Chapter 4

Targeting the Nucleosome Core Particle with Polyamide Dimers

The work described in this chapter was done in collaboration with Karolyn Luger (Colorado State University), Joel Gottesfeld (The Scripps Research Institute), and David Chenoweth.

Abstract

The Nucleosome Core Particle (NCP) is the minimal unit in eukaryotic chromatin and a unique target for DNA-binding polyamides. The ability to modulate the stability of nucleosomes would be a useful tool for molecular biology and gene regulation. A series of polyamide clamps designed to differentiate between the NCP and B-form DNA were synthesized and analyzed using DNase I footprinting. In addition, a new polyamide dimer designed to bind to and stabilize the NCP was synthesized. Finally, the NCP was used to perform the DNA-templated ligation of alkyne- and azide-functionalized polyamides.

4.1 Introduction

In order for DNA-binding molecules to be useful tools for gene regulation, they must be able to access DNA in biological systems. DNA in eukaryotic organisms exists in a highly condensed form as chromatin. The nucleosome core particle (NCP) represents the most basic unit of the higher order structure chromatin. The NCP consists of 147 bp of DNA wrapped twice around an octamer of histone proteins. The histone octamer contains two copies each of histones H2A, H2B, H3, and H4. One of the most interesting features of the NCP is the alignment between the major and minor-grooves of the two gyres of DNA. These aligned minor-grooves separated by a small gap between the gyres of DNA create a "super groove", consisting of 14-16 bp of DNA that is accessible for recognition. Each of these supergrooves bring sequence elements that are 80 bp apart in the linear DNA strand into close spatial proximity. As a result, supergrooves create a recognition platform that exists solely in the context of the NCP.

The ability of polyamides to bind to B-form DNA has been well characterized.[1] As discussed in the previous two chapters, polyamides are also capable of binding to the DNA present in 2-dimensional nanostructures.[2, 3] The NCP is similar to these structures in that it creates a nanometer scale DNA architecture. Initial biochemical studies showed that polyamides are capable of binding to sites on the NCP facing away from or even partially towards the histone proteins.[4] High resolution crystal structures were also obtained, showing polyamides in complex with the NCP.[5]

It was later shown that by linking two polyamides together as a dimer, with a linker of sufficient length to extend over the gap, one can specifically bind to a super groove on the NCP.[6] Interestingly, the polyamide dimer was shown to greatly improve

the stability of the NCP, presumably by preventing unraveling and dissociation of the DNA from the histone octamer. As a result, the linked polyamide dimer has been referred to as a "clamp" because of its ability to effectively clamp the DNA together around the histone core.

The ability to stabilize the NCP could prove useful for a variety of purposes. Since polyamide clamps can stabilize the NCP *in vitro*, they may also be able to stabilize chromatin in cells, a feat which would have important ramifications for gene regulation. Additionally, it was observed that incubation with the polyamide clamps led to the growth of larger crystals in the previous X-Ray crystallography studies. This is presumed to a result of the increased stability of the NCP. As such, polyamide clamps could prove to be extremely useful tools for generating high quality crystal structures of the NCP. Since the original clamp site is located opposite the ends of the DNA it still allows partial dissociation of the DNA from the protein core. We hypothesized that it would be possible to further increase the stability of the NCP by creating a new polyamide clamp targeted to a supergroove located closer to the ends of the DNA. This new "end clamp" could potentially prevent partial dissociation and result in further stabilization of the NCP.

In addition, we initially sought to determine whether one could design a clamp that would be selective for the NCP versus regular B-form DNA. The ability to specifically target DNA solely when in the NCP could potentially be used as a tool for either detecting or stabilizing chromatin, providing a useful tool for modifying gene expression. The key design element was to take advantage of the gap between the gyres of DNA in a supergroove which would not be present in B-form DNA. This gap should be able to accommodate increases in the size of the linker that connects the two

polyamides, unlike the minor groove which would create a steric clash with larger linkers presumably leading to abrogated binding.

Finally, linked dimers provide an interesting motif for targeting extended DNA sequences. A standard eight-ring polyamide has specificity for 6 bases while linked polyamides can target sequences up to 10 bp in length.[7-9] For future gene regulation projects, it may be desirable to target longer sequences to prevent binding at other undesired locations in the genome. However, the uptake properties of these larger conjugates are generally thought to be worse than eight-ring polyamides due to their larger size. A potential solution which has been suggested to this problem is to generate the larger polyamide dimers inside the cell. To this end, prior work has demonstrated how DNA-templated ligation chemistry can be used to link polyamides once they are bound in close proximity in the minor-groove of DNA.[10] Having demonstrated the ability for both halves of a polyamide dimer to bind to adjacent minor-grooves in the supergroove of the NCP, we postulated that it may be possible to use the super groove as a unique template for DNA-templated ligation. This reaction should be specific to the NCP as the two half sites are separated by 80 bp on the DNA strand and are only brought into close spatial proximity when wrapped around the histone octamer.

4.2 Results and Discussion

Engineering a Clamp specific for the NCP versus B-form DNA

The original NCP clamp contains two identical polyamides linked by a short PEG linker. Using this as a starting point, a small library of homodimer polyamide clamps was synthesized using the synthetic scheme shown in Figure 4.1. The advantage of this

scheme is that it is amenable to the use of any diacid as a linker and requires just a single coupling step followed by HPLC purification. Therefore, a small library of diacids can be rapidly converted to polyamide dimers. Several diacids, selected on the basis of their presumed rigidity, appropriateness of length, and steric bulk were chosen and the corresponding dimers were synthesized. The small library is shown in Figure 4.2.

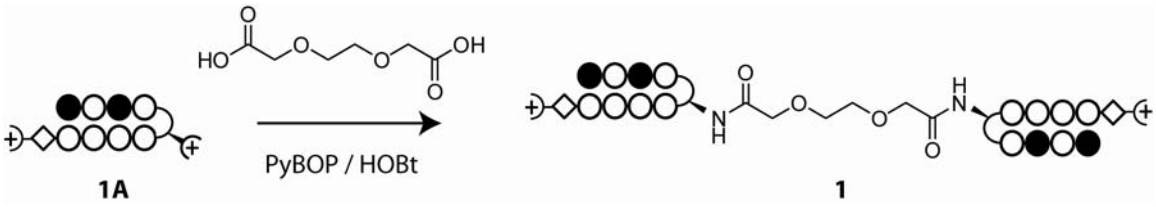


Figure 4.1 Synthetic scheme for the synthesis of polyamide homodimers. The synthesis of compound **1** is shown. The synthesis can be readily adapted for the use of any diacid linker.

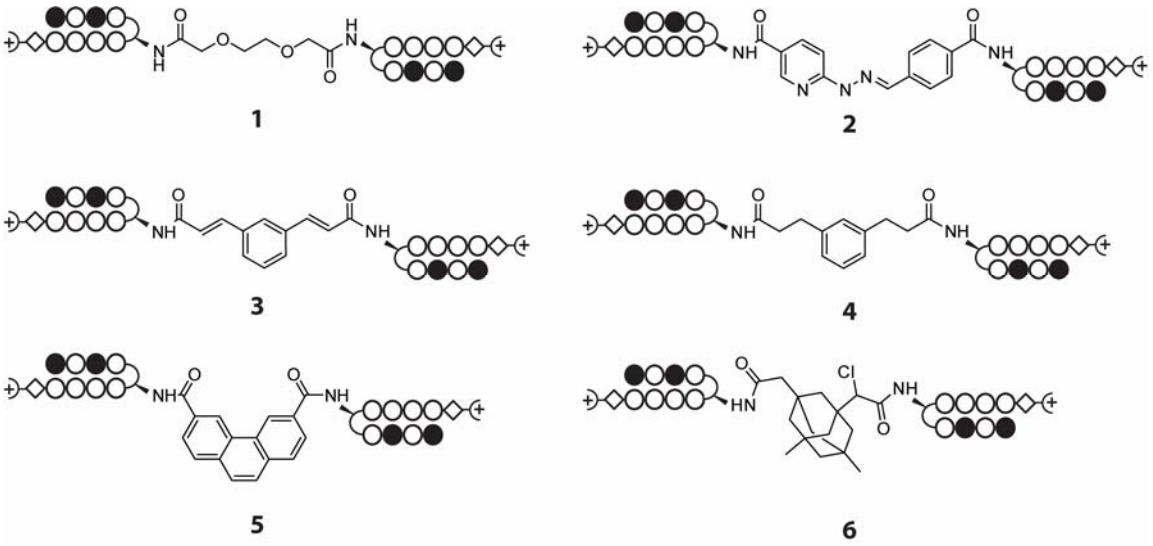


Figure 4.2 Small library of polyamide clamps containing sterically bulky linkers.

DNAse I footprinting was performed to determine whether the library of clamps retained the ability to bind to B-form DNA. The DNA sequence that they were footprinted on was designed to contain three clamp binding sites. Each dimer site consists of two hairpin binding sites separated by 0, 1, or 2 bps respectively. With the exception

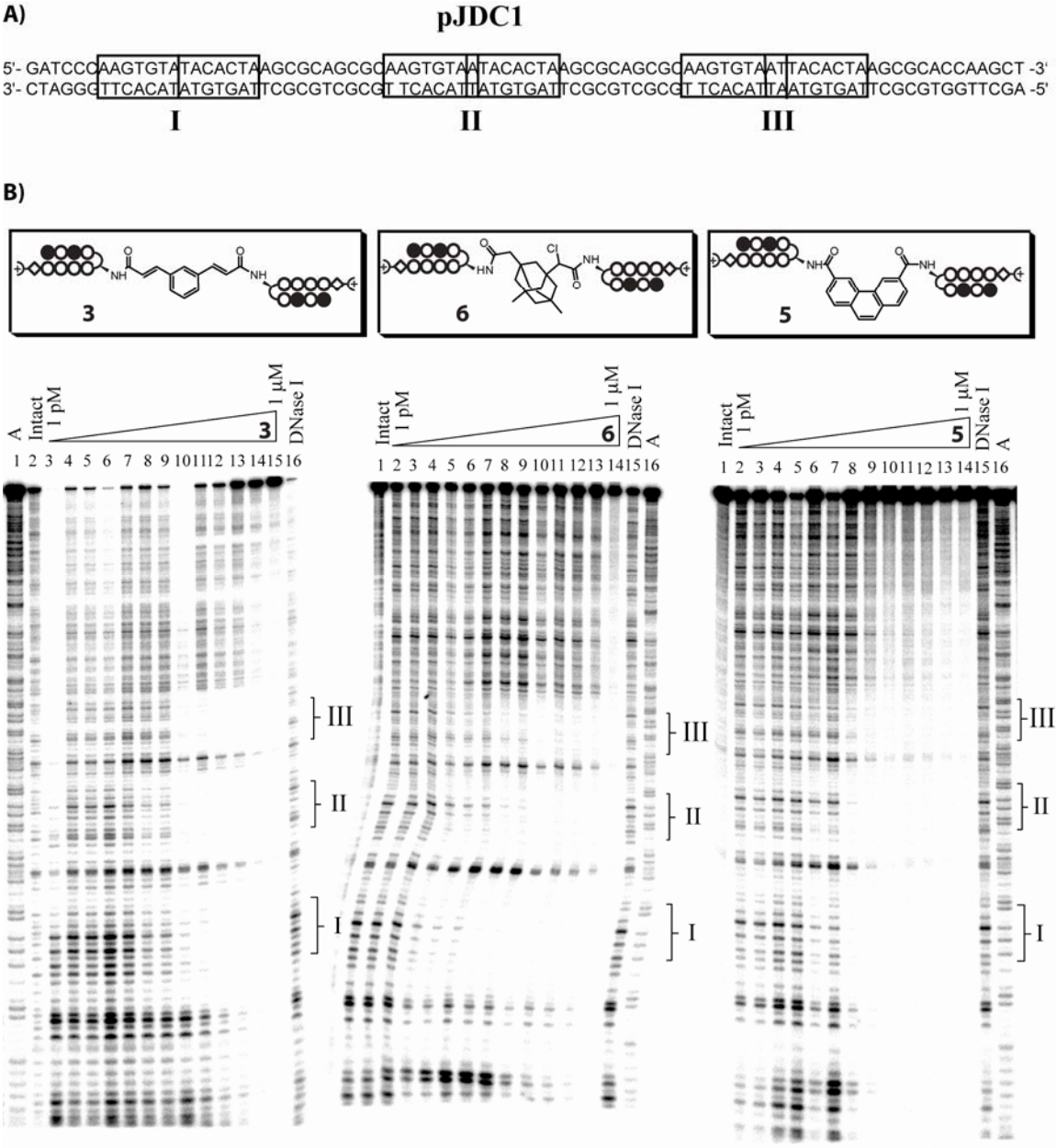


Figure 4.3 DNAse I footprinting experiments with proposed NCP-specific clamps. A) Sequence of the insert from plasmid pJDC1 used for footprinting. The insert contains three binding sites with 0, 1, and 2 bp spacing between the two half sites. B) Representative gels are shown for polyamide dimers **3**, **5**, and **6**. As shown, both **3** and **6** show binding at all three target sites. Polyamide dimer **5** shows non-specific DNA binding.

of **5**, all clamps that were tested showed good binding affinity for all three sites. **5** showed high affinity binding with almost no selectivity across the entire insert, consistent with nonselective DNA coating (Figure 4.3B).

The ability of all library members to bind to DNA was unexpected, as it was presumed that the use of larger linkers, especially in compounds **5** and **6** would prevent binding to B-form DNA for steric reasons. One potential explanation for the observations is that only one half of the dimer is binding specifically while the other half is prevented from interacting with the DNA by the linker. Because we are seeing a statistical mixture of fragments, this binding mode would be indistinguishable by standard DNAse I footprinting. In the original work done by Phillip Weyermann characterizing turn-to-turn hairpin dimers,[9] it was found that double mismatches were tolerated if both mismatches occurred under just one polyamide, whereas a mismatch under both polyamides of the dimer created a 10 fold decrease in binding affinity. This is consistent with the proposed mode of binding that we observed.

Following the work done using a small library approach, we attempted to use rational design to approach the problem of finding an appropriate linker that would give us the desired specificity between the NCP and B-form DNA. Going back to the original PEG linker in **1**, it was hypothesized that by building outwards while keeping the same basic structure and length, it would be possible to perform a more detailed investigation of the role of the linker while eliminating the clamp's ability to bind B-form DNA.

By adding methyl groups to the polyethylene glycol linker, one can slowly add steric bulk while keeping the same linker distance. The synthetic scheme shown in Figure 4.4, which had previously been described was used to synthesize the diacids.[11]

Conveniently this synthetic scheme allows for the use of any diol as a starting material and numerous potential linkers can be synthesized with relative ease.

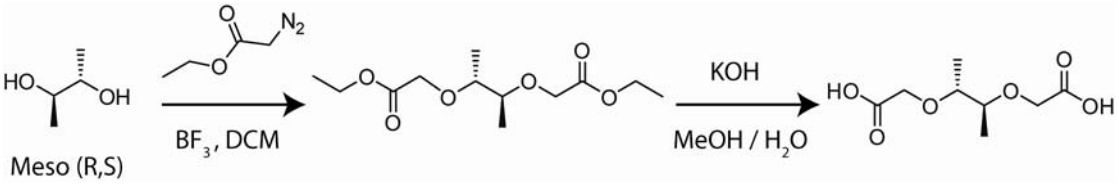
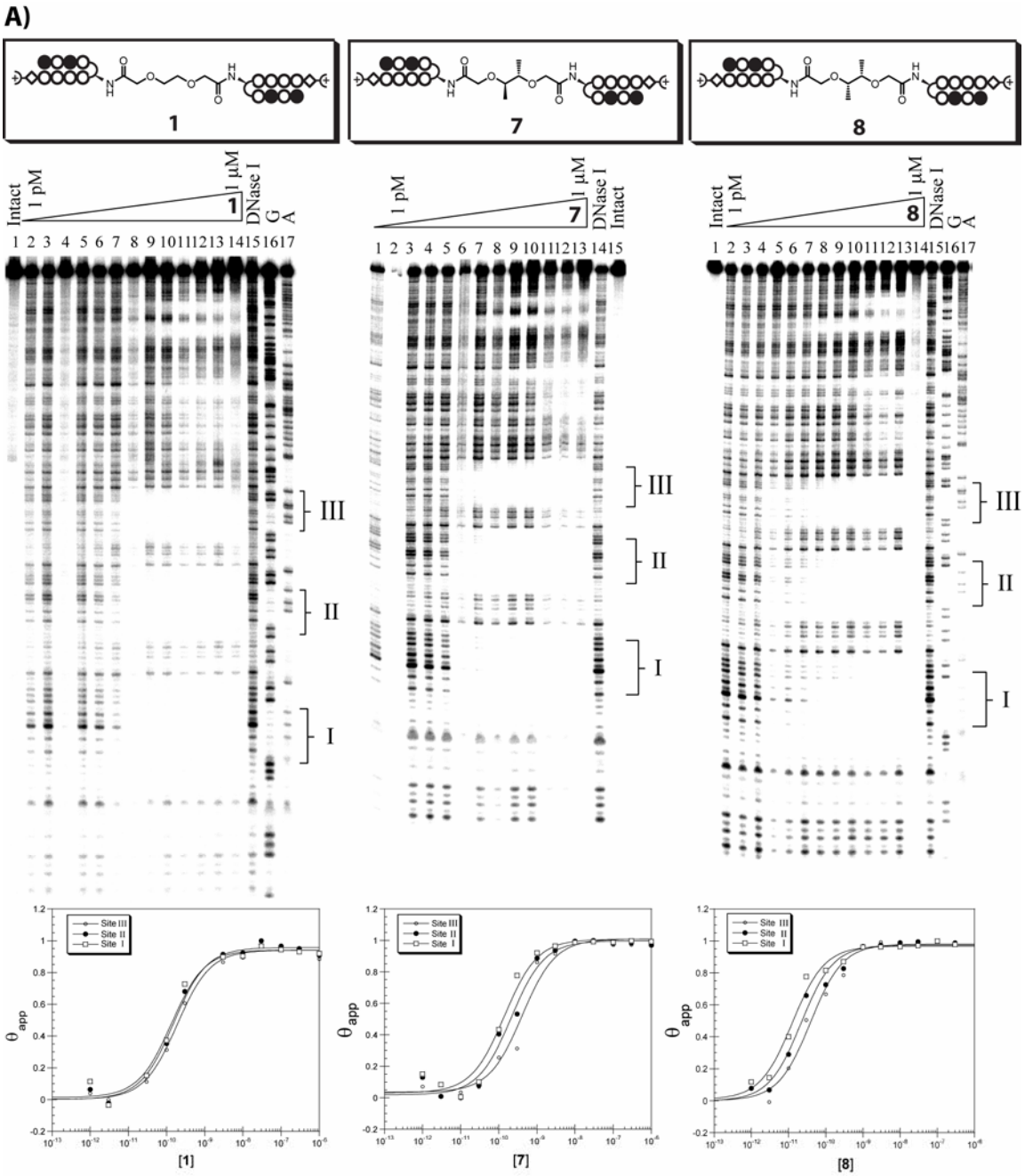


Figure 4.4 Synthetic scheme for synthesizing PEG diacids from starting diols.[11]

In order to investigate the role of the stereochemistry of the methyl groups, both the meso and R,R diols were used to make the corresponding diacids. Both methylated PEG diacids were then used to make clamps **7** and **8**. Quantitative DNAse footprinting of was performed using the previously described DNA sequence and the K_a values were determined at all three binding sites. The data for clamps **7** and **8** is shown next to the results using the original clamp **1** in Figure 4.5. As shown, the addition of the methyl groups did not interfere with binding, providing further evidence that the linker was not sequestered in the floor of the minor-groove.

Based upon these results, several new experiments to further investigate what effect the linker has upon clamp binding were performed. In order to dissect the interactions that were occurring, control compound **9**, the polyamide with an acetylated turn, was synthesized. Similarly, the parent polyamide was coupled to the PEG linker and then capped using methylamine to make control compound **10**. Finally, to investigate whether the clamp was truly binding as a dimer, or with only one polyamide at a time, a new plasmid with a modified insert was constructed. Plasmid pJDC2 was designed so that the hairpin binding sites for the clamp are separated by 0, 3, and 4 bps. The 0 bp



B)

Equilibrium Association Constants, K_a (M^{-1})

Compound	Site I 0 space	Site II 1 space	Site III 2 space
1	$7.5(\pm1.6)\times10^{9}$	$6.6\ (\pm1.0)\times10^{9}$	$5.3\ (\pm.8)\times10^{9}$
7	$7.2\ (\pm1.6)\times10^{9}$	$4.8\ (\pm1.0)\times10^{9}$	$2.5\ (\pm.6)\times10^{9}$
8	$7.9\ (\pm1.7)\times10^{10}$	$4.6\ (\pm1.0)\times10^{8}$	$2.6\ (\pm.5)\times10^{10}$

Figure 4.5 DNase I footprinting of compounds **1**, **7**, and **8**. (A) The structure of each clamp is shown along with the DNAse I footprinting gel. Footprinting was performing using plasmid pJDC1 whose sequence is

shown in Figure 4.3A. The binding isotherms for each compound for each of the three binding sites are shown at bottom. (B) Table showing equilibrium association constant for each compound at each of the three sites.

site was retained as a control to be able to compare results between the two plasmids. Based upon calculations and computational modeling it should be impossible for the clamp to bind as a dimer at the 3 and 4 bp sites because the linker cannot extend over a 3 bp gap in binding sites. For binding to occur at these sites, at least one of the polyamides would have to be localized over a mismatch sequence.

DNAse footprinting was done on the newly designed plasmid insert using compounds **1**, **9**, and **10**. The results are shown in Figure 4.6. As shown, the affinities of the original clamp **1** and compound **10** which has just a single polyamide attached to the linker, are nearly indistinguishable. This indicates that the second polyamide present in **1** does not appear to contribute to binding. It is interesting that **9,** the only compound that does not contain the PEG linker has higher affinities than either **1** or **10**. This indicates that the PEG linker may at least partially interfere with binding. Finally, the similar affinities of compound **1** across all three designed binding sites, despite the fact that the linker is not long enough to cross the 3 and 4 bp gap between the half sites, indicates that the dimer is likely binding in the previously discussed mode where only one of the polyamides is bound sequence specifically onto its match site and the other is bound non-specifically over a mismatch site adjacent to it.

Although unsuccessful in creating a dimer capable of discriminating between the NCP and B-form DNA, a large amount of information regarding the binding mode of dimeric polyamides was gained through this study. As demonstrated, polyamide turn-to-turn dimers are capable of binding to a match site for one of their component polyamides,

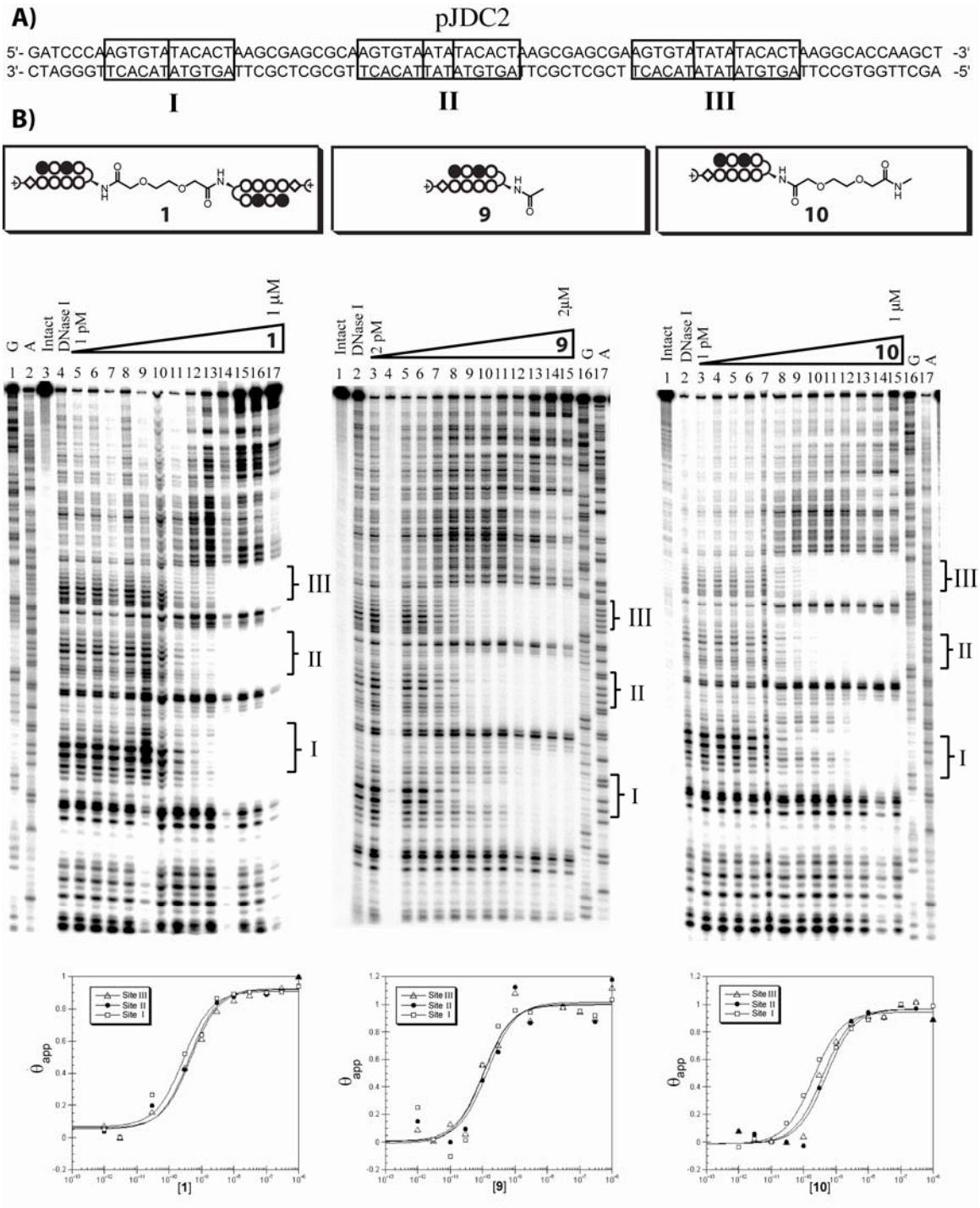


Figure 4.6 DNAse footprinting of compounds **1**, **9**, and **10**. A) The sequence of plasmid Insert pJDC2 containing three binding sites labeled I, II, and III which contain 0, 3, and 4 bp respectively between polyamide binding sites. B) Structures of polyamides **1**, **9**, and **10**. Representative footprinting gels and binding isotherms for each of the three polyamides are shown below.

Table 4.1 Equilibrium association constants, K_a (M^{-1}) on plasmids pJDC1 and pJDC2.

	Plasmid pJDC1			Plasmid pJDC2		
Polyamide	0 bp	1 bp	2 bp	0 bp	3 bp	4 bp
7	$7.2 (\pm 1.6)\times 10^{9}$	$4.8 (\pm 1.0)\times 10^{9}$	$2.5 (\pm .6)\times 10^{9}$	N.D.	N.D.	N.D.
8	$7.9 (\pm 1.7)\times 10^{10}$	$4.6 (+1.0)\times 10^{10}$	$2.6 (\pm .5)\times 10^{10}$	N.D.	N.D.	N.D.
1	$7.5(\pm 1.6)\times 10^{9}$	$6.6 (\pm 1.0)\times 10^{9}$	$5.3 (\pm .8)\times 10^{9}$	$4.1(\pm .9)\times 10^{9}$	$2.9 (\pm .6)\times 10^{9}$	$2.5 (\pm .4)\times 10^{9}$
9	N.D.	N.D.	N.D.	$8.7(\pm 2.1)\times 10^{9}$	$8.1 (\pm 1.9)\times 10^{9}$	$1.0 (\pm .2)\times 10^{10}$
10	N.D.	N.D.	N.D.	$5.4(\pm 1.3)\times 10^{9}$	$2.5 (\pm .8)\times 10^{9}$	$2.5 (\pm .8)\times 10^{9}$

despite the presence of a mismatch sequence under the second polyamide. This is in agreement with previous reports.[9] This could have important implications for the design of future clamp molecules as well has for the design of polyamide motifs that can bind to increasingly longer sequences. Additionally, the loss in affinity upon incorporation of the PEG linker indicates that the currently used PEG linker may be interfering with binding. Additional studies are needed to determine whether a linker that does not negatively affect binding can be developed.

Creating a NCP End Clamp

The original NCP-binding clamp **1** targets a symmetrical "supergroove" and as such consists of two copies of the same polyamide joined by a short linker. To create a clamp that could further stabilize the NCP, we first examined the two sites that make up the super groove which is located closest to the ends of the DNA. Unlike the supergroove targeted in the previous studies the supergoove at this location contains two distinct half sites, making it necessary to synthesize heterodimer clamps containing two different polyamides.

To determine which polyamides were optimal for targeting the two half sites the five polyamide-EDTA conjugates shown in Figure 4.7 were synthesized, building upon previous knowledge about targeting the selected DNA sequences. The 5'-TCCACCT-3' site is designed to be targeted by polyamide **15**. As shown previously, a 2-β-2 motif using ImIm-β-ImIm without a chiral turn binds duplex DNA with a K_a of 7.6 x 10^9 M^{-1}.[12] Interestingly, affinity cleavage experiments showed that the polyamide preferentially binds in a reverse orientation, aligning C to N along the 5' to 3' strand of DNA.

Polyamide **12** which targets the 5'-TGATGGA-3' sequence has been previously characterized and was shown to have a K_a of 3.0 x 10^8 M^{-1} and also to prefer binding in the reverse orientation when it had an acetylated chiral turn.[13] In addition, polyamide **13** is a characterized forward-binding polyamide that has a K_a of 8.6 x 10^9 M^{-1} when the turn is acetylated.[13] In addition to these, an additional polyamide **14** was designed based on the hypothesis that polyamides too far back along the sequence may sterically clash with the histone proteins. This polyamide was designed to target a region slightly more exposed on the NCP. Polyamide **14** incorporates 3-methoxythiophene in the n-terminal position which has been shown to have improved affinity for the T·A base pair,[14] and was used to target the 5'-TTTGAT-3' sequence. Finally, as a positive control for the affinity cleavage studies, polyamide **11** was used as it was well-behaved and previously characterized in this assay.[4] These five polyamide-EDTA conjugates were synthesized and affinity cleavage experiments were done to examine the ability of each compound to bind to the NCP.

The NCP was reconstituted using previously established protocols.[4] Samples of reconstituted histone octamer as well as the 146 bp DNA were obtained from Karolin

Luger at Colorado State University. In general, reconstitution occurred in greater than 90% purity. Affinity Cleavage experiments were done at The Scripps Research Institute with the assistance of Joel Gottesfeld following established protocols.[4] The expected binding sites as well as weak mismatch sites are shown in Figure 4.7. All five EDTA compounds were tested for cleavage on both the NCP as well as the linear 146 bp DNA.

As shown in Figure 4.8, polyamide **11** showed very little binding to the NCP in contrast to previous reports.[4] As a result we cannot rule out the possibility of false negatives in our results. The cleavage observed with **11** in the DNA lane would indicate

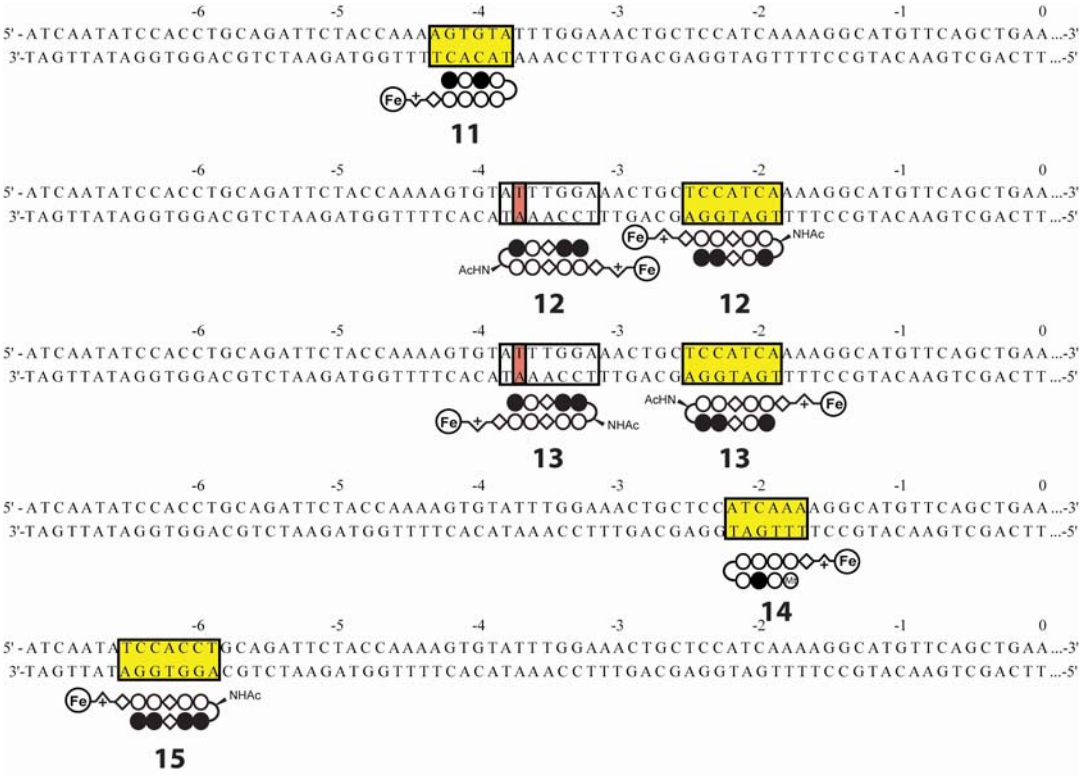


Figure 4.7 Binding sites for polyamide-EDTA conjugates **11 - 15** on the NCP. Half of the palindromic 146 bp sequence is shown. Match sites for each polyamide are boxed in yellow. A weak mismatch site for polyamides **12** and **13** is boxed and the mismatched base pair highlighted in red. Polyamide **15** targets the first half-site of the supergroove while **12 – 14** target the second half-site. Polyamide **11** is included as a positive control for affinity cleavage. Mt = 3-methoxythiophene.

that the polyamide has not degraded. No conclusions were made regarding compound **15** as the location of the presumed binding site is located too close to the end of the DNA to be visible in our experiments. Polyamide **14** showed little cleavage and presumably did not bind well in these experiments. Cleavage was observed for **12** and **13** at both match site IV and mismatch site III although the amount of cleavage appears to be much weaker

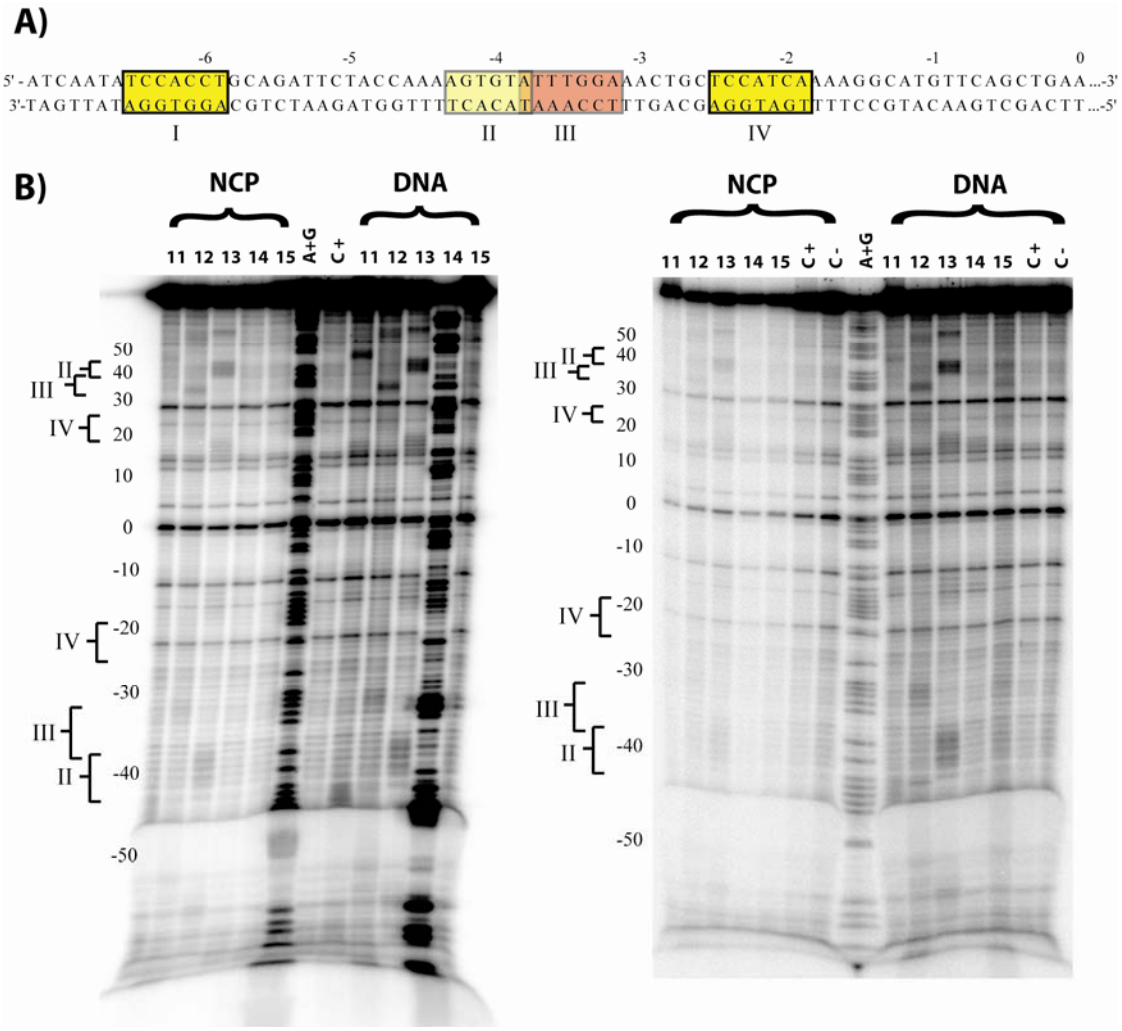


Figure 4.8 Affinity cleavage on the NCP. A)The 146 bp palindromic DNA strand is shown with the polyamide binding sites shown in Figure 4.9 boxed in yellow. The mismatch binding site for **12** and **13** is boxed in red. B) The results of two affinity cleavage experiments that were done on the NCP and the linear 146 bp DNA with compounds **11 – 15**. 50 nM of each polyamide was used. Lanes denoted A+G are for sequencing. Lanes marked C+ had ferrous ammonium sulfate, used to initiate the cleavage reaction, added to them although they did not contain any polyamide, while lanes marked C- did not contain polyamide or the ferrous ammonium sulfate solution.

for the NCP in comparison to linear DNA. **12** binds predominantly at location III while **13** shows cleavage at both location III and IV. The amount of cleavage observed also appears to be strongest for compound **13** which is consistent with its parent having a 20-fold higher affinity than the parent of **12** as discussed previously. As such, our design efforts for an end clamp focused on this compound.

Given these results, we wished to further characterize binding at both of the half sites before making a potential end clamp. A series of melting temperature experiments were performed to determine the ability of polyamides **16** - **19**, the parent compounds of EDTA conjugates **15** and **11**, to bind to their respective half sites. A series of 13 bp duplexes were designed based upon the sequences found in the NCP. As a mismatch

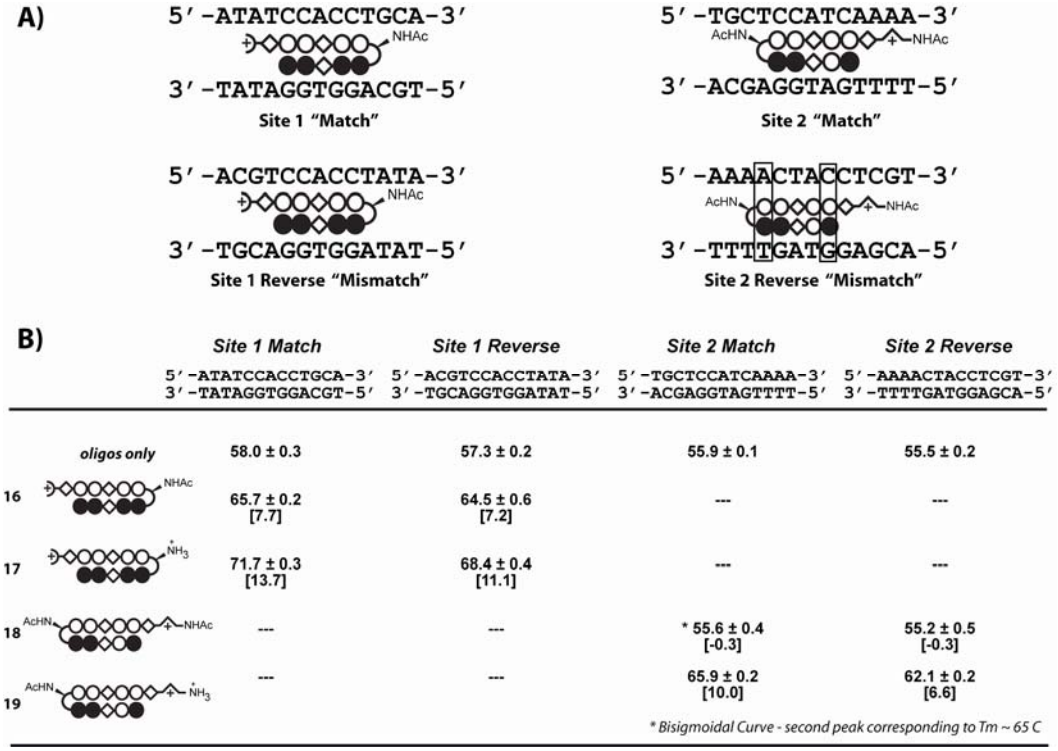


Figure 4.9 Melting temperature analysis for end clamp polyamides. A) The design of the match and mismatch DNA duplexes is shown. The most likely binding modes for the mismatch duplexes which contain the reversed DNA sequence are shown. B) Melting temperatures for each of the compounds. The numbers in brackets are the difference in melting temperature observed when compared to duplex only controls.

control, DNA duplexes where the DNA sequence was reversed were used as well (Figure 4.9A). As shown, both polyamides **16** and **17** bind to the site 1 match site. Binding at the site 1 mismatch was also observed due to the symmetric nature of the binding site, which allows the polyamide to bind at the same sequence with only slightly different flanking base pairs. (see Figure 4.9A, Site 1 match and reverse mismatch). As expected the charged compound **17** had a significantly higher melting temperature consistent with the higher binding affinities generally observed with the free amine present on the turn. Compounds **18** and **19** showed similar binding profiles for the Site 2 match and mismatch sequences. Compound **18** had a bisigmoidal curve at the match site which complicated the analysis. However it appears that both of these compounds bind at the match site and have significantly decreased binding affinities at the mismatch sites. This is consistent with the analysis of mismatch binding in Figure 4.9A where either the polyamide must bind with the chiral turn facing into the minor groove or over a double base pair mismatch site, either of which would lead to an extreme decrease in binding. As was seen for compounds **16** and **17**, the free amine on polyamide **19** resulted in a larger increase in melting temperature than the acylated compound **18**. Based upon this analysis, the polyamide cores chosen appear to bind to the sequences present in the supergroove with high affinity.

As a final result of these studies, end-clamp **20** was synthesized as outlined in Figure 4.10. The longer 3-oxa PEG linker was chosen as it should allow the most freedom to bind, and previous studies had shown that increasing linker length had no deleterious effects on clamp binding.[6] HPLC purification was done after each coupling to purify the intermediates. The second coupling went in 51% yield and produced 120 nmol

of product **20** after purification. It should be noted that in the course of this study and similar experiments, yields for these types of couplings after HPLC purification rarely exceed 30% and often times were less than 10%. The relatively low yields made the synthesis of heterodimer **20** and other heterodimers not discussed here extremely challenging. As a result investigations were done to optimize the yield of the coupling steps. Numerous peptide coupling reagents were tested on small scale for use in this reaction. These include DCC, EDC, NHS, HBTU, HATU, TFFH, DMTMM, and DMAP. However, PyBOP was found to be the optimal activation reagent for these reactions. The 120 nmol of **20** was shipped to the Karolyn Luger's group at the University of Colorado for crystallization trials. It is hoped that this clamp will have an improved ability to stabilize nucleosome core particles.

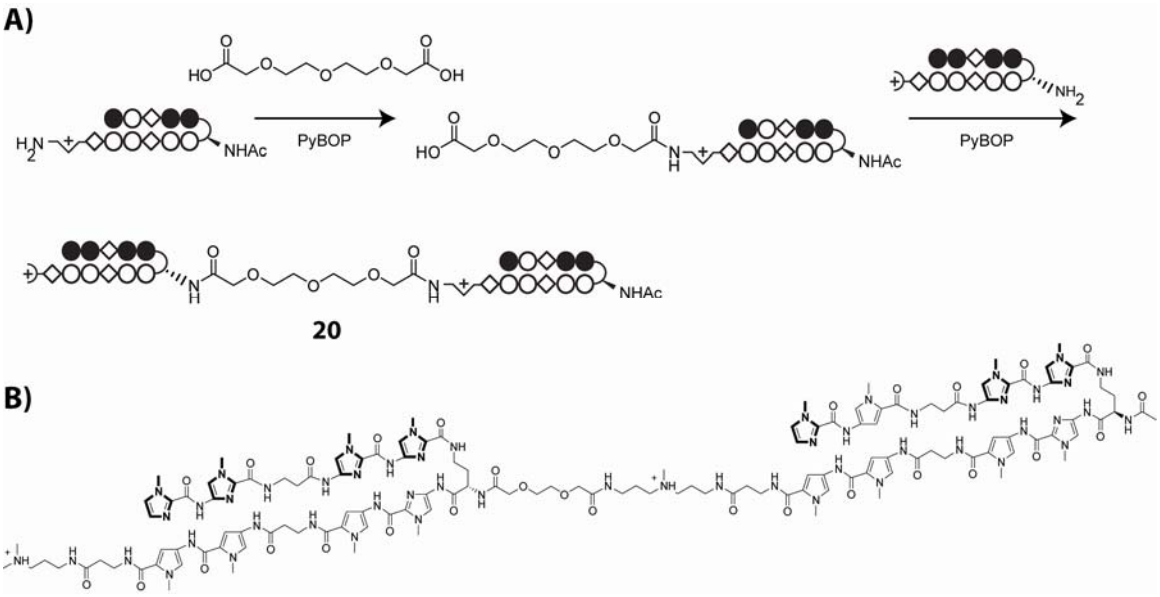


Figure 4.10 Synthesis of End Clamp **20**. A) The heterodimer end clamp was synthesized by two coupling steps. B) Chemical structure of **20**.

NCP Templated Ligation

Previous work has shown how DNA can be used as a template to direct reactions between two functionalized polyamides.[10] Two polyamides targeted to adjacent sites in the minor groove of duplex DNA, one functionalized with an azide and the other with an alkyne were able to undergo a 1,3-dipolar cycloaddition reaction, forming a covalent

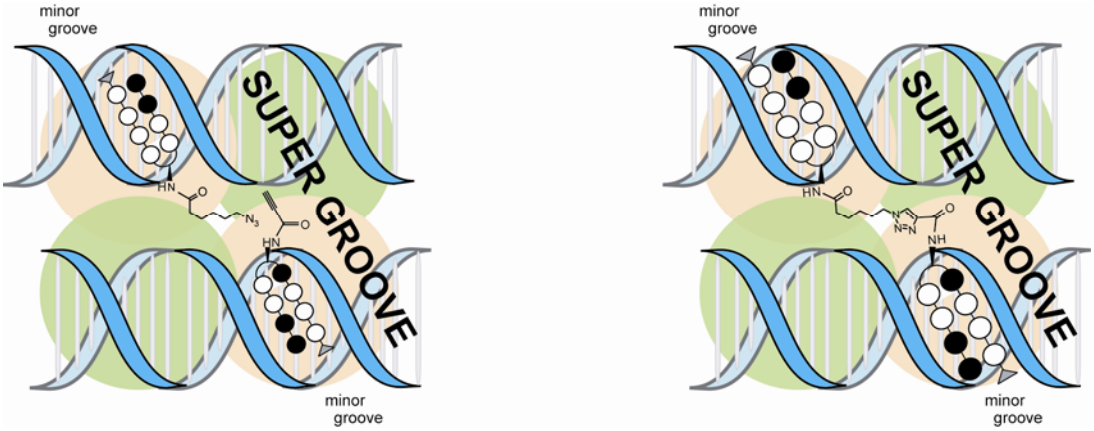


Figure 4.11 Schematic for NCP-mediated ligation. Two polyamides functionalized with an alkyne and azide are able to react in the supergroove of the NCP.

triazole linkage.[10] In this section we examine whether the analogous templated reaction could be performed on the NCP to generate a polyamide dimer.

In order to ascertain whether the NCP could be used to template ligation reactions, a series of three azide-containing polyamides and two alkyne-containing polyamides were synthesized. Several linker lengths were used in order to examine the distance dependence of the reaction. Examination of the crystal structure of the polyamide clamp bound to the NCP,[6, 15] and computational modeling showed that only the azide containing the longest linker was expected to be capable of reacting. In addition, the previous study of DNA ligation had demonstrated that the activated alkyne present in **24** reacts over 100 times faster than alkyl alkyne in **25**.

The NCP was reconstituted as outlined in the previous section. The same ligation reaction was done using all three azides and both of the alkynes. In separate tubes, 200 pmol of each polyamide-alkyne was reacted with 200 pmol of each polyamide-azide and 40 pmol of the reconstituted NCP. The reaction was incubated for 5 hrs at 37°C

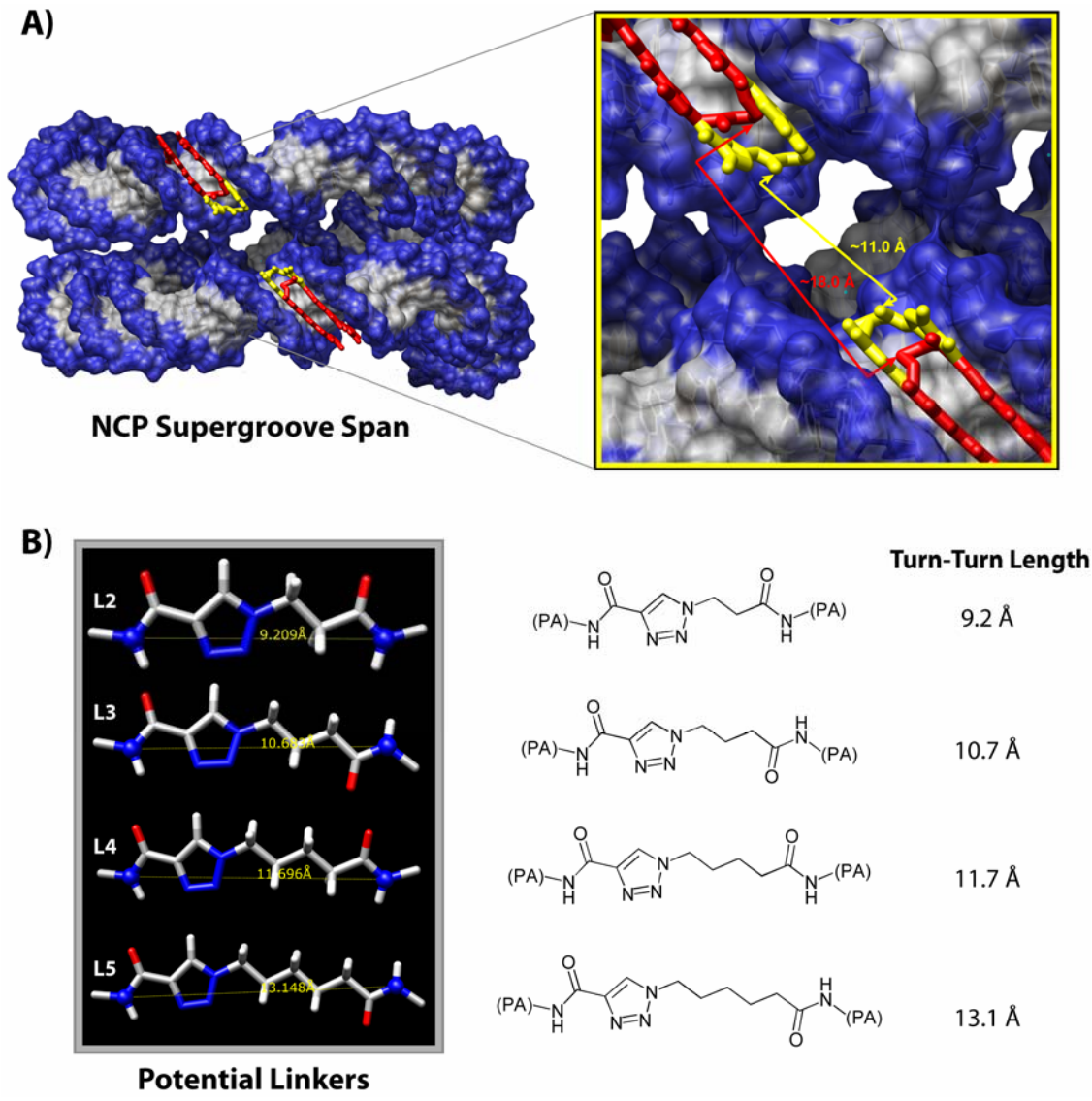


Figure 4.12 Analysis of the linker dependence of NCP templated ligations. A) Crystal structure view of the supergroove. The predicted distance between the amines of the turn are shown in yellow for binding at the original site, and in red if the sites are moved one base pair back. B) Modeling of the linkers and the calculated turn-to-turn length.

followed MALDI-TOF mass spectrometry to analyze the reaction. Of the six potential ligation reactions, we only observed product in the case of compounds **23** and **24**. This agrees with both the computational work that indicated the linkers of compounds **21** and **22** were too short to react, as well as with the prior work that showed that the alkyne used in **25** was significantly less reactive. In the case of **23** and **24** we observed masses at 2664.88 and 2687.77 that correspond to calculated values for the $[M+H]^+$ and $[M+Na]^+$ peaks.

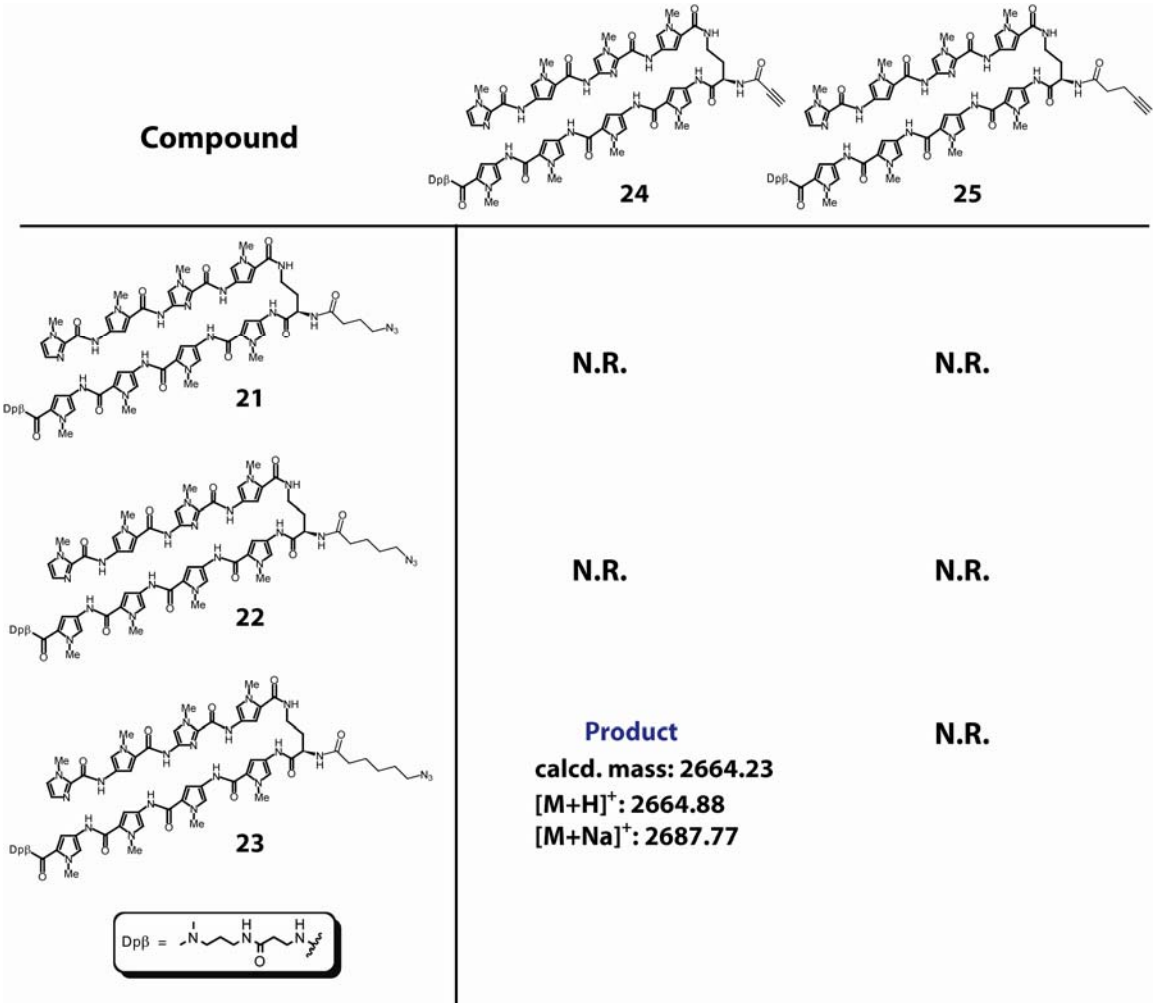


Figure 4.13 NCP-templated ligation between different length azides and alkynes. The reaction product was only detected following reaction with polyamides **23** and **24**. N.R. = no reaction.

We then checked whether control samples run under the same conditions but containing either the DNA, histone octamer, or buffer instead of the NCP would result in product formation. In all three cases the target mass was not observed, indicating that the ligation reaction was dependent upon the presence of the NCP. The lack of product in the presence of just DNA is promising, as the reaction appears to only take place when the

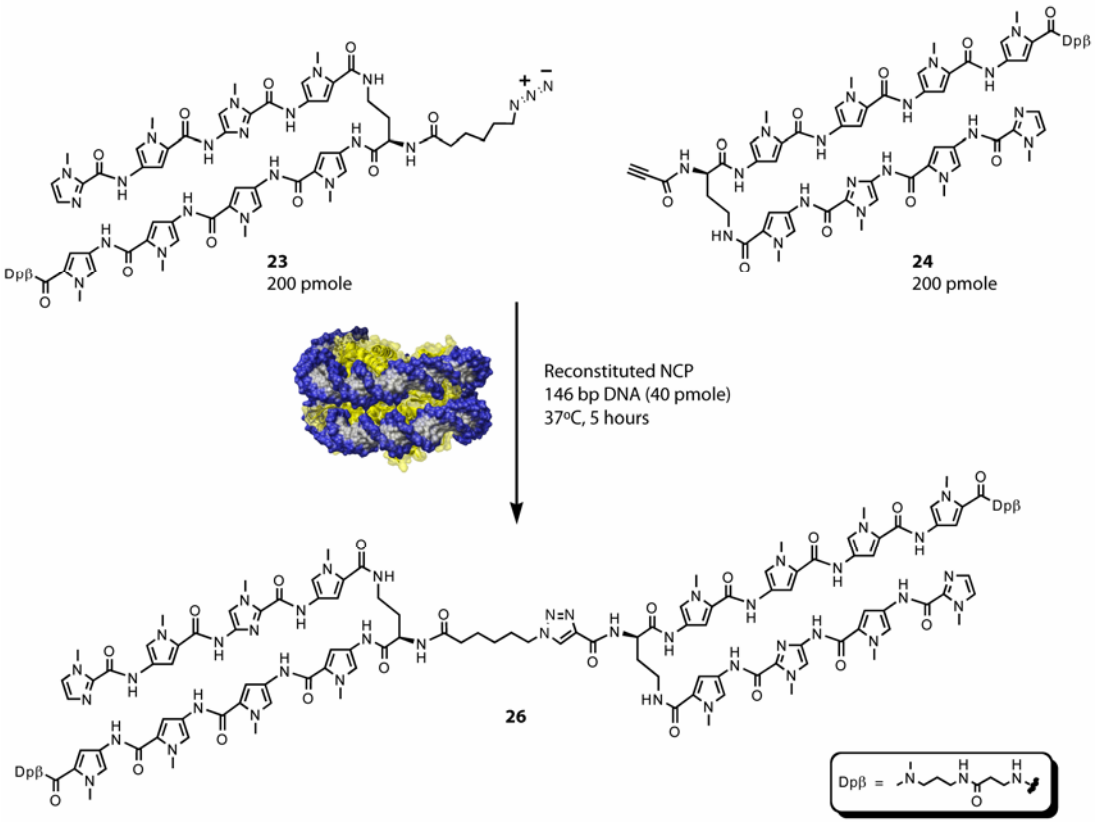


Figure 4.14 DNA templated ligation on the NCP. The reaction of **23** and **24** with the NCP leads to dimer **26** which was observed by MALDI mass spectrometry.

two binding sites which are spatially separated on the linear DNA are brought into close proximity by NCP formation to generate a functional template.

These results demonstrate the ability to perform templated ligation reactions on the NCP. Although we were able to detect product formation by MALDI, the small scale

of our reactions prevented us from quantitating the yields of the reaction. Future studies will be needed to determine the efficiency of the cycloaddition reaction. It should be noted that the polyamides uses for both the azide and alkyne conjugates contained the same parent polyamide core due to the symmetric nature of the targeted supergroove. As a result, the reaction rate of the ligation may be slowed by non-productive binding

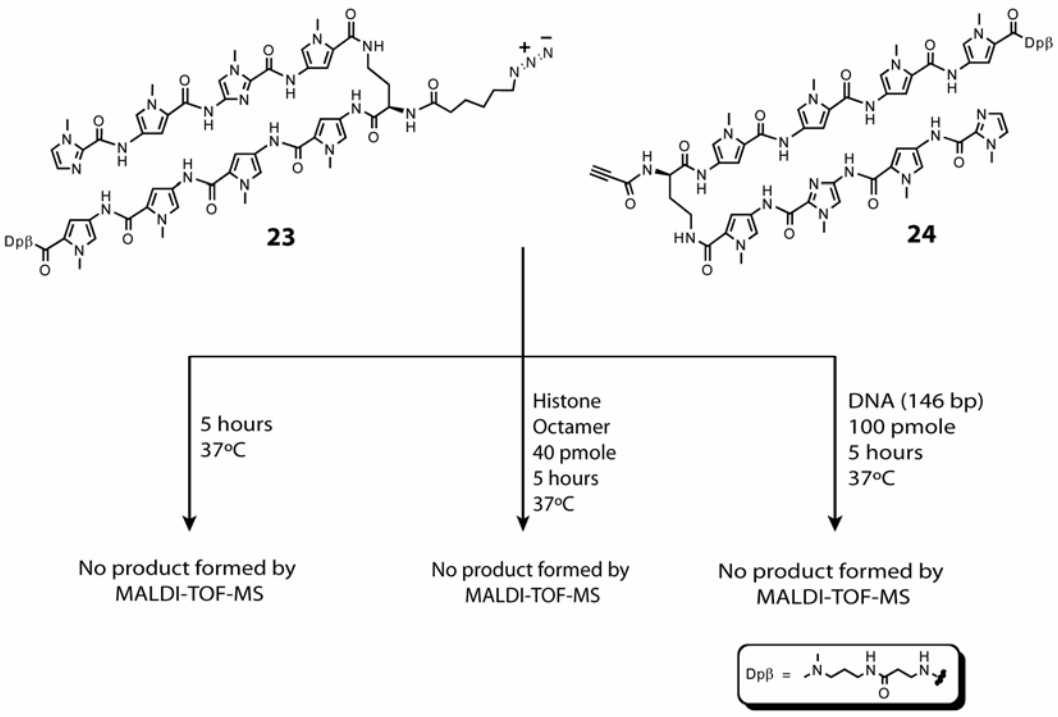


Figure 4.15 Control reactions for the NCP-templated ligation reaction. The ligation reaction does not occur in the presence of only buffer, histone octamer, or the linear 146 bp DNA.

of two alkynes or two azides to a single NCP. Future work should examine the use of an asymmetric supergroove, such as the one targeted in the previous section for the creation of an end clamp, where a unique binding site would be present for both the alkyne and azide conjugates. The use of an asymmetric template should improve the reaction efficiency. Future work using highly sensitive HPLC-based approaches, pro-fluorescent polyamide reagents, or mass spectrometry quantitation of yields will be necessary for

further understanding of the ligation reaction and to determine the feasibility of performing templated reactions in an intracellular setting.

4.3 Conclusions

The use of polyamides as unique molecular tools for performing reactions and manipulating the NCP has been described in detail. The efforts in designing an NCP-specific clamp have led to greater understanding of the binding mode of polyamide dimers. The design and synthesis of a new polyamide dimer end clamp may prove to be a useful tool for X-ray crystallography. Finally, the NCP templated ligation studies have demonstrated the feasibility of this approach for generating linked polyamide dimers. One key question is whether the knowledge gained through these studies can be applied to intracellular systems. The use of polyamide clamps to stabilize the NCP, and the ability to generate these clamps via a templated ligation reaction inside the cell, has exciting potential applications for gene regulation.

4.4 Experimental Details

Materials and General Methods. Dicyclohexylcarbodiimide (DCC), 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), N-Hydroxybenzotriazole (HOBt), Fmoc-Dab(Boc)-OH and Boc-β-Ala-Pam resin were purchased from Peptides International. Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) was purchased from Novabiochem. Flouro-N,N,N’,N’-tetramethylformamidinium hexafluorophosphate (TFFH) was purchased from Advanced ChemTech. O-(7-Azabenzotriazol-1-yl)-N,N,N’,N’,-tetramethyluronium

hexafluorophosphate (HATU), 4-(dimethylamino)-pyridine (DMAP), N-hydroxysuccinimide (NHS), N,N-dimethylformamide (DMF), N-methylpyrrolidinone (NMP), N,N-dimethylpropylamine (Dp), N,N-diisopropylethylamine (DIEA), ethylene diamine, piperidine, and other miscellaneous chemicals were purchased from Aldrich and used without further purification. All other solvents were purchased from EM Sciences and were reagent grade. Trifluoroacetic acid (TFA) was purchased from Halocarbon.

^{1}H NMR spectra were recorded using a 300MHz General Electric-QE NMR spectrometer. $CDCl_3$ was obtained from Cambridge Isotope Laboratories. UV spectra were recorded in water using a Beckman Coulter DU 7400 Spectrophotometer. Matrix-assisted LASER desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was performed using an Applied Biosystems Voyager DE Pro Spectrometer. Electrospray ionization (ESI) mass spectrometry was performed by the Protein and Peptide Microanalytical Facility at the California Institute of Technology. Analytical High-Pressure Liquid Chromatography (HPLC) was performed with a Beckman Gold system using a Varian Microsorb-MV 100 C_{18} column (5 μm particle size, 250 x 4.6mm). Preparative HPLC was performed using a Beckman Gold system with either a Waters Bondapak C_{18} column (15-20 μm particle size, 25 x 100mm) or a Phenomenex Gemini C_{18} column (5 μm particle size, 250 x 21.2 mm). For both HPLC systems Solvent A was 0.1% (v/v) aqueous TFA and solvent B was acetonitrile. Analytical HPLC was done using a gradient of 1.85%/min of Solvent B starting from 0% over 35 min with a flowrate of 1.5 mL/min. Preparative HPLC was typically done using a gradient of 1%/min of Solvent B for 20 min followed by a gradient of .3%/min for an additional 100 min at a

flowrate of 8 mL/min. Radioactive gels were imaged using a Molecular Dynamics 400S PhosphorImager.

Restriction endonucleases, deoxyribonucleotide triphosphates, DNAse I, Polynucleotide kinase (PNK), and glycogen were purchased from Roche. [α-^{32}P]-Thymidine-5'-triphosphate (≥ 3000 Ci/mmol) and [α-^{32}P]-Deoxyadenosine-5'-triphosphate (≥ 6000 Ci/mmol) were purchased from Perkin Elmer. [γ-^{32}P]-Adenosine-5'-triphosphate (≥ 7000 Ci/mmol) was purchased from MP Biomedicals. Water was purified from a Millipore Mill-Q purification system for general use. Ultrapure RNAse/DNAse free water from USB was used for biological work. All buffer reagents used were molecular biology grade. Buffers were sterilized using a Nalgene .2μm cellulose filtration device.

Plasmids pJDC1 and pJDC2 were constructed using 80mer oligonucleotides purchased from Integrated DNA Technologies. pUC19 plasmid was purchased from Sigma. JM109 Competent Cells (>10^8 cfu/μg) were purchased from Promega. A Rapid DNA ligation kit was purchased from Roche. Purification was done using a Promega Wizard *Plus* Midipreps DNA purification system. Sequence analysis was performed by the Sequence Analysis Facility at the California Institute of Technology.

Polyamide Synthesis. Polyamide synthesis was performed as previously reported.[16] All polyamides were synthesized using Boc-β-Ala-PAM resin (~.59 meq/g). The resin was initially swelled in DMF for 5 min in a glass reaction vessel fitted with a glass filter and stopcock. The vessel was drained and the resin washed twice with DCM. Deprotection with 80% TFA:DCM was performed for 20 min while the resin was shaken. Following deprotection, the resin was washed 2 x DCM, 1 x 4:1 DMF:DIEA, and 1 x DMF.

Coupling of the Boc-Py-OBt pre-activated ester was performed using 1.8 eq of monomer in ~1 mL of NMP. Coupling of Boc-Im-OH, Boc-PyIm-OH, and other acids was done by first preactivating 1.8 eq of the monomer with 1.7 eq of HBTU, 5.4 eq of DIEA in ~2mL of NMP. The activation mixture was shaken for 20min before being filtered and added to the resin. Couplings were allowed to proceed for 2hrs except in the case of Im-OH which was allowed to react overnight. After each coupling step the resin was washed 2 x DMF and then 2 x DCM before the next deprotection step. Polyamides were cleaved from resin using 1.5 mL of Dp for 200 mg of resin at 55°C for eight hours. Crude products were purified by preparative HPLC.

1A: MALDI-TOF-MS calculated $[M+H]^+$: 1237.59, found 1237.5

1: MALDI-TOF-MS calculated $[M+H]^+$: 2617.20, observed 2617.3

2: MALDI-TOF-MS calculated $[M+H]^+$: 2724.23, observed 2724.3

3: MALDI-TOF-MS calculated $[M+H]^+$: 2657.21, observed 2657.2

4: MALDI-TOF-MS calculated $[M+H]^+$: 2661.24, observed 2661.1

5: MALDI-TOF-MS calculated $[M+H]^+$: 2705.21, observed 2704.9

6:MALDI-TOF-MS calculated $[M+H]^+$: 2753.28, observed 2753.3

7: MALDI-TOF-MS calculated $[M+H]^+$: 2645.23, observed 2645.2

8: MALDI-TOF-MS calculated $[M+H]^+$: 2645.23, observed 2645.8

9: MALDI-TOF-MS calculated $[M+H]^+$: 1279.60, observed 1279.58

10: MALDI-TOF-MS calculated $[M+H]^+$: 1410.66, observed 1410.62

11: MALDI-TOF-MS calculated $[M+H]^+$: 1539.7, observed 1540.01

12: MALDI-TOF-MS calculated $[M+H]^+$: 1739.8, observed 1740.0

13: MALDI-TOF-MS calculated $[M+H]^+$: 1739.8, observed 1739.8

14: MALDI-TOF-MS calculated $[M+H]^+$:1571.7, observed 1571.89

15: MALDI-TOF-MS calculated $[M+H]^+$: 1740.8, observed 1740.8

16:MALDI-TOF-MS calculated $[M+H]^+$: 1423.66, observed 1423.8

17: MALDI-TOF-MS calculated $[M+H]^+$: 1381.65, observed 1381.86

18: MALDI-TOF-MS calculated $[M+H]^+$: 1507.72, observed 1508.63

19: MALDI-TOF-MS calculated $[M+H]^+$: 1465.71, observed 1465.71

20: MALDI-TOF-MS calculated $[M+H]^+$: 3033.41, observed 3032.9

21: MALDI-TOF-MS calculated $[M+H]^+$: 1348.63, observed 1348.54

22: MALDI-TOF-MS calculated $[M+H]^+$: 1362.65, observed 1362.61

23: MALDI-TOF-MS calculated $[M+H]^+$: 1376.66, observed 1376.73

24: MALDI-TOF-MS calculated $[M+H]^+$: 1289.58, observed 1289.67

25: MALDI-TOF-MS calculated $[M+H]^+$: 1317.61, observed 1317.6

26: MALDI-TOF-MS calculated $[M+H]^+$: 2664.23, observed 2664.88

Synthesis of diacids. The following diacids were synthesized for use in homodimer clamp synthesis.

3,3' – (1,3-Phenylene)bisacrylic Acid The diacid was synthesized as previously described.[17] From 500mg (3.7mmol) of isopthaldehyde was yielded 631mg (78%) of the diacid as an off white solid. ^{1}H NMR (DMSO-d_6) δ 12.4 (s, 1H), 8.05 (s, 1H), 7.7 (d, 2H), 7.6 (d, 2H), 7.41 (t, 1H), 6.61 (d, 2H).

3,3' – (1,3-Phenylene)bispropanoic Acid The diacid was synthesized as previously described.[17, 18] From 400mg (1.83mmol) of the bisacrylic Acid from above was yielded

335mg (82.4%) of the diacid as a white solid. ^{1}H NMR ($CDCl_3$) δ 7.2 (m, 4H), 2.9 (t, 4H), 2.65 (t, 4H). ESI MS m/e: $[M-H]^-$ 221.0 calc'd: 221.08.

(R,S) 4,5-dimethyl-3,6-dioxaoctanedioic diethyl ester The compound was prepared as previously described.[11] From 500mg (5.55mmol) of meso 2,3-butanediol was yielded 1.3g (90%) of the diethyl ester. ^{1}H NMR ($CDCl_3$) δ 4.2 (m, 8H), 3.7 (m, 2H), 1.3 (t, 6H) 1.2 (d, 6H). ESI MS m/e: $[M+H]^+$ 263.0 calc'd: 263.15.

(R,S) 4,5-dimethyl-3,6-dioxaoctanedioic acid The compound was prepared as previously described.[11] From 178mg (.68mmol) of **1B** was yielded 32mg (23%) of the diethyl ester. ^{1}H NMR ($CDCl_3$) δ 4.33 (s, 1H), 4.28 (s, 1H), 4.13 (s, 1H), 4.08 (s, 1H), 3.7 (m, 2H), 1.2 (d, 6H). ESI MS m/e: $[M-H]^-$ 204.8.0 calc'd: 205.07.

(R,R) 4,5-dimethyl-3,6-dioxaoctanedioic diethyl ester The compound was prepared as previously described.[11] From 500mg (5.55mmol) of meso 2,3-butanediol was yielded 1.5g (99%) of the diethyl ester. ^{1}H NMR ($CDCl_3$) δ 4.2 (m, 8H), 3.6 (m, 2H), 1.28 (t, 6H) 1.2 (d, 6H). ESI MS m/e: $[M+Na]^+$ 284.8 calc'd: 285.13.

(R,R) 4,5-dimethyl-3,6-dioxaoctanedioic acid The compound was prepared as previously described.[11] From 847mg (3.2mmol) of **1B** was yielded 600mg (90%) of the diethyl ester. ^{1}H NMR ($CDCl_3$) δ 4.35 (s, 1H), 4.3 (s, 1H), 4.15 (s, 1H), 4.08 (s, 1H), 3.45 (m, 2H), 1.15 (d, 6H).

Synthesis of Polyamide Dimers. All homodimers were synthesized using the same general protocol. Briefly, the diacid was first activated in 200 μL of anhydrous DMF with 5 eq. of PyBOP and 10 eq. of HOBt. 100 μL of DIEA was added, followed by addition of Polyamide **1A** in slight excess. All reactions were shaken for 2 hrs and shown to be

complete by analytical HPLC. Preparative HPLC workup was used to isolate the final product.

Heterodimers were synthesized using a similar protocol. The first polyamide was initially coupled using the same conditions as previously described for homodimers except that the linker diacid was added in >10 fold excess to afford solely the mono-coupled product. Reactions were monitored by analytical HPLC and worked up after 2 hrs using a preparative HPLC. Next, the polyamide with linker attached is taken up in 100 μL anhydrous DMF and reactivated with 5 eq. of PyBOP and 10 eq. HOBt. 100 μL of DIEA is added to the mixture, followed by addition of the second polyamide in slight excess. Reactions were monitored by HPLC and in most cases were allowed to react overnight as the observed rate of product formation was slow in comparison to the initial coupling. In some cases another addition of PyBOP was necessary to get the reaction to proceed. Products were isolated using a preparative HPLC as previously described.

EDTA couplings. Polyamide-EDTA conjugates were prepared by heating 50 mg of EDTA dianhydride in 1 mL of DMSO / NMP for 5 min. Approximately 1.5 umol of polyamide was added with 500 μL of DIEA and the sample heated for 37 min at 60°C. 3 mL of .1 M NaOH was added to the sample followed by heating for an additional 15 min. Samples were then purified by HPLC.

Preparation of 3' ^{32}P-labeled DNA. Plasmid pJDC1 was initially digested using the restriction enzymes *PvuII* and *EcoRI*. The 3' overhang regions were filled in with [α-^{32}P]-dATP and [α-^{32}P]-dTTP using the Klenow fragment of DNA polymerase or

Sequenase. Labeled oligonucleotides were then purified by band extraction from a 7% non-denaturing polyacrylamide gel.

Preparation of 5' ^{32}P-labeled DNA 5' end labeling was accomplished by incubation of a PCR forward primer, PNK, and an aliquot of [γ-^{32}P]-ATP in PNK buffer at 37°C for 30 min. Labeled primer was then used in a PCR reaction to amplify the selected region on either plasmid pJDC1 or pJDC2. The radiolabelled PCR product oligonucleotides were then purified by band extraction from a 7% non-denaturing polyacrylamide gel.

Affinity determination by quantitative DNase I footprinting. Reactions were carried out in a volume of 400 μL in aqueous TKMC buffer according to published protocols.[19] Standard molecular biology techniques were used to insert a 75bp DNA sequence into the *Bam*HI/*Hind*III restriction site of pUC19.[20] This plasmid was used to generate a 5' ^{32}P labeled 283bp *Eco*RI/*Pvu*II restriction fragment which was used for all footprinting experiments. Developed gels were imaged using storage phosphor autoradiography using a Molecular Dynamics 400S Phosphorimager. Equilibrium association constants were determined as previously described.[19]

Reconstitution of the NCP. The NCP was reconstituted following established protocols. Briefly, 10 μg of the 146 bp DNA was 5' radiolabelled as described above using Polynucleotide Kinase and purified using a Chroma Spin STE 10 column from BD Biosciences. Histone octamer which was obtained from Karolin Luger's lab at Colorado State University was diluted to 1 μg / μL using 2 M NaCl TE Buffer. 5 μL of the DNA

was added to 5 µL of 4M NaCl, 10 mM Tris, 1 mM EDTA solution to make a sample of 10 µL of DNA in 2 M NaCl TE buffer. Five samples were made up and 0, .6, .8, 1.0, or 1.2 µL of octamer (1 µg / µL) was added to each sample. The total volume was brought to 12 µL using 2 M NaCl TE buffer. The samples were incubated for 1 hr before the addition of 12 µL, 6 µL, 6 µL, and 84 µL of dilution buffer (TE) with each addition followed by a 1 hr incubation. The sample was heated to 37°C for 2 hrs before a final addition of 120 µL of dilution buffer to give a final NaCl concentration of .1 M. Reconstituted NCP samples were stored at 4°C. Non-denaturing PAGE was done on a 6% polyacrylamide TBE gel. 1 µL of each sample was diluted to 10 µL with 10 mM Tris, 20 mM NaCl, 10% glycerol, and .1% Igapal for loading. The gel was run at 150 V for 20 min, dried and imaged.

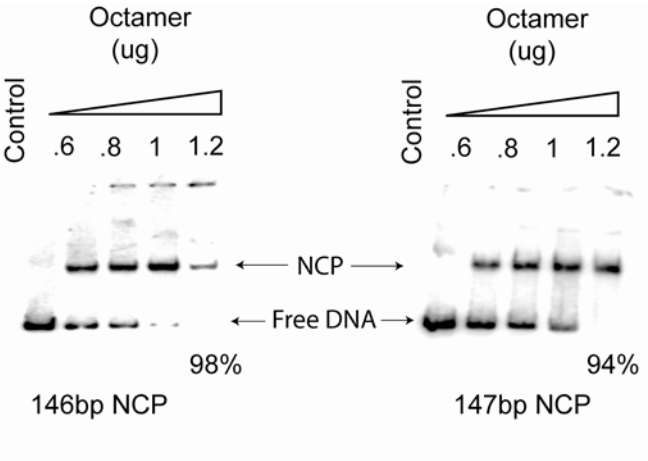


Figure 4.16 Representative gels of NCP reconstitution. The amount of histone octamer is titrated into a constant amount of the 146 or 147 bp radiolabelled DNA. Non-denaturing polyacrylamide gels are used to determine the yield from reconstitution.

Affinity Cleavage on the NCP. Affinity cleavage experiments were performed at The Scripps Research Institute under the guidance of Joel Gottesfeld. The previously published protocols were used.[4]

Melting Temperature Determinations. The melting temperature analysis was done following previously reported protocols.[21] The analysis was performed using a Varian Cary 100 spectrophotometer equipped with a thermocontrolled cell holder and possessing a cell path length of 1 cm. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 was used as the analysis buffer. The buffer conditions were designed to mimic those of DNase I footprinting experiments as closely as possible. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM. Samples were initially heated to 90 °C and cooled to a starting temperature of 25°C with a heating rate of 5 °C/min for each ramp. Two ramps were performed during each experiment and each experiment was run at least twice. Denaturation profiles were recorded at 260 nm from 25 to 90 °C with a heating rate of 0.5 °C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile.

NCP Templated Ligation Reactions. The NCP was reconstituted as described above with the sole difference that non-radiolabelled DNA was used. For each ligation reaction 200 pmol of each polyamide was lyophilized into an eppendorf tube and 40 pmol of the reconstituted NCP was added. Samples containing only the 146 bp DNA, the histone octamer, or buffer alone were prepared as negative controls. The reaction was incubated for 5 hrs at 37°C before MALDI-TOF analysis.

4.5 References

1. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides *Curr. Opin. Struct. Biol.* **2003,** 13, 284-299.

2. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., Addressing single molecules on DNA nanostructures *Angew. Chem.-Int. Edit.* **2007,** 46, 7956-7959.

3. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., Programming multiple protein patterns on a single DNA nanostructure *J. Am. Chem. Soc.* **2008,** 130, 402-403.

4. Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides *J. Mol. Biol.* **2001,** 309, 615-629.

5. Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands *J. Mol. Biol.* **2003,** 326, 371-380.

6. Edayathumangalam, R. S.; Weyermann, P.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Molecular recognition of the nucleosomal "supergroove" *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 6864-6869.

7. Herman, D. M.; Baird, E. E.; Dervan, P. B., Tandem hairpin motif for recognition in the minor groove of DNA by pyrrole - Imidazole polyamides *Chem.-Eur. J.* **1999,** 5, 975-983.

8. Kers, I.; Dervan, P. B., Search for the optimal linker in tandem hairpin polyamides *Bioorg. Med. Chem.* **2002,** 10, 3339-3349.

9. Weyermann, P.; Dervan, P. B., Recognition of ten base pairs of DNA by head-to-head hairpin dimers *J. Am. Chem. Soc.* **2002,** 124, 6872-6878.

10. Poulin-Kerstien, A. T.; Dervan, P. B., DNA-templated dimerization of hairpin polyamides *J. Am. Chem. Soc.* **2003,** 125, 15811-15821.

11. Ammann, D.; Bissig, R.; Guggi, M.; Pretsch, E.; Simon, W.; Borowitz, I. J.; Weiss, L., Preparation of neutral ionophores for alkali and alkaline-earth metal cations and their application in ion-selective membrane electrodes *Helv. Chim. Acta* **1975,** 58, 1535-1548.

12. Trauger, J. W. Ph. D. Thesis. CalTech, 1998.

13. Rucker, V. C.; Melander, C.; Dervan, P. B., Influence of beta-alanine on hairpin polyamide orientation in the DNA minor groove *Helv. Chim. Acta* **2003,** 86, 1839-1851.

14. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., Shape selective recognition of T center dot A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues *Bioorg. Med. Chem.* **2003,** 11, 4333-4340.

15. Edayathumangalam, R. S.; Weyermann, P.; Dervan, P. B.; Gottesfeld, J. M.; Luger, K., Nucleosomes in solution exist as a mixture of twist-defect states *J. Mol. Biol.* **2005,** 345, 103-114.

16. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids *J. Am. Chem. Soc.* **1996,** 118, 6141-6146.

17. Fukuda, Y.; Seto, S.; Furuta, H.; Ebisu, H.; Oomori, Y.; Terashima, S., Novel seco cyclopropa[c]pyrrolo [3,2-e]indole bisalkylators bearing a 3,3 '-arylenebisacryloyl group as a linker *J. Med. Chem.* **2001,** 44, 1396-1406.

18. Schimelpfenig, C. W., Synthesis of oxometacyclophanes with Dieckmann condensation *J. Org. Chem.* **1975,** 40, 1493-1494.

19. Trauger, J. W.; Dervan, P. B., Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes *Methods Enzymol.* **2001,** 340, 450-466.

20. Sambrook, J.; Fritsh, E. F.; Maniatis, T., *Molecular Cloning: Standard Protocols for DNA Manipulation. A laboratory Manual 2nd ed.* Cold Spring Harbor Laboratory: Plainview, NY, 1989.

21. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit *J. Am. Chem. Soc.* **2008,** 130, 6859-6866.