Molecular Dynamics Simulation of DNA Stretching Is Consistent with the Tension Observed for Extension and Strand Separation and Predicts a Novel Ladder Structure

Michael W. Konrad* and Joel I. Bolonick

Contribution from GeneVue, Inc., 1199 Camino Vallecito, Lafayette, California 94549, and NZYM, Inc., 1933 Davis Street, Suite 226, San Leandro, California 94577

Received May 23, 1996. Revised Manuscript Received August 12, 1996[⊗]

Abstract: Molecular dynamics simulations were used to model the response of several double-stranded dodecamers to gradually increasing tension applied to the opposing 3′ ends of the two polynucleotide strands. At forces between 0.80 and 1.45 nN, depending on sequence, the strands separated completely. The separation force for one of the dodecamers studied has been measured and is close to that seen in the simulation. Before strand separation, at forces between 0.065 and 0.090 nN, again depending on sequence, there was an abrupt extension and transition to a novel ladder structure in which bases of one strand were stacked on those of the other strand. Sudden extensions have been observed in very long DNA molecules at forces similar to those seen in the simulations. After the abrupt extension but before strand separation, the ladder structure became more regular, and the phosphate backbones became more linear. Throughout the entire molecular extension, most hydrogen-bonded base pairs remained intact.

Introduction

As the cost of computer hardware decreases while becoming more powerful and molecular simulation algorithms and software becomes more sophisticated (for comprehensive reviews, see refs 1 and 2), increasingly ambitious projects become practical. Molecular dynamics simulation has recently been used to examine the process of pulling a biotin molecule from the biotin binding protein strepavidin.[3] The simulation was consistent with the observed separation force but also predicted changes in the protein structure during the process. Molecular dynamics simulations have also been applied to nucleic acids for the purpose of predicting the most probable structure consistent with NMR constraints or predicting conformational variations in time of a known structure at a fixed temperature. In the following paper, we start with the established B-helix structure of DNA and simulate the process of progressive elongation and eventual strand separation caused by increasing tension applied at the opposing ends of the helix.

The extensibility of DNA was first proposed in 1951 by Wilkins et al.,[4] based on altered optical properties of stretched fibers. More recently, a sudden 2.1-fold increase in the length of single large DNA molecules stretched with a receding meniscus was interpreted as a transition from a helix to a discrete new structure.[5] Subsequent experiments using laser tweezers[6,7] or single glass fibers[8] to measure tension while stretching single DNA molecules have produced quantitative force versus length curves, confirming the presence of a sharp transition. Specific linear ladder structures for DNA, qualitatively consistent with the non-helical nature of diffraction patterns from bulk stretched DNA,[9] have been proposed by Subirana[10] and by Yagil and Sussman.[11] However, Cluzel et al.[8] have proposed a stretched helical model for the extended form. At higher forces, the strands of single short double-stranded oligonucleotides can be pulled apart, and the forces have been measured using atomic force microscopy.[12]

Molecular Dynamics Simulations

Molecular dynamics simulations were performed on a Silicon Graphics Indigo[2] computer using the Insight II and Discover (version 2.3.5) software modules from BioSym, Inc. (San Diego, CA). All dynamics simulations were preceded by an initialization phase in which velocities were rescaled to achieve a temperature of 300 K. The integration algorithm used was that of Verlet[13] with a time step of 1 fs.

An implicit solvent representation was implemented by modifying the Discover module, replacing the linear distance-dependent dielectric function with the sigmoidal dependence originally proposed by Hingerty et al.,[14] developed and simplified by Ramstein and Lavery[15] and by Mazur and Jernigan.[16] This sigmoidal function (with $n = 1$), a bulk dielectric constant of 80, and a charge of −1 on phosphates simulates the electrostatic environment of the typical aqueous solvent in which B-form DNA helices are stable. An implicit solvent representation using a sigmoidal distance-dependent dielectric has been successfully used by Fritch et al.[17] in simulation studies of DNA. While a complete and accurate explicit representation of DNA bathed in a large body of

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1996.

(1) Brooks, C. L., III; Karplus, M.; Pettit, B. M. *Proteins: A theoretical perspective of dynamics, structure, and thermodynamics*; Wiley: New York, 1988; Vol. 71.
(2) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: Oxford, 1987.
(3) Grubmuller, H.; Heyman, B.; Tavan, P. *Science* **1996**, *271*, 997−999.
(4) Wilkins, M. H. F.; Gosling, R. G.; Seeds, W. E. *Nature (London)* **1951**, *167*, 759−760.
(5) Bensimon, D.; Simon, A. J.; Croquette, V.; Bensimon, A. *Phys. Rev. Lett.* **1995**, *74*, 4754−4757.
(6) Smith, S. B.; Cui, Y.; Hausrath, A. C.; Bustamante, C. *Biophysical J.* **1995**, *68*, A250.
(7) Smith, S. B.; Cui, Y.; Bustamante, C. *Science* **1996**, *271*, 795−798.
(8) Cluzel, P.; Lebrun, A.; Heller, C.; Lavery, R.; Viovy, J.-L.; Chatenay, D.; Caron, F. *Science* **1996**, *271*, 792−794.
(9) Fornells, M.; Campos, J. L.; Subirana, J. A. *J. Mol. Biol.* **1983**, *166*, 249−252.
(10) Subirana, J. A. *FEBS Symp.* **1970**, *21*, 243−253.
(11) Yagil, G.; Sussman, J. L. *EMBO J.* **1986**, *5*, 1719−1725.
(12) Lee, G. U.; Chrisey, L. A.; Colton, R. J. *Science* **1994**, *266*, 771−773.
(13) Verlet, I. *Phys. Rev.* **1967**, *159*, 98−103.
(14) Hingerty; Ritchie; Ferrell; Turner *Biopolymers* **1985**, *24*, 427−439.
(15) Ramstein, J.; Lavery, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7231−7235.
(16) Mazur, J.; Jernigan, R. L. *Biopolymers* **1991**, *31*, 1615−1629.
(17) Fritsch, V.; Ravishanker, G.; Beveridge, D. L.; Westhof, E. *Biopolymers* **1993**, *33*, 1537−1552.

S0002-7863(96)01751-9 CCC: $12.00

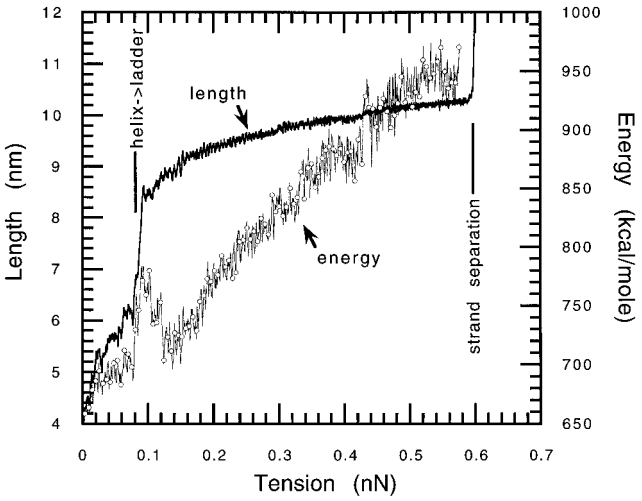

Figure 1. Length and internal energy versus the tension applied to the double-stranded DNA $(ACTG)_3$.

water plus counterions must be, by definition, a more faithful representation, attempts to accomplish this using a modest number of solvent molecules (approximately 400 water and one counterion per nucleotide) have been reported to require extensive periods of stepwise heating and equilibrium to avoid extreme conformational changes.[16] Thus, in this first survey of DNA stretching, we have used an approximation consistent with our computing resources.

The Amber force field, as supplied by BioSym, was used with all charges at their default values. No constraints to enforce helical geometry were added, and the cutoff for non-bonded interactions was set to 1.50 nm. All double-stranded 12-mers were initially constructed as canonical B-form helixes, and each helix was then subjected to 40 ps of molecular dynamics to allow sequence-dependent conformational changes before any stretching force was applied. Average conformations were then computed using the equilibrated portion of the trajectory.

The application of tension was achieved by incrementally increasing the scaling coefficient, CSCALE, of the distance constraint function, CONSTRAIN, usually used to implement NMR constraints. Two anchor points were placed on the axis of the helix, 200 nm from each end, and stretching was implemented by periodically increasing two equal and opposing forces between each anchor and the 3′ hydroxyl carbon at the closest end of the helix. A force increment of 11 pN was found to be sufficiently small to avoid strand separation due to inertial effects. Simulation at constant tension for 40 ps was required for the structure to achieve complete equilibrium. The forces required for the helix to ladder transitions were determined with these parameters. However, the computational time required to achieve strand separation by incrementing the force every 40 ps was prohibitive. Thus, simulations were repeated with the tension incremented every 10 ps. However, these fast simulations to obtain strand separation, which typically required 5 days of CPU time, produced structure transition forces that were only 10−15% higher than seen in the slow simulations. Base stacking energies were computed with the same force field and electrostatic parameters used during simulations but with selected groups of atoms. Representative base stacking for the B-helix was taken from average structures of the dodecamers used as starting structures in the simulations. Base stacking for the S-ladders was taken from average structures of a portion of the simulation in which very little change was occurring. Energies were determined directly from the stacked bases in the average structures without any further minimization or dynamics.

Results

A plot of extension and internal energy versus tension for the double-stranded DNA with sequence $(ACTG)_3$ is seen in Figure 1. This extension curve can be divided into four distinct phases:

Phase 1. The length increases progressively to 1.5 times that of the relaxed helix as the tension approaches 0.08 nN.

Phase 2. A sudden increase in length to 2.1 of that of the relaxed helix occurs at about 0.085 nN.

Phase 3. The molecule stretches an additional 10% as the tension increases to 0.6 nN.

Phase 4. The strands separate at 0.6 nN. This is a rapid and irreversible transition.

The structures of the oligonucleotide at four key stages of the extension are shown in Figure 2.

Structure a. The initial B-helix with the line of tension vertical is through the 3′ OH of T_{12} at the top and the 3′OH of G_{12} at the bottom. Since there are about 10 base pairs per complete turn in a B-helix, the strands of this 12-mer trace out slightly more than one turn.

Structure b. The partially unwound helix is at a tension of 0.077 nN, just before the sudden extension. There is now less than one complete turn. The base pairs are intact but are tilted, but not consistently in the direction of tension. All bases are stacked on bases of the same strand, the normal B-helix pattern.

Structure c. A new, novel structure, the S-ladder is seen at a tension of 0.094 nN, just after the extension. There is essentially no helical twist, most H-bonded base pairing is intact, but most base stacking is now between strands. There are bases in the middle that are still involved in intrastrand stacking. The variation in molecular structure along the length of the DNA is greatest at this stage, which correlates with the peak in the curve of internal energy seen in Figure 1.

Structure d. The S-ladder at a tension of 0.220 nN, when all stacking is between strands (although the nucleotide at the bottom, C_1, is not stacked or paired). The structure is now more regular than in the preceding panel; however, there is variation in the rotation of base pairs relative to the phosphodiester backbone, so they are not all seen edge on in one view.

The integrity of H-bonded base pairs can be monitored as the distance between the two ring nitrogens (purine N_1−pyrimidine N_3) in the pair. The distances for the 10 internal base pairs are plotted versus tension in Figure 3. All N−N distances are maintained at the normal 0.28 nm value except for the middle two. The T_7:A_6 pair is recovered a while after the transition, while the C_6:G_7 pair is never re-established. The lack of H-bonds in this region can be seen in structure c, Figure 2. Disruption of both hydrogen bonding was observed at the ends of the helix as the transition occurred. However, when the simulation was repeated with terminal base pair hydrogen bonding artificially stabilized, transition to the S-ladder occurred within a few percent of the original force. Thus, disruption of hydrogen bonding is neither necessary nor a characteristic part of the transition.

The formation of the new base stacking pattern is associated with a pronounced and sudden decrease in the angle between the plane of the bases and the line of tension, which is also seen as a decrease in distance between the ring N_1 or N_3 and the phosphate toward which it tilts. This distance is plotted in Figure 4 for G_7 and T_8. This parameter emphasizes the most discrete characteristic of the helix to ladder structural transition, the change in base stacking.

An enlarged detail of a segment of the S-ladder structure is seen in Figure 5. This structure was obtained by averaging over 10 ps near the end of the length versus force plateau when length was essentially constant. The top 5′ to 3′ chain, with ...TGAC... visible, has been pulled to the right relative to the bottom 3′ to 5′ chain by slightly more than one nucleotide unit. The two phosphodiester chains and ribose sugars form a flat ribbon, with the base pairs perpendicular to the plane of the ribbon and inclined 22° to the backbone. Nucleotides occur every 0.74 nm along the ladder, but stacked bases are separated by 0.36

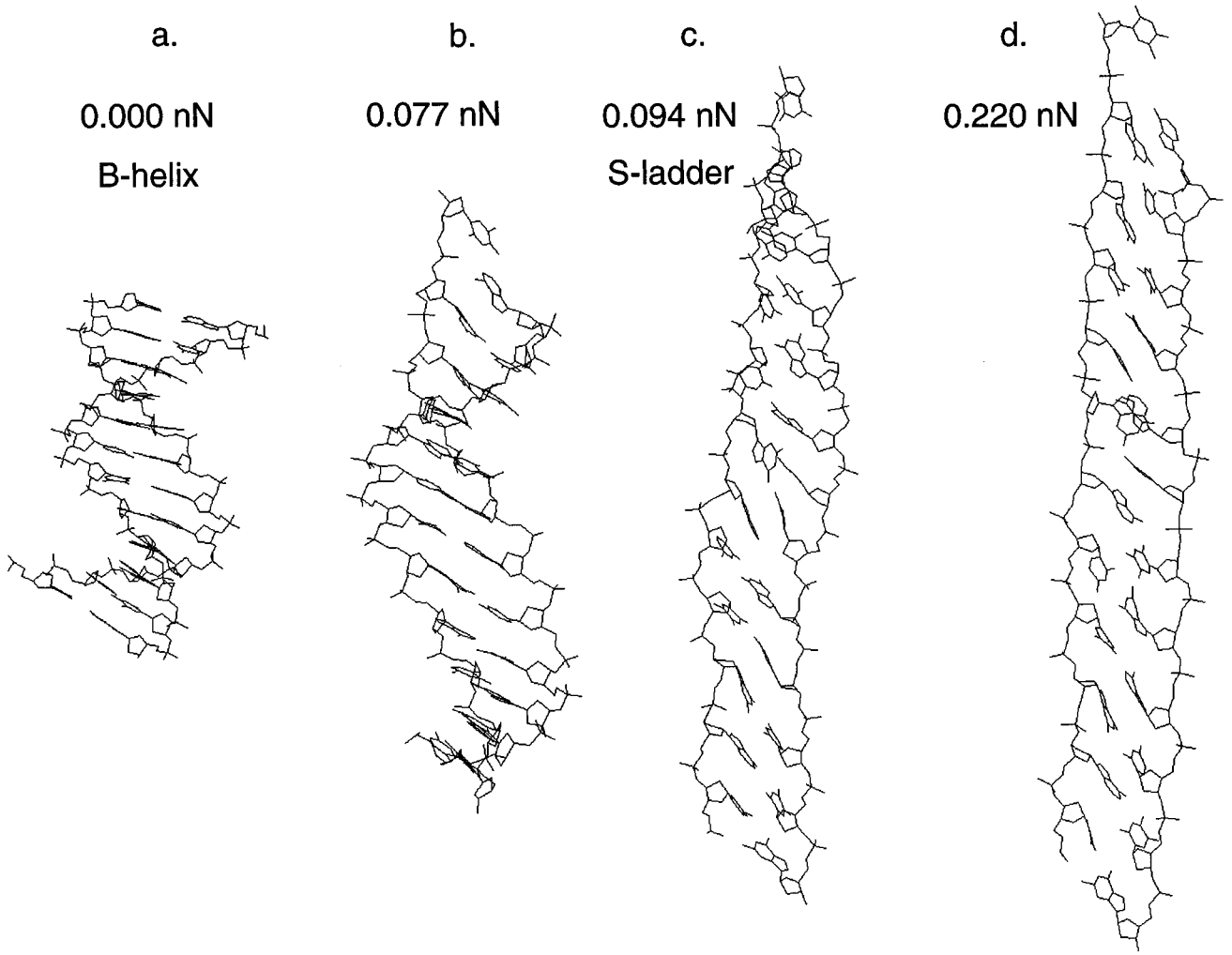

Figure 2. DNA at four characteristic stages of elongation.

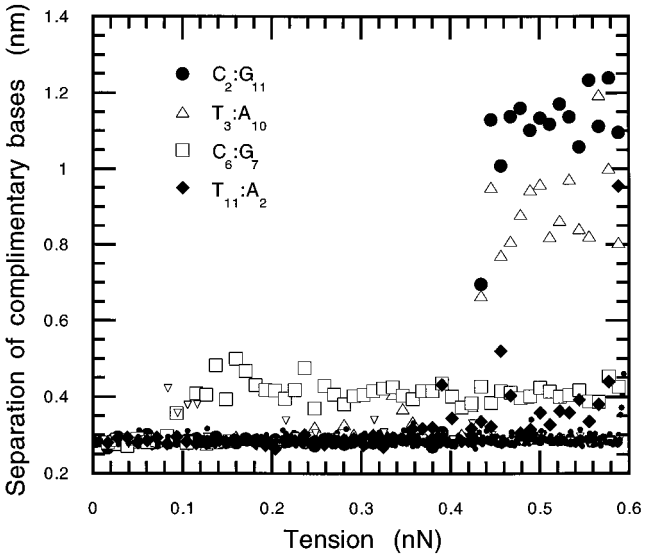

Figure 3. Separation between the paired complimentary bases (atoms $N_1-N_{3)}$ versus tension. Values for the middle 10 base pairs are plotted, and the legend identifies the four pairs for which the normal hydrogen bonding distance is exceeded during extension.

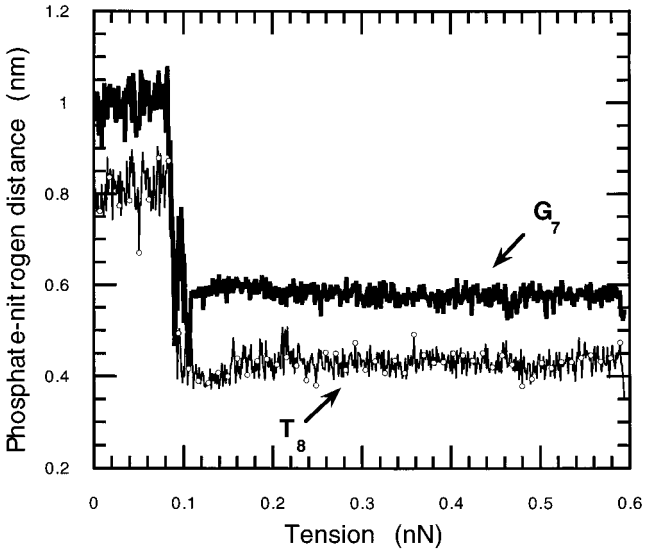

Figure 4. Phosphate-purine N_1 (of nucleotide G_7) and phosphate-pyrimidine N_3 (of nucleotide T_8) separation versus tension.

nm normal to the plane of the bases. The distance between phosphates on opposite chains is 1.08 nm, but phosphates are not exactly opposite each other, since the two chains have slipped in the direction of the tension 0.94 nm, 0.20 nm more than the nucleotide repeat distance. As suggested by Smith et al.,[7] most sugars (19 of 24) in this stretched form are in the *exo*, specifically O_4'-*exo*, configuration as opposed to the *endo* conformation of the unstrained B-helix.

Figure 6 shows, in the plane of the bases, an adenine stacked on guanine in the original B-helix and a similar stacking pair in the S-ladder. In both structures there is extensive overlap, but the relation between the bases is very different. Since the

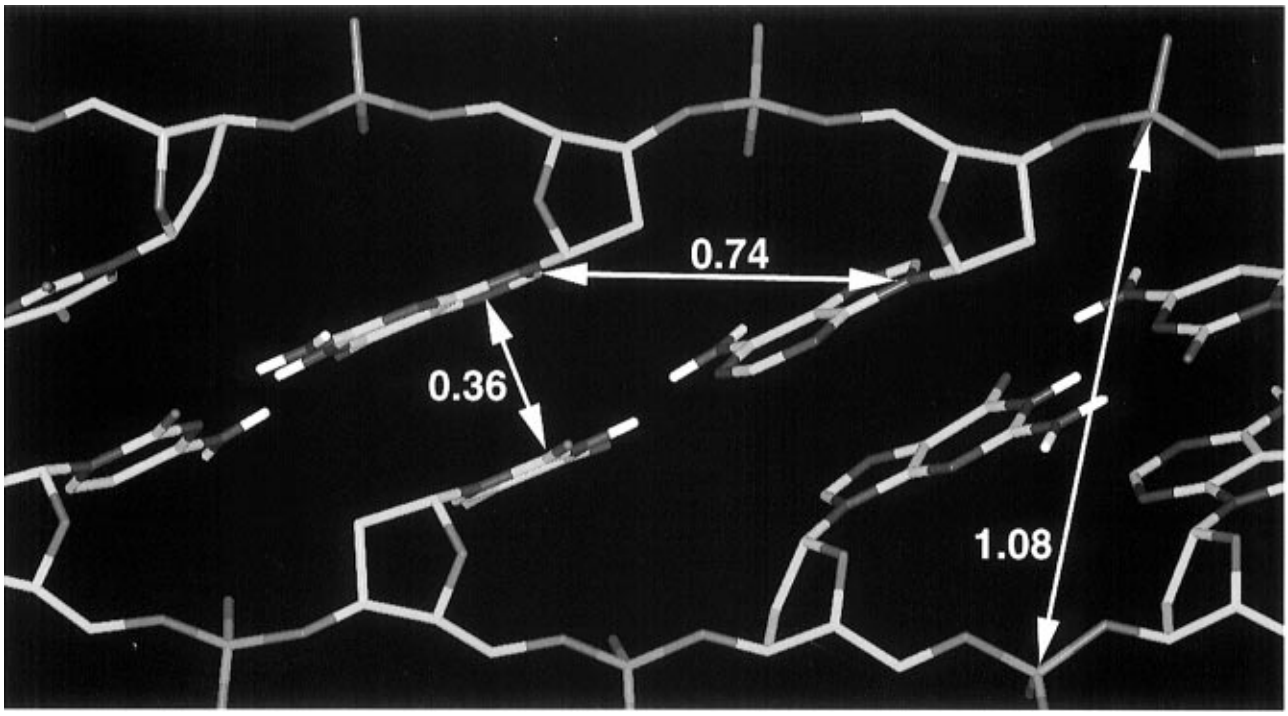

Figure 5. Typical segment of an S-ladder at a tension of 0.4 nN.

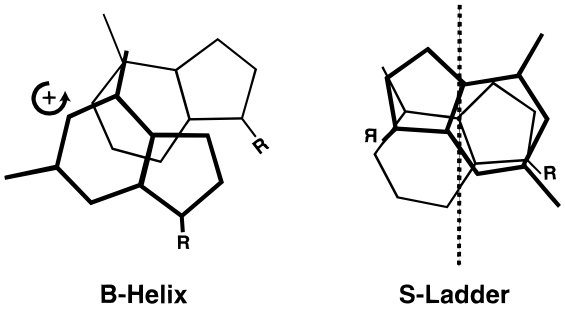

Figure 6. Guanine stacked on adenine in a B-helix and S-ladder.

two purines in the helix are on the same backbone, they are related by a rotation of 34° about the axis of the helix. In the ladder, the adenine and guanine are attached to different, antiparallel backbones and thus are related to each other in the same way as two opposing hydrogen-bonded bases are in the helix: a 180° rotation about a pseudo-C_2 axis. A translation of 0.063 nm toward each other along this axis superimposes the structures. The bonds to the ribose C1′ atoms make an angle of approximately 88° to each other, which is only a little larger than the corresponding angle of 77° seen for opposing bases in the B helix. Thus, there has been only a slight rotation of the bases relative to each other as the new stacking was established.

The stacking energy for the 10 possible base pairs are presented in Table 1 for both the B-helix and the S-ladder. The energies of the individual pairs are generally different in the two structures, which is not surprising since the geometries of the overlaps are quite different. About half are larger in the ladder and half are smaller, thus the average for many sequences could be quite similar. However, in the helix there are two stacked pairs for every nucleotide repeat, while for the ladder there is only one, since each base stacks only with the base on the opposite strand.

Simulations of response to tension were carried out on four additional helixes, each 12 nucleotides long but with different sequences. All simulations produced a force-extension curve qualitatively similar to Figure 1, but with forces required for transition and for strand separation that were characteristic of

Table 1. Computed Stacking Energies (kcal/mol) of Bases in a B-Helix and S-Ladder

pair	E_B (helix)	E_S (ladder)	$E_S - E_B$
AA	−12.0	−11.3	+0.7
AC	−8.2	−9.6	−1.4
AG	−13.7	−12.5	+1.2
AT	−10.1	−9.0	+1.1
CC	−8.4	−8.4	0.0
CG	−13.7	−10.5	+3.2
GG	−12.2	−11.1	+1.1
TC	−9.0	−9.1	−0.1
TG	−10.7	−11.2	−0.5
TT	−7.8	−8.0	−0.2

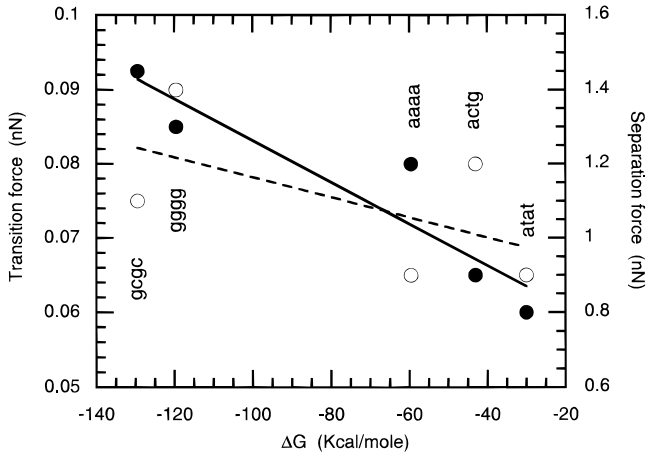

Figure 7. Dependence of transition and strand separation force on nucleotide composition and sequence.

the sequence. In Figure 7, both forces are plotted versus the change in Gibbs free energy, ΔG, for helix denaturation.[18]

The dashed line is the linear least squares fit to the transition forces, with a correlation coefficient of 0.58, and the solid line is fit to separation forces, with a correlation coefficient of 0.93. The correlation of separation force to ΔG is appreciable,

(18) Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746−3750.

considerably higher than the correlation of transition force to ΔG. There is of course no fundamental, direct relation between separation force and ΔG; they do not even have the same units. However, since the extension−tension curves are similar for the helixes, with strand separations at about the same extensions, one might expect the force at strand separation to be correlated with the energy of denaturation when all interstrand interactions are abolished. In contrast, the helix to ladder transition is more complex, involving the substitution of one base stacking pattern with another, keeping hydrogen bonding intact. The helix that deviates most from the trend line is polyA:polyT, which is known to have a unique B-helix structure.[19]

Discussion

Several groups have studied stretched forms of DNA from bacteriophage λ, which is about 50 000 base pairs long with an AT/GC ratio of 0.99. Individual molecules of λ DNA with half the linear density of a B-helix have been observed by atomic force microscopy.[20] Tension on λ DNA produced by a receding meniscus was found to produce a transition to a form 2 times the normal length at forces of 0.084−0.16 nN.[5] Detailed force versus length curves for λ DNA have recently been obtained by two groups, with an abrupt increase in length at 0.08 nN seen using a glass fiber,[8] and a similar sharp transition observed at 0.065 nN using laser tweezers[7] to generate the stretching tension.

We believe the transition at 0.085 nN produced by the computer simulation of stretching the $(ACTG)_3$ helix is the same transition that occurs at essentially the same force during stretching of the much longer λ DNA. In the λ experiments tension was the dependent variable, while in simulations length plays that role (thus the vertical and horizontal axes of the graphs in this report have to be interchanged to be compared to the experimental work). More fundamentally, the 15 μm long λ DNA molecule undergoes a smooth geometric transition from a random coil to a asymptotically straight rod as the tension is increased, which is superimposed on the extension of contour length caused by unwinding of the helix. The dodecamers are far too short to display such an effect. In addition, the behavior at the ends of a small helix, where the bases are not stacked and sometimes not part of a H-bonded pair, plays a much larger part of the total extension as compared to λ DNA. The quantitative relationship between the length versus tension curve and the number of base pairs in the DNA is beyond the scope of this introductory study.

The S-ladder structure generated by the simulations is defined by tension, the covalent bond geometry of the polynucleotide chains and interaction forces from base−base stacking, complimentary base pair hydrogen bonds, and repulsion of the charged phosphate groups on the two chains. However, the final length is defined only by the phosphodiester−ribose bond geometry and not by forces between the nucleotides and thus not by the nature of the force fields used in the simulations. Thus, the observed maximum extension of single- and double-stranded DNA is the same.[7] The value of the tension at which the transition occurs is of course a function of the Amber force field used in the simulations. The S-ladder is a very open molecule, in which the entropic force favoring contraction is balanced only by the externally applied tension. Thus, while there is considerable periodicity in structure along the ladder, which increases as the tension approaches the limit before strand separation, there is also variability with time and position. Changes in the plane of the ribose rings along the ribbon is evident in panel d of Figure 2. The transient increase in internal energy just after the length transition is correlated with incomplete base stacking, particularly in the middle of the molecule. Increasing tension drives the conformation over this barrier during extension; however, if tension were then decreased, the process of contraction might pause before passing back over this energy barrier. Attempts to simulate relaxation, not reported in detail here, became trapped at this stage, although it is not possible to estimate the real clock time expected for this pause. In fact, such a hysteresis effect was observed during relaxation of λ DNA, although the degree varied among individual molecules and was at least partially ascribed to the presence of a variable number of single-strand breaks in the DNA.[7]

Since the covalent structure of the two strands is known and the extreme degree of extension of the strands puts strong limits on possible structures, we propose that the interstrand base stacking pattern is the defining feature of the S-ladder proposed here. The change in stacking between helix and ladder would be both simple and dramatic for homopolymers. Thus, stacking for poly(AT):poly(TA) would change from AT and TA to AA and TT. While this should produce changes in many optical properties, perhaps single molecules for which the tension can be precisely controlled can best be examined by changes in emission spectra, e.g., fluorescence. The effects of base stacking on fluorescence have been studied, and there are clear differences between purine−purine stacking and purine−pyrimidine stacking.[21]

Several ladder structures for non-helical (linear) DNA or DNA under tension have been proposed, either based on general geometric constraints or on energy minimization structural computations. However, the present paper is the first to use dynamical simulation, in which tension, inertia, atomic interactions, and geometric constraints are explicitly included and the evolution of structure with time is obtained. The S-ladder of Figure 2 is actually close to the formal linear structure in Figure 2.4 of Calladine and Drew.[22] However, their base pair repeat spacing of 0.60 nm gives a tilt of 57°, considerably less than the 68° seen here in Figure 5. The energy-minimized structures of Yagil and Sussman have even smaller tilt angles of about 40°, presumably because energy was defined only by atomic interaction energies and did not include tension. In the modeling described by Cluzel et al.,[8] DNA was constrained to stretch as a helix and thus could not progress to the structure seen in Figure 5. Their simulation did produce a plateau in the force versus length curve, but at 0.24 instead of the 0.07 nN seen in their experiments. Smith et al.[7] proposed that the stretched form of double-stranded DNA could not have appreciable helical content from the fact that the length asymptotically approached that seen in stretched single-stranded DNA. However, they did not propose a specific molecular structure for the stretched form.

The force necessary to separate the $(ACTG)_3$ helix has been measured by Lee et al.[12] using the atomic force microscope (these experiments motivated our work). They observed a separation force of 0.83 ± 0.11 nN, which is indistinguishable from that seen in the simulation. Only separation events, in

(19) Saenger, W. *Principles of nucleic acid structure*; Springer-Verlag: New York, 1984.
(20) Thundat, T.; Allison, D. P.; Warmack, R. J. *Nucleic Acids Res.* **1994**, *22*, 4224−4228.
(21) Browne, D. T.; Eisinger, J.; Leonard, N. J. *J. Am. Chem. Soc.* **1968**, *90*, 7302−7316.
(22) Calladine, C. R.; Drew, H. R. *Understanding DNA, the molecule and how it works*; Academic Press: London, 1992.

which the AFM probe jumped typically 20 nm away from the binding surface as the force fell to zero, were described in the report. Thus, there was no evidence presented for a helix to ladder transition, which would be seen as a displacement of only 2−4 nm at essentially constant force.

The most obvious example of a stretching force being directly applied to DNA in the cell occurs during cell division. The forces exerted on chromosomes during prometaphase have been measured in grasshopper spermatocytes[23] and vary from 0.3 to 0.8 nN. Since these forces are sufficient to cause a helix to ladder transition in free DNA, it is possible that such a structural transition occurs, particularly near the centromere region, and that the altered form of the DNA plays a role in control of cell division.

(23) Nicklas, R. B. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 431−449.

Acknowledgment. M.W.K. was supported in part by a Small Business Innovative Research grant (R43 DK49891-01) from the National Institutes of Health, U.S. Public Health Service. We also acknowledge the hospitality of The Cancer Research Fund of Contra Costa.

JA961751X