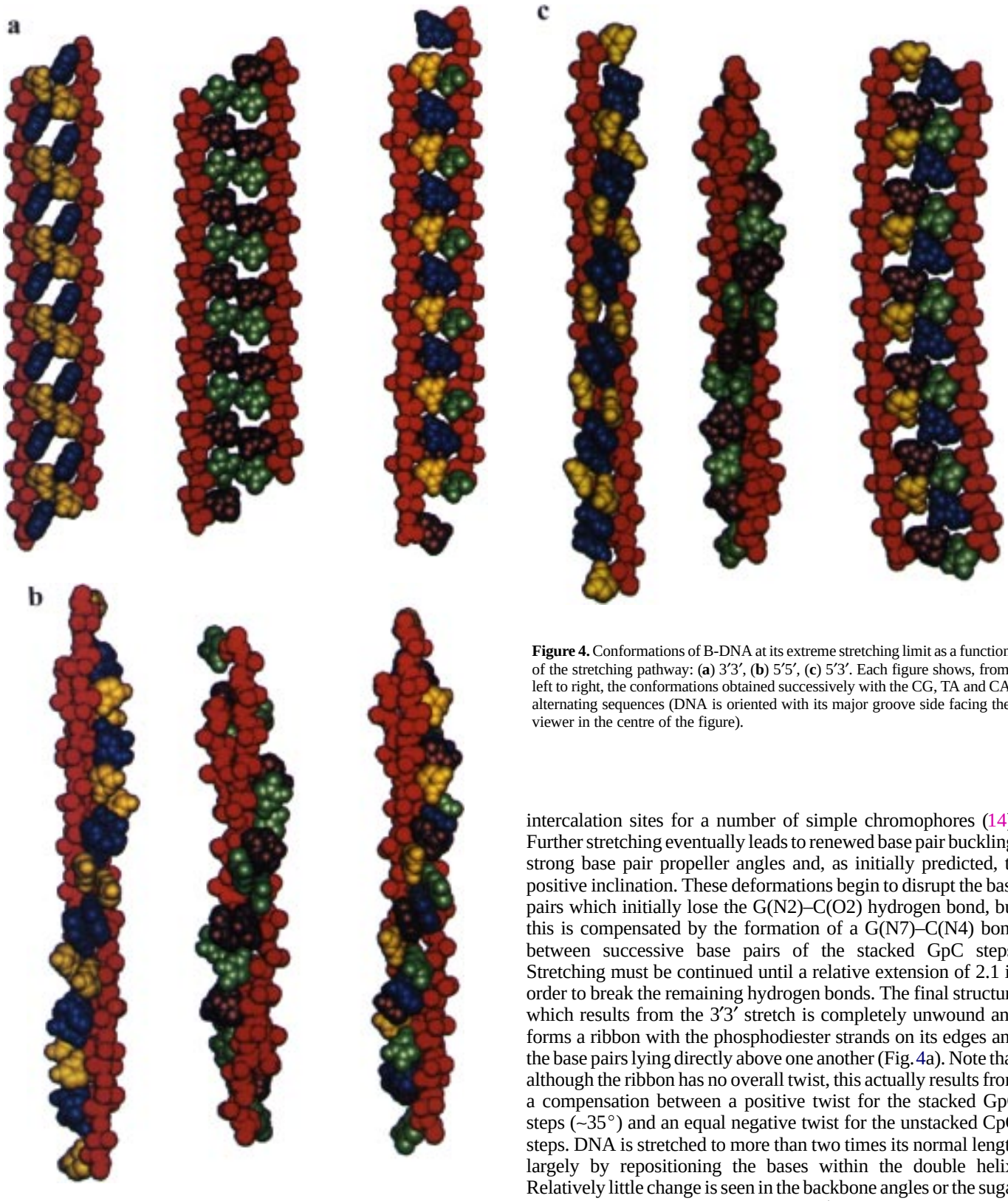

Figure 4. Conformations of B-DNA at its extreme stretching limit as a function of the stretching pathway: (**a**) 3′3′, (**b**) 5′5′, (**c**) 5′3′. Each figure shows, from left to right, the conformations obtained successively with the CG, TA and CA alternating sequences (DNA is oriented with its major groove side facing the viewer in the centre of the figure).

preferentially at the steps which have a lower stacking energy in the relaxed DNA duplex (CpG: –6.3 Kcal/mol versus –13.8 GpC Kcal/mol). These steps are also well known as the preferential intercalation sites for a number of simple chromophores (14). Further stretching eventually leads to renewed base pair buckling, strong base pair propeller angles and, as initially predicted, to positive inclination. These deformations begin to disrupt the base pairs which initially lose the G(N2)–C(O2) hydrogen bond, but this is compensated by the formation of a G(N7)–C(N4) bond between successive base pairs of the stacked GpC steps. Stretching must be continued until a relative extension of 2.1 in order to break the remaining hydrogen bonds. The final structure which results from the 3′3′ stretch is completely unwound and forms a ribbon with the phosphodiester strands on its edges and the base pairs lying directly above one another (Fig. 4a). Note that although the ribbon has no overall twist, this actually results from a compensation between a positive twist for the stacked GpC steps (~35°) and an equal negative twist for the unstacked CpG steps. DNA is stretched to more than two times its normal length largely by repositioning the bases within the double helix. Relatively little change is seen in the backbone angles or the sugar puckers, apart from a tendency of the ξ dihedral to move from g^- towards t and an αγ transition $g^-g^+ \rightarrow tt$ for the GpC steps at the end of the stretching pathway (of the type observed in certain oligonucleotide crystal structures, 15). The only other notable change involves the guanidine glycosidic angle which moves to a syn conformation as the structure deforms.

Stretching the 5′ ends of duplex DNA

If we now consider 5′5′ stretching shown in Figure 3b, early stages of the stretch again induce base pair buckling coupled with partial unwinding. However, the inclination of the base pairs is now negative, as expected, and the minor groove width is reduced with extension. As stretching continues, no intercalation sites appear, the buckling remains and the bases pairs adopt an increasingly negative inclination. In contrast to the 3′3′ pathway, only moderate unwinding occurs (~5°–10° per base pair step). Beyond a relative extension of 1.7, the base pairs break open towards the major groove side of the duplex, losing successively their G(O6)–C(N4) and G(N1)–C(N3) hydrogen bonds. This brings together the phosphodiester strands on the minor groove side of the duplex and improves the interstrand stacking of the guanines. Finally, clusters of CGGC bases are formed, two intrastrand CG (or GC) stacks being coupled with an interstrand GG stack. This deformation leads to a very narrow duplex with bases thrust outwards on the side of the major groove and almost aligned with the axis of the duplex, quite unlike the flat ribbon obtained by the 3′3′ stretch (Fig. 4b). Once again, changes in backbone conformation are limited to more negative values of ξ, an $\alpha\gamma\, g^-g^+ \rightarrow tt$ transition at both CpG and GpC steps and changes in the glycosidic angles which become more negative, again showing a strong correlation with inclination. (Negative inclination however pushes the nucleotides further into the anti domain and can consequently continue without resistance from either purines or pyrimidines.) We can see from these results that the intrinsic handedness of the DNA duplex does indeed strongly modify deformation as a function of the stretching pathway.

Stretching a single strand of duplex DNA

Lastly, what happens when we only stretch one strand of the duplex? The conformation obtained for the alternating CG sequence (Fig. 4c) in fact strongly resemble those along the 5′5′ pathway and consequently lead to the narrow fibre-like conformation with negatively inclined bases. Even the energy jumps occurring along the stretching pathway, corresponding to sudden conformational rearrangements, are almost identical for the 5′3′ and 5′5′ energy profiles, apart from an offset of roughly RE = 0.1 (Fig. 2a). Since this type of deformation, and notably the negative inclination, is not imposed by pulling a single strand, this result confirms the energetic preference for this pathway, at least for relative extensions ≤1.5 (Fig. 2a).

Base sequence effects

The deformation energy curves for the AT alternating sequence are shown in Figure 2b and have a similar appearance to those obtained with the alternating CG sequence. However, the total energy needed to fully stretch DNA drops to ~15 Kcal/mol, only half the value found for the alternating CG sequence. The 3′3′ pathway again leads to a ribbon-like structure, similar to that obtained with the CG sequence, although unwinding is still not complete at a relative extension of 1.9 (Fig. 4a). This conformation however has very different hydrogen bonding. Relatively early in the stretching pathway, alternate AT base pairs become very distorted and finally rupture. In consequence, alternate thymines interact with a total of three other bases—their usual paired adenine, via a single A(N6)–T(O4) hydrogen bond, and with both of the bases of the disrupted neighbouring pair—thymine via a T(O4)–T(N3) bond and adenine through an intrastrand A(N6)–T(O2) hydrogen bond. Despite these deformations, backbone modifications are again limited to more negative ξ and χ values and no $\alpha\gamma$ transitions are observed.

The 5′5′ and 5′3′ pathways again resemble one another with strong negative base pair inclination leading to a narrow fibre with the bases on the major groove side and the phosphodiester strands close together on the minor groove side (Fig. 4b and c). These conformations are dominated by interstrand stacking, involving TT and AA pairs on the 5′5′ pathway and TAAT clusters on the 5′3′ pathway, very similar to the CGGC clusters seen with the CG alternating sequence. Also in common with the CG sequence, the hydrogen bond on the major groove side of the base pairs, here an A(N6)–T(O4) bond, is the first to break, allowing the bases to open outwards to the major groove. The fact that the 5′3′ and 5′5′ pathways resemble one another is again in line with the lower deformation energy of the latter pathway compared to the formation of the 3′3′ ribbon structure (Fig. 2b).

Lastly, we consider the alternating sequence CA whose stretched conformations are also shown in Figure 4. This sequence behaves similarly to the others studied for both 3′3′ and 5′5′ stretching. The final 3′3′ conformation is once more a flat ribbon, although it is narrower than those of either the CG or TA sequences due to the formation of interstrand AG and CT stacking. The 5′5′ pathway, as expected, forms a narrow fibre with negatively inclined bases, involved in stacked CAGT clusters. The surprise with this sequence comes from the 5′3′ pathway which, unlike the former sequences, forms a ribbon conformation most closely resembling the 3′3′ ribbon of the TA alternating sequence. This result, which can be related to the relative stability of the 3′3′ stretching pathway for small relative extensions (results not shown), demonstrates that pulling on a single strand of the DNA duplex can lead to either of the characteristic conformations obtained by pulling both 3′ or 5′ extremities.

Calculations without symmetry constraints

In order to test the effect of symmetry constraints upon these calculations, we have firstly repeated the various stretching pathways for the alternating CG and TA sequences replacing the normal dinucleotide constraint with a decanucleotide constraint. This change, while allowing us to continue to simulate an infinitely long polymeric DNA, increases the number of degrees of freedom in the system by a factor of five. The results obtained in this way yield very similar conformations for both the 3′3′ and 5′5′ stretching pathways, whichever base sequence is considered. The deformation energies per base pair are however smaller, for any given relative extension, and the energy curves are smoother, since sudden conformational changes are not obliged to occur at the same moment within every dinucleotide step. For the 5′3′ pathway, the results with the TA sequence are again very similar to the dinucleotide symmetry results, however, the CG sequence now goes to a ribbon conformation almost identical to that of the 3′3′ pathway. This once more emphasises that stretching a single strand of the duplex does not predetermine the type of deformation pathway that will be followed. Since the deformation energy curves are relatively smooth under decanucleotide symmetry constraints it is possible to make a polynomial fit and to derive these polynomials to obtain the force necessary to stretch the DNA. The resulting curves either show a force plateau or a sinusoidal variation of the force in the region

between relative extensions of ~1.3–1.7 with forces varying between 150 and 350 pN. These values are higher than the recent experimental result of 70 pN (4), but exact agreement cannot be expected given the sensitivity of the force to the exact shape of the deformation energy curve and the limitations of the simplified model presently being used.

In order to take a final step towards simulating a long unconstrained DNA fragment with an inhomogeneous base sequence, we have carried out calculations on a 15 base pair oligomer with the sequence GCGTATATAAAACGC. This oligomer was stretched to a relative extension of 2.0 using both the 5′5′ and 3′3′ stretching pathways. The conformations obtained along these pathways again resemble respectively the fibre and ribbon structures discussed above although the deformation develops inhomogeneously, the GC-rich ends deforming first along the 5′5′ pathway and the central AT-rich segment deforming earlier along the 3′3′ pathway. There is also some irregularity in the fully stretched structures, with partial base pair breakage being limited the 3–5 central nucleotide pairs. Energetically, the 5′5′ stretching pathway is preferred, leading to a deformation energy per nucleotide pair of 2 Kcal/mol at RE = 1.6, which is the same as the value obtained by fitting to the experimental results (4). In contrast, 3′3′ stretching to the same relative extension requires 5.7 Kcal/mol (and leads to a force plateau at ~300 pN). The energetic preference for the fibre structure is again confirmed by the fact that stretching a single strand of this oligomer again leads to the fibre-like conformation.

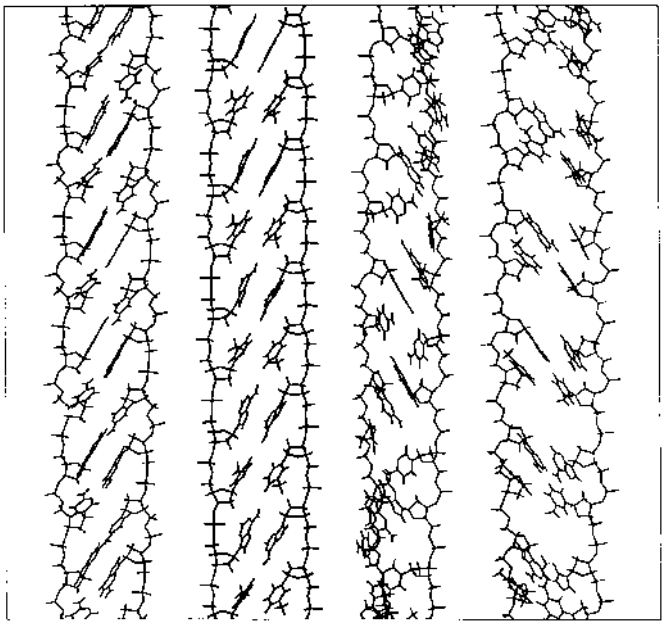

Figure 5. Conformations of A-DNA at its extreme stretching limit as a function of the stretching pathway and sequence. From left to right: 3′3′, CG alternating; 3′3′, TA alternating; 5′5′, CG alternating; 5′5′, TA alternating. (DNA is oriented with its major groove side facing the viewer in the centre of the figure.)

Stretching A-DNA

It is worth briefly considering whether different allomorphic forms of the double helix respond in the same way to this type of stress. We have addressed this question with some preliminary calculations on A-DNA. Studies were made of both CG and TA alternating sequences by the 3′3′ and 5′5′ pathways. The final conformations of these four simulations (Fig. 5) show that changing the allomorphic form indeed affects stretching deformations. The 3′3′ structures differ principally by an increase in positive inclination, while the 5′5′ pathway leads to smaller negative inclination and less unwinding. The bases remain between the phosphodiester strands rather than being pushed out to form a fibre-like structure. These change can be linked to the natural preference of A-DNA for a positive inclination, which apparently continues to be felt in the stretched forms. Inclination is once again directly coupled to the glycosidic torsions and the high positive values at the end of the 3′3′ pathway lead the purine nucleotides of both sequences to enter the syn domain. Lastly, both pathways also show $\alpha\gamma$ transitions ($g^-g^+ \rightarrow tt$), partial $g^- \rightarrow t$ transitions in ξ and, for the CG sequence, 5′5′ stretching forces the cytidine sugar to change from C3′-endo to C2′-endo. All these modifications help in lengthening the phosphodiester backbone of the A-form which is initially ~1 Å shorter per nucleotide step than that of B-DNA.

DISCUSSION

The modelling we have carried out shows that DNA can be stretched, largely by base pair reorientation, to roughly two times its original length before its base pairs must be broken. This result is in good agreement with the maximal extension observed in so-called DNA combing experiments, where strand breakage occurred beyond a relative extension of 2.14 ± 0.2 (2,3). It is remarked that the ~1.7 times stretching limit of recent nanomanipulation experiments on single DNA molecules (4) is due to breaking the molecular ‘scotch-tape’ holding DNA, rather than to the rupture of the DNA molecule itself.

The energetics of stretching show only a minor dependence on the way the duplex is pulled, but base sequence does have an influence, GC pairs leading to greater resistance than AT pairs. The force versus extension curves deduced from our modelling correlate well with the experimental measurements on single DNA molecules and show a force plateau between relative extensions of ~1.3→1.7. The height of the theoretical plateau decreases when helical symmetry is relaxed, but remains at least twice the experimental value of 70 picoNewtons.

The most striking result of these studies is that there are two distinct conformational pathways for stretching duplex DNA. Which path is followed depends on which ends of the DNA duplex are pulled: pulling the 3′ ends leads to a flat ribbon conformation, whereas pulling the 5′ ends leads to a narrow fibre with strongly negatively inclined bases. Still more surprisingly, pulling both ends of a single strand of the duplex also leads to one of these two forms and the same is true when both strand of the duplex are pulled simultaneously. The latter result can be explained by the fact that the relatively rigid base pairs forming the centre of the duplex transmit stress very effectively from one strand to the other. However, in order to understand why there are only two routes for stretching, it is worth considering a little geometry.

We have already remarked that stretching occurs largely by repositioning the base pairs within the duplex and that relatively little deformation is seen within the phosphodiester strands. If, to a first approximation, we assume that the inter-phosphate distance L within each strand remains constant then, as Figure 6 shows, there are indeed two ways to increase the rise R of the

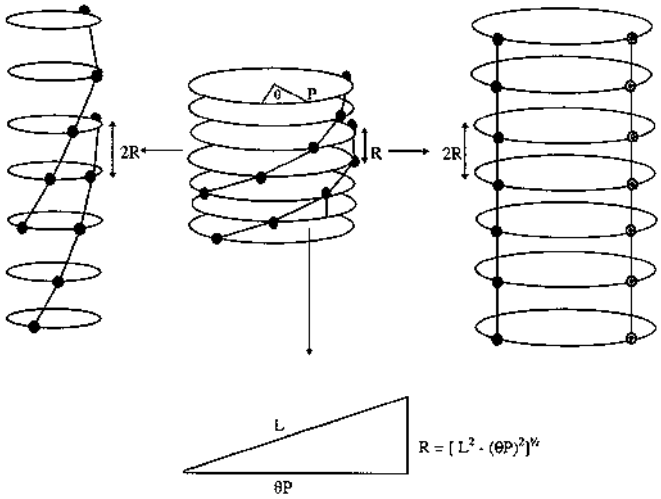

Figure 6. Schematic model of DNA stretching: maintaining a constant inter-phosphate distance L within each strand of the duplex, stretching to twice the normal rise R can be achieved by reducing the twist angle θ or by reducing the radius P of the duplex.

duplex in order to double its length. This requires reducing the base of the triangle θP roughly to zero so that R becomes equal to L. (In canonical B-DNA, L is ~7 Å and R is ~3.4 Å. Setting R = L therefore implies stretching the duplex by slightly more than two times.) θP can be made zero either by unwinding the helix (θ→0) or by reducing its radius (P→0).

In practice, when the 3′ ends of a right-handed duplex DNA are pulled, the bases tend towards positive inclination and the groove widths increase (16). However, as pulling continues further inclination is opposed by coupled changes in the glycosidic angles which lead into the syn domain, unfavourable for pyrimidine nucleotides. Due to this restriction, unwinding takes over, leading finally to the flat ribbon conformation shown on the right of Figure 6. In contrast, pulling on the 5′ ends of the duplex causes strong negative inclination coupled to a reduction of groove width (16). This inclination causes a reduction in the helix radius which continues, with little change of twist, until the narrow fibre conformation shown on the left of Figure 6 is reached. Lastly, when one strand of the helix is pulled (or when both strands are pulled simultaneously) either the ribbon or the fibre form are created as a function of base sequence. This implies that hybrid conformations mixing reduction in helix radius (via base pair inclination) and untwisting are disfavoured. (This finding is not however applicable to A-DNA where both 5′ and 3′ stretching leads to inclination and unwinding).

Our calculations suggest that there is a small energetic preference for the narrow fibre form of stretched DNA over the ribbon form, since it maintains better base stacking. However, the fibre form also has short interstrand phosphate–phosphate distances leading to electrostatic repulsions which may not be correctly treated by our simple solvent/counterion model. [It should nevertheless be remarked that although the phosphodiester backbones do approach one another rather closely in the fibre form, the shortest inter- and intrastrand P–P distances remain ~7.5 Å, which is ~0.5 Å more than the intrastrand P–P distance in canonical B-DNA. We have also found that preliminary calculations with Poisson-Boltzmann electrostatics, using Delphi (17), show no destabilisation of the fibre form.]

It would, in principle, be possible to stabilise the ribbon form by filling the gaps created between successive base pairs and this is indeed what happens when intercalating molecules interact with DNA. Stabilisation can also come from interaction with another macromolecule and there are indeed cases of protein–nucleic acid complexes which contain locally stretched DNA resembling our ribbons. Such conformations are induced by the minor groove binding proteins, SRY (18) and TBP (19,20), which both unwind and stretch the duplex over 4 to 5 base pairs, leading to local relative extensions of 1.83 and 2.16 respectively (close to the known limiting values). It is also interesting to note that in the case of the TATA box, an analysis of the bound DNA (21) shows both a transition to A-like sugar puckers and a strong positive inclination in line with our findings for 3′3′ stretched A-DNA. SRY also induces C3′-endo sugars at its binding site. These deformations finally also involve the intercalation of protein side chains between the separate bases of the deformed DNA (22). (A more detailed discussion of local stretching and its role in protein binding will be published elsewhere.)

Since proteins can induce such local structures resembling the ribbon form of DNA one must conclude that the large forces and deformations associated with extreme stretching are not incompatible with biological activity. Since, in addition, our modelling suggests that the energetics of ribbon and fibre form stretching are similar, it seems reasonable to suppose that the fibre form may also occur in biological environments. (It is finally interesting to add that the existence of the fibre-like form appears to be supported by the early experiments of Wilkins *et al*. on stretched DNA fibres (23) which observed an important reduction in fibre diameter and also provided spectroscopic evidence of base pair inclination). [Note: The coordinates of the stretched DNA conformations (in the form of .pdb files) are available from the authors upon request (rlavery@ibpc.fr).]

ACKNOWLEDGEMENTS

The authors wish to thank the Institut de Développement et des Resources en Informatique Scientifique (Orsay, France) for their allocation of computer resources and the Association for International Cancer Research (St Andrews, UK) for their support of this research.

REFERENCES

1 Rao,B.J., Chiu,S.K., Bazemore,L.R., Reddy,G. and Radding,C.M. (1995) *Trends Biol. Sci.* **20**, 109–113.
2 Bensimon,A., Simon,A., Chiffaudel,A., Croquette,V., Heslot,F. and Bensimon,D. (1994) *Phys. Rev. Lett.* **74**, 4754–4758.
3 Bensimon,D., Simon,A., Croquette,V. and Bensimon,A. (1995) *Science* **265**, 2096–2098.
4 Cluzel,P., Lebrun,A., Heller,C., Lavery,R., Viovy,J.-L., Chatenay,D. and Caron,F. (1996) *Science* **271**, 792–794.
5 Smith,S.B., Cui,Y., Hausrath,A.C. and Bustamente,C. (1995) *Biophys. J.* **68**, A250.
6 Smith,S.B., Cui,Y. and Bustamente,C. (1996) *Science* **271**, 795–799.
7 Lavery,R., Zakrzewska,K. and Sklenar,H. (1995) *Comp. Phys. Commun.* **91**, 135–158.
8 Lavery,R. (1994) *Adv. Comp. Biol.* **1**, 69–145.
9 Hingerty,B., Richie,R.H., Ferrel,T.L. and Turner,J.E. (1985) *Biopolymers* **24**, 427–439.
10 Jayaram,B., Swaminathan,S., Beveridge,D.L., Sharp,K. and Honig,B. (1990) *Macromolecules* **23**, 3156–3165.

11 Fritsch,V. and Westhof,E. (1991) *J. Am. Chem. Soc.* **113**, 8271–8277.
12 Lavery,R. and Hartmann,B. (1994) *Biophys. Chem.* **50**, 33–45.
13 Dickerson,R.E., Bansal,M., Calladine,C.R., Diekmann,S., Hunter,W.N., Kennard,O., Lavery,R., Nelson,H.C.M., Olson,W.K., Saenger,W. *et al.* (1989) *J. Mol. Biol.* **205**, 787–791.
14 Saenger,W. (1984) *Principles of Nucleic Acid Structure.* Springer-Verlag, New York, pp. 350–367.
15 Heinemann,U., Lauble,H., Frank,R. and Blöcker,H. (1987) *Nucleic Acids Res.* **15**, 9531–9550.
16 Boutonnet,N., Hui,X. and Zakrzewska,K. (1993) *Biopolymers* **33**, 479–490.
17 Gilson,M.K., Sharp,K.A. and Honig,B. (1987) *J. Comp. Chem.* **9**, 327–335.
18 Werner,M.H., Huth,J.R., Gronenborn,A.M. and Clore,G.M. (1995) *Cell* **81**,705–714.
19 Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) *Nature* **365**, 512–520.
20 Kim,J.L. and Burley,S.K. (1994) *Nature Struct. Biol.* **1**, 638–653.
21 Guzikevich-Guerstein,G. and Shakked,Z. (1996) *Nature Struct. Biol.* **3**, 32–37.
22 Werner,M.H., Gronenborn,A.M. and Clore G.M. (1996) *Science* **271**,778–784.
23 Wilkins,M.H.F., Gosling,R.G. and Seeds W.E., *Nature* **167**, 759 (1951).