Mechanics of DNA Sticky End Joints

by

Ehsan Ban

A Thesis Submitted to the Graduate
Faculty of Rensselaer Polytechnic Institute
in Partial Fulfillment of the
Requirements for the degree of
MASTER OF SCIENCE
Major Subject: MECHANICAL ENGINEERING

Approved by:

_____
Professor Catalin R. Picu, Thesis Adviser

_____
Professor Assad A. Oberai, Member

_____
Professor Mark S. Shephard, Member

Rensselaer Polytechnic Institute
Troy, New York

July 2012
(For Graduation August 2012)

CONTENTS

LIST OF TABLES iii

LIST OF FIGURES iv

ACKNOWLEDGMENT vi

ABSTRACT vii

1. INTRODUCTION 1

    1.1 DNA structure 1

    1.2 Self-assembled DNA constructs 3

        1.2.1 3D DNA crystals 4

    1.3 DNA sticky end joints 5

2. MOLECULAR DYNAMICS SIMULATION OF DNA 7

3. THE STRETCH OF A 20 BASE PAIR DOUBLE STRANDED DNA 12

4. STRETCH OF NICKED DNA STRUCTURES 21

    4.1 Dissociation of 2 base pair-long sticky end joints subjected to axial loading . 23

    4.2 Dissociation of 4 base pair-long sticky end joints subjected to axial loading . 26

    4.3 Dissociation of 6 base pair-long sticky end joints subjected to axial loading . 29

5. CONCLUSIONS 33

REFERENCES 34

LIST OF TABLES

Table 4.1, List of the DNA sequences used in simulations 22

LIST OF FIGURES

Figure 1.1, Structure of DNA 2
Figure 1.2, Watson-Crick base pairing in DNA 3
Figure 1.3, Self-assembly of DNA strands into novel constructs 3
Figure 1.4, Applications of DNA self-assemblies 4
Figure 1.5, A 3D DNA crystal self-assembled from triangular units 5
Figure 1.6, Sticky end joints in the 3D DNA crystal 6
Figure 2.1, DNA model in a water box surrounded by ions 8
Figure 2.2, Internal coordinates for bonded interactions 10
Figure 3.1, Double stranded DNA model 12
Figure 3.2, Force-displacement curve and snapshots of a 20 base pair double stranded DNA in stretch 13
Figure 3.3, Validation of the force-displacement curve from simulations against the existing literature 14
Figure 3.4, Orientation of bases in stretch of the double stranded DNA 15
Figure 3.5, Dimensions of different regions of the double stranded DNA in stretch 16
Figure 3.6, Rise per nucleotide step at different stages of deformation of the double stranded DNA 17
Figure 3.7, Breakdown of potential energy of the double stranded DNA during stretch 18
Figure 3.8, Variation of the dihedral energy and applied force on the dsDNA during stretch demonstrating the correlation between the two quantities 19
Figure 3.9, Breakdown of dihedral energies in stretch of the double stranded DNA 20
Figure 3.10, Breakdown of the potential energy of the bases of the double stranded DNA during stretch 20
Figure 4.1, Model of a DNA sticky end (SE) 21
Figure 4.2, Force-displacement curve and snapshots of the 2 base pair-long AT sticky end <5> during stretch 23
Figure 4.3, Comparison of changes in the conformational (blue) and nonbonded (red) energies of the 2 base pair long sticky end <5> in stretch 24

Figure 4.4, Force-displacement diagrams of 2 base pair-long sticky ends of different sequences, <3> to <7> 25
Figure 4.5, Force-displacement diagram of the 4 base pair DNA sticky end <8> 26
Figure 4.6, Force-displacement curve of two trajectories of a 4 base pair long sticky end joint <8, 9> compared to the 2 base pair AT sticky end joint <5> and the 20 base pair double stranded DNA <1> 27
Figure 4.7, Configuration and structural evolution of the strong intermediate complex formed during the stretch of the 4bp sticky end joint <8> 28
Figure 4.8, Correlation of the dihedral energy of the 4 base pair long sticky end joint <8> with the force experienced in stretch 29
Figure 4.9, Comparison of force-displacement diagrams of different 6bp sticky ends <10-12> and the 4bp sticky end <8> 30
Figure 4.10, Correlation of the dihedral energy of the 6bpGCGCGC sticky end joint <12> with the external force experienced in stretch 31
Figure 4.11, Formation of multiple helical domains in sticky end molecules in stretch . 32

ACKNOWLEDGMENT

I’m sincerely grateful of my advisor Professor Catalin R. Picu for advising, inspiring and supporting me. I’m also grateful of my parents, brother and sister for their support and encouragement. Also I’m thankful of all the good teachers I’ve had throughout my education.

ABSTRACT

Self-assembled DNA structures are promising means of programmed assembly of nanoscale components. For example 3D DNA crystals with sizes of few hundred microns have been self-assembled from single stranded DNA. These crystals have applications in different areas including X-ray crystallography of biomolecules and directed assembly for nanoelectronics. These self-assemblies are held together through molecular configurations similar to DNA sticky ends. Understanding and enhancing the mechanics of DNA sticky ends helps engineer more stable self-assembled structures. In this work, the mechanics of DNA sticky ends is studied by molecular modeling and simulation and by performing uniaxial stretch tests of DNA molecules connected by sticky ends. Sticky ends of different lengths and base sequences are tested. They are divided into two classes: weak and strong sticky ends. The evolution of the macromolecules during stretch and up to failure is studied in detail and a cohesive intermediate complex inhibiting dissociation is identified in the case of the strong sticky ends.

1. INTRODUCTION

Deoxyribonucleic acids (DNA) are well-known for carrying genetic information through generations. Information is stored in DNA in form of a sequence of nucleobases (bases), Adenine (A), Thymine (T), Cytosine (C) and Guanine (G) [1]. The large variability of possible sequences of bases makes DNA a programmable molecule. The ability to be programmed along with the cheap synthesis of arbitrary sequences makes DNA a great candidate for self-assembly of versatile nanostructures and nanoscale machines. These structures have versatile applications. For example they can be used to form arrays of biomolecules for X-ray crystallography, or as templates for the self-assembly of nanoparticle arrays [2]. Structures similar to sticky end (SE) joints and Holliday junctions (HJ) appear in these self-assembled structures. Joints are the weak parts of these structures. So, understanding their mechanics can lead to making more stable self-assembled structures.

In this thesis the mechanics of DNA sticky ends is discussed. Also, results of stretching a double stranded DNA (dsDNA) molecule is presented and set as a benchmark for reference. Further, the effects of the variation of the sticky end length and base sequence on its strength are studied.

This chapter introduces the DNA molecule, DNA self-assemblies, and sticky end joints.

1.1 DNA structure

Watson and Crick suggested a double helical structure for DNA in 1953 [3]. The structure of DNA consists of two strands twisted about an axis. Each strand is made by sugar rings and phosphate groups connected by ester bonds. To each sugar ring, one of the four bases, A, T, C or G, is connected. A base, the sugar ring connected to it and the phosphate group attached to the sugar make up a nucleotide. Nucleotides are the basic units for making single strands of DNA. In double stranded DNA the two strands are connected by pairing of bases via hydrogen bonds. These are shown in Figure 1.1a. The distance between each two consecutive base pairs is called the rise per base step. Along the helix axis there are regions where bases are more exposed. Such regions are called

the major grooves. In contrast, in minor grooves the bases are covered by the backbone. Each strand makes turns about the helix axis as it travels along. The distance over which a strand makes a full turn is called the helix pitch. These are illustrated in Figure 1.1b.

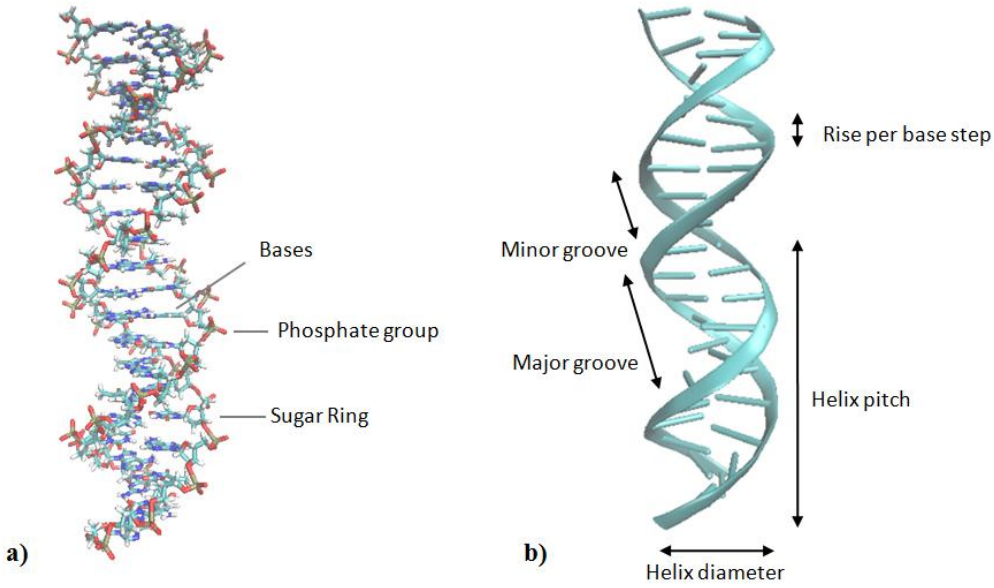

Figure 1.1, Structure of DNA.(a) Licorice view of double stranded DNA structure. Double stranded DNA consists of two strands having a phosphate-sugar backbone connected by paired bases in between. (b) Cartoon view of the DNA structure. Two turns of a double stranded DNA are shown and helix dimensions are marked.

Four types of bases are found in DNA structures, A, T, C and G. Among these A and G consist of two aromatic rings while T and C consist of one ring. A pairs with T and C pairs with G. In an AT pair the two bases are connected by 2 hydrogen bonds while in a CG pair they are connected by 3 hydrogen bonds. Each two consecutive bases interact through stacking forces (excluded volume interactions) [1]. For an illustration, see Figure 1.2.Also when two sequences have bases that pair one by one they are called complimentary sequences.

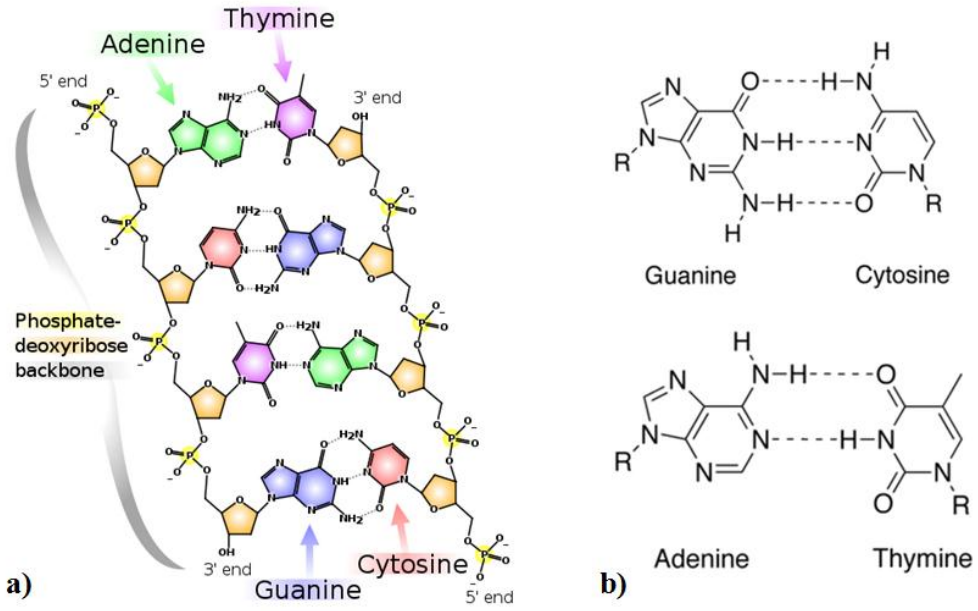

Figure 1.2, Watson-Crick base pairing in DNA. (a) Pairing of complimentary bases connects two DNA strands. Adenine (A) pairs with Thymine (T) and Guanine (G) pairs with Cytosine (C) (image from [4]). (b) The nitrogenous bases pair via hydrogen bonds (image from [5]).

1.2 Self-assembled DNA constructs

DNA of an arbitrary sequence can be artificially synthesized. This enables programming DNA strands to pair at certain locations and in certain ways. With designing a suitable set of complimentary DNA sequences one can expect to produce certain structures by mixing single strands of DNA. For examples see Figure 1.3.

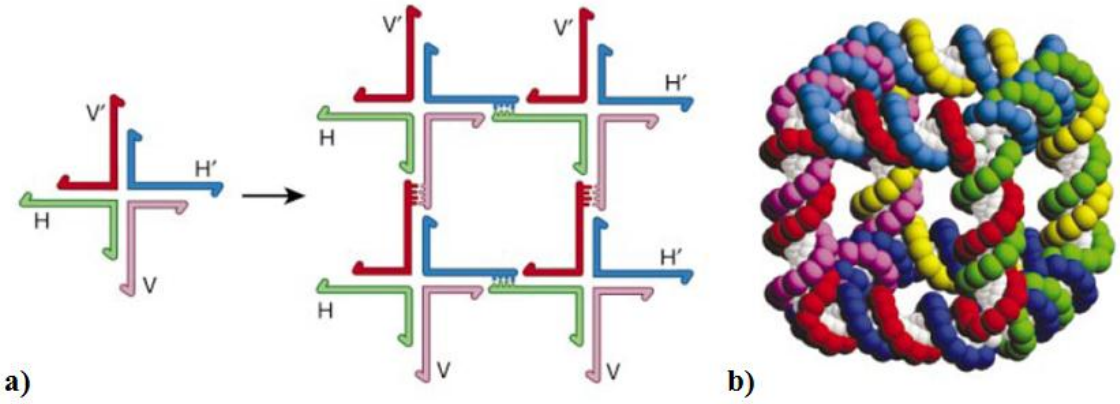

Figure 1.3, Self-assembly of DNA strands into novel constructs.(a) Self-assembly of a four way DNA junction into a two dimensional lattice. The sequence of the H overhang is complimentary to H' and

the sequence of V is complimentary to V'. The cohesive complimentary overhangs can connect to form a sticky end joint between two four armed junctions and produce a 2D lattice. (b) A self-assembled cube made up of six interlocked single strands. (Image from [2] with permission from Nature Publishing Group)

The self-assembled DNA structures have applications in crystallography and nanoelectronics. The big pores in these structures can be used as filters, to form arrays of aligned biomolecules which can then be observed through X-ray crystallography. Also DNA nanotechnologies can be used to construct scaffolds for directing the assembly of nano sized electrical circuits. These applications are depicted in Figure 1.4.

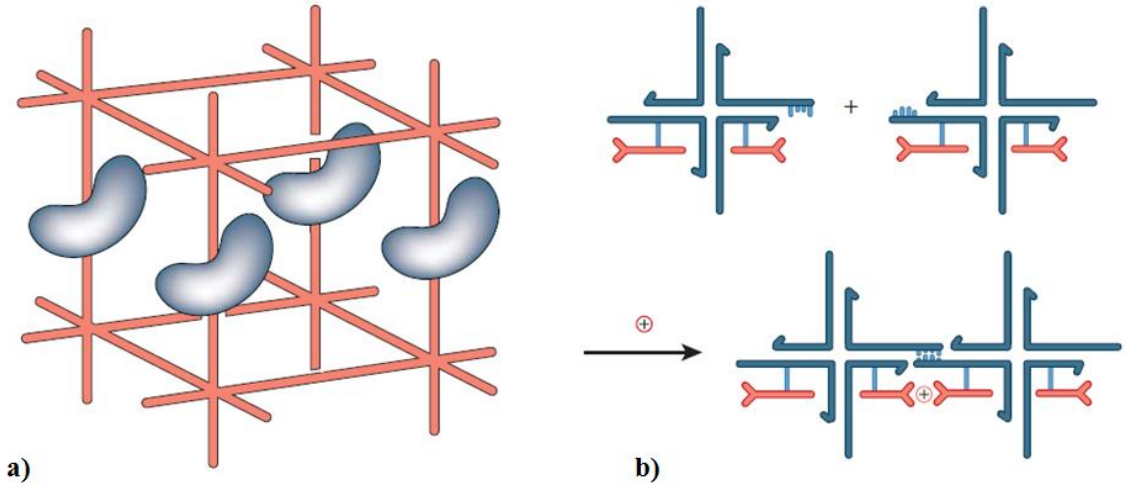

Figure 1.4, Applications of DNA self-assemblies.(a) DNA self-assemblies (red) can be engineered to have pores big enough to trap biological macromolecules (blue). This can be used to form arrays of parallel biomolecules. (b) DNA four way junctions connecting via sticky ends (blue) can be used as scaffolds to direct the assembly of components of nanoscale electrical circuits (red). These components are then stabilized by the addition of positively charged ions (red circles). (Image from [2] with permission from Nature Publishing Group).

1.2.1 3D DNA crystals

3D DNA crystals are self-assembled from single strands of DNA that form triangular units. These crystals have a rhombohedral shape and a size of up to few hundred microns. Three motifs are used to make triangles that self-assemble to produce the crystal structure. The triangles connect to each other through DNA sticky ends and

form structures similar to Holliday junctions at the corners. This is shown in Figure 1.5. These crystals have pores of up to about 1000 nm^3 in volume. Applications in nanoelectronics and crystallography are suggested for these structures [6].

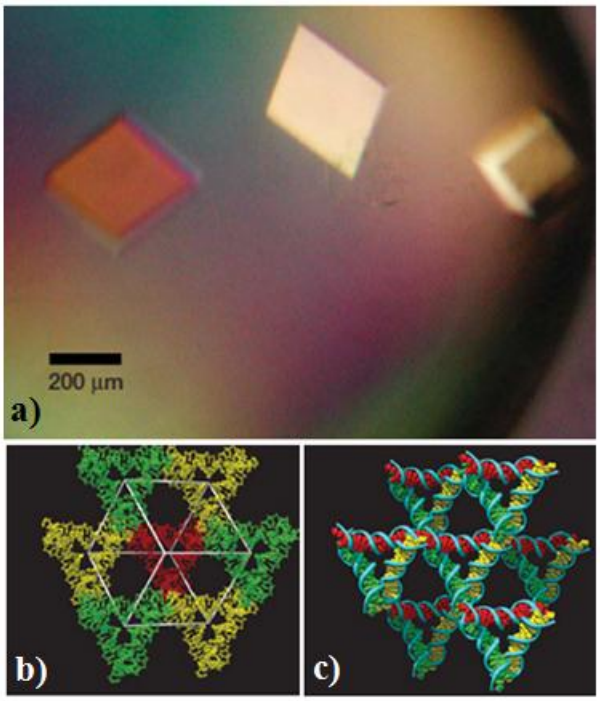

Figure 1.5, A 3D DNA crystal self-assembled from triangular units.(a) An optical image of the rhombohedral crystal structure. (b) Rhombohedral cavity formed by eight triangular units located at the corners of a rhombohedron. The red triangle is located at the far corner of the rhombohedron (farthest from the viewer). The yellow ones are at the corners in the middle and the green units are at the nearest corner to the page (closest to the viewer). (c) Surroundings of one triangle. The triangle in the middle is surrounded by six others. Each side of the triangle in the middle is along one of the three crystalline directions of the self-assembled crystal structure (Image from [6] with permission from Nature Publishing Group).

1.3 DNA sticky end joints

The connection between two triangular units of the 3D DNA crystal is shown in Figure 1.6. The joint between the two triangles is similar to a dsDNA with carries two nicks in opposite strands.The nicks are not at the same level in the two strands, since in this case the structure would lose stability. Rather, they are separated by a number of nucleotides, which defines the length of the sticky end. In Figure 1.6 two bases are

located in the region between the nicks. We call this a two base pair long sticky end. This is the weak spot of the crystal structure.

Structures containing DNA sticky ends have been characterized experimentally under different crystal confinement conditions [7]. In this work we test the mechanics of these structures by performing Molecular Dynamics simulations of stretch tests. The simulation method is described in the next chapter.

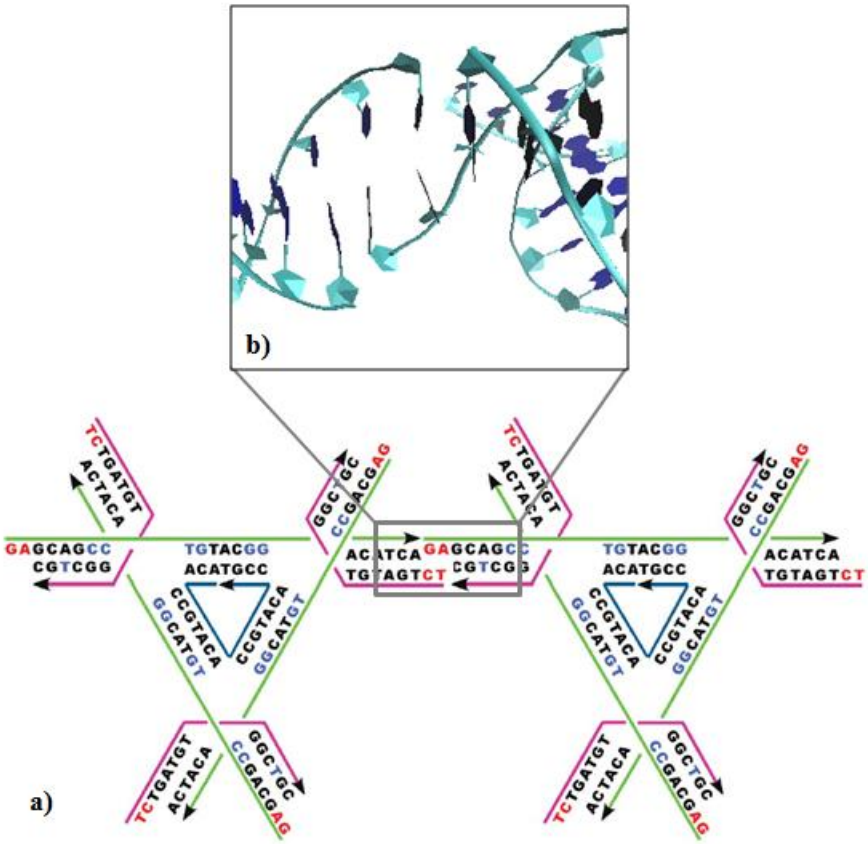

Figure 1.6, Sticky end joints in the 3D DNA crystal.(a) Two copies of the triangular unit connected via a cohesive sticky ends. The complementary GA and CT bases from the two triangles connect establishing a sticky end joints between the two units. (Image from [6] with permission from Nature Publishing Group) (b) Molecular view of the sticky end joint region. DNA backbone is shown as a cyan ribbon. Bases are shown in navy blue and sugars in cyan polygons.

2. MOLECULAR DYNAMICS SIMULATION OF DNA

Molecular Dynamics (MD) is a method used to advance in time (simulate the mechanics) discrete models of matter [8]. In this work an all atomistic classical MD model of DNA is used to study the mechanics of the DNA sticky end joints. This chapter explains how MD is performed and used to simulate DNA structures.

To start an all atomistic MD simulation one needs to know the coordinates of all the atoms making up the system to be studied. These coordinates define the initial structure of the molecule/simulation trajectory. A force field is needed to describe the bonding of the system. Then Newton's equations of motion are then integrated over time to find the new position of all the atoms at each time step.

We use the haddock internet server to produce initial structures of arbitrary sequences [9]. Also we use the CHARMM27 force field [10, 11] to describe the interaction between atoms. We prepare the simulations, carry them out and post-process the results using VMD [12] and NAMD [13]. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign.

After obtaining the desirable DNA structure from haddock we use psfgen in VMD to make protein structure files (psf). Then we solve the DNA structure in a box of water. The box dimensions are 6nm x 6nm x 30nm. The box also contains ions in order to neutralize the simulation cell; specifically, it contains 15mM of KCl. The simulation box is shown in Figure 2.1. TIP3 water model is used in NAMD and the simulations are conducted using periodic boundary conditions in all directions.

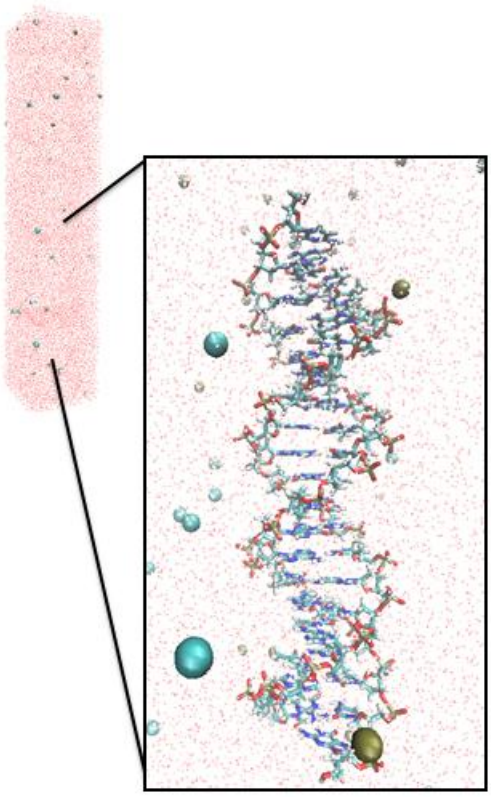

Figure 2.1, DNA model in a water box surrounded by ions. Water molecules are shown as small red and white dots. Potassium and Chloride ions are shown respectively by ochre and cyan spheres. A 20 base pair double stranded DNA is shown in the middle of the model.

The force acting on each atom is calculated using the CHARMM27 forcefield at each time step of the simulation. Forces are calculated as

$$\boldsymbol{F}_i = -\nabla_{\mathbf{r_i}} \mathrm{U} \tag{1}$$

where U is the sum of conformational (bonded), van der Waals, and Columb potentials. The simulations are not reactive meaning that the bonds are fixed throughout the simulation (no chemistry). Hydrogen bonds are an exception to this rule as may break and reform during the simulation. They are represented using special van der Waals and Columb components of the potential. The total potential entering eq. (1) is a function of the bond lengths, bond angles, dihedral angles. In other words

$$U = U_{bond} + U_{angle} + U_{dihedral} + U_{vdW} + U_{Columb} \tag{2}$$

where

$$U_{bond} = \sum_{bondsi} k_i^{bonds}(r_i - r_{0i})^2,$$

$$U_{angle} = \sum_{anglesi} k_i^{angle}(\theta_i - \theta_{0i})^2,$$

$$U_{dihedral} = \sum_{dihedralsi} k_i^{dihedral}[1 + \cos(n_i\phi_i - \gamma_i)], \quad (3)$$

$$U_{vdW} = \sum_i \sum_{j>i} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right],$$

$$U_{Columb} = \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}.$$

Here r_i, θ_i, and $\phi_i$are the bond lengths, bond angles and the dihedral angles in the DNA structure. These coordinates are depicted in Figure 2.2.

Values of k_i^{bonds}, k_i^{angle}, $k_i^{dihedral}$, r_i, θ_{0i}, n_i, γ_i,σ_{ij}, and ε_{ij} are potential constants specific to each type of atom. These constants are set in a way that the simulations reproduce structural and vibration (spectral) parameters close to those measured experimentally. Also, $r_{ij}$is the distance between two atoms i and j. q_i is the partial charge of atom i and ε_0 is the dielectric constant. Like bonds, partial charges are also defined when setting-up the simulation and remain constant during it. It should be noted that the columbic force is a long range force and is costly to calculate for all pairs of atoms. So, NAMD uses the PME method to calculate it.

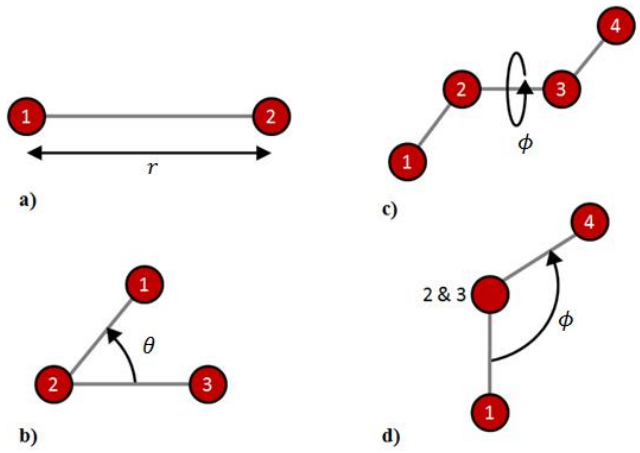

Figure 2.2, Internal coordinates for bonded interactions.(a) r is the distance between two atoms and governs bond stretch. (b) θ is the angle formed by two bonds connected to one atom. (c) and (d) ϕ is the dihedral angle formed by one plane containing atoms 1, 2 and 3 and another plane containing atoms 2, 3 and 4 where consecutive atoms are bonded.

Unlike other bonds in the structure, hydrogen bonds are not fixed during the simulation. Also they don't have to be defined before the simulation starts. They can be formed or broken as the simulation proceeds. These bonds are accounted for by the van der Waals and Columb parts of the potential shown in eq (3). A 12 Å cut-off is used for calculating the van der Waals forces.

MD simulations in the NVE ensemble (constant number of atoms, volume and energy) are performed by stepwise time integration of eq (1), i.e.

$$\boldsymbol{F}_i = -\nabla_{\mathrm{r_i}}\mathrm{U} = \mathrm{m_i}\ddot{\boldsymbol{r}}_i = \dot{\boldsymbol{p}}_i. \qquad (4)$$

Here $\boldsymbol{F}_i$, $\boldsymbol{r}_i$, $\mathrm{m_i}$, and $\boldsymbol{p}_i$are respectively the force acting on, the coordinate of, the mass of and the momentum of particle i. The time integration leads to finding theparticle coordinates in the next time step using the coordinates at the current step.

In this work simulations are also carried out in the NPT ensemble, i.e. keeping constant the number of atoms, the pressure and the temperature. The temperature is 300K and the pressure is 1 atm. To simulate a system at constant temperature and pressure NAMD uses a time integration method inspired by the Langevin dynamics:

$$\dot{\boldsymbol{r}}_\iota = \frac{\boldsymbol{p}_i}{m_i} + \frac{p_\varepsilon}{W}\boldsymbol{r}_i,$$

$$\dot{\boldsymbol{p}}_{\iota} = -\nabla_{\mathrm{r_i}}\mathrm{U} - \frac{p_{\varepsilon}}{W}\boldsymbol{p}_i - \gamma \boldsymbol{p}_i + \boldsymbol{R}_i,$$

$$\dot{\mathcal{V}} = \frac{d\mathcal{V}p_{\varepsilon}}{W}, \tag{5}$$

$$\dot{p_{\varepsilon}} = d\mathcal{V}(X - P_{ext}) - \gamma_p p_{\varepsilon} + R_p,$$

$$X = \frac{1}{d\mathcal{V}}\left[\sum_{i=1}^{N}\frac{\boldsymbol{p_i^2}}{m_i} + \sum_{i=1}^{N}\boldsymbol{r}_i.\boldsymbol{F}_i\right] - \frac{\partial}{\partial\mathcal{V}}U(\boldsymbol{r}^N, \mathcal{V}).$$

Here p_{ε}, W, $\mathcal{V}$, γ_p, and R_p are respectively momentum, mass, volume, friction coefficient and the random force associated with the barostat. Also, $\boldsymbol{R}_i$, and γ_p are the random force and friction coefficient associated with the thermostat. Further $P_{ext}$is the external pressure. Further details of the method are provided in [14].

After the initial is structure is defined and before it can be stretched, the system needs to be equilibrated. The system energy is first minimized to insure the stability of the initial configuration. This is done while all atoms of DNA are fixed in place using a hard spring connecting the respective atom to the background. Then the temperature is gradually increased from 0K to 300K in 150ps, gradually. At the same time, the stiffness of the springs holding the atoms in place is gradually reduced to zero. Finally the system is relaxed at 300K for 2ns.

Now the equilibrated system is ready for mechanical testing. Stretching the DNA molecule is performed by moving the C5' atom at one end while holding the C5' atom at the other end fixed in place. This is done by use of two dummy atoms (see also Fig. 3.1). Each of the dummy atoms are connected to one of the C5' atoms through a spring. The spring constant is 100 (kcal/mol/Å^2). At each step of deformation the moving dummy atom is displaced 1 Å and the system is relaxed for 100ps. The purpose of this relaxation is to allow the deformation imposed at the ends to "propagate" inside the molecule. The relaxation time is selected such to allow a wave to propagate several times across the structure. To evaluate the force at the imposed deformation, we calculate the net force acting on the dummy atoms connected with the C5' ends (the atoms defining the applied deformation) and take an average of this (highly fluctuating) force over the last 20 ps of each 100 ps relaxation step of the simulation. During each relaxation step the two C5' ends are fixed in all directions.

3. THE STRETCH OF A 20 BASE PAIR DOUBLE STRANDED DNA

Double stranded DNA (dsDNA) is stretched to verify the simulation results against the existing literature and to establish a benchmark for the stretch of similar structures containing sticky ends. This chapter presents the behavior of a short dsDNA subjected to stretch and the microstructural changes associated with this deformation.

The molecule shown in Figure 3.1 is used for testing. This sequence is chosen identical to that used in [15].

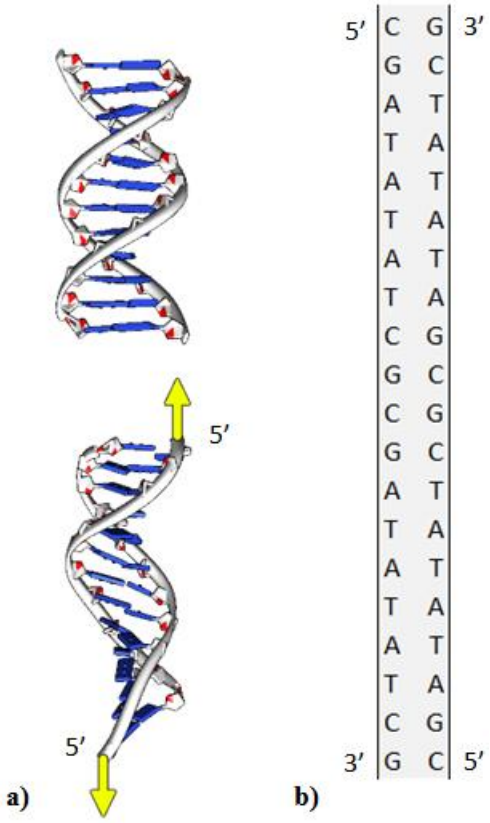

Figure 3.1, Double stranded DNA model. (a) The double stranded DNA is stretched axially by imposing displacements to dummy atoms connected to C5' atoms at opposite ends of the molecule (image created using Chimera [16]). (b) Base sequence of the 20mer double stranded DNA used in the stretch test.

When the dsDNA is stretched through this procedure, "un-peeling" takes place at the ends. At later stages, untwisting is observed. Further, the dsDNA forms a ladder like structure which then transforms to a zipper-like structure [17]. This corresponds to the peak in the force-displacement curve. Beyond this point the bases loose engagement

with each other and turn outwards; the force starts to decrease. The two strands become unloaded and pass each other. This evolution is shown in Figure 3.2.

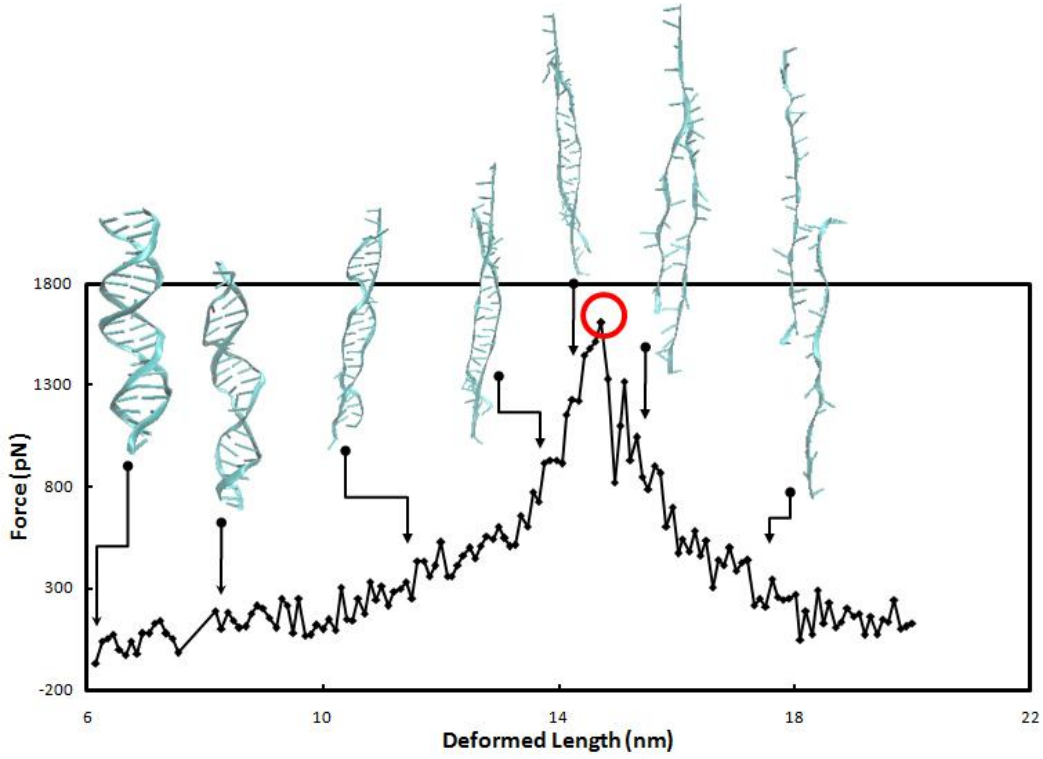

Figure 3.2, Force-displacement curve and snapshots of a 20 base pair double stranded DNA in stretch. The snapshots show the changes in the microstructure of the molecule as it deforms. Also the red circle marks the maximum force. At the early stages of deformation there is unpeeling from the ends. As deformation progresses the structures adopts a ladder-like state and then to a zipper-like structure. Then the force reaches the peak and under the maximum force the stacked bases are forced to face outward. Finally the strands pass each other and the force decreases gradually as the strands are unloaded.

We have compared the force displacement curve from our simulation to that obtained by similar methods in [15], Fig. 3.3. The curves have similar shapes and nearly reach the same maximum force at the same stage. The major difference between our simulations and those reported in [15] is that instead of performing the sequence of dummy atom displacement followed by 100 ps relaxation described in Chapter 2, the authors of Ref. [15] move the dummy atoms continuously, as they perform MD. Consequently, our simulations are significantly longer and more relaxed. This does not have an effect on the loading part of the curve, when the structure is essentially intact.

However, it affects the unloading part, as in our simulations unloading is more gradual and this allows us to observe the structural changes described in the previous paragraph.

It is useful to compare these results with stretch results from experiments. The forces reported here and in [15] are significantly larger than the experimental forces. This is primarily due to the fact that in experiments one stretches a very long DNA double strand which is initially coiled. Small forces are need to uncoiling, while much larger forces are needed if one would stretch a straight molecule. Due to compute time limitations, it is not possible to simulate large, coiled molecules and hence we only capture the last part of the experimental curves [18]. An additional factor which leads to discrepancies between models and experiment is the loading rate, which, as usual, is much higher in the simulation. In addition to this salt conditions have tremendous effects on how DNA behaves in stretch. So the results vary with use of different salt concentration and types [19].

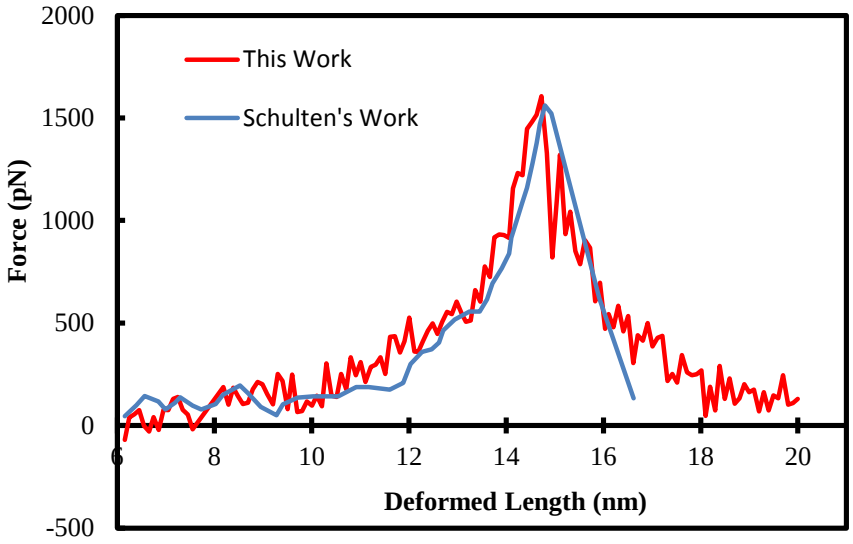

Figure 3.3, Validation of the force-displacement curve from simulations against the existing literature. Force against deformed length from this work (red curve). Force against deformed length for a 20 base pair double stranded DNA from [15] (blue curve). The two curves have similar shape and close values for peak force and peak deformation. However the curve from this work decreases slowly after the peak force while in the other curve the force drops abruptly after the peak. This is because in this work the structure is equilibrated after imposing the deformation in each time step. This gives the structure more time find a relaxed configuration compatible with the imposed end

displacement. From a comparison of the trajectories it's evident that in [15] the two strands do not interact as much as they do in this work.

It is also observed that beyond the maximum force point the structure is significantly distorted and the bases start to face outward. Snapshots before and after the peak force and the number of unpaired bases facing outward are shown in Figure 3.4.

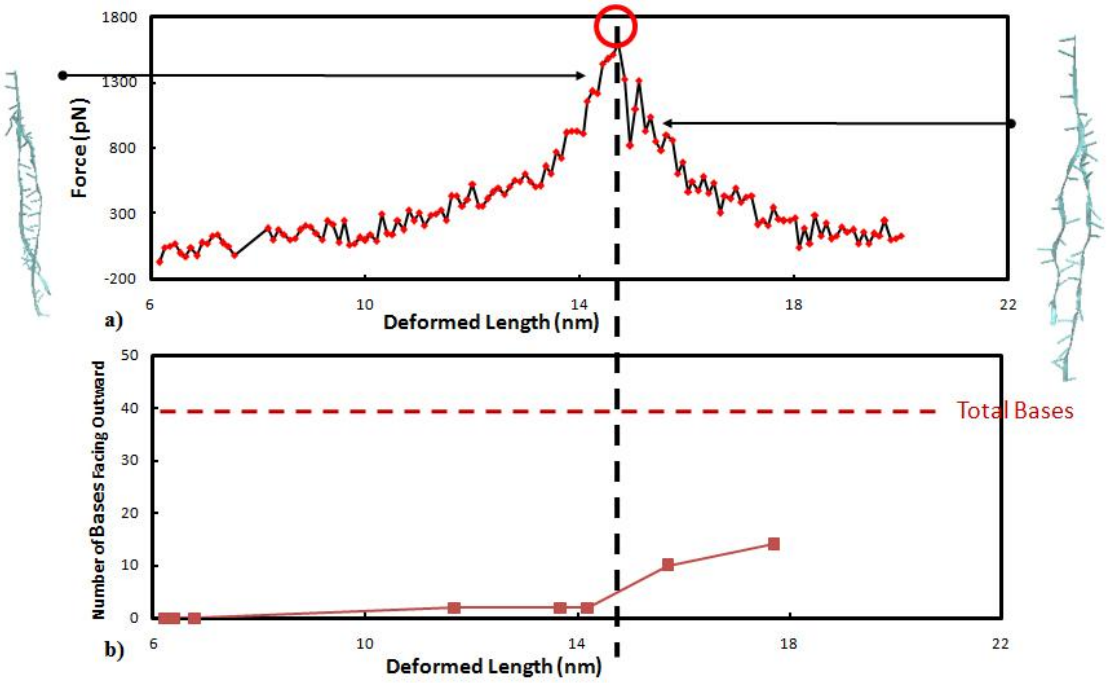

Figure 3.4, Orientation of bases in stretch of the double stranded DNA. After the peak force bases rotate and face outward. (a) Snapshots of the DNA structure before and after the peak force. The red circle and the black dashed line indicate the maximum force. (b) Number of bases facing outward as a function of the deformed length.

It is interesting to characterize the deformation by monitoring the length of the two strands during loading and the length of the region where they overlap (i.e. bases of facing strands are still connected). These measures are shown in Fig. 3.5. Note that up to peak force the length of the strands increases, primarily due to elastic untwisting. Beyond the peak force, the strands start to slide relative to each other. This reflects in a sharp decrease of the overlap length and associated unloading of each strand.

Interestingly, the process of relative slide of the two strands does not begin at peak load, rather at a slightly larger stretch.

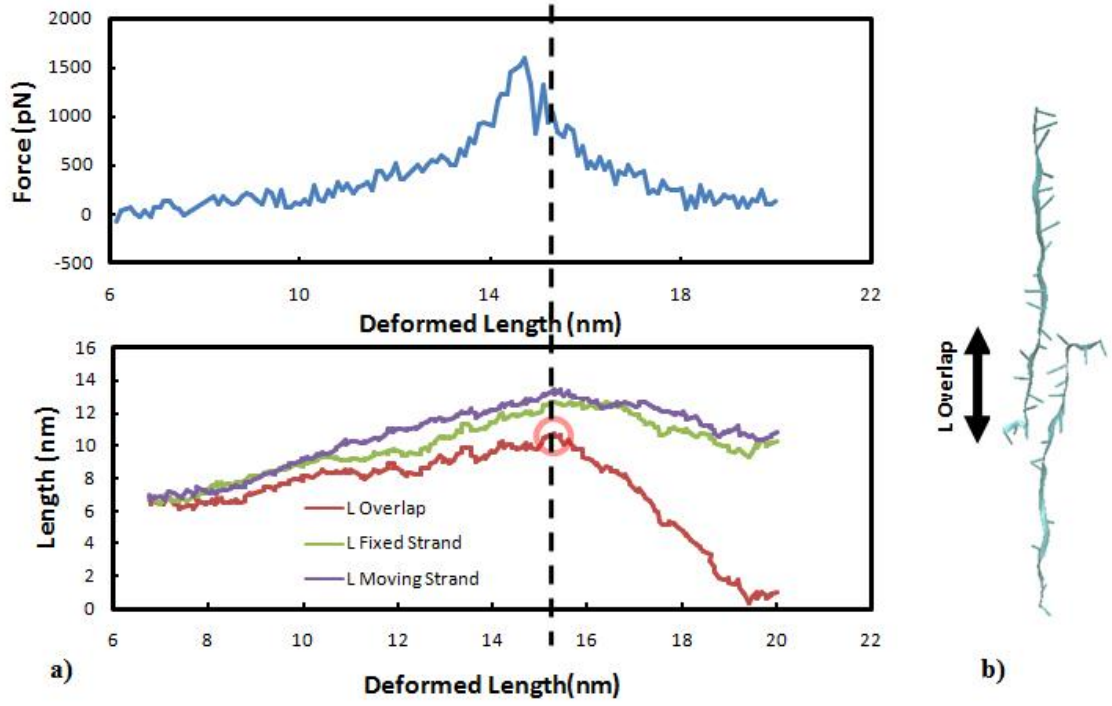

Figure 3.5, Dimensions of different regions of the double stranded DNA in stretch. (a) Length of fixed strand (green), moving strand (violet), and the overlapping region (red) against the deformed length. (b) The overlap region is the length of one strand that's still interacting with the opposite strand. This snapshot is taken at the deformed length of 18.4 nm.

At different stages of dsDNA stretch the parts of DNA near the controlled ends are more deformed. This can be the cause for unpeeling of the structure at the ends. This effect was shown in experiments by fluorescence imaging [20]. To elucidate this from the simulation results the distance between each adjacent C3' atoms are measured at different stages of stretch. Results are drawn in Figure 3.6. As can be seen the nucleotide step is greater at the controlled ends. This is more evident at larger deformations.

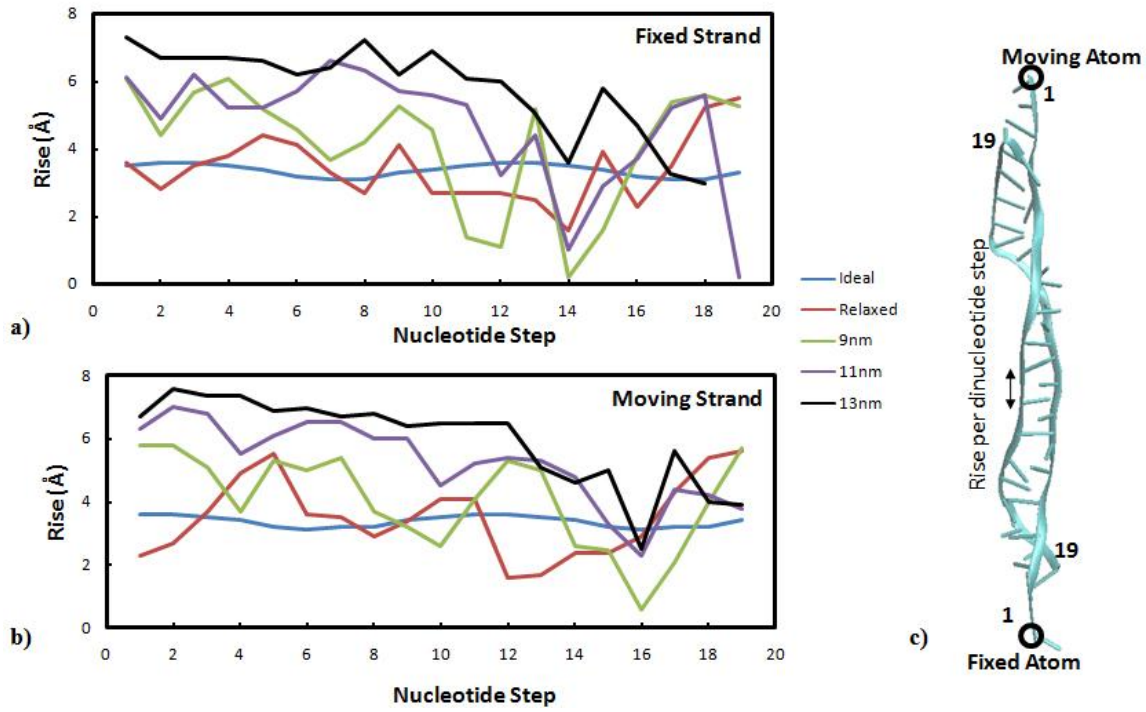

Figure 3.6, Rise per nucleotide step at different stages of deformation of the double stranded DNA. (a) Rise per nucleotide step in the fixed strand. (b) Rise per nucleotide step in the moving strand. (c) The rise per nucleotide is shown along the strand at various steps during stretch (corresponding to different colors) when the imposed displacement is 9, 11 and 13 nm. The rises of the initial relaxed structure ("Relaxed") and of the ideal structure before relaxation are shown. The fixed atom is at the beginning of the fixed strand and the moving atom is at the beginning of the moving strand. Note how the rise increases during deformation, but also varies along the strand. The rise is lower at the free ends of the two strands.

Deformation of the dsDNA is accompanied by changes in potential energy and entropy of the dsDNA molecules and the surrounding solvent. Potential energy is the energy of the system described by the force field. Breakdown of the potential energy shows that deformation of dsDNA in stretch is accompanied by significant changes in both bonded (conformational) and nonbonded parts of the potential energy. Further breakdown of the conformational part shows that the change in this energy is mostly due to change in the dihedral energy. These energies are shown in Figure 3.7. Also it's observed that the change in the dihedral energy is correlated to the observed force, as expected. This is shown in Figure 3.8. In comparison to the other internal coordinates dihedral angles change with less increase in the potential energy.

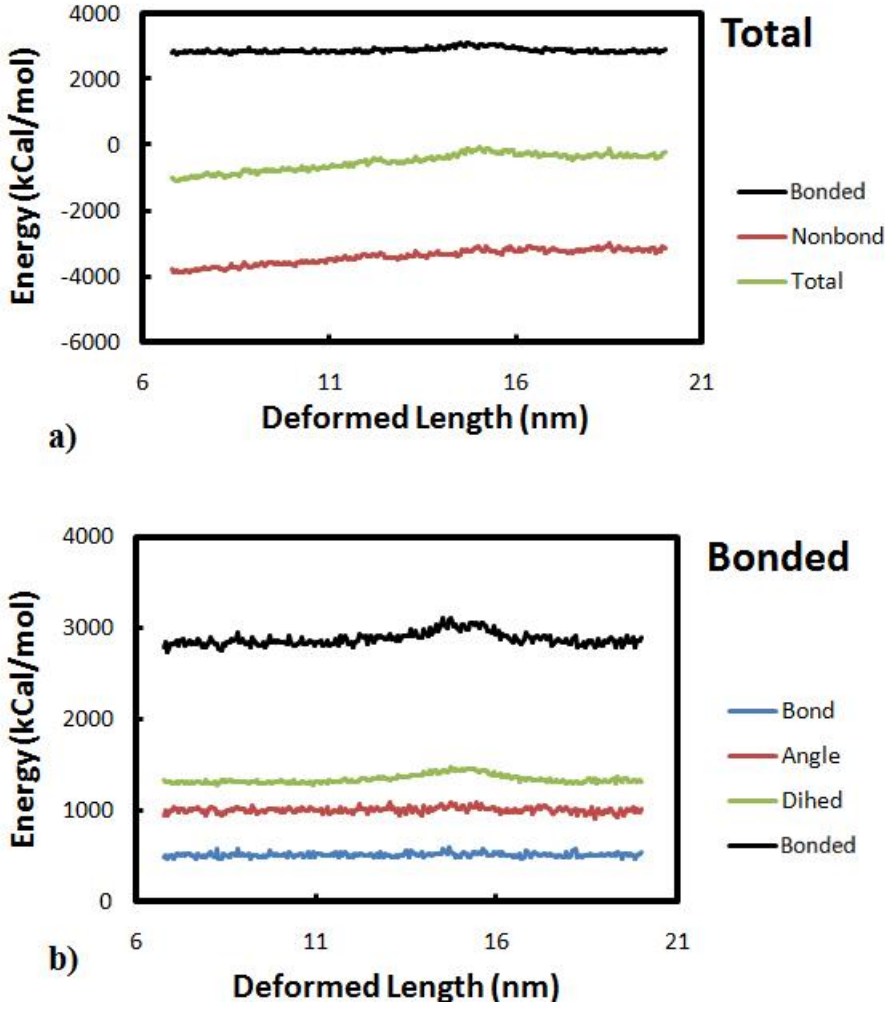

Figure 3.7, Breakdown of potential energy of the double stranded DNA during stretch. (a) Breakdown of the total energy (green) of deformation into bonded (black) and nonbonded (red) contributions. As shown, both bonded and nonbonded energies contribute to the change of the potential energy during deformation. (b) Breakdown of the bonded energy (black) into dihedral energy (green), angle energy (red), and bond energy (blue). The contribution of the dihedral energy is more significant than that of the other components.

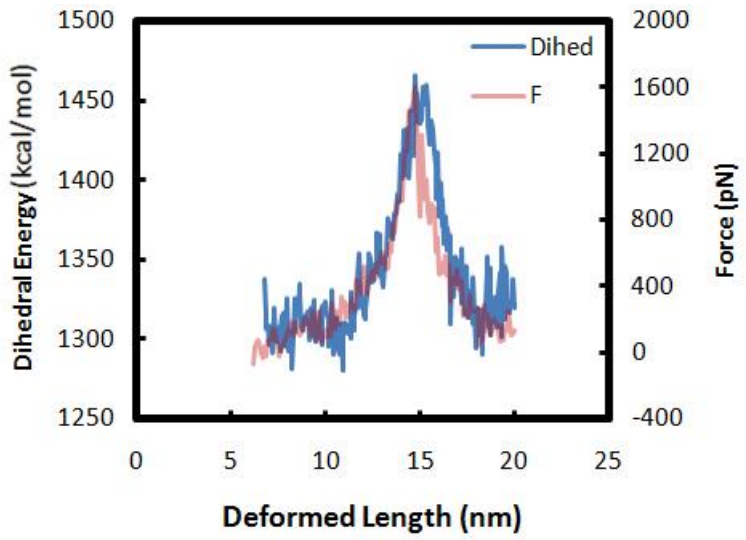

Figure 3.8, Variation of the dihedral energy and applied force on the dsDNA during stretch demonstrating the correlation between the two quantities.

The six dihedral angles along the backbone are denoted by α to ζ. Also the dihedral angle between the base and the sugar is denoted by χ. These angles are shown in Figure 3.9a. If a dihedral angle is defined by the four atoms 1-4 connected by three bonds from 1 to 2, from 2 to 3 and from 3 to 4 the dihedral angle is thought of as the twist about the bond 2-3. This is shown in Figure 3.9b.

The breakdown of dihedral energies for the dihedral angles α to χ is shown in Figure 3.9c. The change in energy due to the change in χ is further divided into change in χ from purines connected to sugars and pyramidines connected to sugars. The breakdown shows that the energy change due to change in a dihedral angle decreases as the angle is closer to the sugar and farther from the phosphate group. This is accompanied by greater changes in the dihedrals closer to the phosphorous atom. Phosphate group is softer than the sugar and the bases. This suggests extension of the backbone in a manner similar to extension of a telephone cord. The extension happens in a way that has the least contribution from change in bond lengths or angles. Further the breakdown of potential energy for the bases is drawn in Figure 3.10. The diagram indicates that the bonded potential energy for the bases does not change significantly. So they can be considered as rigid parts of the structure.

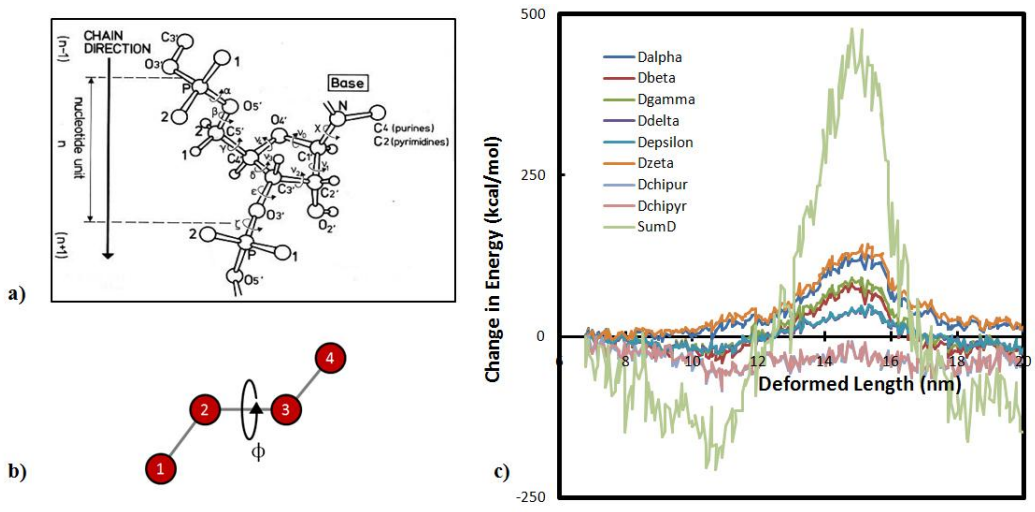

Figure 3.9, Breakdown of dihedral energies in stretch of the double stranded DNA. (a) The dihedral angles α to ζ along the backbone and the dihedral angle χ connecting the sugar ring to the bases are shown (image by Christoph Schneider from [21]). (b) Dihedral angle is the angle between the plane on which atoms 1, 2 and 3 reside and another plane on which the atoms 2, 3 and 4 reside when consecutive atoms are bonded. (c) Breakdown of dihedral energies to each of the ten depicted dihedral angles in (a) against the deformed length. Also the sum of all these energies is plotted (green). As can be seen these dihedral angles contribute to the deformation potential energy at different extents. Contribution of the dihedral angles to the total dihedral energy decreases as we move from the phosphate groups towards the sugar and bases. This suggests an unwinding mechanism mediated by rotation about bonds in the phosphate group.

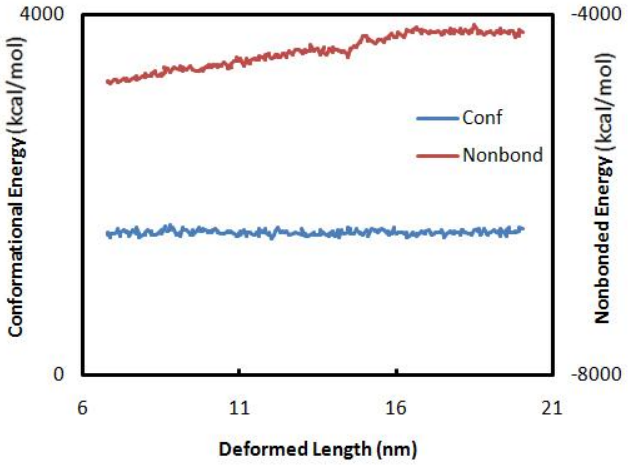

Figure 3.10, Breakdown of the potential energy of the bases of the double stranded DNA during stretch. The nonbonded energy (red) changes more significantly than the conformational (bonded) energy (blue). This indicates that bases remain relatively rigid during stretch of the double stranded DNA.

4. STRETCH OF NICKED DNA STRUCTURES

This chapter presents the stretch of DNA structures containing sticky end joints (SE).SEs are made by introduction of two opposite nicks in the middle of a double stranded DNA (dsDNA) structure. This is indicated in Figure 4.1 using red lines. In this work stretch of 9 different sequences of DNA was undertaken. The sequence introduced in the previous chapter (Fig. 3.1) is taken as reference and other structures are obtained starting from it. Even with use of massively parallel machines each of these tests take days and it's not feasible to simulate all possible combination of bases.

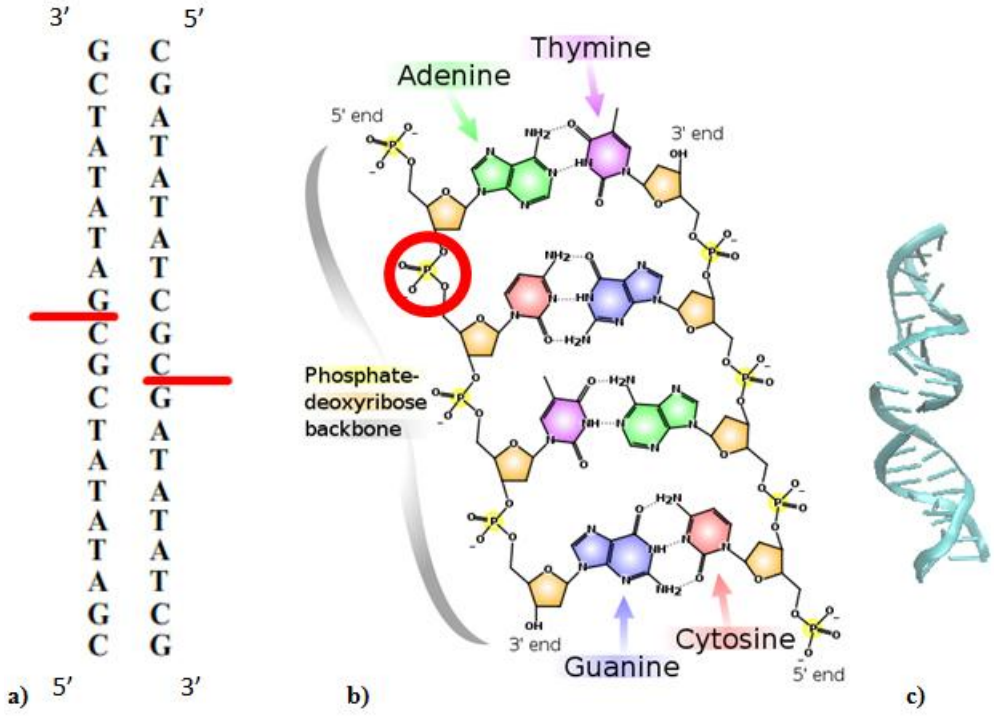

Figure 4.1, Model of a DNA sticky end (SE). (a) One of the sequences used for stretching DNA SEs. Red bars indicate position of the nicks in the backbone. (b) Nicks are introduced by removing one phosphate groups from the DNA backbone. One of the phosphate groups is marked by a red circle (image from [4]). (c) the resulting SE contains four chains and two nicks. This SE has a length of two as there are two base pairs located between the two nicks (both are of CG type).

The SE sequences are distinguished by their length and sequence. Here the part of the sequence between the two nicks is called the SE core and the number of base pairs (bp) located inside the core is considered as the length of the sticky end. For example the SE shown in Figure 4.1c is a 2 base pair sticky end. Further the part of sequence adjacent

to the nicks is called the “SE edge”. SEs differ by the sequence of bases inside the core and at edges. The notation used here is as follows: the SE shown in Figure 4.1a is called 2 base pair GC sticky end or 2bpGC SE. We have performed to replicas of the stretch simulation for this SE. We denote them by 2bpGC and 2bpGC_2.

Table 4.1 shows a list of all sequences tested in this work. Each sequence is given an abbreviated name similar to that indicated above. Also each test is given a unique ID. In the rest of this chapter the different nicked DNA structures are referred to using the notation shown in this table.

Table 4.1, List of the DNA sequences used in simulations. Some of the sequences have two names and IDs. This indicates that two replicas of the same structure subjected to the same loading conditions have been run. The red vertical bars indicate the positions of the nicks in the backbone. Also the highlighted sequence is the part of the sequence that forms the cohesive complex during the dissociation of the stronger sticky ends.

ID	Name	Sequence
<1>	dsDNA (20mer)	5' C G A T A T A T C G C G A T A T A T C G 3' 3' G C T A T A T A G C G C T A T A T A G C 5'
<2>	nicked	5' C G A T A T A T C G C G A T A T A T C G 3' 3' G C T A T A T A G C | G C T A T A T A G C 5'
<3,4>	2bpGC & 2bpGC_2	5' C G A T A T A T C G C | G A T A T A T C G 3' 3' G C T A T A T A G | C G C T A T A T A G C 5'
<5,6>	2bpAT &2bpAT_2	5' C G A T A T A T C A T | G A T A T A T C G 3' 3' G C T A T A T A G | T A C T A T A T A G C 5'
<7>	2bpATAT	5' C G A T A T A T A T A | T A T A T A T C G 3' 3' G C T A T A T A T | A T A T A T A T A G C 5'
<8,9>	4bp & 4bp_2	5' C G A T A T A T C G C G | A T A T A T C G 3' 3' G C T A T A T A | G C G C T A T A T A G C 5'
<10>	6bp	5' C G A T A T A T C G C G A | T A T A T C G 3' 3' G C T A T A T | A G C G C T A T A T A G C 5'
<11>	6bpCGEdgeATCore	5' C G A T A T C G A T A T C | G A T A T C G 3' 3' G C T A T A G | C T A T A G C T A T A G C 5'
<12>	6bpGCGCGC	5' C G A T A T A G C G C G C | T A T A T C G 3' 3' G C T A T A T | C G C G C G A T A T A G C 5'

4.1 Dissociation of 2 base pair-long sticky end joints subjected to axial loading

The first SE stretched is the 2bpAT, ID <5> in Table 4.1.It is observed that the SE is much weaker than the corresponding dsDNA structure. The peak force observed is about 200 pN which is much less than the force observed for the dsDNA subjected to similar conditions, i.e. about 1600 pN. Consider that the SEs dissociate at the peak force while for dsDNA the peak force occurs when sliding of the strands relative to each other starts, i.e. before final failure. The force displacement diagram of the 2bpAT, <5> SE is shown in Figure 4.2. Because the forces observed in stretch of SEs are low they are closer to the amplitude of the noise in these simulations. To evaluate it, one needs to perform time averaging as indicated in Fig. 4.2.

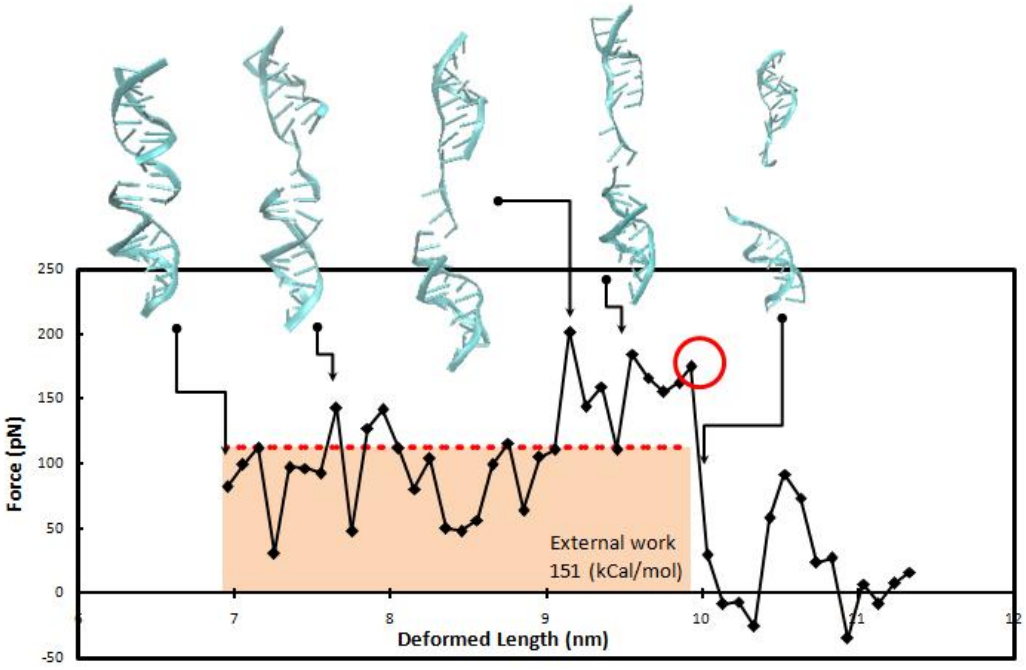

Figure 4.2, Force-displacement curve and snapshots of the 2 base pair-long AT sticky end <5> during stretch. Snapshots show the evolution of a sticky ends microstructure during axial loading. The red circle indicates the point at which the molecule dissociates. The red dashed line indicates the mean force up to dissociation. The area under this line is the external work required to dissociate the sticky end.

The 2bp SE splits into two helical domains at the two sides of the SE. This is shown in snapshots in Figure 4.2. Unlike the dsDNA there is not much deformation to each part of the SE. Calculating the potential energy of the SE in stretch shows that the change in conformational energies (bonded energies) is negligible in comparison to the change in the nonbonded energies. The energy-displacement plot in Figure 4.3 shows this effect. This indicates that the internal coordinates in the SE structure do not change significantly and the two helical domains are not distorted very much. So dissociation happens after the hydrogen bonding of paired bases and the base stacking forces are overcome.

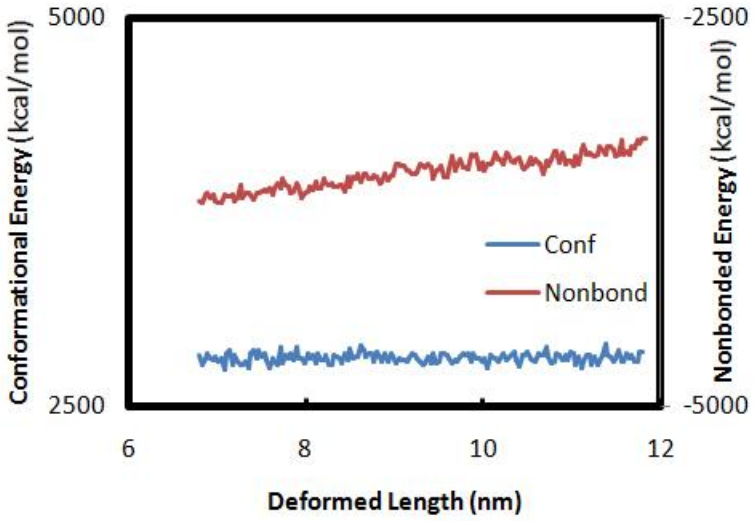

Figure 4.3, Comparison of changes in the conformational (blue) and nonbonded (red) energies of the 2 base pair long sticky end <5> in stretch. Only the nonbonded energy has a significant contribution to the change of the total potential energy of DNA. This quantitatively shows that distortions of the structure and untwisting do not play a role in the dissociation of this sticky end joint.

Further stretch of 4 more SE specimens is performed. They differ only with respect to the sequence of bases in the SE region, both in the core and at the edges. The 2bpGC <3, 4> sequence results by modifying the 2bpAT <6, 7> sequence by substituting A and T bases in the core with G and C bases. Also the 2bpATAT <7> sequence results by changing the G and C bases adjacent to the nicks into A and T.

The stretch of the second replica of the 2bpAT produces the same results as 2bpAT. Also stretching of 2bpGC and 2bpGC_2 indicate that the change of the AT base pairs inside the core to GC base pairs does not make the SE stronger. The 2bpGC and 2bpGC_2 are almost as weak as the 2bpAT SE. Further the 2bpATAT SE is weaker than the 2bpAT case and dissociates at a much lower force. The force displacement diagrams for these SEs are shown in Figure 4.4.

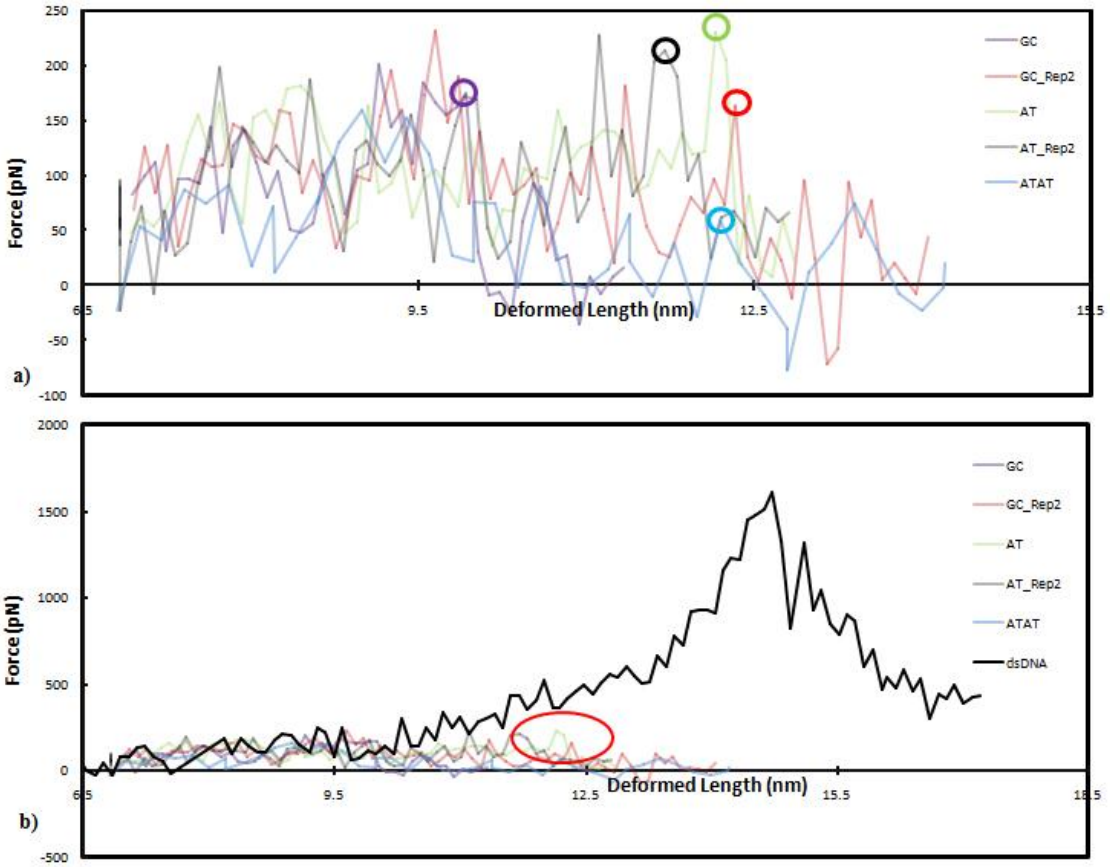

Figure 4.4, Force-displacement diagrams of 2 base pair-long sticky ends of different sequences, <3> to <7>. (a) Comparison of force displacement curves of sticky end joints with different base sequences. Note that the 2bpATAT sequence shows almost no strength. This sequence does not include any G or C bases at or near the sticky end joint (b) All 2 base pair-long sticky ends are two weak in comparison to a perfect double stranded DNA (thick black line in b). They all dissociate either in the region indicated with the red ellipse or before reaching it.

4.2 Dissociation of 4 base pair-long sticky end joints subjected to axial loading

Two replicas of the same sequence are tested: 4bp and 4bp_2 <8, 9> (Table 4.1). This sequence is produced by inserting nicks at a distance of 4 base pairs in the middle of the original dsDNA sequence tested.

The 4 base pair-long SEs appear much stronger than the 2 base pair ones. They dissociate at a peak force of about 500 pN. This is almost one third of the peak force for dsDNA and 2.4 times the force at dissociation of 2 base pair SEs. The force-displacement diagram and snapshots of the 4bp <8> are shown in Figure 4.5. Also the force-displacement of the 4bp SE, the 2bpAT SE and the dsDNA are shown together in Figure 4.6.

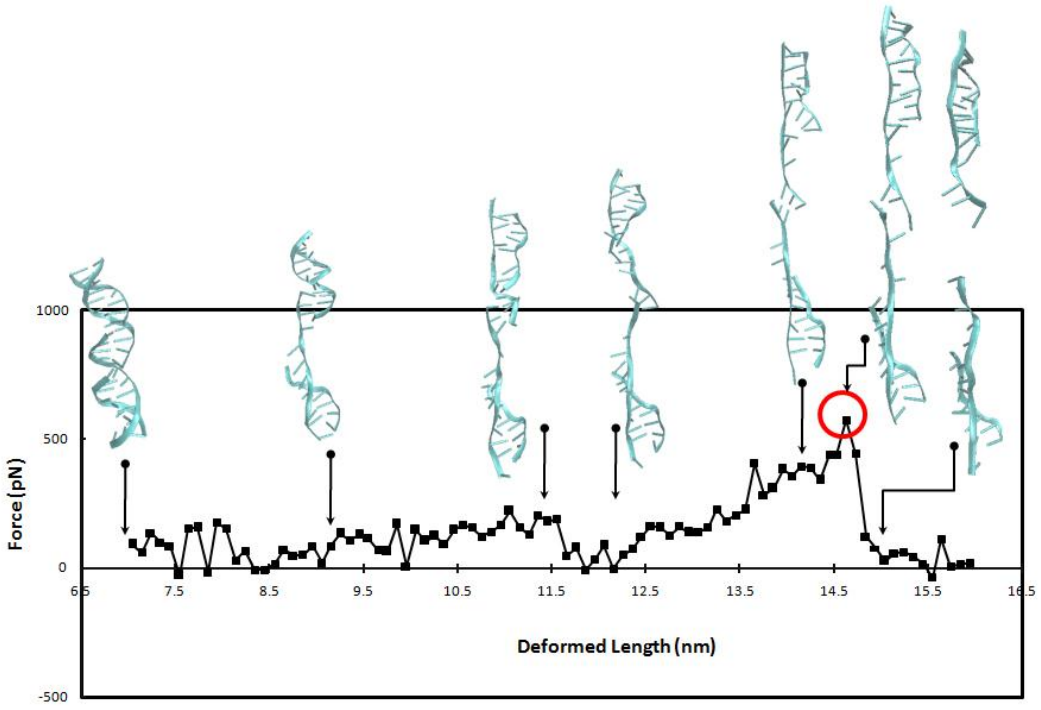

Figure 4.5, Force-displacement diagram of the 4 base pair DNA sticky end <8>.The red circle indicates the dissociation force.

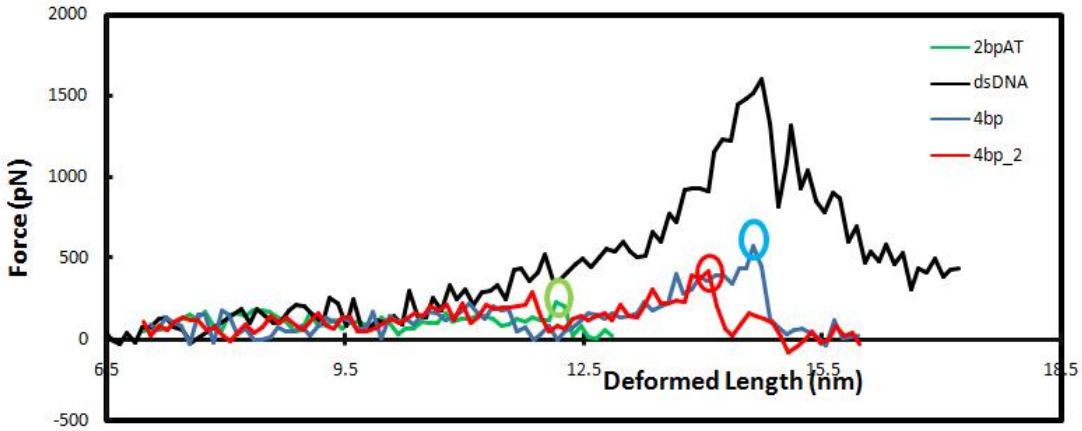

Figure 4.6, Force-displacement curve of two trajectories of a 4 base pair long sticky end joint <8, 9> compared to the 2 base pair AT sticky end joint <5> and the 20 base pair double stranded DNA <1>.

In the early stages of stretching the 4bp SE the structure splits into two helical domains and later a third helical domain is formed in the middle. In late stages of stretch a cohesive complex forms in the middle helical domain. This complex includes a total of 6 bases, 2 GC base pairs and two G bases supporting them. These bases are shaded in the sequence in Table 4.1. The supporting G bases are unpaired. In relation to the loading direction they are located in a specific way. They're located in a way that sliding of strands in the direction of loading will result in two mismatched GG and one mismatched CC base pairs. The cohesive complex retards the dissociation of the SE. In the early stages of its formation it is inclined relative to the loading direction. But it turns gradually and just before dissociation of the SE it aligns itself with the loading in a way that the hydrogen bonds in the structure align with the loading direction. The cohesive complex and its evolution with loading are shown in Figure 4.7.

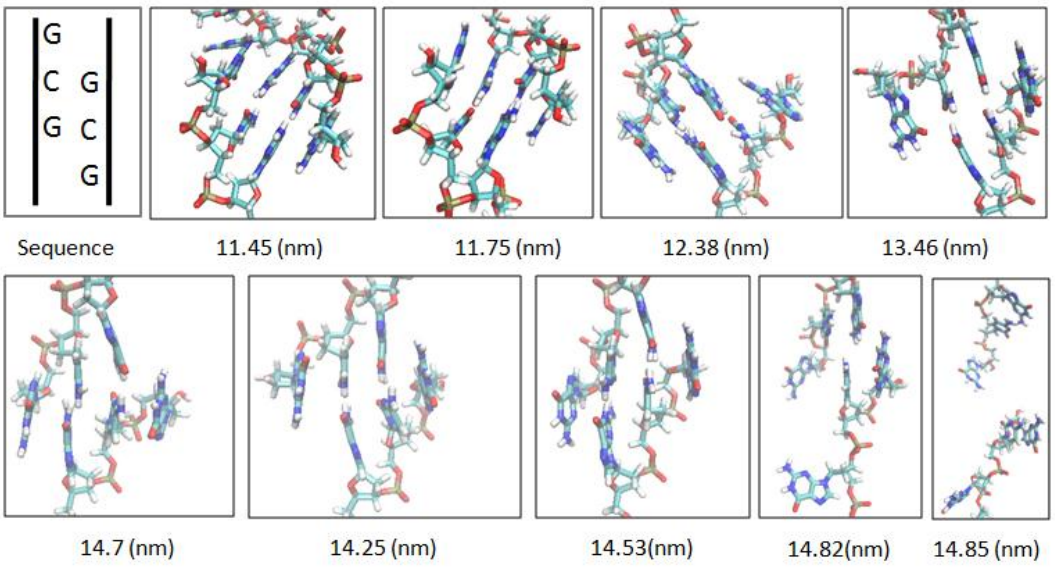

Figure 4.7, Configuration and structural evolution of the strong intermediate complex formed during the stretch of the 4bp sticky end joint <8>. On top left a schematic of this cohesive complex and its sequence is shown. Each snapshot shows the structure of the complex at a certain deformation step (total imposed displacement is indicated).The direction of loading is vertical. The strand in the bottom is fixed and the one on top is mobile.

At late stages of the 4bp SE stretch the structure is distorted. This is shown by plotting the changes of the conformational energy and nonbonded energy portions of the potential energy during stretch. The variation of the conformational energy is a significant part of the energy change and is mostly due to change in the dihedral angles energy, as seen before for the dsDNA. Further there is a correlation between the dihedral energy variation and the variation of the force, as expected (Figure 4.8).

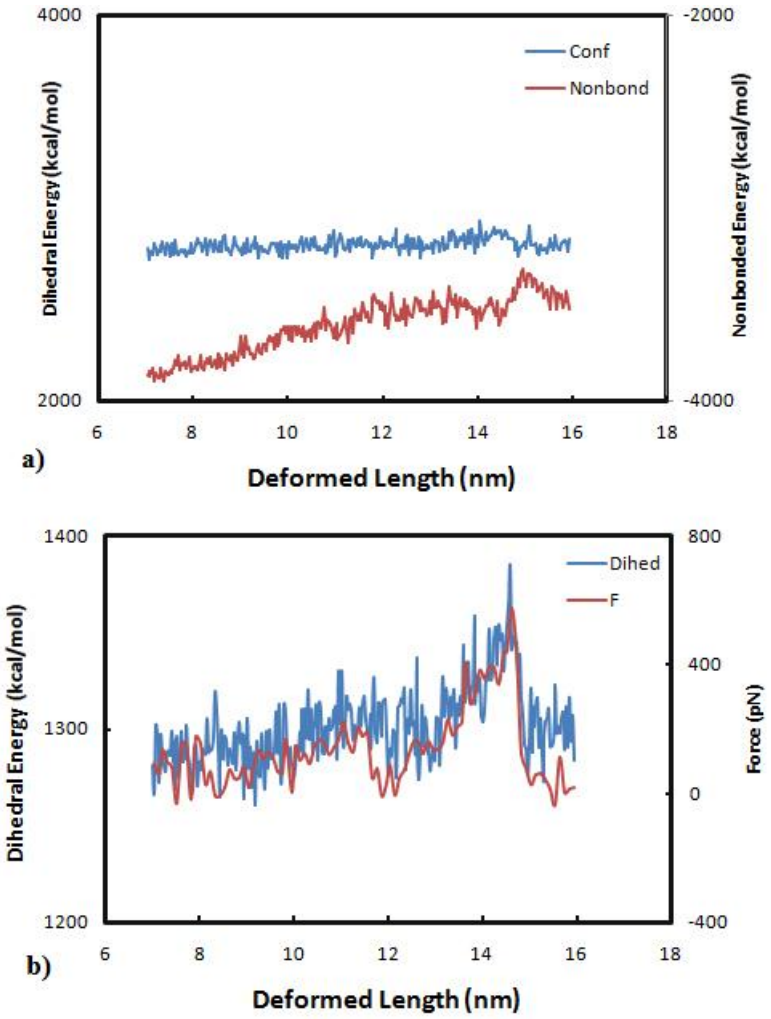

Figure 4.8, Correlation of the dihedral energy of the 4 base pair long sticky end joint <8> with the force experienced in stretch. (a) The conformational (blue) and nonbonded (red) parts of the potential energy of the SE in deformation. (b) Correlation between the dihedral energy (blue) and force (red) of the 4bp SE in stretch.

4.3 Dissociation of 6 base pair-long sticky end joints subjected to axial loading

Six base pair sticky ends with three different sequences are tested. 6bp <10> has the same sequence as the dsDNA. The 6bpGCEDGEATCORE <11> is made by replacing the GC base pairs in the core with AT pairs and the AT base pairs at the edges by GC

base pairs (Table 1). Further the 6bpGCGCGC <12> is made by replacing the AT base pairs in the core by GC base pairs.

The first two sequences are as weak as the 2 base pair long sticky ends. But the last sequence, 6bpGCGCGC is way stronger. It dissociates at a peak force of about 700 pN. The responses of these sequences are compared to that of the 4bp sequence in Figure 4.9.

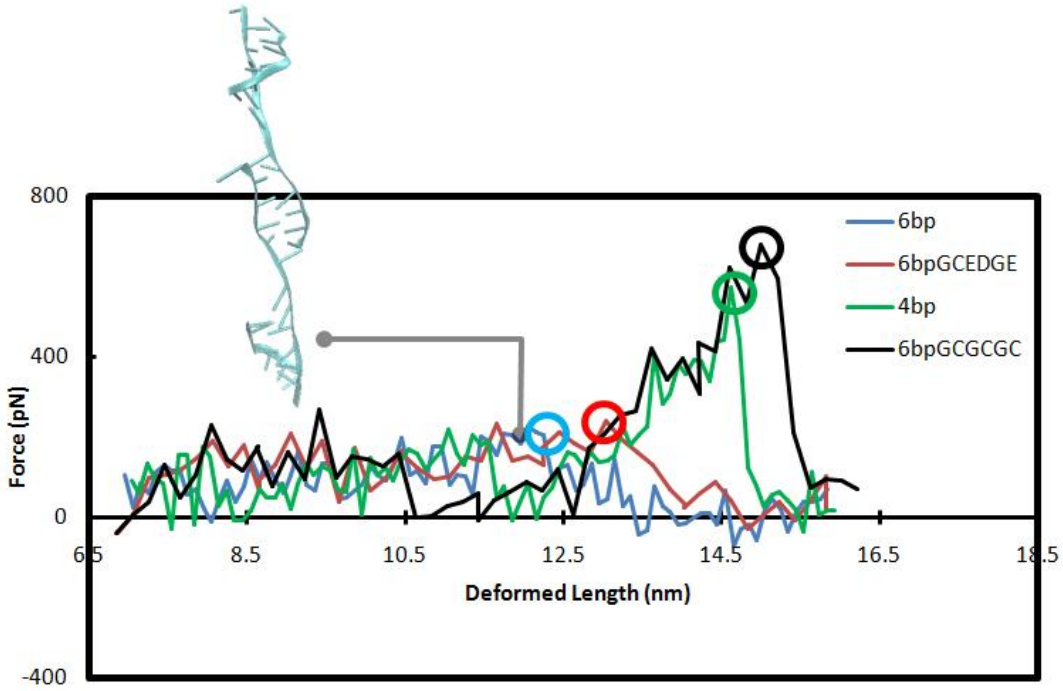

Figure 4.9, Comparison of force-displacement diagrams of different 6bp sticky ends <10-12> and the 4bp sticky end <8>. The snapshot is related to the 6bp sticky end <10> at a displacement of about 12 nm.

The strong 6bpGCGCGC sequence dissociates in a similar fashion to the 4bp SE. The SE first splits into two helical domains and in later stages into three helical domains. The same cohesive complex shown in Figure 4.7 appears in the middle helical domain during the stretch. The sequence producing the cohesive complex is similar in the two cases and is shaded in Table 4.1. Similarly the structure is distorted at late stages of deformation. The correlation between the dihedral part of the potential energy of the SE before dissociation and the force shows this. The significant change in dihedral energy and its correlation with the force are shown in Figure 4.10.

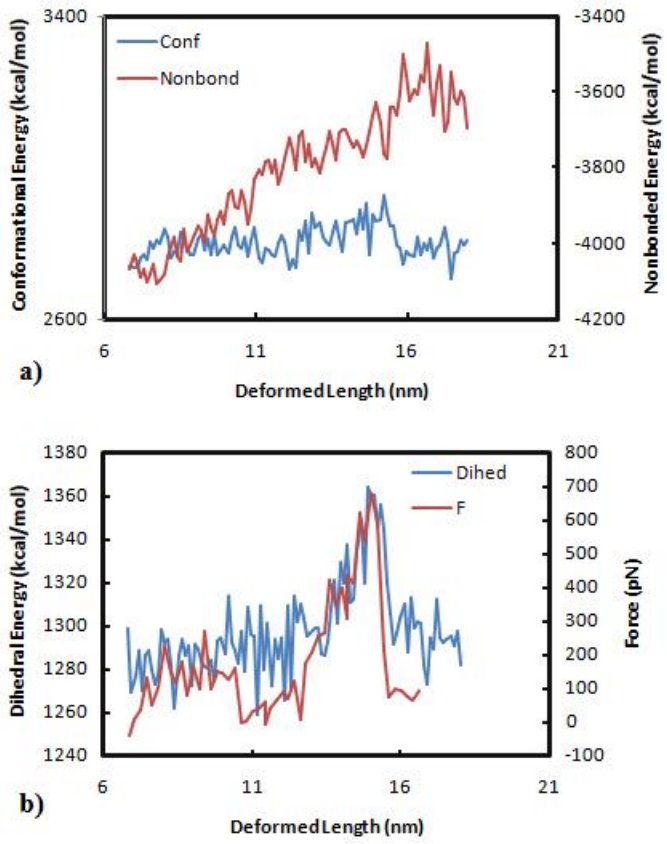

Figure 4.10, Correlation of the dihedral energy of the 6bpGCGCGC sticky end joint <12> with the external force experienced in stretch.(a) The conformational (blue) and nonbonded (red) parts of the potential energy of the SE in deformation. (b) Correlation between the dihedral energy (blue) and force (red) of the 4bp SE in stretch.

The strength of the 6bp SE changes greatly by changing the sequence. High dissociation forces in stretch of the SEs are achieved by the formation of the cohesive complex described. The cohesive complex is observed in stretch of the two 4bp and one of the 6bp SEs <12>, but it was not observed in any of the other 6 base pair SEs <10> and <11>. Also it was not observed during stretching of any of the 2 base pairs SEs<3-7>.

It is observed that during the stretch of the 2 base pair sticky ends, the SE splits into two helical domains in the early stages of stretch. The resulting structure is shown in Figure 4.11a. Formation of the cohesive complex structure is not possible after this point. Note that the cohesive complex is stabilized because if sliding happens two GG and one CC mismatched pairs occur. For a sticky end of length 2 if sliding happens only

two bases will remain in contact. In addition, SEs of lengths 4 and 6 form three helical domains in the late stages of stretch. This is shown in Figure 4.11b and Figure 4.11c. Many bases are engaged between the two strands in the helical domain in the middle and the cohesive complex occurs in this region.

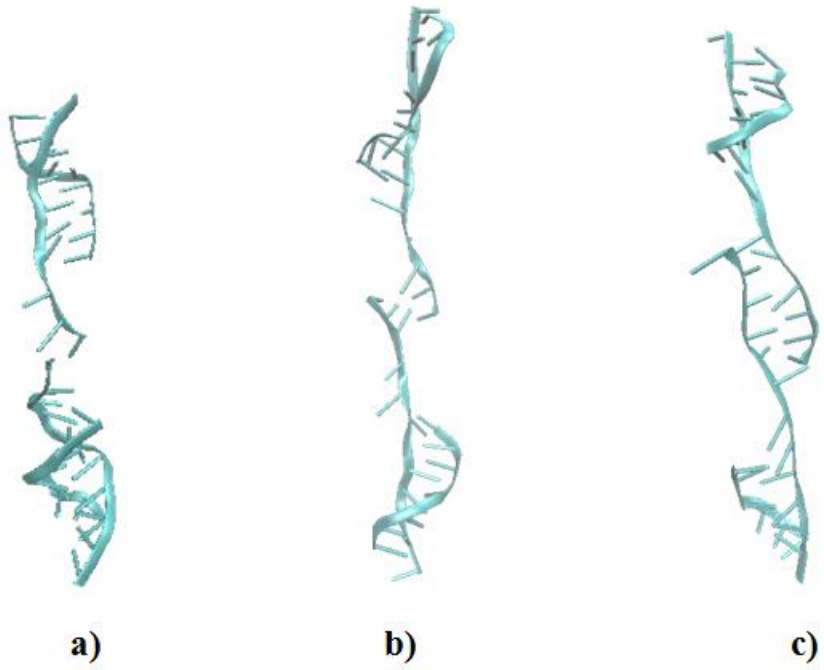

Figure 4.11, Formation of multiple helical domains in sticky end molecules in stretch. (a) Formation of two helical domains in stretch of the 2bp SE at the deformed length of 10 nm. (b) Formation of three helical domains in stretch of the 4bp SE at the deformed length of 12.5. (c) Formation of three helical domains in stretch of the 6bp SE at the deformed length of 12.5.

5. CONCLUSIONS

The mechanical behavior of DNA sticky ends (SEs) of different lengths and base sequences is studied in this thesis. Stretch of a double stranded DNA (dsDNA) is investigated to provide a reference and to verify the model and simulations against the existing literature. Microstructural changes of the structures are studied through measurements of different base parameters and potential energies. Also a cohesive intermediate complex of bases is identified, that forms during the stretch of molecules with certain base sequences. These molecules also exhibit the highest strength.

A variety of base sequences are tested for 2, 4 and 6 base pairs (bp) sticky ends. The 2 bpSEs fail under much lower values of force and extension compared to the dsDNA and the 4bp SEs. Among the three 6bp SEs two are as weak as the 2bp SEs and one is as strong as the 4bp SEs. So under stretch the SEs can be categorized into two groups, weak and strong.

The failure of the weak SEs can be described as the dissociation of two helical domains which remain essentially rigid during stretch. Only the SE region deforms and eventually separates. However this is not the case for the strong SEs. These first split into two and then three helical domains. In later stages of deformation a cohesive complex of two GC base pairs and two unpaired supporting G bases forms in the helical domain in the middle. This complex stabilizes the SE and postpones the dissociation of the SE. So in stretch of the stronger SEs at late stages of dissociation the helical domains deform and that is accompanied by changes in the dihedral angle part of the energy. This is similar to the behavior of the dsDNA in stretch before the peak force is reached.

At early stages of its formation, the cohesive complex is inclined with respect to the direction of loading. Then it gradually rotates so that before dissociation the hydrogen bonds of the base pairs are aligned with the direction of loading. The six base pairs included in the cohesive complex are arranged in a way that sliding of the bases in the direction of loading will result in a complex of two mismatched GG base pairs and one mismatched CC base pair. This keeps the strands from sliding and for the failure of the entire structure, all pairs in the SE have to dissociate together. This structure is not observed in 2bp SEs.

REFERENCES

[1] R. Sinden, *DNA Structure and Function*. San Diego: Academic Press, 1994.

[2] N. C. Seeman, "DNA in A Material World." *Nature*, vol. 421, no. 6921, pp. 427–31, Jan. 2003.

[3] J. D. Watson and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, vol. 171, no. 4356, pp. 737–738, Apr. 1953.

[4] M.P. Ball, http://en.wikipedia.org/w/index.php?title=File:DNA_chemical_structure.svg. Date last accessed 10/23/2012.

[5] http://en.wikipedia.org/wiki/File:GC_base_pair_jypx3.png, and http://en.wikipedia.org/wiki/File:AT_base_pair_jypx3.png. Date last accessed 10/23/2012.

[6] J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginell, C. Mao, and N. C. Seeman, "From Molecular to Macroscopic via the Rational Design of A Self-Assembled 3D DNA Crystal.," *Nature*, vol. 461, no. 7260, pp. 74–7, Sep. 2009.

[7] H. Qiu, J. C. Dewan, and N. C. Seeman, "A DNA Decamer with A Sticky End: the Crystal Structure of d-CGACGATCGT.," *Journal of Molecular Biology*, vol. 267, no. 4, pp. 881–98, Apr. 1997.

[8] M. Allen and D. Tildesley, *Computer Simulation of Liquids*. New York: Oxford University Press, 1987.

[9] M. van Dijk and A. M. J. J. Bonvin, "3D-DART: a DNA Structure Modelling Server.," *Nucleic Acids Research*, vol. 37, no. Web Server Issue, pp. W235–9, Jul. 2009.

[10] N. Foloppe and A. D. Mackerell, "All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data," *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 86–104, 2000.

[11] A. D. Mackerell and N. K. Banavali, "All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution," *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 105–120, 2000.

[12] W. Humphrey, A. Dalke, and K. Schulten, “VMD: Visual Molecular Dynamics,” *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, Feb. 1996. http://www.ks.uiuc.edu/Research/vmd/ (Date last accessed 10/23/2012).

[13] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, “Scalable Molecular Dynamics with NAMD,” *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–802, Dec. 2005. http://www.ks.uiuc.edu/Research/namd/ (Date last accessed 10/23/2012).

[14] D. Quigley and M. I. J. Probert, “Langevin Dynamics in Constant Pressure Extended Systems,” *The Journal of Chemical Physics*, vol. 120, no. 24, pp. 11432–41, Jun. 2004.

[15] P. M. D. Severin, X. Zou, H. E. Gaub, and K. Schulten, “Cytosine Methylation Alters DNA Mechanical Properties.,” *Nucleic Acids Research*, vol. 39, no. 20, pp. 8740–51, Nov. 2011.

[16] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, “UCSF Chimera—A Visualization System for Exploratory Research and Analysis,” *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.

[17] A. Balaeff, S. L. Craig, and D. N. Beratan, “B-DNA to Zip-DNA: Simulating a DNA Transition to A Novel Structure with Enhanced Charge-Transport Characteristics,” *The Journal of Physical Chemistry. A*, vol. 115, no. 34, pp. 9377–91, Sep. 2011.

[18] C. Bustamante, Z. Bryant, and S. B. Smith, “Ten Years of Tension: Single-Molecule DNA Mechanics,” *Nature*, vol. 421, no. 6921, pp. 423–427, Jan. 2003.

[19] P. Hagerman, “Flexibility Of Dna,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 17, no. 1, pp. 265–286, Jan. 1988.

[20] J. van Mameren, P. Gross, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G. J. L. Wuite, and E. J. G. Peterman, “Unraveling the Structure of DNA During Overstretching by Using Multicolor, Single-Molecule Fluorescence Imaging.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 43, pp. 18231–6, Oct. 2009.

[21] C. Schneider, “Nucleic Acid Nomenclature and Structure.” http://www.fli-leibniz.de/ImgLibDoc/nana/IMAGE_NANA.html. Date last accessed 10/23/2012.