Nucleation of Bubbles in DNA

Kathleen Hummel
Department of Physics, The College of New Jersey, Ewing, NJ 08628

Yan Zeng and Giovanni Zocchi
Department of Physics and Astronomy, University of California Los Angeles,
Los Angeles, CA 90095-1547

Abstract

Understanding the thermal denaturation process of DNA provides insight into other biological processes that involve denaturation. Examining bubbles in DNA investigates the intermediate states that exist in this important melting transition. Information can be gathered by studying the impact the location of a defect that will nucleate a bubble has on the amount of bubble states that form. A relationship exists where the closer the defect is to the middle, the less bubble states that form. Further investigation found that the nature of the defect appeared to impact the amount of bubble states that formed.

Introduction

DNA is a polymer composed of two single strands that form a double helix. The double helix is made up of the bases: A, T, C and G. The single strands complement each other such that, A pairs with T, and C pairs with G. When the DNA double helix is heated to sufficiently high temperature, the base pairs will completely separate into single strands of DNA or denature. Depending on the composition of the DNA sequence, it may form intermediate states that contain bubbles. Bubbles are single stranded regions bound by double stranded regions. This melting process illustrates a phase transition, since the double-stranded and single-stranded DNA are two different types of polymers. The double stranded DNA is stiff and the single-stranded DNA is flexible.

DNA is a favorable polymer to study because sequences can be synthesized such that, the exact structure is known. Furthermore, defects can be inserted into the sequence so the location and type of the defect is known. This can be accomplished by creating a site where the bases are not complementary. Frequently with polymers, defects can not be controlled in this manner. By inserting a single mismatch into a sequence that is known to form bubble intermediate states, it ensures the bubble will nucleate there. In this way, the bubble intermediates can be studied in more depth.

The study of bubble intermediates and the thermal denaturation process will lead to a better understanding of denaturation in general. This is important to the understanding of biological processes such as duplication and transcription.

Results

To begin with, the same defect located at different points along the sequence is examined. This is followed by a study of a different defect in the same location.

Bubble Location

To study the effect the location of the defect has on the nucleation of bubbles, five different samples were used. Each sample was based on a 40mer sequence and contained a G-G defect at different positions along this sequence. These samples contained mismatches at 7, 14, 22, 25, and 29. This data was then compared to the same sequence without a defect.

Aside from one outlying point (mis29), the remaining samples illustrated that the location of the defect produced a fairly linear trend (Figure 2b). As the bubble was placed toward the middle, the statistical weight of intermediate states decreased. Similarly, when the defect was located near the end of the sequence, more bubble intermediates were present.

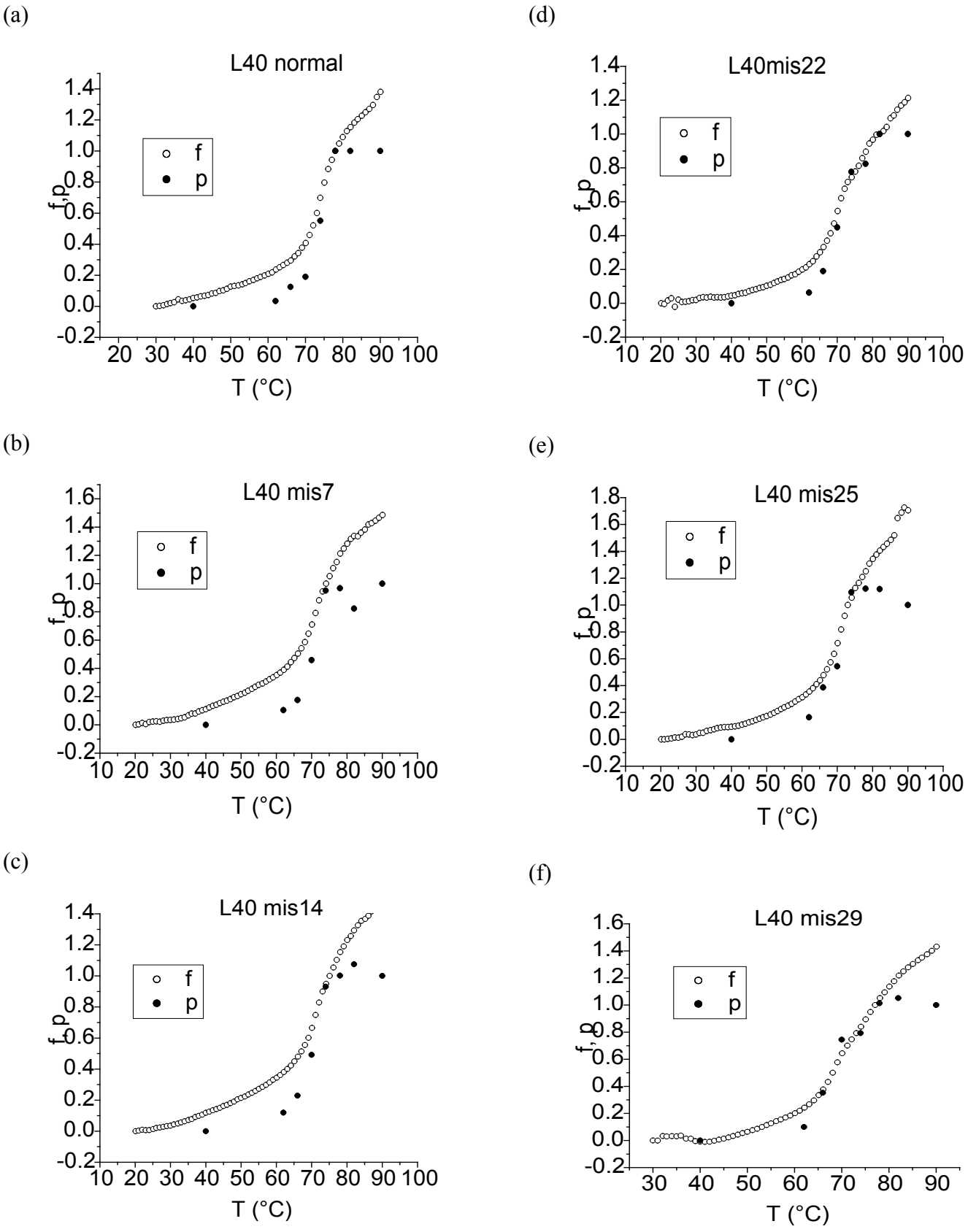
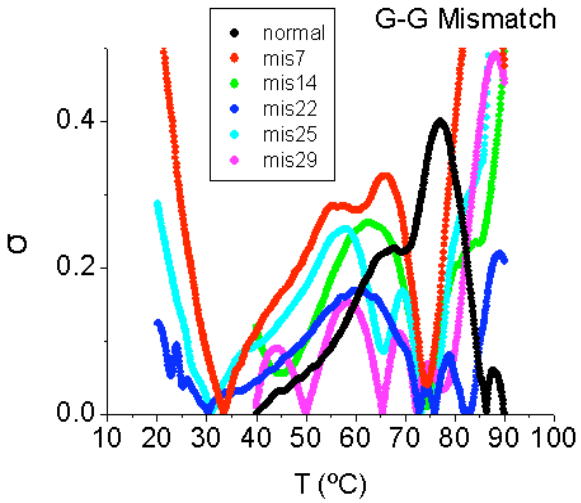


Figure 1. Melting curves, f (UV) and p (gel) for the 40mer without a mismatch (a), and the 5 sequences with a G-G mismatch (b-f).

(a)



(b)

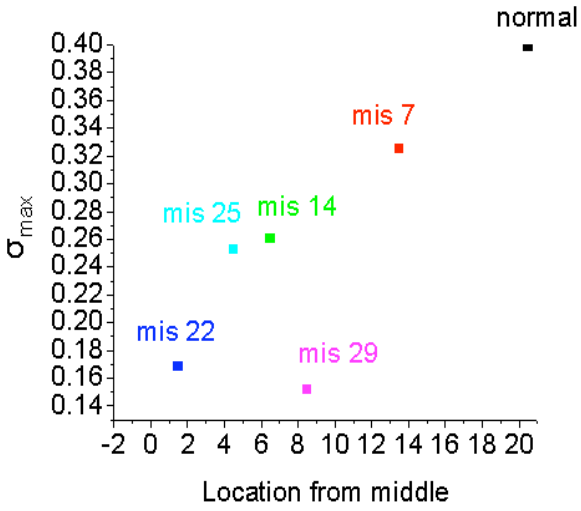


Figure 2. (a) _ = f-p, where _ is the fraction of the bases in a bubble state for each of the 5 sequences with a G-G defect, as compared to the sequence without a defect. (b) The maximum _ for each of the sequences plotted against the number of bases the mismatch is located from the middle of the sequence (20.5).

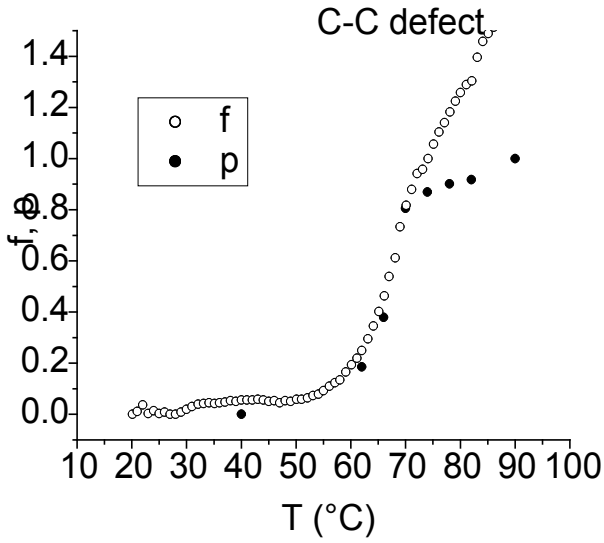
In comparison to the sequence without a defect, or the normal sample, the samples with a bubble near the end behaved similarly to the normal sample. In addition, the normal sample has the greatest amount of bubble intermediates out of all of the samples.

Type of Defect

To examine the relationship between the nature of the defect and the nucleation of bubbles, two samples were used. The same 40mer sequence was used as before, but the type of defect was modified. Both a G-G defect and C-C defect were synthesized at the 22 location to be used for comparison.

The G-G defect appears to form more bubble intermediate states than the C-C defect in the same location. In fact, the C-C defect is closer to a two state process, where it seems to form no bubble intermediates. The G-G defect is closer to the normal strand.

(a)



(b)

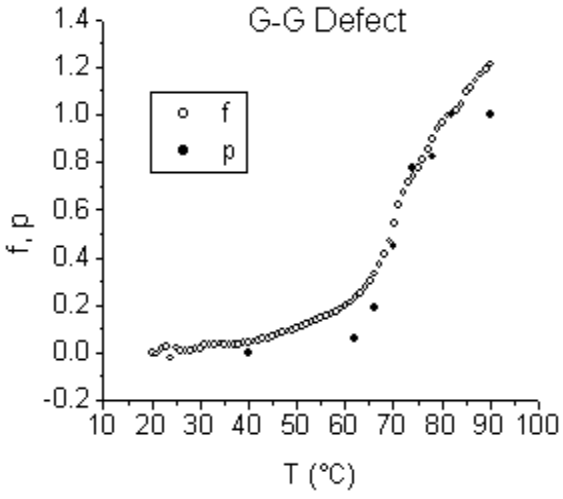


Figure 3. For the type of defect study (a) C-C defect at 22 location. (b) G-G defect at 22 location.

Discussion

The data from the bubble location study supports previous results that a defect at the end of a sequence will behave similarly to the sequence with no defect at all. In this case, the mismatch at location 7 is the closest to the end, and has the largest amount of base pairs involved in a bubble state, other than the normal strand. This can be explained by the concept of "fraying." Since the end is already a defect, the generation of an additional defect near the end will increase the instability of the already existing defect. This is why the mismatch 7 sample is the most similar to the normal strand.

As the defect and the bubble nucleation site is located in the center of the sequence (mis 22 or mis 25), the sequence now has 3 distinct defects, the one at each end and the one in the middle. This is generating a sample that would be comparable to two smaller sequences of 22 bp and 19 bp (mis22), or 25 bp and 16 bp (mis 25), only influenced by end affects. From previous study, there is a linear relationship between the length of the sequence and the number of bubble states it forms. The longer the sequence the more bubble states it will form. Therefore, if the longest double helix structure present is 22 bp, or 25 bp long, this will have fewer bubble states than the 40mer, or even the 34mer represented by the mis 7. The maximum number of base pairs involved in a bubble state is directly correlated to the distance the defect is from the middle.

The results of the type of defect study suggest that the nature of the defect is significant to the amount of bubble states that will form when the sequence denatures. This could be explained by the difference in chemical composition, size, or nearest-neighbor interactions. Further study needs to be conducted to distinguish the cause of these results or whether this is an anomaly.

Materials and Methods

Synthetic, salt-free, DNA oligomers were purchased from Operon. The sequences were generated so they are partially self-complementary. This ensures they can fold over to form hairpins if rapidly cooled. The single strands were annealed to their complements by heating them together in a 90ºC water bath for five minutes. The samples are then cooled down to room temperature overnight.

Sequences for Bubble Location study:
L40r: CCGCGCCACTTGGCCTGCCT CCG TCCCGCGGGACGGAGGC
L40rmis7: CCGCGCGACTTGGCCTGCCT CCGTCCCGCGGGACGGAGGC
L40rmis14: CCGCGCCACTTGGGCTGCCT CCGTCCCGCGGGACGGAGGC
L40rmis22: CCGCGCCACTTGGCCTGCCT CGGTCCCGCGGGACGGAGGC
L40rmis25: CCGCGCCACTTGGCCTGCCT CCGTGCCGCGGGACGGAGGC
L40rmis29: CCGCGCCACTTGGCCTGCCT CCGTCCCGGGGGACGGAGGC
Sequences for Type of Defect study:
L40mis22(G-G): CCGCGCCACTTGGCCTG CCTCGGTCCCGCGGGACGGAGGC
L40mis22(C-C): GCCTCCGTCCCGGGGACC GAGGCAGGCCAAGTGGCGCGG

Quenching Method

As the DNA double helix is heated to higher temperatures, the strands will begin to separate. When the DNA is rapidly cooled, most of the completely open single strands will fold over to form hairpins, whereas, the partially open DNA will reform a double helix.

To execute this method, each of the samples is diluted to a concentration of 1 pmol/_L and divided into 8, 20 _L aliquots. Each aliquot is heated in a water bath of a different temperature for three minutes. It is then placed in an ice water bath (~0ºC) for quenching. These aliquots are studied using gel electrophoresis. A 3% agarose gel is stained with ethidium bromide and run at 120 V for 65 minutes. The resulting gel is digitally photographed under UV light and the relative intensities of the hairpin and duplex bands are interpreted using an image processing program, called Scion Image.

If greater resolution is required, Polyacrylamide Gel Electrophoresis (PAGE) is employed. The samples remain at a 1 pmol/_L concentration, but are divided into 15, 10 _L aliquots. This gel is typically run at 160 V for 45 minutes, followed by 40 minutes of staining in a SYBR Gold solution. The rest of the procedure remains the same.

Once the band intensities for the hairpin (hp) and duplex (ds) are obtained from the program, the fraction of hairpins can be determined using hp/(hp+ds), and will provide the p curve. At low temperatures, the ratio should go to zero, to indicate that only duplex exists. If this is not the case, the amount of hairpin present at the lowest temperature is subtracted from the subsequent temperatures. Similarly, at sufficiently high temperatures, this ratio should go to one, to indicate all of the duplexes have separated. When some duplex still exist at high temperature, this is corrected by the factor, _, in the following way:

$$p = \frac{hp}{hp + ds} \frac{1}{1 - g(hp)} \qquad (1)$$

UV Absorption

UV absorption detects the amount of open base pairs present in the sample. Once this melting curve is normalized between 0 and 1, it will provide the fraction of open base pairs present at a particular temperature. This is the f curve.

When the p curve is subtracted from the f curve _ is obtained, where _ is the fraction of base pairs involved in the bubble.

Acknowledgements

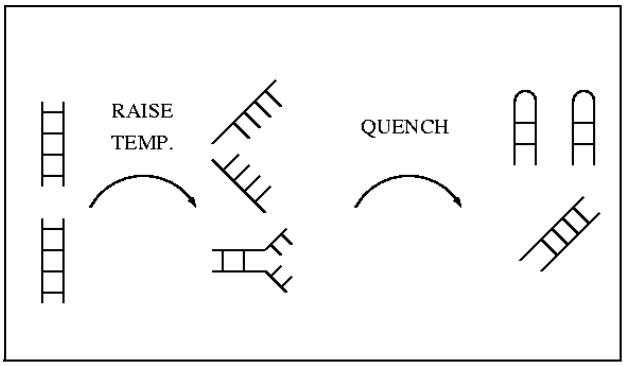
I would like to thank Dr. Giovanni Zocchi and Yan Zeng for mentoring this research. I would also like to thank Francoise Queval for coordinating the program and NSF for funding this REU program.

References

[1] Zeng, Y., Montrichok, A., & Zocchi, G. (2004). Bubble nucleation and cooperativity in DNA melting. *J. Mol. Biol.* **339**, 67-75.

[2] Montrichok, A., Gruner, G. & Zocchi, G. (2003). Trapping intermediates in the melting transition of DNA oligomers. *Europhys. Letters*, **62**, 452-458.

(a)



(b)

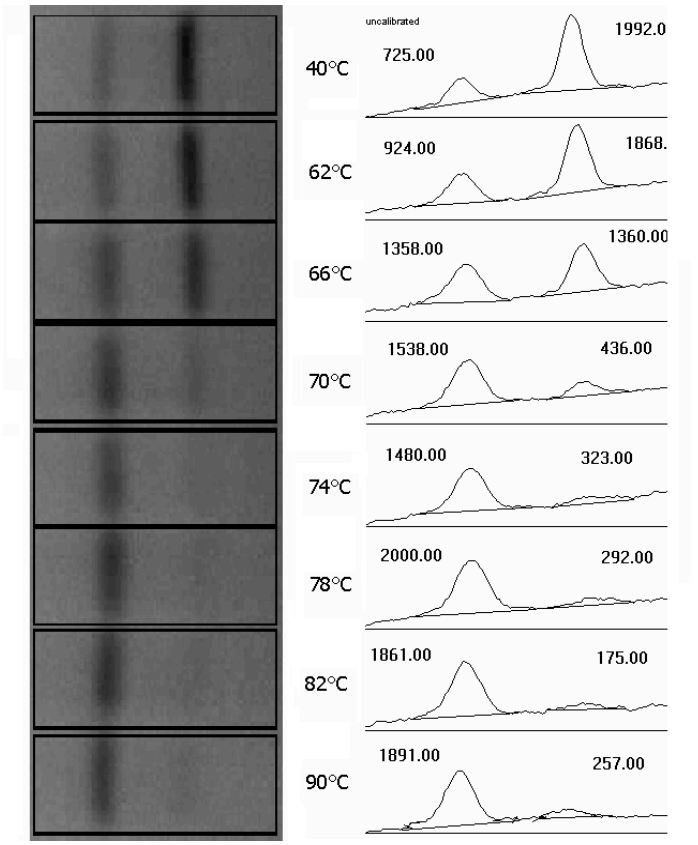


Figure 4. (a) Diagram of the quenching method. (b) Example of UV picture of an agarose gel and the corresponding intensity plot from Scion Image.