Published in final edited form as:
Nano Lett. 2010 November 10; 10(11): 4697–4701. doi:10.1021/nl102986v.

Peptide Nucleic Acids as Tools for Single-Molecule Sequence Detection and Manipulation

Hagar Zohar[a], **Craig L. Hetherington**[b], **Carlos J. Bustamante**[b,c], and **Susan J. Muller**[a]

[a]Department of Chemical Engineering, Biophysics Graduate Group, California Institute for Quantitative Biosciences (QB3), and Howard Hughes Medical Institute, University of California, Berkeley, California 94720

[b]Department of Physics and Jason L. Choy Laboratory of Single-Molecule Biophysics, Biophysics Graduate Group, California Institute for Quantitative Biosciences (QB3), and Howard Hughes Medical Institute, University of California, Berkeley, California 94720

[c]Department of Molecular and Cellular Biology, Department of Chemistry, Biophysics Graduate Group, California Institute for Quantitative Biosciences (QB3), and Howard Hughes Medical Institute, University of California, Berkeley, California 94720

Abstract

The ability to strongly and sequence-specifically attach modifications such as fluorophores and haptens to individual double-stranded (ds) DNA molecules is critical to a variety of single-molecule experiments. We propose using modified peptide nucleic acids (PNAs) for this purpose and implement them in two model single-molecule experiments where individual DNA molecules are manipulated via microfluidic flow and optical tweezers, respectively. We demonstrate that PNAs are versatile and robust sequence-specific tethers.

Keywords

DNA; Peptide Nucleic Acid; Optical Tweezer; Microfluidic Cross-Slot; Single-Molecule; Sequence-Specific

Peptide nucleic acids (PNAs) are synthetic molecules composed of a peptide backbone and nucleic acid bases, which confer upon them DNA sequence recognition. They form strong and specific bonds to sequences in dsDNA, via four different binding modes, depending on their design.[1] Due to a neutrally-charged backbone, their binding to dsDNA deviates substantially from that of conventional oligonucleotides. These differences are exploited in the many applications of PNAs, which are covered in a number of excellent works.[1–11]

While a myriad of PNA applications have been developed, we seek to focus on the promising application of PNAs as tools for single-molecule experiments and to highlight the advantages of single-molecule experiments for optimizing PNA-DNA binding conditions. Many single-molecule dsDNA experiments require a method to attach a desired modification to a specific, non-terminal location on dsDNA. The modification is usually a small molecule such as a hapten[12–14] or fluorophore.[15–17] The modification can facilitate sequence-specific detection or attachment of other components, such as nanospheres.

muller2@berkeley.edu .

Supporting Information Available: Additional background on PNAs, detailed materials and methods, and supporting figures are available free of charge via the Internet at http://pubs.acs.org.

Ultimately, such methods are being explored for a range of single-molecule biophysics experiments, for single-molecule genotyping, and for sequence-based separation methods. Since PNAs are currently commercially available with a choice of modifications including haptens, such as biotin, and fluorophores, such as TAMRA, they are excellent tools for attaching sequence-specific modifications to dsDNA.

There are several notable examples of single-molecule DNA experiments successfully employing PNAs, primarily to locate sequences on genomic-length DNA using various approaches. [12,15,16,18,19] While there are other commonly used techniques to introducing site-specific modification for single-molecule DNA experiments,[17,20–26] PNAs are advantageous because they are compatible with cofactors such as Mg^{2+} and non-destructive to dsDNA. Furthermore, they do not require ligation and bind with high specificity and yield.[15] While PNAs have many favorable attributes, their binding conditions require careful optimization.

Here we describe two model experiments that demonstrate the virtues of PNAs as tools for single-molecule DNA experiments. In the first, individual PNA-DNA complexes are studied via fluorescence microscopy. The DNA backbone is fluorescently stained, the bound PNAs are labeled with fluorescent nanospheres, and the complexes are stretched either on a slide or in a stagnation-point extensional flow, which produces controlled, precise extension of the DNA (Figure 1). In the second, individual PNA-DNA complexes are characterized via optical tweezers. A sphere attached to one end of the PNA-DNA complex is manipulated with optical tweezers while a second sphere, attached to the bound PNA, is held stationary by suction on a micropipette (Figure 2). Collectively, these single-molecule experiments with 8 bp bisPNAs demonstrate that (i) PNA-DNA binding optimization can be performed simultaneously across all target sites for an entire genomic-length DNA molecule (ii) PNAs can be used to locate specific dsDNA sequences on individual DNA molecules (iii) PNAs can serve as sequence-specific tethers for optical tweezer setups, and (iv) the PNA-DNA bond can sustain forces of roughly 60 pN on average.

The two bisPNAs used here, referred to as flPNA and twPNA for fluorescence microscopy and optical tweezer experiments, respectively, are functionalized with TAMRA at the N-terminus and biotin at the C-terminus (see Supporting Information). The target binding sites for flPNA and twPNA are GAGAAGGA and AAGAGAAA, respectively. PNA-DNA binding conditions were first optimized for specificity because, in addition to binding to target sequences, PNAs are known to bind to sequences that have single-end mismatches (SEMM) or double-end mismatches (DEMM).[12,15,16] The traditional approach to determining the extent of PNA binding and specificity is a bulk gel-shift assay. Subsequent to hybridizing DNA with PNA at a given set of conditions, the mobility of fragments of DNA containing either a target site or mismatch site is compared. A representative example of this optimization process for flPNA is presented in the Supporting Information. While the widely-used bulk gel-shift assay is informative, it is limited in that the non-specific binding sites must be anticipated and, thus, such sites will only be discovered if the appropriate fragments are selected in advance.

To simplify the binding optimization procedure and decrease the opportunities for non-specific binding, PNA sequences should be selected such that the occurrences of SEMM and DEMM sites, particularly when clustered, are minimized. The locations of the target, SEMM, and DEMM sites are easily determined by searching a known DNA sequence. When all of these sites are binned in the same way as the experimental data, the histograms provide the locations of potential erroneous peaks. In Figure 3a, the predictive binding map for the binding of flPNA on λ-DNA, there is a cluster of DEMM sites between 30 and 40 kbp. There are also clusters consisting of both SEMM and DEMM sites between 40 and 50

kbp. This type of non-specific binding, where a cluster of low-likelihood binding events appears as a substantial peak in the binding location histogram, is made obvious by the predictive binding map (Figure 3a) and would not have been detected with the traditional bulk gel-shift assay approach. When the binding reaction is not followed by additional heating to optimize binding, the slide-stretching assay indeed reveals extensive non-specific binding (Figure 3b).

Using a simple single-molecule slide-stretching method, entire molecules, up to several Mbp in size,[27] can be assayed simultaneously for non-specific binding. The specificity of binding conditions can be evaluated by stretching the PNA-DNA complexes created at those conditions on a slide and determining the positions of the labels. The presence of a fluorescent label, as opposed to a single fluorophore, simplifies this approach by not requiring a TIRF setup or too much attention to reducing background fluorescence. The flPNA and λ-DNA complexes (flPNA-DNA) were formed in a binding reaction that was incubated at 37 °C for 25 hours. Drop dialysis was performed to remove excess flPNA. To optimize binding specificity, the recovered volume was heated to 63 °C for 15 minutes following dialysis. Immediately after heating, the NaCl concentration was brought to 100 mM to inhibit any further binding. The flPNA-DNA complexes were labeled with 40 nm NeutrAvidin-coated fluorescent polystyrene spheres and stained with YOYO-1 prior to slide-stretching. A 5 μL drop of stained, labeled flPNA-DNA in slide-stretching buffer was deposited onto an untreated glass slide (see Supporting Information). After 1 minute, a poly-L-lysine-coated cover slip was deposited onto the drop. Individual DNA molecules were stretched onto the positively charged surface by the receding meniscus. For both slide-stretching and flow-stretching (discussed below), the locations of the labels on individual DNA molecules are readily determined from intensity profiles along each DNA backbone, imaged via fluorescence microscopy (Figure 4). The most compelling advantage of this slide-stretching approach is that mismatch sites do not need to be anticipated and all types of mismatches are detected. This is particularly important when clusters of low likelihood mismatches are present and can result in erroneous peaks.

When the binding conditions are optimized, slide-stretching reveals only binding to the target sites (Figure 5a). The target site positions can be visualized by fitting Gaussians to the two apparent subpopulations. Note that both the predicted and experimental maps of binding locations are modified such that all sites are on one half of the molecule. This approach is consistent with procedures used in other single-molecule studies[20,23,26,28] and is necessary because, when there is only one label on a DNA molecule, the orientation of the DNA molecule is not known. Furthermore, both target binding sites for flPNA are known to be on the same half of the λ-DNA molecule. The label locations determined via slide-stretching (Figure 5a), which indicate the locations of flPNA binding sites on λ-DNA, are 23.6 ± 3.3 and 40.5 ± 3.0 kbp for the middle and end peaks, respectively. Using slide-stretching, absolute error for the middle peak is 0.8 kbp (target binding site at 24.4 kbp) and for the end peak is 1.1 kbp (target binding site at 41.6 kbp).

After the binding conditions of the flPNA-DNA complexes were optimized for specificity using the slide-stretching assay, the flPNA-DNA binding was more rigorously determined using a flow-stretching assay, which allows for better fidelity in linearizing the DNA and better localizes the binding locations. The stained, labeled flPNA-DNA complexes were hydrodynamically trapped and stretched in a stagnation-point extensional flow generated in a microfludic cross-slot geometry (Figure 1), and the positions of the labels were then determined in the same manner as for slide-stretched molecules. Such microfluidic devices have been used extensively to study single DNA molecules.[29–33] The channel cross-section of the PDMS flow cell was 800 μm by 120 μm and the fabrication is described elsewhere.[28] As illustrated in Figure 5, when the same flPNA-DNA complexes are analyzed via slide-

stretching and flow-stretching, the latter produces a more accurate and precise map. Flow-stretching increases the accuracy and precision of the label position measurements relative to slide-stretching because it results in a narrower distribution of DNA extension and the flow eliminates “false positives” that occur in slide-stretching when free fluorescent nanospheres are coincidentally adsorbed to the slide near a stretched DNA molecule. In addition, in the present experiments, the flow-stretching conditions allow for somewhat greater mean extension of the DNA molecules. Comparing the distribution of DNA extensions, represented as percentage of the DNA contour length, slide-stretching results in 89% ± 6% while flow-stretching results in 96% ± 3%. The same label locations determined via flow-stretching (Figure 5b) are 24.0 ± 1.5 and 42.4 ± 1.6 kbp for the middle and end peaks, respectively. Using flow stretching, absolute error for the middle peak is 0.4 kbp and for the end peak is 0.8 kbp. Thus, utilizing labeled PNAs, the single-molecule flow-stretching assay correctly localizes target sequences to within 1 kbp, which is comparable to other single-molecule sequence detection methods.[23,26,28,34] This flow assay demonstrates the ability of PNAs to serve as indicators of specific sequences on dsDNA.

To demonstrate that PNAs can also be applied as tethers for optical tweezer experiments, the strength of the PNA tether was evaluated. Ideally, tethers for nucleic acids in optical tweezer experiments should be able to sustain forces at least on the order of tens of pN. Forces in this range are characteristic of the folding of RNA and some proteins[35–37] and many molecular motors that act on DNA and RNA.[38–42] As these optical tweezer experiments required distinct attachment sites to two beads, a digoxigenin-modified DNA, half-λ-DNA-dig, was first formed as a precursor. TwPNA-DNA-dig complexes, which have a bound biotinylated PNA, were formed from twPNA and half-λ-DNA-dig in a binding reaction incubated at 45 °C for 5 hours, then cooled to 4 °C over 30 minutes. The NaCl concentration was brought to 50 mM to inhibit any further binding and the reaction volume was heated to 50 °C for 10 minutes to enhance binding specificity. Note that since twPNA and flPNA bind to different target sequences, their binding conditions were optimized separately, resulting in different optimal binding conditions.

Individual twPNA-DNA-dig complexes were manipulated via two spheres attached to the biotinylated PNA and the dig modification, respectively (Figure 2). In addition, similar control molecules that were not tethered via twPNA, called biotin-DNA-dig (see Supporting Information), were subjected to the same routine. Control molecules did not exhibit any rupture, indicating that rupture of the twPNA-DNA-dig complex is due to the twPNA-DNA bond. The twPNA-DNA-dig complexes were repeatedly stretched and relaxed between 0 and 90 pN at a rate of 175 nm per second. Several tw-PNA-DNA-dig complexes were robust to the repeated application of 90 pN forces. For the twPNA-DNA-dig complexes that ruptured at forces below 90 pN, there was a unimodal distribution of rupture forces (Figure 6) with a mean rupture force of 57.3 pN and a standard deviation of 14.5 pN (sample size of 12 molecules). While this rupture force distribution directly characterizes only the specifically-bound twPNA used here, we expect the values to be of the same order for other 8 bp bisPNAs with similar base content. Note that we only consider here those complexes where the twPNA was specifically bound. DNA molecules with specifically-bound twPNAs are easily distinguishable in the optical tweezer experiment. When the DNA is stretched to its contour length, the distance between the two beads is known and corresponds to the distance between the end of the DNA and the twPNA binding location. The binding properties of PNAs make them attractive candidates for single-molecule dsDNA recognition and manipulation. Typical nucleic acid studies use biotin-streptavidin bonds and the weaker dig-anti-dig bonds. The mean rupture force of the PNA-DNA bond compares favorably with the rupture force of a single dig-anti-dig bond, roughly 20 pN at these loading rates,[43] making PNA-mediated tethers another excellent alternative to biotin-mediated tethers for

nucleic acid manipulation. The strength of interaction for a typical 8 bp bisPNA with DNA is well-suited for optical tweezer experiments.

In conclusion, while the versatility of PNAs and their many uses have been discussed widely in the literature, little direct attention has been paid to their potential specifically as tools for single-molecule experiments. They are commercially available and can be synthesized with fluorophores or haptens, their binding to dsDNA is strong and specific, and they are compatible with cofactors required in many single-molecule biological assays. We present both design considerations for PNAs and an approach to optimizing their binding to dsDNA. We discuss their desirable attributes for single-molecule experiments and demonstrate this capacity with both fluorescence microscopy and optical tweezer single-molecule DNA studies. PNAs are amenable to labeling and can be used to locate specific dsDNA sequences, to act as obstacles at specific DNA sequences, or as strong sequence-specific tethers for optical or magnetic tweezer setups. With careful attention to PNA design and optimization of binding conditions, PNAs have great potential as versatile tools for single-molecule research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by NIH awards R21HG4342 to S.J.M, GM32543 and GMOZ1552 to C.J.B. H.Z. gratefully acknowledges an NSF Graduate Research Fellowship.

References

(1). Nielsen PE. Current Opinion in Biotechnology. 2001; 12:16–20. [PubMed: 11167067]
(2). Nielsen PE. Current Opinion in Biotechnology. 1999; 10:71–75. [PubMed: 10047504]
(3). Nielsen PE. Methods in Enzymology. 2001; 340:329–340. [PubMed: 11494857]
(4). Nielsen PE. Methods in Molecular Biology. 2007; 208:1–24.
(5). Nielsen PE, Haaima G. Chem Soc Rev. 1997; 26:73–78.
(6). Lundin KE, Good L, Stromberg R, Graslund A, Smith CIE. Adv Genet. 2006; 56:1–51. [PubMed: 16735154]
(7). Demidov VV. Trends in Biotechnology. 2003; 21:4–7. [PubMed: 12480343]
(8). Larsen HJ, Bentin T, Nielsen PE. BBA. 1999; 1489:159–166. [PubMed: 10807005]
(9). Paulasova P, Pellestor F. Annales de Genetique. 2004; 47:349–358. [PubMed: 15581832]
(10). Shakeel S, Karim S, Ali A. J Chem Technol Biot. 2006; 81:892–899.
(11). Ray A, Norden B. Faseb J. 2000; 14:1041–1060. [PubMed: 10834926]
(12). Demidov VV, Cherny D, Kurakin A, Yavnilovich M, Malkov V, Frank-Kamenetskii M, Sonnichsen S, Nielsen PE. Nucleic Acids Research. 1994; 22:5218–5222. [PubMed: 7816609]
(13). Bryant Z, Stone MD, Gore J, Smith SB, Cozzarelli NR, Bustamante C. Nature. 2003; 424:338–341. [PubMed: 12867987]
(14). Gore J, Bryant Z, Stone MD, Nollmann MN, Cozzarelli NR, Bustamante C. Nature. 2006; 439:100–104. [PubMed: 16397501]
(15). Chan EY, et al. Genome Res. 2004; 14:1137–1146. [PubMed: 15173119]
(16). Phillips KM, Larson JW, Yantz GR, D'Antoni CM, Gallo MV, Gillis KA, Goncalves NM, Neely LA, Gullans SR, Gilmanshin R. Nucleic Acids Research. 2005; 33:5829–5837. [PubMed: 16243782]
(17). Xiao M, et al. Nucleic Acids Research. 2007; 35:e16. [PubMed: 17175538]
(18). Kim J, Hirose T, Sugiyama S, Ohtani T, Muramatsu H. Nano Letters. 2004; 4:2091–2097.

(19). Singer A, Wanunu M, Morrison W, Kuhn H, Frank-Kamenetskii M, Meller A. Nano Letters. 2010; 10:738–742. [PubMed: 20088590]

(20). Géron-Landre B, Roulon T, Desbiolles P, Escudé C. Nucleic Acids Research. 2003; 31:e125. [PubMed: 14530458]

(21). Kuhn H, Frank-Kamenetskii MD. Nucleic Acids Research. 2008; 36:e40. [PubMed: 18344522]

(22). Jo K, Dhingra DM, Odijk T, de Pablo JJ, Graham MD, Runnheim R, Forrest D, Schwartz DC. PNAS. 2007; 104:2673–2678. [PubMed: 17296933]

(23). Ebenstein Y, Gassman N, Kim S, Antelman J, Kim Y, Ho S, Samuel R, Michalet X, Weiss S. Nano Letters. 2009; 9:1598–1603. [PubMed: 19290670]

(24). Oana H, Ueda M, Washizu M. Biochem Bioph Res Co. 1999; 265:140–143.

(25). Taylor JR, Fang MM, Nie S. Analytical Chemistry. 2000; 72:1979–1986. [PubMed: 10815954]

(26). Xiao M, Gordon MP, Phong A, Ha C, Chan T, Cai D, Selvin PR, Kwok P. Hum Mutat. 2007; 28:913–921. [PubMed: 17443670]

(27). Aston C, Mishra B, Schwartz DC. Trends in Biotechnology. 1999; 17:297–302. [PubMed: 10370237]

(28). Dylla-Spears R, Townsend JE, Jen-Jacobson L, Sohn LL, Muller SJ. Lab on a Chip. 2010; 10:1543–1549. [PubMed: 20358051]

(29). Smith DE, Chu S. Science. 1998; 281:1335–1340. [PubMed: 9721095]

(30). Perkins TT, Smith DE, Chu S. Science. 1997; 267:2016–2021. [PubMed: 9197259]

(31). Schroeder CM, Babcock HP, Shaqfeh ESG, Chu S. Science. 2003; 301:1515–1519. [PubMed: 12970560]

(32). Schroeder CM, Shaqfeh ESG, Chu S. Macromolecules. 2004; 37:9242–9256.

(33). Tanyeri M, Johnson-Chavarria E, Schroeder C. Applied Physics Letters. 2010; 96:224101–224103. [PubMed: 20585593]

(34). Yu H, Schwartz DC. Analytical Biochemistry. 2008; 380:111–121. [PubMed: 18570883]

(35). Liphardt J, Onoa B.; Smith SB, Tinoco I, Bustamante C. Science. 2001; 292:733–737. [PubMed: 11326101]

(36). Kellermayer MS, Smith SB, Bustamante C, Granzier HL. Biophysical Journal. 2008; 80:852–863. [PubMed: 11159452]

(37). Cecconi C, Shank EA, Bustamante C, Marqusee S. Science. 2005; 309:2057–2060. [PubMed: 16179479]

(38). Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C. Nature. 2001; 413:748–752. [PubMed: 11607035]

(39). Wuite GJL, Smith SB, Young M, Keller D, Bustamante C. Nature. 2000; 404:103–106. [PubMed: 10716452]

(40). Yin H, Wang MD, Svoboda K, Landick R, Block SM, Gelles J. Science. 1995; 270:1653–1657. [PubMed: 7502073]

(41). Stone MD, Bryant Z, Crisona NJ, Smith SB, Vologodskii A, Bustamante C, Cozzarelli NRP. Natl Acad Sci Usa. 2003; 100:8654–8659.

(42). Wen J, Lancaster L, Hodges C, Zeri A, Yoshimura SH, Noller HF, Bustamante C, Tinoco I. Nature. 2008; 452:598–604. [PubMed: 18327250]

(43). Neuert G, Albrecht C, Pamir E, Gaub HE. FEBS Lett. 2006; 580:505–509. [PubMed: 16388805]

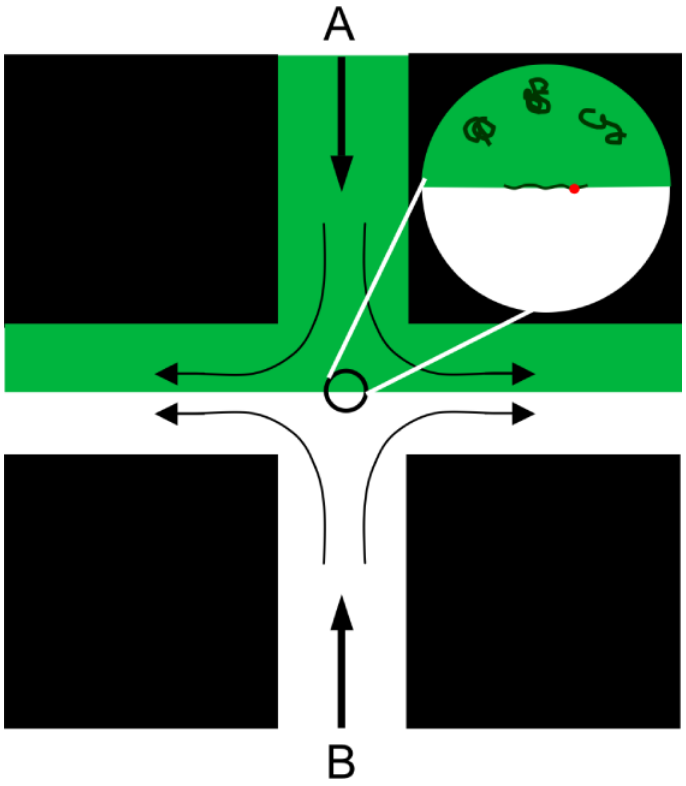

Figure 1.
Schematic of the cross-slot flow device for single-molecule flow experiments. The green area is occupied by fluid from the top inlet arm (A) and the white area is occupied by fluid from the bottom inlet arm (B). An extensional stagnation point flow is created at the center of the device. Inset: an enlarged view of the stagnation point where a labeled, stained DNA molecule is trapped and stretched.

Figure 2.
Optical tweezer manipulation of the twPNA-DNA-dig complex, which is attached by one end to an anti-digoxigenin-coated sphere, manipulated by the optical trap, and by the other end to a streptavidin-coated sphere, sucked onto a micropipette.

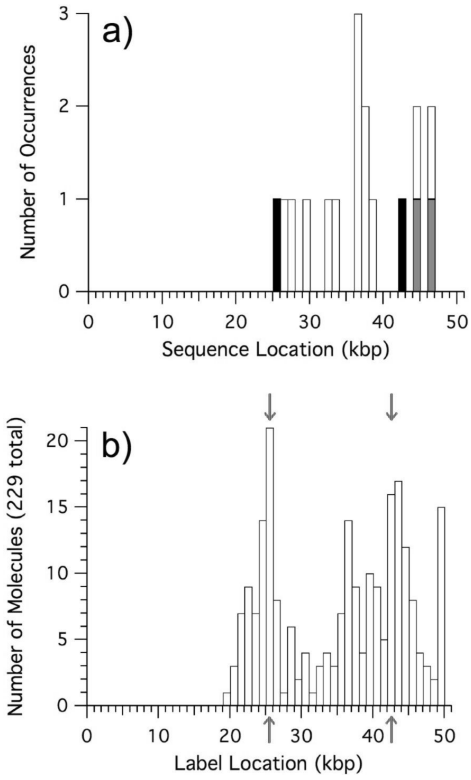

Figure 3.

The predictive binding map (a) indicates the locations of target, SEMM, and DEMM sequences for flPNA on λ-DNA with black, gray, and white bars, respectively. A histogram of label locations (b) determined via slide-stretching (229 molecules) reveals that flPNA-DNA complexes formed at non-optimal binding conditions exhibit non-specific binding. Arrows point to locations of target sequences.

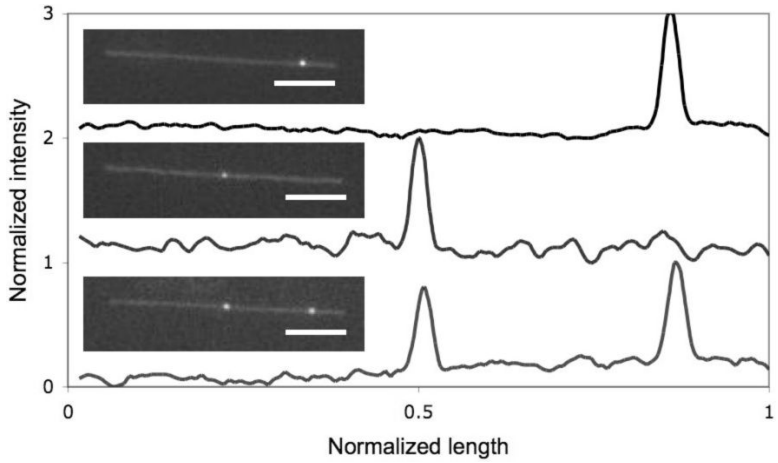

Figure 4.

Intensity profiles of images (inset) of stained, labeled PNA-DNA complexes stretched in the cross-slot flow device labeled at each of the flPNA target binding locations. The DNA molecules are extended to 94% of their contour length. Plots are offset for clarity. All scale bars are 5μm.

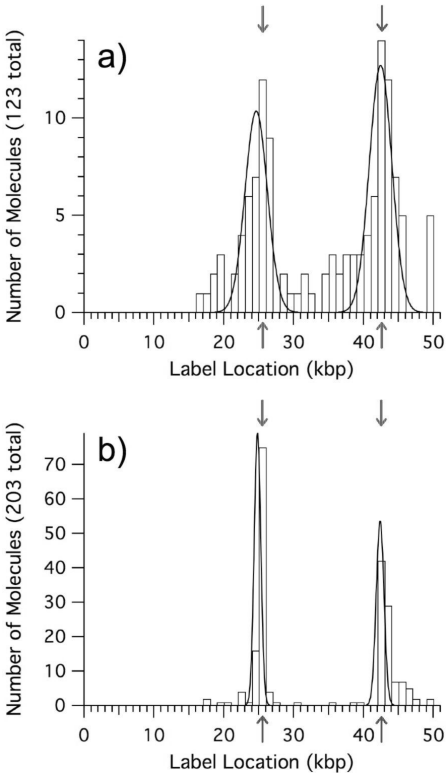

Figure 5.

Histograms of label locations on flPNA-DNA complexes formed at optimal binding conditions determined via (a) slide-stretching (123 molecules) and (b) flow-stretching (203 molecules) where the average DNA extension is 96% of the DNA contour length. Arrows point to locations of target sequences.

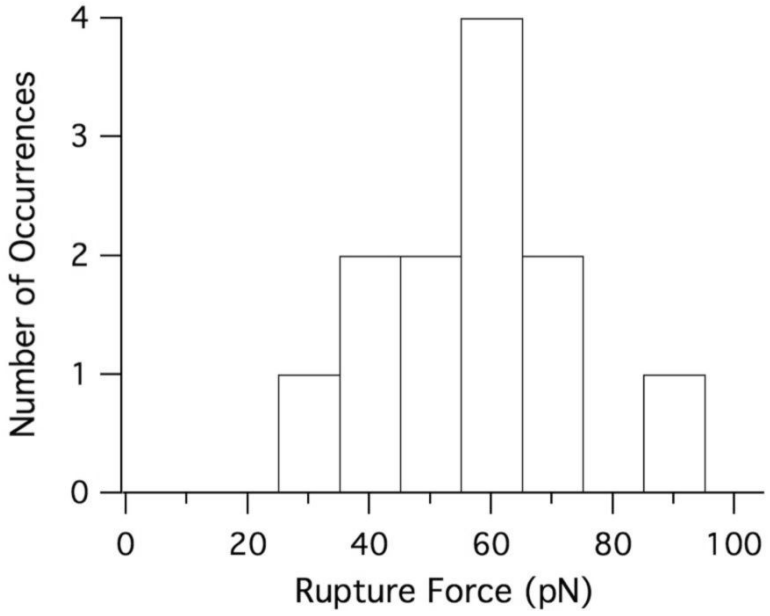

Figure 6.
Optical tweezer manipulation of the twPNA-DNA-dig complex produces a distribution of rupture forces for those complexes that did not sustain the maximum applied force of 90 pN.